UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA DEPARTAMENTO DE INFORMÁTICA Y SISTEMAS



TESIS DOCTORAL

RECONSTRUCCIÓN DE OBJETOS A PARTIR DE VISTAS BIDIMENSIONALES Y SU RECONOCIMIENTO MEDIANTE MOMENTOS 3D. DESARROLLOS TEÓRICOS Y APLICACIONES

JOSÉ MANUEL GÁLVEZ LAMOLDA

Las Palmas de Gran Canaria, 1990

ersidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2004

DEPARTAMENTO DE INFORMATICA Y SISTEMAS FACULTAD DE INFORMATICA UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

TESIS DOCTORAL

RECONSTRUCCION DE OBJETOS A PARTIR DE VISTAS BIDIMENSIONALES Y SU RECONOCIMIENTO MEDIANTE MOMENTOS 3D DESARROLLOS TEORICOS Y APLICACIONES

AUTOR: José Manuel Gálvez Lamolda DIRECTOR: Manuel Cantón Garbín



Universidad de Las Palmas de Gran Canaria

DEPARTAMENTO DE FISICA

D. MANUEL CANTON GARBIN, Profesor Titular de Universidad del Area de Conocimiento de Física Aplicada de la Universidad de Las Palmas de Gran Canaria (ULPGC),

CERTIFICA:

Que D. José Manuel Gálvez Lamolda, Licenciado en Ciencias Físicas, ha efectuado bajo mi dirección en los Departamentos de Física y de Informática y Sistemas de las Universidades de La Laguna y de Las Palmas de Gran Canaria respectivamente, durante los años 1986-1990, los trabajos conducentes a la realización de su Tesis Doctoral, que presenta en esta Memoria titulada: "Reconstrucción de Objetos a partir de Vistas Bidimensionales y su Reconocimiento mediante Momentos 3D. Desarrollos Teóricos y Aplicaciones", para optar al grado de Doctor en Informática.

Las Palmas, 16 de Mayo de 1990

EL DIRECTOR DE LA TESIS

Manuel Cantón Garbín

A mi madre

y a la grata memoria de mi padre

AGRADECIMIENTOS

Este estudio ha sido realizado bajo la dirección del Profesor Dr. Manuel Cantón Garbín, al que quiero expresar mi profundo agradecimiento por haber inspirado y orientado acertadamente mi trabajo, así como por el constante apoyo y estímulo recibidos durante la elaboración de esta memoria.

A todos los compañeros de los Departamentos de Informática y Sistemas y de Física de las Universidades de Las Palmas de Gran Canaria y de La Laguna respectivamente, agradezco la ayuda que me dispensaron en aquellos momentos en que precisé de ella; de manera especial a Félix Herrera Cabello y Santiago Candela Solá.

También quiero expresar el más sincero reconocimiento a todos cuantos, personal o institucionalmente, han hecho posible que este trabajo sea una realidad. En particular debo testimoniar la valiosa colaboración prestada por el personal del Centro de Cálculo de la Universidad de La Laguna, así como las facilidades recibidas en el Departamento de Física de esta Universidad, gracias a las cuales ha sido posible realizar gran parte de este trabajo. A los miembros del grupo de Automática y Control del Departamento de Física les recuerdo con especial cariño por la hospitalidad con la que me trataron durante mi estancia allí. Gracias a todos ellos.

A la Escuela Universitaria de Informática y al Departamento de Informática y Sistemas de la Universidad de Las Palmas de G.C. agradezco el inestimable apoyo material proporcionado. Igualmente agradezco las palabras de aliento recibidas de mis compañeros del Grupo de Tenerife en la etapa final de este trabajo.

Estoy particularmente agradecido al Profesor Dr. Roberto Moreno Díaz cuya labor ha posibilitado la presentación de esta Tesis.

Finalmente deseo manifestar mi gratitud a Teresa, por su inagotable paciencia, comprensión, y apoyo que en todo momento me ha brindado. Gracias a su constante estímulo ha sido posible completar este trabajo.

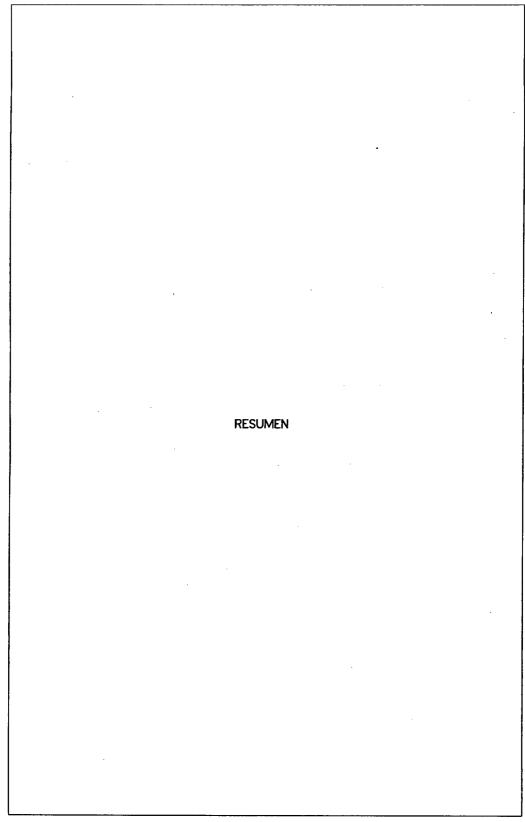
ANTECEDENTES Y OBJETIVOS

Este trabajo tiene su origen en los programas de investigación que se han desarrollado desde hace años en el Departamento de Informática y Sistemas de la Universidad de Las Palmas de Gran Canaria fundamentalmente, y en la contribución que a ellos se realizó durante cierto tiempo desde el Departamento de Física de la Universidad de La Laguna.

En el año 1980, y dentro de los programas mencionados, se comenzó una investigación sobre el significado, propiedades y aplicación de los momentos generalizados bidimensionales (2D), fundamentalmente en el reconocimiento de imágenes, obteniéndose resultados altamente satisfactorios en el tratamiento de dicho problema. Como consecuencia de este estudio, se pensó en la aplicación de momentos 3D al reconocimiento de objetos tridimensionales. Hasta el presente, los trabajos de reconocimineto de objetos 3D mediante momentos, se han realizado haciendo uso de momentos 2D calculados sobre un número considerablemente elevado de imágenes del objeto (del orden de 500), adquiridas desde diferentes posiciones. Esto se traduce en sistemas de reconocimiento poco eficientes desde un punto de vista práctico.

Con objeto de solucionar este problema, pensamos en el desarrollo de un proceso de reconocimiento basado en el uso de momentos 3D, que constaría de dos etapas: en primer lugar, había que desarrollar un método de reconstrucción de la superficie tridimensional del objeto a reconocer, f(x,y,z), a partir de un número reducido de vistas f(x,y,z) del mismo (en nuestro trabajo el valor máximo de i ha sido 10). En segundo lugar, tras la normalización de los momentos 3D frente a traslaciones, homotecias y rotaciones, se calcularían éstos sobre f(x,y,z) y se representaría a cada objeto, en el espacio de características, mediante un vector de momentos normalizados 3D. De esta forma se conseguiría una reducción drástica en el número de imágenes a adquirir y en el número de características a evaluar, lo que conduciría a un método más eficiente de reconocimiento de objetos 3D.

Este ha sido por tanto, el objetivo fundamental de este trabajo de doctorado, que se enmarca dentro de las áreas de visión artificial y reconocimiento de formas.



RESUMEN

La presente memoria está estructurada en cinco capítulos agrupados en tres bloques. En el primero de ellos, que contiene a los capítulos I y II, se presenta el campo de Visión Artificial como una de las áreas de investigación con métodos propios dentro de la Inteligencia Artificial. Se exponen los fines y métodos seguidos así como su relación con el Reconocimiento de Formas. Se revisan en este primer bloque las técnicas más representativas de recuperación de información tridimensional de una escena, proponiéndose además un procedimiento de reconstrucción de la superficie de objetos tridimensionales a partir de los contornos de múltiples imágenes bidimensionales del mismo.

El reconocimiento de formas tanto bidimensionales (2D) como tridimensionales (3D) con independencia de su posición, tamaño y orientación, es de primordial interés en Visión Artificial. Uno de los problemas básicos en la aplicación de las técnicas de Reconocimiento de Formas en el reconocimiento de objetos 3D es la definición de un conjunto apropiado de características normalizadas. En el capítulo III de esta memoria, único del segundo bloque, se definen los momentos 3D como características globales de una forma tridimensional evaluados sobre su superficie. Su utilización dentro de este trabajo puede enfocarse desde dos puntos de vista diferentes:

- Como herramienta en la normalización de formas 3D que permite obtener una versión canónica o estándar para cada objeto 3D en estudio.
- 2) Como componentes del vector de características, o descriptores empleados en una etapa posterior de reconocimiento.

En el capítulo IV, y dentro del último bloque, se presenta el desarrollo de los métodos de reconstrucción y normalización anteriormente indicados. Las aplicaciones en reconocimiento de objetos 3D se llevan a cabo sobre un banco de trece prototipos, exponiéndose los resultados más destacados en el capítulo V.

El capítulo I presenta una visión general de la disciplina de Visión Artificial haciendo énfasis en los distintos esquemas de representación que permiten cubrir la gran separación existente entre la imagen en bruto de una escena y lo que en ella se ve.

Nuestro interés en este trabajo se centra básicamente en los esquemas de representación geométrica de las formas u objetos 3D presentes en una escena. La recuperación y posterior decripción de la forma externa de un objeto a partir de la información contenida en imágenes del mismo, posibilitará la incorporación adicional, sobre esta estructura geométrica, de otro tipo de información como color, orientación local de su superficie, etc.. Tras presentar el modelo computacional de la visión, propuesto por Marr, se describen algunas técnicas de recuperación de la geometría 3D de una escena, así como sus ventajas e inconvenientes. El capítulo I finaliza con una revisión de los contenidos en Reconocimiento de Formas empleando métodos estadísticos y estructurales.

El capítulo II se ocupa del problema de reconstrucción de la estructura tridimensional de objetos a partir de múltiples imágenes. Se organiza en tres secciones. En la primera se describen los fundamentos, características y restricciones impuestas al sistema de reconstrucción propiamente dicho. En la siguiente sección se presentan las características de algunos esquemas de representación de formas 3D. Finalmente se describe el problema de triangulación y la solución aportada para pasar de un esquema de representación basado en líneas de contorno a otro esquema, también superficial, en el que los elementos básicos de descripción son superficies planas triangulares.

La técnica empleada en el sistema de reconstrucción propuesto se basa en el método de intersección de volúmenes, admitiendo que las distintas imágenes se han adquirido bajo la aproximación de proyección ortográfica. De la información contenida en cada imagen sólo se hace uso del contorno de la silueta del objeto, lo que limita el alcance del sistema al no ser posible detectar, o resolver, cierto tipo de concavidades presentes en su superficie.

Una de las características más notables del método de reconstrucción empleado es la eliminación del problema de correspondencia, común en otras técnicas también basadas en múltiples vistas. La utilización de, únicamente, los contornos hace innecesaria la localización de una misma característica en la secuencia de imágenes, problema éste que se ve agravado por la posibilidad de oclusiones debidas a otros objetos o autooclusiones ocasionadas por el movimiento relativo objeto-cámara.

El método de reconstrucción consta de dos partes bien diferenciadas:

- 1) Reconstrucción inicial a partir de los contornos de dos vistas distintas.
- 2) Refinamiento de la anterior y sucesivas reconstrucciones mediante los contornos de nuevas vistas.

El resultado de los procesos anteriores conduce a una familia de líneas de contorno que aproximan la superficie reconstruido. Dicha reconstrucción no goza de las propiedades de invarianza deseables en tareas de reconocimiento, ya que el número y orientación de las líneas de contorno es función de la posición y orientación de las dos primeras vistas empleadas en la reconstrucción inicial. Un atributo invariante y que define fielmente la geometría de un objeto es su superficie. Por este motivo se precisa transformar la anterior descripción en otra, de manera tal que la estructura 3D reconstruida constituya una aproximación acertada de la superficie real del objeto, con independencia de cuales hayan sido las dos vistas iniciales.

El paso de una descripción a otra mediante conexión de secciones contiguas a través de elementos de superficies triangulares, se conoce en la literatura como "problema de triangulación". En la última sección del capítulo II se define el problema y se muestran varias alternativas a su solución que proporcionan resultados usualmente satisfactorios en situaciones simples. No obstante, si la distancia entre secciones adyacentes es grande y si las formas de éstas son complejas, el problema presenta serias dificultades.

Las líneas de contorno de las distintas secciones se suelen representar mediante polígonos. Así los aspectos críticos en el problema de triangulación son tres:

- ¿Cómo seleccionar los contornos -uno de cada sección- a conectar mediante triángulos?.
- 2) Una vez elegidos dos contornos, ¿cómo encontrar el par inicial de vértices -uno de cada contorno- para iniciar la triangulación?.
- 3) Dado un par de vértices, ¿cómo elegir el tercer vértice que define la superficie elemental triangular?.

En un intento de eliminar estas dificultades se propone un método alternativo que, en situaciones simples, actúa igual que las técnicas existentes, pero que aporta una solución aproximada al problema de emparejamiento entre contornos de secciones contiguas cuando el número de éstos es superior a la unidad en alguna de las secciones.

El capítulo III, núcleo de este trabajo, se organiza en dos grandes apartados, siendo su objetivo mostrar la utilidad de los momentos en dos y tres dimensiones como características apropiadas para acometer tareas de reconocimiento. La primera parte se dedica al estudio de momentos 2D, mientras que los momentos 3D son tratados en la siguiente.

Los momentos 2D son una de las características globales más útiles que pueden extraerse de una imagen, pudiendo gozar de la propiedad de invarianza frente a traslaciones, cambios de tamaño e intensidad, y rotaciones de la imagen. En este primer apartado se definen los momentos 2D ordinarios así como otros tipos de momentos bidimensionales -Momentos de Legendre, Momentos de Zernike, Momentos Complejos- resumiéndose algunas de sus propiedades más destacadas y su relación con los primeros.

La versatilidad de los momentos 2D en problemas generales de manipulación y análisis de imágenes es motivo de que en la actualidad

se esté convirtiendo en un problema relativamente bien estudiado. En cualquier caso, la utilidad del método de los momentos proviene de dos factores:

- Los momentos presentan propiedades de transformación muy convenientes cuando la imagen se somete a cambios de tamaño, contraste, traslaciones y rotaciones.
- Una imagen tiene una única reconstrucción en términos de un conjunto finito de sus momentos.

Tras mostrar las ventajas de los momentos centrales de bajo orden en la localización del centroide de la imagen y la estimación de medidas acerca del tamaño y orientación, se presentan dos estrategias en la obtención de descriptores normalizados de una imagen mediante el empleo de momentos:

- Obtención de momentos invariantes, es decir combinaciones de momentos que bajo las transformaciones citadas permanezcan invariantes.
- Normalización de la imagen transformándola en su versión estándar o canónica.

En la última parte del capítulo III se definen los momentos 3D evaluados sobre la superficie de objetos tridimensionales, mostrándose la validez de estas magnitudes como herramienta en la obtención de una versión estándar, o normalizada frente a traslaciones, rotaciones y cambios de escala para cada forma tridimensional. La evaluación de los momentos 3D sobre la superficie de un objeto, y no sobre líneas de su contorno o sobre su volumen, se justifica por el hecho de que la superficie es la parte visible y describe a un objeto sin ambigüedad en la mayoria de los casos.

Esta sección comienza con la obtención analítica de las expresiones que proporcionan los valores de los momentos 3D sobre cada una de las superficies planas triangulares que aproximan, mediante un

poliedro, a la superficie reconstruida de un objeto según se indicó en el capítulo II. A continuación se establecen las relaciones que entre sí guardan los momentos ordinarios y los momentos resultantes después de aplicar transformaciones del tipo traslación, rotación y amplificación a la forma en cuestión.

Finalmente se trata el problema de normalización como paso previo al reconocimiento de objetos 3D. En términos geométricos, el procedimiento de normalización se puede reducir a encontrar las operaciones de traslación, rotación y cambio de escala que aplicadas a cada objeto lo transforman en su versión canónica. Cada versión normalizada viene caracterizada por unos valores específicos para los parámetros de posición, orientación y tamaño. El capítulo III concluye con un análisis detallado relativo a la determinación de las transformaciones anteriores, poniendo especial énfasis en la obtención de una única matriz de rotación para conseguir la orientación estándar.

La elección más frecuente para la posición estándar consiste en situar el centroide de todos los objetos en el origen de coordenadas. Para la obtención del tamaño estándar se proponen dos criterios; uno basado en la igualdad de áreas para la superficie de todos los objetos, y el otro basado en la igualdad de tamaños para el semieje mayor del elipsoide de inercia asociado a cada forma tridimensional. En el estudio de la normalización frente a rotaciones, se generaliza el método clásico de los ejes principales de 2D a 3D. En 2D esta técnica no es directamente aplicable debido a la ambigüedad existente para definir una única orientación de la forma bidimensional. En el caso de objetos 3D se demuestra la existencia de ocho orientaciones diferentes de un mismo objeto al imponer los criterios clásicos de normalización, consistentes en la diagonalización de la matriz de inercia. Como aportación original de este trabajo se propone una elimina la anterior degeneración, heurística que solución seleccionando una única orientación de entre las ocho posibles. Los resultados obtenidos al aplicar este procedimiento son de calidad suficientemente buena como para justificar su uso.

En el capítulo IV se recogen las aplicaciones realizadas en reconstrucción y normalización de formas 3D a partir de proyecciones 2D, en base a las propuestas de los capítulos II y III. El proceso de reconstrucción a partir de múltiples vistas ha sido simulado sobre un ordenador VAX 11/780. Nuestro dominio de formas 3D lo hemos limitado a objetos poliédricos de caras planas. En este sentido cabría mencionar que los mecanismos básicos en el reconocimiento de objetos simulados debieran ser similares a los involucrados en la interpretación de escenas naturales. Además, el trabajar con datos sintéticos presenta la ventaja de poder determinar el grado de actuación de cada proceso en la etapa de reconstrucción. Asimismo, permite trabajar bajo las mejores condiciones al no tener en cuenta problemas adicionales derivados de las condiciones de iluminación, características de las superficies en observación, calibración de las cámaras, ruido, etc.

En orden a recuperar información 3D a partir de proyecciones 2D es esencial el conocimiento explícito de modelos de la transformación de proyección 3D \longrightarrow 2D, así como del sistema de formación de imágenes. El capítulo IV se inicia con una descripción y comparación de las transformaciones de proyección central, paralela u ortográfica y esférica. En nuestro trabajo se ha optado por la proyeción paralela puesto que es la que más sencillez presenta desde el punto de vista matemático al elaborar los algoritmos de reconstrucción, sin que ello disminuya la potencia del esquema de recuperación de información 3D.

A continuación se describen y comentan las características más sobresalientes de los algoritmos y programas diseñados en la implementación del método de reconstrucción descrito en el capítulo II. Para ello se simulan, previamente, los procesos de adquisición de las distintas vistas de un objeto bajo proyección ortográfica, especificando la posición y orientación de las diversas cámaras. El banco de objetos empleado consta de trece formas poliédricas con distinto grado de complejidad. Para cada una de ellas se adquieren diez vistas desde posiciones y orientaciones completamente aleatorias. Las dos primeras se usan en la reconstrucción inicial para obtener una estructura de paralelogramos, que se ve mejorada al incorporar la información proporcionada por las ocho vistas restantes. Los

resultados obtenidos en términos de una familia de polígonos paralelos y equiespaciados contenidos en planos consecutivos, constatan la bondad y eficiencia del método, poniéndose también de manifiesto que en la mayoria de los casos sería suficiente la utilización de un número de vistas menor a diez para conseguir resultados análogos.

A continuación se comenta en detalle el algoritmo diseñado en el proceso de triangulación entre "rodajas", mostrándose igualmente los resultados obtenidos cuando se aplica a las reconstrucciones 3D previas. En la siguiente sección se presentan los algoritmos que llevan a cabo la normalización de las reconstrucciones ya trianguladas transformando cada una de ellas en su versión normalizada en lo relativo a posición, orientación y tamaño. En la siguiente sección se describen los programas desarrollados para la obtención de los momentos 3D sobre cada una de las reconstrucciones normalizadas, de acuerdo a lo expuesto en el capítulo III. Finalmente se recogen una serie de consideraciones prácticas acerca de cada uno de los programas generados, indicándose también las utilidades desarrolladas para la visualización del banco de objetos, sus proyecciones 2D, sus contornos y reconstrucciones en los dos esquemas de descripción.

El capítulo V recoge las aplicaciones en reconocimiento de formas 3D en función de las características extraidas, no sobre proyecciones bidimensionales, sino sobre modelos 3D reconstruidos. Con el fin de que los momentos evaluados sobre cada reconstrucción 3D permitan discriminar entre las distintas clases de objetos, y ser comparados entre sí, se extraen sobre su versión canónica. De esta forma se posibilita distinguir entre las distintas categorias de formas 3D con independencia de cual sea su tamaño, orientación y posición espacial.

En la primera parte del capítulo se presenta la estrategia de reconocimiento basada en los métodos de discriminación lineal. Después de establecer la muestra de formas 3D que constituyen el conjunto de aprendizaje, cada nueva reconstrucción 3D desconocida se asigna al grupo con el que presenta una mayor similitud. La medida de similitud se establece evaluando la distancia que separa a cada individuo desconocido del centro de todas las clases. Se describe igualmente el

procedimiento de paso a paso empleado en la selección de un conjunto reducido de momentos 3D normalizados que asegure una buena separación entre clases, y por tanto reducir las tasas de asignación errónea en la etapa de clasificación. Finalmente se exponen las características más destacadas del programa de análisis discriminante, MAHAL3, utilizado para discernir entre tres ó mas categorias de formas distintas.

En la segunda y última parte del capítulo V se analizan y comentan algunas de las pruebas más representativas en la identificación de distintas versiones trasladadas, rotadas y amplificadas, de los objetos que componen el banco de formas 3D analizado en este trabajo.

Además de los cinco capítulos previos, la memoria incluye un apartado en el que se reflejan las conclusiones y principales contribuciones de esta tesis, así como algunas posibles direcciones futuras de investigación. En los apéndices A y B se recogen, respectivamente, la obtención de la matriz de rotación que alinea dos referenciales ortonormales y las expresiones que relacionan las coordenadas de un mismo punto respecto de dos bases distintas.



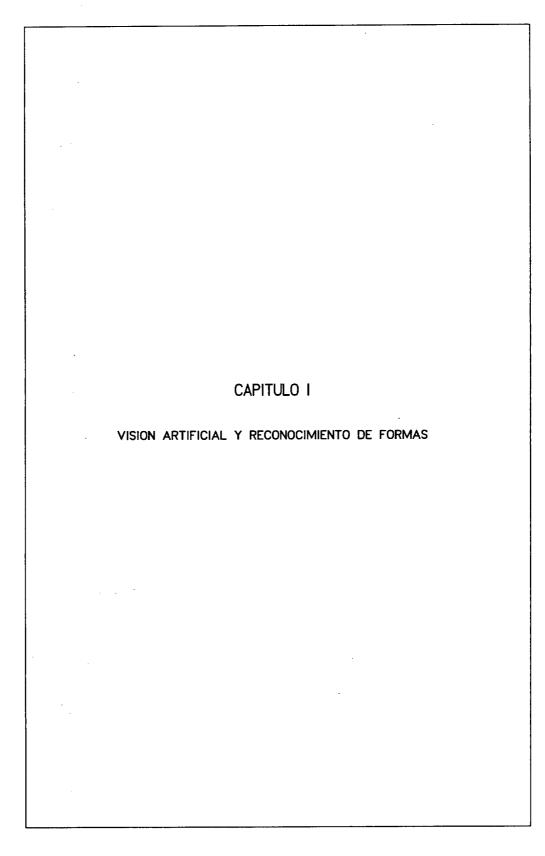
CAPITULO I: VISION ARTIFICIAL Y RECONOCIMIENTO DE FORMAS
I.1 La Naturaleza de la Inteligencia Artificial
I.1.1 Representación del Conocimiento
I.1.2 Aprendizaje6
I.1.3 Sistemas Expertos
I.1.4 Comprensión del Lenguaje Natural8
I.1.5 Desarrollo de Sistemas Robóticos8
I.2 <u>Visión Artificial</u>
I.2.1 Una Visión General y Evolución Histórica de la Visión
Artificial12
I.2.2 Relación con Otros Campos
I.2.2.1 Procesamiento de Imágenes
I.2.2.2 Clasificación de Formas16
I.2.2.3 Gráficos por Ordenador17
I.2.3 Esquemas de Representación en Visión Artificial17
I.2.3.1 Imágenes Generalizadas20
I.2.3.2 Imágenes Segmentadas21
I.2.3.3 Representaciones Geométricas22
I.2.3.4 Modelos Relactonales25
I.2.4 La Teoría de Marr en Visión Humana y Visión Artificial26
I.2.5 Sistemas y Aplicaciones
I.2.5.1 Sistemas de Propósito General28
I.2.5.2 Aplicaciones
I.2.5.3 Requerimientos de Hardware
I.3 <u>Técnicas de Visión</u> <u>3D</u> 35
I.3.1 Técnicas Basadas en Sensores Activos de Profundidad37
I.3.2 Técnicas Basadas en el Análisis de Imágenes Monoculares38
I.3.2.1 Análisis Local de Intensidades
I.3.2.2 Métodos Basados en Luz Estructurada41
I.3.3 Técnicas Basadas en el Empleo de Múltiples Imágenes42
I.3.3.1 Visión Estereoscópica42
I.3.3.2 Estereoscopía Fotométrica45
I.3.3.3 Análisis de Escenas Dinámicas48
I.4 Reconocimiento de Formas
I.4.1 Métodos Estadísticos53
I.4.1.1 Extracción y Selección de Características54
I A 1 2 Clasificación y Aprendiza io

I.4.2 Mé	todos Estructurales o Sintácticos57
CAPITULO II:	RECONSTRUCCION Y REPRESENTACION DE LA SUPERFICIE DE OBJETOS TRIDIMENSIONALES A PARTIR DE MULTIPLES
	VISTAS
II.1 <u>Introdu</u>	<u>ucción</u> 63
II.2 Algorit	mo de Reconstrucción66
II.2.1	Reconstrucción Inicial68
	II.2.1.1 Determinación de Líneas Generadoras de
	Contornos69
	II.2.1.2 Intersección de las Líneas Generadoras
	de Contornos para Dos Vistas Distintas72
II.2.2	Etapa de Refinamiento
,	Ejemplo de Reconstrucción79
II.3 Represe	<u>entación de Formas 3D</u> 83
	Características de los Distintos Esquemas
•	de Representación84
11.3.2	Descripción Superficial de las Reconstrucciones
	3D Obtenidas a Partir de Múltiples Vistas86
II.4 Descrip	oción de la Superficie de un Objeto Mediante
Triangu	<u>llación de Puntos en el Espacio 3D</u> 89
	Definición del Problema90
II.4.2	Método de Keppel93
	Método de Fuchs, Kedem y Uselton95
	Método de Christiansen y Sederberg96
	Método Propuesto100
CAPITULO III	: CARACTERIZACION DE FORMAS TRIDIMENSIONALES MEDIANTE MOMENTOS GENERALIZADOS 3D
III.1 Introd	l <u>ucción</u> 109
III.2 Moment	os <u>Bidimensionales</u> 111
III. 2.	1 Definición y Teorema de Unicidad111
III.2.	2 Momentos Centrales. Propiedades de los Momentos de
	Bajo Orden113
III.2.	3 Momentos de Legendre115
III. 2.	4 Momentos de Zernike
III.2.	5 Momentos Complejos
III. 2.	6 Transformaciones de Momentos Ordinarios 2D120

III.2.6.1 Transformación Frente a Traslaciones120
III.2.6.2 Transformación Frente a Rotaciones121
III.2.6.3 Transfromación Frente a Amplificaciones122
III.2.7 Normalización de Formas Bidimensionales Basada en
Momentos 2D123
III.2.7.1 Momentos Invariantes125
III.2.7.2 Normalización de la Imagen
III.2.7.3 Normalización de Momentos
III.3 Momentos Tridimensionales
III.3.1 Momentos 3D Evaluados Mediante Integrales de
Superficie137
III.3.1.1 Evaluación de Momentos Ordinarios Sobre
una Superficie Triangular139
III.3.2 Transformaciones de Momentos Ordinarios144
III.3.2.1 Transformación Frente a Traslaciones144
III.3.2.2 Transformación Frente a Amplificaciones146
III.3.2.3 Transformación Frente a Rotaciones149
III.3.3 Normalización de Formas Tridimensionales Mediante
Momentos 3D153
III.3.3.2 Estándar de Orientación
III.3.3.2.1 Elipsoide de Inercia. Ejes
Principales de Inercia157
III.3.3.2.2 Orientación de Formas 3D162
III.3.3.3 Estándar de Tamaño
CAPITULO IV: DESARROLLO ALGORITMICO DE LOS PROCESOS DE RECONS- TRUCCION Y OBTENCION DE MOMENTOS NORMALIZADOS
IV.1 Proyección Paralela. Comparación con Otras Transformaciones
<u>de Proyección</u> 175
IV.2 <u>Construcción de Modelos Para Objetos 3D</u>
IV.2.1 Modelado Geométrico182
IV.2.1.1 Información Topológica
IV.2.1.2 Información Geométrica
IV.3 <u>Desarrollo del Proceso de Reconstrucción</u> <u>3D</u>
IV.3.1 Proceso de Adquisición de Vistas de Objetos 3D187
IV.3.2 Extracción de Contornos
IV.3.2.1 Algoritmo de Extracción de Contornos193

	IV. 3. 3 Rasterización de las Dos Primeras Vistas
	IV.3.3.1 Construcción de las Líneas de Rasterización198
	IV.3.3.2 Obtención de los Segmentos de Rasterización200
	IV.3.3.3 Criterios de Interioridad200
	IV.3.3.4 Algoritmo de Rasterización203
	IV.3.4 Obtención de la Estructura de Paralelogramos Asociada
	a la Reconstrucción 3D Inicial205
	IV.3.5 Refinamiento de la Reconstrucción 3D208
	IV.3.5.1 Procedimiento Refinamiento210
	IV.3.5.2 Intersección Entre Dos Polígonos213
	IV.3.5.2.1 Características del Algoritmo
	Desarrollado215
	IV. 3.5.2.1.1 Conexiones Entre Segmen-
	tos de la Intersección215
	IV.3.5.2.1.2 Especificación de los Po-
	lígonos Resultantes de la
	Intersección216
	IV.3.5.2.2 Algoritmo INTER_POLIGONOS220
IV. 4	Etapa de Normalización225
	IV. 4.1 Triangulación Entre Rodajas Consecutivas226
	IV.4.1.1 Características del Proceso de Triangulación228
	IV. 4. 1. 2 Procedimiento Triangular
	IV. 4.1.3 Triangulación Entre Dos Polígonos232
	IV. 4. 1. 3. 1 Elección del Segmento Inicial de Co-
	nexión Entre Polígonos233
	IV. 4.1.3.2 Generación de Triángulos234
	IV. 4. 1. 3. 3 Algoritmo de Triangulación 234
	IV.4.2 Obtención de la Versión Normalizada para la
	Reconstrucción 3D236
	IV. 4. 2. 1 Características del Proceso de Normalización 237
	IV. 4.2.2 Algoritmo STANDARD239
IV. 5	Obtención de Características: Momentos 3D Normalizados243
	IV.5.1 Procedimiento MOMENTOS244
	IV.5.2 Subrutina MOMENTOS_OR245
IV.6	Consideraciones Prácticas en la Etapas de Reconstrucción,
	Normalización y Obtención de Características247
	IV.6.1 Programa PROYECCION247

IV.6.2 Programa CONTORNO	. 248
IV.6.3 Programas RASTERIZAR y PARALELOGRAMOS	
IV.6.4 Programa REFINAMIENTO	. 250
IV.6.5 Programa TRIANGULAR	
IV.6.6 Programa STANDARD	. 252
IV.6.7 Programa MOMENTOS	. 260
IV.6.8 Material y Utilidades	. 261
CAPITULO V: IDENTIFICACION DE OBJETOS 3D. APLICACIONES	. 263
CAPITOLO V: IDENTIFICACION DE OBSETOS OB. AL DIONOTONAS	
V.1 Introducción	
V.2 <u>Discriminación:</u> <u>Definición</u>	. 266
V.2.1 Discriminación Desde un Punto de Vista Descriptivo	. 266
V.2.2 Identificación	. 266
V.3 <u>Método</u> <u>de</u> <u>Asignación</u>	. 267
V.3.1 Notación	. 267
V.3.2 Definición del Criterio	. 268
V.4 Procedimientos de Paso a Paso	. 270
V.4.1 Criterio de Selección de Variables	
V.4.2 Maximizar TRAZA (T ⁻¹ B)	. 272
V. 4.3 Caso de Dos Grupos	. 274
V.4.4 Resumen del Criterio MAX TRAZA (T ₀ ⁻¹ B ₀)	. 275
V.5 <u>Descripción</u> <u>de</u> <u>MAHAL3</u>	. 276
V.5.1 Muestra de Aprendizaje	. 277
V.5.2 Muestra de Prueba	. 277
V.5.3 Individuos Anónimos	
V.5.4 Desarrollo del Programa	. 278
V.6 Consideraciones Prácticas en la Identifiación de Formas 3D	. 279
V.6.1 Pruebas Realizadas y Discusión de los Resultados Obtenidos	. 281
V.6.1.1 Prueba 1	
V.6.1.2 Prueba 2	287
V.6.1.3 Prueba 3	290
V.6.1.4 Prueba 4	295
CONCLUSIONES Y PRINCIPALES APORTACIONES	299
APENDICE A	
APENDICE B	
REFERENCIAS	



1.1 LA NATURALEZA DE LA INTELIGENCIA ARTIFICIAL

La inteligencia artificial (I.A.) es una disciplina interesada en el uso de los computadores en tareas que requieren normalmente del conocimiento, percepción, razonamiento, aprendizaje, comprensión y facultades cognitivas similares [DUDA-79a]. Así, la meta en I.A. consiste en una expansión de las capacidades de un ordenador, es decir el desarrollo conceptual de nuevas aplicaciones del ordenador en tareas que, hasta fechas recientes, sólo habían sido desarrolladas por seres inteligentes. Existen muchas formas de definir el campo de la I.A. [CHAR-85], [STEV-85]; la propuesta por Patric Henry Winston [WINS-84] es la siguiente:

"La inteligencia artificial contempla el estudio de las ideas que facilitan a los computadores el ser inteligentes".

Actuar inteligentemente -pensar- significa algo más que representar, calcular y combinar letras y números. Pero ¿qué es la inteligencia?

¿Es la habilidad de razonar?

¿Es la habilidad de adquirir y aplicar el conocimiento?

¿Es la facultad de percibir y manipular cosas en el mundo físico? ¿Es la facultad de combinar reglas de actuación e incluso la intuición o sentido común para proporcionar respuestas apropiadas frente a algún problema?

Probablemente todas estas habilidades son parte de lo que es la inteligencia, pero no lo son todo. Es por ello que una definición en el sentido usual se presente difícil. Sin embargo las investigaciones en I.A., generalmente, están motivadas por alguno de los dos objetivos siguientes:

- 1. Comprender los principios que hacen posible la inteligencia.
- 2. Convertir los computadores en una herramienta cada vez más útil.

El primer objetivo, el más frecuente en los centros de investigación básica en I.A., explica el vínculo de unión entre la I.A. y la Fisiología, Psicología y Filosofía. El segundo objetivo

conduce a una aproximación más práctica, y por tanto más relacionada con la tecnología del software y del hardware, enlazando con los propósitos y necesidades de la Informática e Ingeniería. No obstante estos dos puntos de vista son complementarios y los avances en un sentido estimulan y refuerzan las investigaciones en el otro.

En sus comienzos, a finales de la década de los 50, la I.A. hizo uso y contribuyó a otras disciplinas afines interesadas en métodos avanzados de procesamiento de información. Así, inicialmente estuvo ligada a problemas de lógica matemática, investigación operativa, teoría de la decisión, teoría de la información, reconocimiento de formas y lingüística matemática. Sin embargo, lo que distingue a la I.A. de estas áreas relacionadas es su interés por todos los mecanismos involucrados con la inteligencia.

Los problemas más interesantes para la I.A surgen en un intento de duplicar las facultades mentales de gente ordinaria, tales como visión y comprensión del lenguaje natural. Esto está en contraposición con hechos tales como que alguna gente considera tareas elementales el ver y hablar, mientras que por el contrario la multiplicación de números de varios dígitos son tareas difíciles de realizar mentalmente. Desde su aparición, el uso de los computadores puso de manifiesto cuán distintas y distantes son las tareas de ver y multiplicar números, por ejemplo. Así, se ha puesto de manifiesto que para un computador, la simple acción de ver requiere millones de multiplicaciones. El hecho de que nosotros podamos efectuar la tarea más complicada -ver- con menos esfuerzo que la más fácil -multiplicar- es meramente un accidente evolutivo [CHAR-85].

El objetivo último perseguido por los investigadores en el campo de la I.A. es construir computadores que actuen de manera análoga al comportamiento inteligente de los seres humanos. Esto presenta serias implicaciones en torno a los seres humanos y su relación con el mundo que los rodea. Este objetivo tan extenso no es alcanzable en un futuro inmediato, o quizás no lo sea nunca. Mas bien los objetivos más realistas van dirigidos, en la actualidad, al desarrollo y creación de máquinas que cumplan propósitos específicos y bien definidos.

Los primeros intentos sistemáticos de resolver problemas en I.A., y más concretamente en la comprensión del lenguaje natural y análisis de escenas visuales, pueden asociarse a la creación de los grupos de investigación en el M.I.T (Massachusetts Institute of Technology) y en la Universidad de Stanford en torno a 1963.

Entre las áreas de interés dentro del, cada vez más amplio, campo de la I.A. podemos citar las siguientes:

- Representación del conocimiento.
- Aprendizaje.
- Desarrollo de sistemas expertos.
- Comprensión del lenguaje natural.
- Desarrollo de sistema robóticos.
- Visión artificial.

1.1.1 REPRESENTACION DEL CONOCIMIENTO

¿Cómo debiera ser adquirido, representado y almacenado el conocimiento de manera que pueda usarse de manera óptima por un computador?.

En contraste con las estructuras de almacenamiento de datos convencionales, los sistemas desarrollados en I.A. requieren una base de conocimientos que contenga información de varios tipos, entre los que se incluyen:

- conocimiento acerca de objetos.
- conocimiento acerca de procesos.
- conocimiento del entorno.
- conocimiento difícil de representar: sentido común, motivaciones, casualidad,...

Un esquema de representación del conocimiento puede pensarse como una combinación de estructuras de datos y procedimientos de interpretación, que si son utilizados correctamente por un programa,

éste puede exhibir un comportamiento inteligente. Entre las cuestiones a plantearse cabe mencionar las siguientes [LOPE-84]:

- ¿Qué tipo de conocimiento es necesario para mostrar un comportamiento inteligente?.
- ¿Qué hechos deben ser conocidos en un momento dado?.
- ¿Cómo codificar las reglas que manipulan una base de conocimientos, con conocimiento puramente explícito para inferir nuevos conocimientos?.
- ¿Cómo tratar los casos de conocimiento dudoso o impreciso?.
- ¿Cómo adquirir nuevos conocimientos y actualizarlos?.

Los sistemas existentes actualmente están basados en representaciones escogidas específicamente para una determinada aplicación, ignorando a menudo sus limitaciones así como las consecuencias que puedan aparecer en etapas posteriores de su desarrollo, debido a decisiones tomadas inicialmente.

1.1.2 APRENDIZAJE

¿Es posible desarrollar programas cuya efectividad en la resolución de problemas se vea automáticamente aumentada a medida que se ejecutan?. Los diseñadores de máquinas se enfrentan continuamente ante la disyuntiva de crear máquinas aplicables a un único problema o máquinas provistas de gigantescas bases de datos con información para atacar distintos tipos de tareas. La máquina ideal sería aquella que estuviera provista de una mínima cantidad de información preinstalada de manera que la información restante fuera asimilada durante la ejecución de distintos cometidos.

1.1.3 SISTEMAS EXPERTOS

Un sistema experto es una máquina que incorpora, en forma

operativa, el conocimiento de una persona experimentada en un dominio particular, de forma que piensa y razona como lo haría un experto en el tema, dando consejos o tomando decisiones inteligentes. Así, cuando diagnostica un caso real no sólo le atribuye los calificativos que considera adecuados, sino que justifica las razones que le llevaron a ello. Además la estructura de estos sistemas permite, como en el caso de las personas, modificar su conocimiento, corrigiendo sus errores o ampliándolos. El estilo de programación que se adopta para conseguir estas características es la programación basada en reglas [FERN-87].

aparición de los sistemas expertos precursores la [DUDA-79a] MACSYMA (MIT) y DENDRAL (Stanford) entre 1965-1970 en las áreas respectivas de Matemática Aplicada (Integración Simbólica) y Química Orgánica (Espectrometría de Masas), el campo de aplicación de este tipo de sistemas se ha extendido considerablemente. Así, en 1976 apareció MYCIN, sistema de asistencia médica en el diagnóstico de infecciones bacterianas y sugerencia de terapias; otro ejemplo lo constituye XCON, sistema experto desarrollado por McDermott en 1981 para la ayuda en la configuración de los ordenadores VAX. Otras áreas de aplicación incluyen la ayuda en prospecciones petrolíferas y predicción de yacimientos minerales, diagnóstico de fallos en máquinas, ayuda en planificación financiera, etc..

En la actualidad los sistemas expertos están aún restringidos a áreas de conocimiento bien limitadas y requieren, para su construcción, una laboriosa interacción con especialistas en la materia y un largo proceso de depuración. Algunos de estos sistemas proporcionan, todavía, soluciones parcialmente satisfactorias en un porcentaje de los casos, siendo incluso incapaces de dar respuesta en otros casos.

El gran auge en el desarrollo de sistemas expertos [CHAR-85] debido a las enormes implicaciones comerciales y económicas que conllevan, ha contribuido paralelamente a un gran avance en las técnicas de adquisición y representación del conocimiento.

1.1.4 COMPRENSION DEL LENGUAJE NATURAL

Para muchos usuarios de sistemas informáticos, la comunicación con sus máquinas resulta, a veces, una experiencia frustrante. Este problema es motivo de que una amplia comunidad de posibles usuarios evite su uso. Los trabajos en inteligencia artificial dentro del área de comprensión del lenguaje natural van dirigidos a lograr no sólo la comprensión automática de subconjuntos de palabras específicas en el contexto de determinados lenguajes de programación, sino de todas las palabras de un lenguaje y de las frases usadas en la comunicación entre humanos. Cuestiones importantes a tener en cuenta son las siguientes:

¿Qué significa el decir que algo se ha comprendido?.

¿Qué partes del discurso son lo suficientemente ricas en contenido de ideas clave?.

¿Qué tipo de representaciones son las adecuadas (acústica, morfológica-sintáctica, semántica, etc.)?

¿Qué cálculos pueden efectuarse sobre cada representación?.

¿Cómo resolver el problema de las ambigüedades presentes en situaciones en las que el orador considera que su interlocutor es conocedor de más información de la que realmente se contiene en la frase?.

1.1.5 DESARROLLO DE SISTEMAS ROBOTICOS

Un área interdisciplinar donde se conjugan distintas técnicas propias de I.A. es el desarrollo de sistemas robóticos. Una definición de Robótica, atribuida a Allen Newell [FAIR-88], es la que se indica a continuación:

"La Robótica es la disciplina en donde se combinan la inteligencia y la acción; es decir, se encarga del control inteligente del movimiento coordinado"

Aunque esta definición presenta deficiencias por sí misma y, en particular, se puede argumentar que esta perspectiva descarta muchos de los sistemas actuales en funcionamiento -específicamante los encargados de ejecutar tareas o acciones repetitivas y previamante definidas- parece reflejar la gran amplitud de este campo aplicado.

En un sentido amplio, es posible considerar un sistema robótico estructurado en tres componentes interrelacionados (Fig. 1.1)

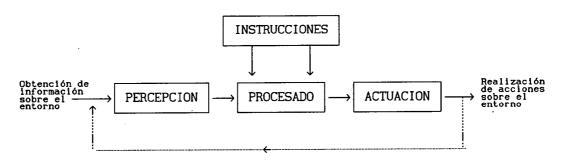


Fig. 1.1. Esquema conceptual de los componentes principales de un sistema robótico basado en información sensorial.

La parte esencial del sistema es el procesado de la información procedente del entorno y que es captada por sensores apropiados. El procesado, en el caso ideal, debiera llevarse a cabo mediante programas inteligentes que, en función del conjunto de instrucciones recibidas, controlasen las acciones que el sistema debe ejecutar.

Cada vez es más frecuente la incorporación de capacidad de procesamiento de información visual y táctil en sistemas robóticos. Además de lo anterior, estos sistemas están equipados con sensores de posición para determinar la posición relativa entre las partes del sistema y el manipulador.

En la actualidad los sistemas robóticos más extendidos son aquellos especializados en la ejecución de tareas muy concretas en entornos fuertemente restringidos. Las aplicaciones más frecuentes caen dentro de alguna de las siguientes categorias: producción, exploración y asistencia. La tendencia actual en este campo se enfoca, fundamentalmente, a dotar de más "inteligencia" a este tipo de sistemas, lo que permitirá una más eficaz interacción con su entorno y una mayor autonomía de funcionamiento.

La ventaja de un sistema robótico frente a una máquina convencional automática es que el primero puede efectuar muchas tareas mientras que la actuación del segundo sistema está limitada a unas pocas funciones. Esto se consigue debido a la disponibilidad del robot para ser programado en la ejecución de cometidos distintos. Desde esta óptica, la tendencia actual es hacia la utilización de un único lenguaje de programación [FERR-86].

1.2 VISION ARTIFICIAL

El campo de percepción artificial intenta dotar a las máquinas manufacturadas por el hombre de la capacidad de percibir su entorno por medios sensoriales similares a los empleados por humanos y animales. Los sistemas biológicos perciben su entorno a través de una variedad de fuentes tales como la luz, el sonido, el tacto, el gusto y el olfato. El proceso de percepción implica la construcción de modelos útiles del entorno a partir de una gran cantidad de datos sensoriales. Desde sus comienzos, uno de los objetivos de la inteligencia artificial ha sido equipar a las máquinas con capacidad para poder tratar datos sensoriales.

Los datos visuales son, probablemente, la más compleja y a la vez más útil información sensorial para los humanos. La visión artificial está interesada en la interpretación de datos visuales por medio de máquinas. Tales máquinas, capaces de percibir su entorno y ejecutar tareas, presentan, obviamente, una gran utilidad en diversas áreas de aplicación tales como inspección y ensamblaje industrial, exploración del espacio planetario, apoyo en diagnóstico médico, teledetección de recursos terrestres mediante sensores remotos, etc.. Podrían, además, asistir en tareas rutinarias, tediosas e incluso peligrosas para los

humanos pero que sin embargo son difícilmente automatizables sin alguna capacidad perceptiva.

La disciplina conocida como visión artificial es parte del más extenso campo de la inteligencia artifical cuyo fin último es la construcción de máquinas con comportamiento inteligente, tanto en el dominio perceptivo como en otros. La visión artificial se encarga de la construcción de descripciones explícitas, y con significado, de objetos físicos a partir de sus imágenes. La interpretación de (Image Understanding) es un campo diferenciado imágenes procesamiento de imágenes, el cual estudia transformaciones del tipo $imagen \longrightarrow imagen$, sin generar descripciones explícitas de las mismas.Las descripciones son prerequisitos necesarios para el reconocimiento, manipulación y emisión de juicios en torno a objetos. Para establecer una distinción con el procesamiento de imágenes, el "interpretación de imágenes" es usado como un sinónimo de visión artificial [NEVA-82].

El sistema visual juega un papel destacado en la inteligencia humana por lo que la comprensión y dotación de capacidad visual en máquinas ayudará al desarrollo de otras formas de comportamiento inteligente en estos sistemas. La percepción visual es tan inmediata, tan cotidiana y comporta tan poco esfuerzo para nosotros, que un no iniciado imaginaría libre de dificultad la tarea de automatizar este proceso. ¿No consiste el proceso de reconocer un objeto en una imagen en una simple tarea de encontrar un patrón particular de intensidades luminosas previamente visto?. Un análisis más detallado nos revela numerosas complejidades y dificultades en la percepción del mundo tridimensional a través de sus proyecciones bidimensionales sobre alguna superficie sensible. Nosotros percibimos el mundo de objetos 3D con muchas propiedades invariantes. Objetivamente, los datos visuales de entrada no presentan la correspondiente coherencia o invarianza; contienen mucha información irrelevante o incluso engañosa. De algún nuestro sistema visual, desde la retina hasta los niveles cognitivos, comprende o impone orden sobre las entradas visuales. Esto lo lleva a cabo empleando información intrínseca extraida fiablemente de los datos de entrada, y también a través de consideraciones y

conocimiento aplicados en los diferentes niveles de procesamiento visual.

En el análisis de imágenes complejas, e incluso poco familiares, necesitamos extraer elementos de información sutiles y complejos y hacer un uso eficiente de nuestro almacen de conocimientos relativos a las estructuras y restricciones sobre los objetos de nuestro mundo perceptivo. Cuestiones importantes a tener en cuenta son:

¿Qué información, relativa a las escenas, puede extraerse de las imágenes 2D usando sólo consideraciones básicas en torno a la física y óptica?.

¿Qué cálculos deben efectuarse explícitamente?.

En el proceso de interpretación de imágenes, ¿a qué etapa ha de incorporarse el conocimiento previo referente a la escena en observación?.

¿De qué forma se han de construir los modelos del mundo externo y la representación del conocimiento?.

El campo de visión artificial es relativamente nuevo y en rápido crecimiento. Los primeros pasos se produjeron en torno a la década de los 50, siendo desarrollados muchos de los conceptos a partir de mediados de la década de los 70. Muchas de las ideas básicas surgieron en áreas tan dispares como I.A., Psicología, gráficos por ordenador, procesado de imágenes, etc.

1.2.1 UNA VISION GENERAL Y EVOLUCION HISTORICA DE LA VISION ARTIFICIAL

Dos grandes aproximaciones a la percepción de formas pueden distinguirse:

- 1. Reconocimiento de formas.
- 2. Análisis de escenas.

La primera se ocupa del problema de asignar una forma, no necesariamente visual, a una de las clases de formas previamente conocidas, basándose para ello en ciertas medidas conocidas como características, calculadas sobre la forma incógnita. Los estudios en esta aproximación se han centrado fundamentalmente en los métodos de clasificación más que en la definición de características apropiadas para llevarla a cabo.

En análisis de escenas se pretende describir una forma en términos de primitivas más simples extraidas de la imagen, efectuándose el reconocimiento mediante análisis de las citadas descripciones. Aquí el énfasis se hace en la selección de las primitivas y sus descripciones, siendo de utilidad el conocimiento de restricciones impuestas en el proceso de formación de imágenes. Las técnicas desarrolladas, a menudo tienden a seguir el comportamiento humano, aunque no es ésta la meta esencial.

El campo de análisis de escenas 3D se inició en los comienzos de la década de los 60 con el trabajo pionero de L.G. Roberts [ROBE-65] sobre reconocimiento de objetos poliédricos simples tridimensionales a partir de una única imagen. Su sistema extraía trazos rectilíneos y sus descripciones a partir de imágenes de TV digitalizadas. El reconocimiento de los objetos presentes en la escena se obtenía contrastando tales descripciones con las proyecciones de objetos conocidos. Podría decirse también que el cálculo de tales proyecciones origen campo de gráficos del 3D por ordenador (Three-Dimensional Computer Graphics).

En los años siguientes, y por razones de sencillez, la atención quedó centrada básicamente en escenas compuestas de objetos poliédricos. En esta etapa el problema central fué el de segmentación, es decir la determinación de las líneas y superficies pertenecientes a objetos simples. En la actualidad este problema está bien comprendido, existiendo técnicas para segmentar escenas poliédricas complejas bajo la consideración de que es posible extraer perfectas descripciones mediante segmentos rectilíneos en cada escena.

Al tratar con imágenes reales, esta consideración última se satisface en pocas ocasiones; consecuentemente se desarrollaron técnicas para trabajar con datos incompletos o defectuosos. En una etapa posterior se diseñaron procesos a distinto nivel, obteniéndose resultados que permiten concluir que en ocasiones el procesado a alto nivel guía el procesamiento a otros niveles inferiores, existiendo una estructura de control retroalimentado entre niveles. Se trata así de procesos con una estructura no jerarquizada.

Intentos por extender las técnicas desarrolladas al trabajar en el dominio de escenas poliédricas a escenas más complejas incluyendo objetos curvados condujeron al desarrollo de nuevos métodos de descripción de formas. La experiencia con escenas más complejas del mundo real indica que las relaciones contextuales entre diferentes objetos son extremadamente útiles.

La segmentación de escenas reales compuestas de objetos 3D, posiblemente ocultos por otros objetos, a partir de los niveles de gris de una imagen presenta serias dificultades. Esto ha motivado que algunos grupos de investigación desarrollaran sistemas telemétricos para obtener directamente la geometría de los objetos presentes en una escena. Sobre la escena se proyecta un plano de luz y simultáneamente se observa con una cámara de televisión. Los primeros trabajos con esta técnica en el reconocimiento de objetos poliédricos desarrollaron en 1972 [SHIR-72]. Posteriormente el tipo de objetos en la escena fué ampliado para incluir objetos con superfices curvadas. Debido al gran esfuerzo requerido en el desarrollo de las técnicas telemétricas, no ha sido sino hasta fechas recientes cuando se ha producido un gran avance en el procesamiento de este tipo de datos. Actualmente, y a nivel comercial, es posible encontrar este tipo de disposiciones y técnicas de procesamiento de datos telemétricos en robots industriales [SHIR-87].

Paralelamente al trabajo en análisis de escenas se desarrolló el campo de análisis de imagen. Esta línea se interesa primordialmente de las técnicas de extracción de descripciones necesarias en los métodos de análisis de escenas a nivel superior, a partir de las imágenes.

Dichas técnicas incluyen, entre otras, el cálculo de brillo y color percibidos, la recuperación parcial o completa de datos tridimensionales de la escena, localización de discontinuidades asociadas a objetos en la escena y caracterización de las propiedades de regiones uniformes dentro de la imagen.

Las capacidades de las máquinas para ver están aún lejos de las encontradas en los humanos, quedando todavía por resolver muchas facetas del procesado y representación. No obstante las técnicas disponibles permiten el diseño de sistemas con posibilidades limitadas, aunque útiles, para ver en aplicaciones de problemas reales. En la construcción de tales sistemas de propósito especial para aplicaciones particulares es frecuente hacer uso del conocimiento previo del entorno de trabajo para simplificar las tares de procesamiento. La organización y utilización de este conocimiento es esencial en este tipo de sistemas.

1.2.2 RELACION CON OTROS CAMPOS

Existen diversos campos que comparten técnicas comunes de procesamiento de datos visuales. Por ejemplo, el procesamiento de imágenes, la clasificación de formas, gráficos por ordenador y la interpretación de imágenes. De entre estas áreas, la última tiene un significado similar al de visión artificial.

1.2.2.1 PROCESAMIENTO DE IMAGENES

El propósito en procesamiento de imágenes es mejorar la calidad de las imágenes para, posteriormente, ser analizadas por observadores humanos (Fig. 1.2). Por este motivo los detalles del procesamiento deben ser especificados por los propios observadores. En este sentido, cabría decir que el sistema de procesamiento no está interesado con lo que la imagen representa. Algunas de las técnicas de procesamiento de imágenes se utilizan en el diseño de los módulos de preproceso en los sistemas de visión artificial.

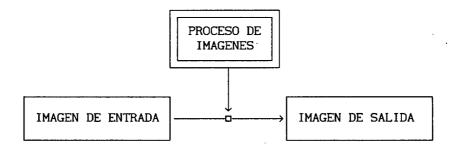


Fig. 1.2. Esquema representativo del procesamiento de imágenes.

1.2.2.2 CLASIFICACION DE FORMAS

En clasificación de formas el objetivo perseguido es clasificar una forma dada, no necesariamente visual, dentro de alguna de las categorias previamente determinadas. Un ejemplo típico es el reconocimiento de caracteres. La clasificación de formas se puede entender como un caso especial de la visión artifical en el sentido de que su resultado último es el nombre o categoría de un patrón de formas (Fig. 1.3).

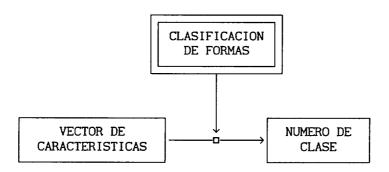


Fig. 1.3. Esquema representativo de la clasificación de formas.

Las formas se clasifican, usualmente, de acuerdo a un conjunto de características extraidas de la imagen de entrada. El conjunto o

compara con de características se las características representativas, y previamente determinadas, de cada categoría. Dicha comparación va acompañada de una decisión en atención a algún criterio. Puesto que el propósito es discriminar entre varias categorias, el sistema de clasificación no está interesado con lo que representa la imagen. La selección de algún conjunto apropiado de cracterísticas puede llevarse a cabo en base a criterios estadísticos. situaciones prácticas, la igualdad entre vectores características para las formas incógnita y patrón, deberá sustituirse por criterios de similitud, y es por ello que se hace uso de teorías de decisión estadística.

1.2.2.3 GRAFICOS POR ORDENADOR

La generación de gráficos mediante un computador es justamente el objetivo contrario al de visión artifical. Partiendo de la lista de coordenadas de un objeto 3D se genera una representación 2D del mismo.

Este campo abarca una gama de aplicaciones en rápida expansión. A menudo, los datos de entrada son descripciones de imágenes. Estas son editadas por un operador humano para emplearlas en la generación de nuevas imágenes. Estas técnicas son importantes en el diseño asistido por ordenador (CAD) para la visualización de los objetos y piezas diseñadas por el usuario. Otro campo típico de aplicación es la generación de imágenes dinámicas y con un elevado grado de realismo en simuladores de vuelo. En análisis de típicamente escenas tridimensionales un problema de interés considerable representación mediante modelos geométricos 3D de las superficie de los objetos presentes en la escena.

1.2.3 ESQUEMAS DE REPRESENTACION EN VISION ARTIFICIAL

La visión artificial tiene por objetivo la automatización e integración de un amplio rango de procesos y representaciones usadas en la percepción de datos visuales. Esta disciplina incluye técnicas

propias de otros campos como procesamiento de imágenes, clasificación de formas, modelado geométrico, etc. (Fig. 1.4)

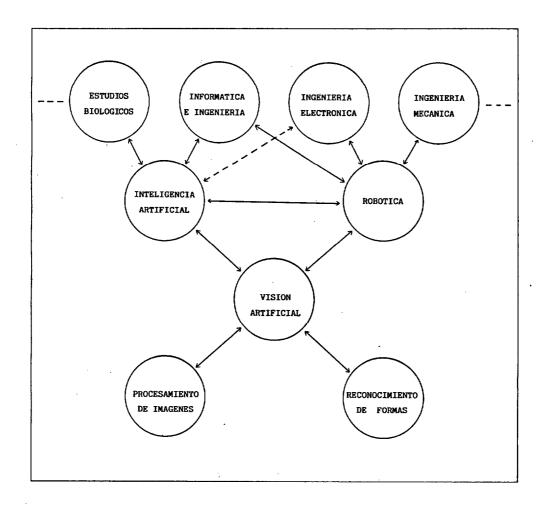


Fig. 1.4. Diagrama de bloques mostrando la relación de Visión Artificial con otras áreas.

En la visión se conjugan procesos cognitivos, modelos geométricos, objetivos a cumplir y planes a ejecutar, siendo estos procesos de alto nivel. Además, la visión requiere de muchas capacidades a bajo nivel en las que no reparamos; así por ejemplo, cuando detectamos los cambios de intensidad definitorios de bordes en una imagen no somos conscientes de la evaluación a bajo nivel de

operadores del tipo "cruces por cero" [GRIM-81]. Por ejemplo la fusión estéreo de las imágenes de cada ojo en visión estereoscópica es una tarea ejecutada a bajo nivel para ordenar por profundidades la percepción tridimensional del mundo. Otra importante tarea a bajo nivel es la percepción de objetos. Con independencia de si esta facultad es innata o aprendida, el hecho es que los sistemas maduros de visión biológica están especializados y acostumbrados a tratar con objetos en su propio entorno. Especializaciones adicionales pueden ser aprendidas. Una capacidad básica en el reconocimiento de objetos es la facultad de discernir o separar objetos del fondo de una escena.

Las capacidades de procesamiento a bajo nivel son difíciles de detectar, se realizan de forma inconsciente y no están bien conectadas a otros sistemas que permiten una introspección directa; por ejemplo, nuestra memoria visual para las imágenes es impresionante y sin embargo nuestras descripciones verbales de ellas son relativamente primitivas. El "hardware" de nuestro sistema biológico visual ha sido desarrollado, perfilado y especializado sobre un largo periodo de tiempo. No obstante, su organización y funcionamiento no está bien comprendido excepto en ciertos niveles y detalles -el comportamiento de pequeños grupos de células corticales en gatos y monos, o el comportamiento de seres humanos en experimentos psicofísicos-.

De esta forma, la visión artificial se enfrenta con un problema difícil; debe idear o redescubrir, ayudándose de técnicas digitales, los más básicos y aún inaccesibles mecanismos de especialización y paralelismo encontrados en el sistema visual biológico. La información de partida con la que ha de trabajar un computador es una representación basada en un array de números, a partir de los cuales un ser humano podría extraer información visual escasa y con un gran esfuerzo [BALL-82].

La visión artificial está interesada en los procesos a bajo y alto nivel así como en la incorporación del conocimiento que en torno a una escena se posea. La percepción visual se ocupa así de la relación entre los datos visuales de entrada y los modelos previamente creados del mundo externo. Existe una gran separación entre una imagen

en bruto y lo que en ella se ve. Este salto no puede cubrirse de un sólo paso, siendo preciso disponer de representaciones intermedias que hagan explícita información relevante e invariante acerca de la escena, tomando como punto de partida las imágenes. Esta tarea se ve dificultada por las diversas condiciones de iluminación, variaciones de posición y orientación de la superficie sensible, ruido, etc.. Asi pués, resulta de interés primordial el diseño de esas representaciones y la implementación de algoritmos para construirlas y relacionarlas entre sí.

Las diferentes representaciones las podemos agrupar en 4 amplias categorias [BALL-82], siendo posible la división de cada una en diferentes subniveles de representación o en varias representaciones cooperativas. Aunque estas categorias están ordenadas de bajo a alto nivel, el flujo de información entre ellas no es unidireccional. Además, no todos los niveles serán necesarios en cada aplicación específica. Las categorias mencionadas son las siguientes:

- Imágenes generalizadas.
- Imágenes segmentadas.
- Representaciones geométricas.
- Modelos relacionales.

1.2.3.1 IMAGENES GENERALIZADAS

El primer paso en el proceso de visión es el de formación de imágenes y éstas pueden ser el resultado de una amplia variedad de tecnologías [ROSE-86]. Los sistemas basados en cámaras de televisión convierten la intensidad de luz en una señal electrónica que posteriormente se digitaliza; otros sistemas usan radiaciones de otro tipo como rayos X, luz láser, ultrasonidos, etc. Generalmente el resultado es una colección o matriz numérica de muestras representativas de algún tipo de energía. Además, los propios dispositivos de adquisición pueden efectuar una gran variedad de operaciones como por ejemplo el suavizado o realce de una imagen.

Mediante el término imagen generalizada se hace referencia a las representaciones analógicas o icónicas de los datos de entrada. Se trata por tanto de un conjunto de entidades similares a una imagen y relacionadas a una escena. pudiendo incluir otro representaciones más útiles para procesamientos posteriores. Este es el caso, por ejemplo, de un array de discontinuidades en los niveles de gris. Las representaciones mediante arrays de alguna propiedad física tal como la orientación de la superficie, se conocen como imágenes intrínsecas. Una imagen intrínseca representa un paso adicional hacia la interpretación de una escena, y generalmente lleva asociada una serie de cálculos importantes. La información necesaria para el cálculo de una imagen intrínseca está contenida en la propia imagen inicial de partida (matriz de valores de intensidad), pudiendo obtenerse invirtiendo el propio proceso de formación de imágenes. Así por ejemplo, la fusión de dos imágenes estéreo da lugar a una imagen intrínseca de profundidades. Algunos algoritmos de recuperación de imágenes intrínsecas se pueden implementar en paralelo simulando los cálculos que tienen lugar al nivel neuronal de procesado biológico de imágenes [GRIM-81].

1.2.3.2 IMAGENES SEGMENTADAS

La idea de segmentación tiene sus orígenes en los trabajos llevados a cabo en Psicología al poner de manifiesto las preferencias mostradas por los seres humanos al agrupar y organizar conjuntos de formas dispuestas en su campo visual. Tales preferencias están basadas en propiedades de proximidad, similitud y continuidad. Otros resultados en este mismo ámbito tienen que ver con la facultad de discriminar un objeto del fondo de la imagen. Desde esta perspectiva se puede establecer que los mecanismos de agrupamiento organizan una escena en entidades con cierto significado, constituyendo así un paso hacia delante en la comprensión de imágenes.

En visión artificial la división de una imagen generalizada en partes que presenten homogeneidad en relación a una o más características, conduce a una imagen segmentada y al proceso empleado

se le conoce como segmentación. El concepto de imagen segmentada amplía al de imagen generalizada en un aspecto crucial: contiene los principios de interpretación dependientes del dominio. Así por ejemplo, la segmentación de una escena que contiene poliedros de caras planas puede conducir a un conjunto de segmentos rectilíneos correspondientes a las aristas del poliedro, o a un conjunto de regiones bidimensionales asociadas a las caras del poliedro. Se pone también de manifiesto que el conocimiento en torno a cada dominio particular de objetos puede proporcionar información útil en el diseño de estrategias de segmentación. En el ejemplo anterior de los poliedros, una información útil es el conocer de antemano que los segmentos deben ser rectilíneos.

El propósito de todo proceso de segmentación es pasar de una descripción de una escena en términos de una colección de pixels o de una imagen generalizada a una representación simbólica. Cuando esto ocurre se han logrado dos fines: de un lado realizar algún tipo de identificación dentro de la imagen y por otro lado alcanzar una compactación de los datos.

Esta fase de un sistema de visión podría ser considerada como un paso intermedio hacia etapas de procesamiento a nivel superior. El análisis de escenas dinámicas así como el análisis de texturas proporcionan unos buenos indicadores para la segmentación de imágenes [BALL-82].

1.2.3.3 REPRESENTACIONES GEOMETRICAS

Uno de los aspectos más interesantes a determinar en una imagen es la "forma" de los objetos que en ella aparecen. La forma es una propiedad intrínseca de los objetos en el espacio tridimensional. En un sentido, la forma es la propiedad primaria para un sistema de visión ya que a partir de ella pueden obtenerse otras características como el límite externo o frontera de ese objeto, la normal en cada punto de su superficie, etc. La forma de un objeto es básica en el sentido de que la definición de un objeto se suele asociar a la de su

forma mas que a su color o a propiedades de reflectancia.

La forma de un objeto puede definirse como aquella cualidad que depende de la posición relativa de todos los puntos que componen su superficie externa. Esta definición pone énfasis en el hecho de que somos conscientes de las formas a través de los perfiles y superficies de los objetos, propiedades ambas que pueden percibirse visualmente.

Las representaciones geométricas se emplean para reproducir fielmente las características relevantes de las formas en 2D y 3D. Tales representaciones deben ser lo suficientemente potentes como para soportar procesamientos complejos sobre ellas, como por ejemplo simular los efectos de diferentes condiciones de iluminación y movimiento.

El reconocimiento de objetos por su forma es, sin lugar a dudas, una de las habilidades más destacables del sistema visual en los mamíferos. La extracción de información 3D a partir de imágenes 2D es un tema en el que se han producido grandes avances en los últimos años, habiéndose desarrollado una gran variedad de técnicas que bajo ciertas restricciones permiten obtener propiedades tridimensionales acerca de los objetos en la escena. Mencionemos, a modo de ejemplo, las técnicas "Shape From ..." [BOYL-88].

Sin embargo, una vez que se conocen ciertas propiedades asociadas a las formas de los objetos 3D presentes en una escena, el problema de la visión artificial no está aún resuelto. Uno de los mayores desafios en visión artificial es la representación de las formas, o de sus características más destacadas, de manera tal que sobre esas reprensentaciones puedan acometerse tareas de aprendizaje, de comparación con otras formas, etc. Esto se ve dificultado por factores como:

Complejidad de las formas

Mientras el color y el movimiento son propiedades que quedan relativamente bien cuantificadas por unos pocos parámetros bien

estudiados, la forma es mucho más sutil. Algunos objetos manufacturados, y sobre todo los naturales, presentan formas increiblemente complejas por lo que la representación explícita de sus superficies requiere un elevado número de parámetros. Más difícil aún resulta el encontrar aquellos aspectos de las formas que son importantes, por ejemplo, en aplicaciones de reconocimiento. Una representación completa y explícita puede resultar computacionalmente intratable en algunas aplicaciones.

La introspección no se ve favorecida

Los seres humanos parecen disponer de una larga fracción de su cerebro dirigida a la simple tarea del reconocimiento de formas. Esta importante actividad está fuertemente establecida por debajo de los niveles permitidos a nuestra introspección consciente. ¿Por qué resulta tan fácil la tarea de reconocimiento de formas y se presenta tan ardua la de su descripción en los seres humanos?. El hecho de que no dispongamos de un lenguaje preciso para la descripción de formas puede ser un argumento a favor de la inaccesibilidad de los algoritmos y estructura de datos para este tipo de procesamiento. Esta falta de influencia cognitiva resulta desalentadora, especialmente al tratar con la complejidad de las formas diarias.

La disciplina es joven

Hasta muy recientemente, la comunicación humana en lo referente a formas complejas se hacía, básicamante, a través de palabras y esbozos bidimensionales. No ha sido sino hasta la aparición de los computadores digitales cuando se ha planteado como tema de interés la necesidad de representar formas complejas de manera que pudieran ser tratadas, manipuladas y representadas en forma gráfica por un ordenador.

Una vez que se ha extraido información de los objetos en una escena y como paso previo a su identificación, será preciso construir un modelo de los objetos en observación. Aún expresado de esta manera

tan simple, el problema está todavia lejos de ser resuelto de forma general [SHIR-87], [BOYL-88]. No podemos decir que exista un único esquema de representación apto para toda clase de formas. A menudo las representaciones buenas para una aplicación no suelen ser aptas para otro tipo de cálculos. Entre los esquemas más usuales se pueden destacar:

- Modelo de alambres (Wireframe).
- Geometría sólida constructiva (CSG).
- Representaciones basadas en la superficie.
- Representaciones volumétricas.
- Cilindros generalizados.

1.2.3.4 MODELOS RELACIONALES

En la interpretación de imágenes se establece una relación entre las estructuras implícitas en los datos visuales de entrada y la estructuras explícitas que ya existen en nuestra representación interna del mundo que nos rodea. Más explícitamente, las operaciones implicadas en la visión deben mantener al dia el conocimiento y modelos relativos al mundo externo.

Para examinar la influencia que sobre la visión ejercen los procesos que tienen lugar a nivel superior, será preciso acometer el mundo no visual y las capacidades de razonamiento de ámbito más global. Aquellos modelos del mundo externo capaces de soportar cálculos avanzados dependientes de cada dominio de aplicación han de poseer una complejidad considerable.

En el contexto de la visión artificial, mediante el término modelos relacionales nos referimos a complejas colecciones de representaciones obtenidas a partir de las imágenes y capaces de soportar procesamientos sofisticados de alto nivel. Es usual, a estos niveles de procesamiento, el empleo de conocimiento y modelos adquiridos con anterioridad a la propia experiencia perceptiva.

Una vez construidos los modelos relacionales de los objetos presentes en la escena, el mecanismo más simple de procesamiento para tomar alguna decisión o emitir algún juicio con el fin de describir lo que en ella se ve, consistirá en comparar estas estructuras relacionales con las de los modelos de objetos previamente conocidos. En el caso de estar considerando sólo una imagen de la escena, estas comparaciones son difíciles de realizar cuando se trata de escenas típicamente tridimensionales; esto se entiende fácilmente si se tiene en cuenta que la imagen sólo muestra la parte visible de los objetos y puede que algunos de ellos estén parcialmente ocultos. Si por el contrario se dispone de varias imágenes de una misma escena pero bajo condiciones, será posible generar representaciones diferentes los objetos que la componen. Bajo estas de geométricas 3D relacionales la comparación de los modelos circunstancias, representativos de la escena con los modelos de objetos conocidos presentará menos dificultades.

En la construcción de los modelos relacionales se hace uso de técnicas propias de otros campos de la inteligencia artificial. Así, el formalismo de las redes semánticas para estructurar conocimientos complejos puede ser empleado en la representación del conocimiento adquirido de una escena. En la toma de decisiones acerca de la interpretación del contenido de una escena podrán utilizarse las [CHAR-85] para manipular de inferencia lógica técnicas representaciones del conocimiento en forma proposicional, o los sistemas de producción [BOYL-88] ampliamente empleados como base de sistemas expertos o en el procesamiento sintáctico del lenguaje.

1.2.4 LA TEORIA DE MARR EN LA VISION HUMANA Y VISION ARTIFICIAL

Uno de los planteamientos pioneros en visión artificial hacia el desarrollo de una teoría efectiva acerca del proceso de visión se debe a David Marr [MARR-76a, 76b, 82].

Marr construyó una teoría computacional estableciendo que la visión puede ser considerada como una tarea de procesamiento de

información, es decir de extracción, a partir de las imágenes, de los diferentes aspectos de la escena observada que son útiles para el observador. En este esquema, el objetivo consiste en disponer de métodos computacionales efectivos para pasar de imágenes a representaciones mediante la imposición de una serie de restricciones. Se establece una secuencia de representaciones cuidadosamente diseñadas, comenzando con descripciones obtenidas directamente a partir de la imagen con el fin de facilitar la posterior recuperación de propiedades físicas cada vez más objetivas sobre la forma de los objetos presentes en la escena.

Cualquier información que pueda obtenerse a partir de representaciones previas de la imagen y que sea útil para la construcción de la siguiente representación ha de expresarse explícitamente en cada etapa. Así, la validez de cada esquema de representación se juzga por dos criterios: la posibilidad de ser viable desde el punto de vista computacional y su idoneidad para niveles superiores de procesamiento [MARR-78a]. Marr y Nishihara [MARR-78a, 82] proponen la existencia de, al menos. representaciones en el curso del procesamiento de los datos visuales:

- Con el término Esbozo Primario (Primal Sketch) se hace referencia a la representación de la imagen obtenida a la salida de los procesos de bajo nivel, haciendo explícita la información relativa a cambios de intensidad, discontinuidades, bordes, etc.
- 2. A la representación obtenida tras los procesos de segmentación, siendo intermedia entre 2D y 3D, le llama Esbozo 2 1/2D. Se decriben las propiedades de la superficie visible de cada objeto para cada localización en la imagen. Aquí se hace explícita la orientación y profundidad de las superficies y contornos visibles respecto de un sistema de coordenadas centrado en el observador.
- 3. Por último, y a partir de la anterior representación, se obtiene un modelo 3D de cada objeto describiendo su forma y organización espacial en un marco de coordenadas centrado en el propio objeto. La descripción en este esquema es modular, empleando como

primitivas elementos de superficie y volumen.

Una implementación de esta teoría computacional a niveles bajos de procesamiento fué llevada a cabo en su grupo de investigación del M.I.T. por Grimson, sirviendo de base para estudios posteriores [GRIM-81]. En este trabajo se investigan las dos representaciones creadas al formar una descripción de aquellas localizaciones en la imagen para las que sus correspondientes localizaciones en la escena experimentan algún cambio en alguna propiedad física; a partir de esta descripción o Esbozo Primario se obtiene una descripción de la geometría de la superficie visible de los objetos en la escena mediante la teoría de visión estereo [MARR-79]. Revisiones y modificaciones más recientes al algoritmo inicial de Grimson pueden verse en [GRIM-86b].

1.2.5 SISTEMAS Y APLICACIONES

Los sistemas de visión artificial pueden ser de propósito general o diseñados para tareas específicas. De un istema de propósito general se espera que tenga capacidades similares a las del sistema visual humano y que pueda tratar una amplia variedad de escenas bajo una gran variedad de condiciones de visión. Así por ejemplo, nosotros somos capaces de generar descripciones de gran calidad incluso para objetos poco familiares. La comprensión de los procesos perceptivos implicados en nuestro sistema de visión es aún pobre, alcanzando los sistemas artificiales un bajo grado de eficacia en comparación con la actuación humana.

Afortunadamente una gran cantidad de aplicaciones prácticas de importancia no requieren esta generalidad ya que el dominio de objetos es pequeño, disponiéndose a priori de un conocimiento significativo en torno a la escena.

1.2.5.1 SISTEMAS DE PROPOSITO GENERAL

Puesto que la información visual es inherentemente ambigua, algún

conocimiento y consideraciones serán precisas para su interpretación. Esas consideraciones pueden conducir, ocasionalmente, a conclusiones incorrectas aunque por fortuna la actuación del sistema visual humano es correcta en casi todas las situaciones de nuestra experiencia diaria. Los sistemas de propósito general y aquellos basados en el conocimiento previo en torno a la escena, difieren no sólo en el tipo de objetos que traten sino también en el tipo de conocimiento que empleen. Los sistemas de propósito general tienden a usar conocimiento genérico mas que conocimiento de tipo específico. El conocimiento de tipo genérico incluye restricciones debidas a fenómenos físicos tales como reflectividad de las superficies, continuidad, coherencia,... Mediante el término conocimiento específico se hace referencia al conocimiento de objetos particulares, probablemente presentes en la escena, sus propiedades y las condiciones particulares de visión. Es de destacar también el hecho de que los sistemas de propósito general tienden a postergar el uso de conocimiento específico, mientras que los sistemas basados en el conocimiento previo acerca de la escena tienden a utilizarlo en las primeras etapas del procesado. Un sistema de propósito general requiere, al menos, de las siguientes facultades:

- Percibir luminosidad y/o color de las superficies bajo una amplia gama de condiciones de iluminación.
- Detectar cambios significativos de intensidad y efectuar segmentación en regiones en el espacio 2D.
- Inferir estructuras 3D de las superficies presentes en una escena a partir de información contenida en vistas monoculares o a partir de estereovisión o de secuencias dinámicas de imágenes.
- Organizar las superficies y regiones extraidas tras la segmentación en objetos de interés.
- Generar descripciones de objetos y reconocerlos entre una amplia gama.
- Realizar inferencias no visuales e inteligentes en torno a la escena en estudio a partir del procesado visual.

Los procesos anteriores pueden ser considerados para formar una estructura jerárquica de niveles de abstracción. Dos métodos de flujo de datos entre niveles son [NEVA-82], [SHIR-87]: un esquema ascendente de abajo hacia arriba (bottom-up) y un esquema descendente de arriba hacia abajo (top-down). En el primer caso la información fluye de un nivel al inmediatamente superior con independencia de cuales sean los fines esperados en este nivel superior. En las estrategias de control descendente el procesado a niveles inferiores está específicamente enfocado a satisfacer objetivos concretos de niveles superiores.

A un nivel superior, la meta puede consistir en verificar si cierto objeto está presente en la escena. Cada vez parece más claro que nuestra habilidad de percepción de escenas poco familiares requiere capacidades de tipo "bottom-up". Sin embargo se hace necesaria una comunicación fluida y extensa entre los distintos niveles por lo que el procesado no es estrictamente bottom-up. Los humanos también efectuamos procesado del tipo top-down como queda indicado por nuestra habilidad de imaginarnos la existencia de objetos de algún tipo en escenas confusas (por ejemplo, la forma de algunas nubes en el cielo nos sugiere o evoca a algún objeto).

En la actualidad no existen sistemas de visión que se aproximen al nivel de actuación del sistema humano de propósito general. No obstante el deseo por alcanzar esa generalidad en el futuro, conduce a estrategias de diseño que no se reducen únicamente a aquellas basadas en el conocimiento previo.

1.2.5.2 APLICACIONES

Las potenciales áreas de aplicación de un sistema de visión artificial de propósito general son numerosas. No obstante, las aplicaciones actuales deben ser cuidadosamente elegidas para estar de acuerdo con las capacidades de las técnicas disponibles. Las aplicaciones con éxito se orientan a tareas específicas bien definidas y dentro de un dominio limitado de objetos y con un adecuado

conocimiento a priori. Sin ser exhaustivos, y dado el creciente número de aplicaciones, podríamos indicar, entre otras, las siguientes grandes áreas de aplicación:

- Aplicaciones industriales.
- Fotointerpretación y detección de cambios.
- Guiado y navegación de vehículos.
- Aplicaciones médicas.

APLICACIONES INDUSTRIALES

Un campo extenso de aplicaciones industriales es el referente a la inspección visual de partes manufacturadas. Las tareas de inspección van desde la detección de grandes defectos, como la ausencia de alguna pieza, hasta la detección de pequeñas imperfecciones, desalineamientos, cambios en el tamaño o color, etc. Para los humanos, esas tareas tienden a convertirse en rutinarias y pesadas, por lo que su repetitividad conduce a una disminución drástica en la eficacia.

Las aplicaciones prácticas eficientes requieren un procesamiento rápido, un hardware económico y una alta fiabilidad. Las condiciones de iluminación se pueden controlar de manera que se obtengan imágenes de elevado contraste posibilitando el proceso de segmentación, o que hagan visibles algunos defectos en piezas, etc. Los éxitos más notables se han logrado en la inspección de tarjetas de circuitos impresos y circuitos integrados [NEVA-82].

FOTOINTERPRETACION Y DETECCION DE CAMBIOS

Las imágenes tomadas desde aviones o por satélites en órbita proporcionan una fuente rica de información para la detección y seguimiento de cambios sobre la superficie de la Tierra. Las aplicaciones incluyen el seguimiento de cultivos, bosques y otras fuentes naturales, polución, vigilancia militar, etc. [MATH-87].

GUIADO Y NAVEGACION DE VEHICULOS

Una parte importante de la visión en humanos y animales tiene como objetivo el contribuir al movimiento en el medio que nos rodea evitando obstáculos y alcanzando las posiciones deseadas. Muchos de los sistemas automáticos de navegación en máquinas, como por ejemplo en aviones comerciales, se basan en el uso de sensores especiales tales como el radar, localizados a lo largo de las rutas deseadas. Sin embargo, algunas aplicaciones -por ejemplo en exploración de planetas lejanos- requieren navegación sin modificación del entorno y con poca, o ninguna, intervención humana y sin conocimiento previo del terreno.

El problema de navegación de vehículos en terrenos desconocidos requiere un sistema de detección de obstáculos y sistemas para evitarlos garantizando así la supervivencia del vehículo. Entre los sistemas de detección se encuentran aquellos que utilizan datos visuales como señales de entrada, siendo necesario resolver los problemas usuales de visión artificial de manera que se posibilite la planificación de estrategias de movimiento para una navegación segura.

APLICACIONES MEDICAS

Muchos procedimientos de diagnóstico médico emplean imágenes radiográficas, fotografías microscópicas de células sanguíneas, imágenes acústicas e imágenes tridimensionales de órganos obtenidas en tomografía computerizada. La interpretación de tales imágenes requiere de personal médico bien entrenado, cuya escasez supone un freno al uso extendido de tales técnicas.

Las imágenes médicas tienen algunas características peculiares. Con frecuencia la iluminación utilizada penetra a los objetos, es decir las superficies de los órganos en observación no son opacas. A veces los defectos a detectar son pequeños y sutiles, caracterizados en ocasiones por cambios suaves en los niveles de gris. También los márgenes de error son pequeños debido a las serias consecuencias que pueden derivarse de un diagnóstico incorrecto. Por otro lado, en estas aplicaciones suele disponerse de una significativa información a

priori proporcionada por la anatomía, siendo además controlables las condiciones de observación.

El análisis de radiografías y de células sanguíneas y cromosomas ha recibido una considerable atención desde inicios de la década de los 70, siendo más reciente el uso de los ordenadores en la obtención de datos 3D del cerebro y otros órganos mediante el empleo de múltiples imágenes 2D. Gran parte de estos trabajos se han concentrado sobre una rápida y exacta reconstrucción de la estructura tridimensional de órganos y el análisis de tales datos con fines de diagnóstico.

1.2.5.3 REQUERIMIENTOS DE HARDWARE

Los grandes requerimientos de procesamiento visual son una limitación en el rango de las aplicaciones prácticas. La resolución de las imágenes es menor a la de la visión humana. La complejidad de los algoritmos empleados vendrá limitada por los requerimientos en tiempo de respuesta. El continuo aumento en la velocidad de los ordenadores de propósito general, unido a la reducción simultánea en costo y tamaño conlleva un avance paralelo en las técnicas empleadas. Sin embargo es poco probable que el procesamiento visual complejo y de alta resolución sea viable solamente con computadores secuenciales de propósito general.

Afortunadamente la estructura del procesamiento de información visual es apropiada para una implementación en paralelo. El procesamiento a bajo nivel tal como la convolución en la detección de bordes, o la elaboración de histogramas para selección de regiones, consume gran parte del tiempo de ejecución. Usualmente en el procesamiento a niveles superiores la cantidad de datos a procesar disminuye drásticamente, con lo que se consume menos tiempo aún en operaciones más complejas. A bajo nivel el procesamiento generalmente es simple, usando información de pequeñas zonas de la imagen lo que facilita la implementación de procesos en paralelo. Además, la sencillez de algunas de las operaciones a bajo nivel permite su

implementación en hardware de propósito específico. Esto asegura un proceso evolutivo en el uso de máquinas para efectuar tareas cada vez más complejas.

Algunos problemas de visión artificial ya han sido resueltos en el laboratorio, sin embargo esas soluciones consumen, con frecuencia, cantidades considerables de tiempo de proceso sobre ordenadores relativamente potentes. A pesar de ello, y debido al aumento de velocidad de procesamiento y disminución del coste, es previsible que cualquier tarea que actualmente se pueda realizar en un laboratorio de investigación será reproducible, tarde o temprano, en tiempo real y con hardware de bajo coste. ¿Qué tipo de arquitecturas de ordenadores serán necesarias para hacer esto posible?. Como se indica en [ROSE-88] la visión involucra varios tipos de datos -matrices de pixels, representaciones geométricas, grafos y bases de conocimiento. Sobre cada uno de estos tipos de datos pueden ser necesarias diversas clases de operaciones. Para algunas de estas operaciones es bien conocida la acelerarlas usando varios tipos de arquitecturas forma de multiprocesadores; sin embargo, para otros tipos de operaciones esto no está tan claro.

Los mayores avances en la consecución de sistemas más potentes aparecerán no como consecuencia única del aumento en los ciclos de reloj por unidad de tiempo sino, más importante aún, como resultado del multiproceso masivo en paralelo. Diversas clases de sistemas multiprocesadores están en fase de elaboración y otros ya comercializados [FREE-88].

En el procesado al nivel array de pixels o similares, está bien comprendido el uso de tales arquitecturas. Desafortunadamente, es mucho menor el conocimiento que se tiene acerca de su empleo en el aumento de la velocidad al acometer operaciones en los niveles de representación geométrica o de estructura de grafos o de representación del conocimiento. El desafío a que hemos de enfrentarnos es, por tanto, cómo diseñar arquitecturas de ordenadores -y los algoritmos para ejecutarse sobre ellas- que puedan compararse y eventualmente exceder al nivel de actuación en humanos.

1.3 TECNICAS DE VISION 3D

Conceptualmente la percepción puede definirse como el proceso de interpretar la información captada por unos receptores. Esta información por si sola recibe el nombre de sensación. Es una experiencia habitual en el entorno humano que la interpretación de una sensación no se basa exclusivamente en la información contenida en la misma, sino que requiere la cooperación de otras fuentes, tales como la experiencia previa, el contexto, etc. Esta es la razón por la cual los sistemas de percepción artificial que se desarrollan en la actualidad utilizan técnicas específicas para la resolución de problemas en dominios restringidos [FERR-86].

No es nuestro objetivo, en este apartado, hacer una clasificación exhaustiva de los diferentes sistemas de percepción artificial, sino más bien indicar algunas de las técnicas más representativas que se usan en la actualidad. Para una descripción detallada del trabajo de algunos de los más destacados grupos de investigación en este campo, véase [ROSE-86].

Básicamente, los sensores empleados en percepción artificial se pueden agrupar en dos amplias categorías: sensores de contacto y sensores a distancia. Los sensores de contacto, que juegan un papel importante en robótica, han recibido menos atención que los del segundo grupo. La información suministrada por este tipo de sensores se refiere fundamentalmente a posición y orientación de puntos sobre la superficie de los objetos, así como a la medida de fuerzas, presiones y temperatura [GAST-84], [GRIM-86a], [BROW-87].

Los sensores de no-contacto, o a distancia, se pueden clasificar a su vez en activos y pasivos. Con el término "activo" se hace referencia al hecho de que el sistema incorpore algún mecanismo de control para la fuente de energía que es enviada a la escena y posteriormente modificada por ésta.

En términos generales, el objetivo final de cualquier sistema de visión es reconocer el contenido de lo que hay en una "imagen". Aunque

en ocasiones es posible identificar los objetos en una imagen a partir de un conjunto reducido de información, esto no es lo usual. El reconocimiento de los objetos presentes en una escena frecuentemente requiere el empleo de modelos. Los modelos 3D presentan gran interés ya que describen a los objetos de manera similar a como ellos son en la realidad. La representación bidimensional de un objeto 3D es una proyección en perspectiva del mismo. Esta descripción bidimensional es incompleta y ambigua al no mostrar las partes ocultas del objeto y no proporcionar la localización y disposición espacial (profundidad) de las partes visibles en la imagen 2D.

Con el fin de conseguir modelos 3D de la geometría de una escena, se precisa de métodos que permitan inferir la naturaleza tridimensional de la misma, de manera que esta información pueda reflejarse en representaciones adecuadas que, como paso último, conduzcan a descripciones invariantes de los objetos en la escena.

Gran parte de las investigaciones desarrolladas en el campo de visión artificial están dirigidas al desarrollo de sistemas que recuperen la información tridimensional asociada a una escena. La actuación de cualquier sistema de análisis de escenas está limitada, en última instancia, por la naturaleza y calidad del sensor que proprociona la información de partida. Estos trabajos pueden agruparse en tres amplios grupos:

- 1. Aquellos métodos que usan información 3D directamente a partir de sensores activos.
- 2. Métodos que emplean solamente información monocular contenida en una imagen (o quizás en varias imágenes desde el mismo lugar pero bajo diferentes condiciones de iluminación).
- 3. Métodos que hacen uso de dos o más imágenes tomadas desde diferentes posiciones y probablemente en instantes de tiempo distintos.

1.3.1 TECNICAS BASADAS EN SENSORES ACTIVOS DE PROFUNDIDAD

Algunos sistemas son intrínsecamente 3D debido a que captan directamente información tridimensional. La mayoría de ellos adquieren información de tipo geométrico como la posición espacial de un conjunto de puntos sobre la superficie de un objeto, o bien la orientación de determinados elementos de su superficie. Este es el caso de los dispositivos basados en telemetría óptica, que determinan directamente la separación de un punto respecto del sensor. En este caso, la medida puntual de profundidades (distancia sensor-objeto) puede efectuarse determinando el tiempo de propagación de un haz luminoso entre un foco emisor y el receptor de la luz reflejada sobre el punto de la escena cuya profundidad se desea medir [AMAT-83], [SHIR-87], [BOYL-88]. Tales sensores activos de profundidad producen matrices de pixels cuyos valores contienen la profundidad de ese punto a la escena, y no valores de intensidad luminosa.

La tecnología predominante para los sensores activos de profundidad está basada en el empleo de láser. La medida de profundidad puede efectuarse mediante medidas del tiempo de ida y vuelta para el haz láser, o bien midiendo la diferencia de fase entre los haces incidente y reflejado. La primera técnica puede proporcionar buenos resultados, sin embargo requiere dispositivos electrónicos de elevada resolución [BOYL-88] (medidas de profundidad con una precisión del orden de 5×10^{-3} m requieren medidas de tiempos del orden de 70 picosegundos). La estrategia basada en medidas de fases por regla general es lenta, necesitando algunos minutos para adquirir una imagen de una escena [JARV-83].

Métodos similares basados en triangulación mediante láser han sido empleados para la reconstrucción 3D y reconocimiento de piezas industriales [BHAN-82, 84], [FAUG-86]. Duda y col. [DUDA-79b], barriendo la escena con un láser, determinan simultáneamente la localización espacial y datos de reflectancia para cada punto de la escena comparando fase y amplitud de los haces incidente y reflejado. Esta información fué posteriormente procesada para descomponer la escena en elementos planos de superficie. Altschuler y col. [ALTS-81]

describen un sistema que analiza una secuencia de patrones de puntos láser proyectados sobre la superficie de un objeto. Dichos patrones son observados por una cámara de televisión y posteriormente procesados para hacer una reconstrucción 3D de su superficie. Para una visión más completa de este tipo de técnicas véase [JARV-83].

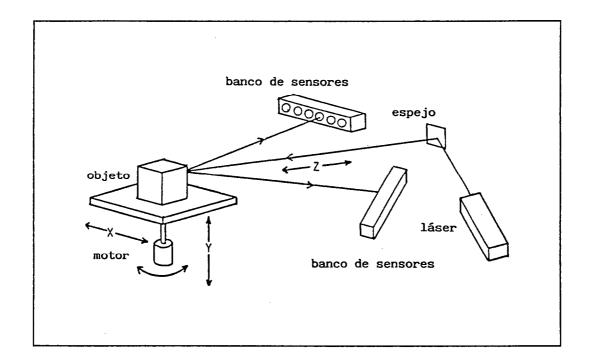


Fig. 1.5. Diagrama esquemático del sistema de adquisición de datos 3D usado en [BHAN-84], [FAUG-86], y basado en el principio de estereoscopía activa mediante láser. Un espejo rotante efectúa un barrido del haz láser a lo largo del eje X. La distancia Z, para cada punto del objeto, se calcula a partir de la localización de la respuesta en cada banco de sensores para el haz reflejado.

1.3.2 TECNICAS BASADAS EN EL ANALISIS DE IMAGENES MONOCULARES

Los sistemas de percepción visual están notablemente desarrollados, y su dominio de apliación es, con mucho, el más extenso. Los sensores más usuales son las cámaras de televisión [ANGU-85]. La información proporcionada -proyección de una escena sobre el plano imagen- se representa por una matriz de valores de

intensidad luminosa (pixels).

Es obvio que toda imagen bidimensional de un objeto contiene información 3D acerca del mismo. No obstante la recuperación de esta información no es una tarea fácil. Especial interés requieren las dos cuestiones siguientes:

- 1. ¿Qué tipo de información puede obtenerse a partir de imágenes bidimensionales?.
- 2. ¿Cuál es el proceso seguido y bajo qué condiciones es operativo?.

La información espacial que puede extraerse de una imagen tal y como es suministrada por una cámara de TV es bien distinta. En principio es imposible recuperar toda la información 3D a partir de una imagen monocular [BAST-86], [SHIR-87]. No obstante si se dispone de algún conocimiento acerca de la escena, una imagen monocular o múltiples images monoculares (tomadas desde la misma posición pero con diferentes condiciones de iluminación) pueden proporcionar, frecuencia, información acerca de la forma y orientación de la superficie de los objetos en la escena. Existen una serie de propiedades que pueden explotarse para obtener la forma, o datos de disposición espacial, que colectivamente dan lugar a una clase de algoritmos conocidos genericamente como "Shape From...". Por ejemplo, es posible realizar un análisis de texturas [WITK-81], [DAVI-83], contornos [MARR-77], [BRAD-84], [IKEU-84], análisis local intensidades [PENT-84] para obtener una descripción espacial limitada acerca de los objetos o puntos característicos de la escena.

1.3.2.1 ANALISIS LOCAL DE INTENSIDADES

Bajo fuertes restricciones impuestas al tipo de objetos que aparecen en la escena y con algún conocimiento a priori en torno a la misma, Roberts [ROBE-65] demostró que a partir de los valores de intensidad en una imagen monocular es posible obtener información 3D.

Bajo otro tipo de restricciones, Horn [HORN-75,77] explora la relación existente entre los valores de intensidad de una única imagen y la forma del objeto generador de la misma. Demostró que los valores de intensidad están fuertemente relacionados con la orientación local de la superficie del objeto. Esta se puede determinar a partir de aquellos resolviendo una ecuación diferencial de 1° orden en derivadas parciales con dos incógnitas -parámetros asociados a la orientación-. La resolución de esta ecuación presupone el conocimiento de las condiciones de iluminación y de las características ópticas de la superficie del objeto (propiedades de reflectancia), además de la suposición de que su superficie no presenta cambios bruscos (es "suave").

En 1981, Ikeuchi y col. [IKEU-81a] reducen el número de restricciones impuestas en la recuperación de la orientación y posición de una superficie a partir del conocimiento de una sola imagen. Bajo condiciones similares de suavidad para la superficie (función de clase C¹) y conocidas sus propiedades de reflectancia, resuelven la ecuación

$$R(p,q) = E(x,y)$$

que relaciona el brillo E(x,y) en cada punto imagen con las posibles orientaciones R(p,q) del correspondiente punto sobre la superficie del objeto. Ikeuchi transforma la ecuación anterior del plano gradiente (p,q) al plano estereográfico. De esta forma posibilita incorporar la información suministrada por la orientación de algunas líneas del contorno [MARR-77] de su superficie en la forma de valores iniciales y condiciones de contorno para la resolución de la ecuación diferencial mediante un algoritmo iterativo.

Las técnicas anteriores requieren de un considerable conocimiento a priori de la escena a analizar; por tanto no pueden aplicarse con éxito a escenas con las que no se está familiarizado. En [PENT-84] se demuestra que con los valores de intensidad de una única imagen y sin restricciones adicionales, no es posible asegurar la unicidad de solución en la recuperación de la estructura 3D. Así mismo, y de

manera elegante, trata la cuestión de cuál es la información que es posible extraer a partir del análisis local de una imagen. Resultados más recientes pueden encontrarse en [FERR-89].

1.3.2.2 METODOS BASADOS EN LUZ ESTRUCTURADA

Otra aproximación en la localización tridimensional de puntos sobre un objeto, o medida de distancias objeto-sensor, fué propuesta y desarrollada por Oshima y col. [OSHI-75,79,83]. Esta técnica, también conocida como de triangulación óptica, (Fig. 1.6) combina un sistema cámara-ordenador para almacenar la posición de puntos iluminados por un plano vertical de luz proyectado oblicuamente a lo largo del campo de visión de la cámara. La localización de cualquier punto iluminado se determina mediante la intersección del plano de iluminación y el rayo desde la escena hasta el punto o pixel en cuestión. La información completa de una escena se adquiere a medida que el plano de luz varia su orientación.

La limitación más importante de este sistema es la posible ocultación de algunas líneas luminosas proyectadas sobre el objeto, debido a formas especiales de su superficie. Otro aspecto a tener en cuenta es el de errores debidos a imprecisiones mecánicas [KAK-86].

Freeman y Potmesil en [POTM-78, 79] presentan un sistema que permite una descripción detallada de la superficie de objetos complejos. El objeto en estudio se ilumina con un patrón o máscara de líneas. La distorsión producida por la superficie sobre el patrón de líneas proyectado, suministra la información necesaria para la descripción 3D del objeto. Este proceso requiere, como paso previo, el cálculo de la matriz que transforma las coordenadas 2D de cada punto del patrón proyectado en sus correspondientes coordenadas 3D sobre la superficie del objeto. Para ello se usan puntos de calibración visibles en cada una de las imágenes que, desde diferentes posiciones, pueden adquirirse de un mismo objeto. Técnicas similares se han empleado en procesos industriales para la detección de defectos en la superficie de piezas industriales [MUND-88], en la determinación de la

orientación de superficies especulares [SAND-88], en aplicaciones médicas [HIER-82], [FROB-83a, 83b] y en aplicaciones topográficas [PERR-79].

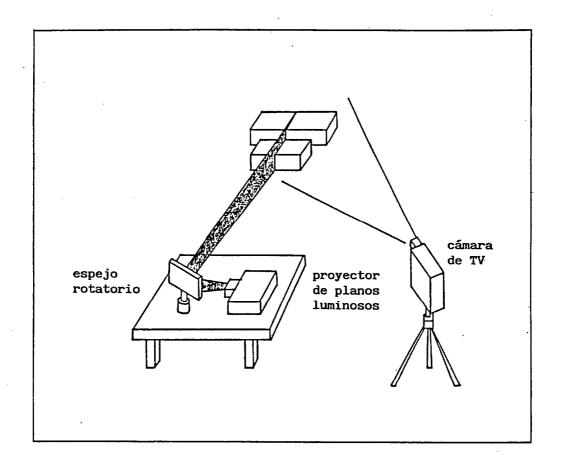


Fig. 1.6. Diagrama esquemático mostrando planos verticales de iluminación junto a una cámara de TV para obtener información 3D acerca de la escena.

1.3.3 TECNICAS BASADAS EN EL EMPLEO DE MULTIPLES IMAGENES

1.3.3.1 VISION ESTEREOSCOPICA

Las restricciones impuestas a la escena pueden reducirse considerablemente mediante el uso de múltiples imágenes de la misma. El caso más simple corresponde al de dos imágenes (Fig. 1.7). Si un

mismo objeto es observado desde dos posiciones diferentes es posible evaluar su localización en el espacio 3D. Esta técnica se conoce como visión estereoscópica o visión binocular. Esta es la manera más importante de captar la profundidad en el sistema visual humano. En este sentido se han realizado gran cantidad de investigaciones dirigidas a encontrar modelos computacionales para la visión estereoscópica [MARR-78b], [GRIM-81, 86b]. En aplicaciones prácticas esta técnica también presenta interés; así por ejemplo, se ha empleado en fotografía aérea para determinar la geometría del terreno.

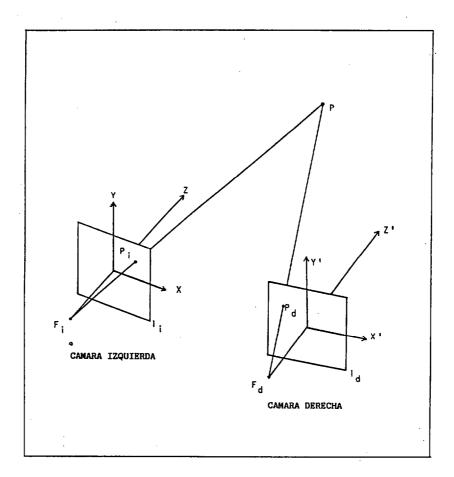


Fig. 1.7. Sistema de estereovisión convencional formado por dos cámaras. Los puntos focales son Fi y Fd. Los planos imagen son Ii e Id, y los ejes de visión son Z y Z'. Un punto P en el espacio 3D proyecta sobre Pi en la imagen izquierda y sobre Pd en la imagen derecha. La disparidad de P es la diferencia en las posiciones de Pi y Pd, y depende de la localización de P en el espacio y de la relación geométrica entre el sistema de cámaras.

Desde un punto de vista amplio, la visión estereoscópica puede definirse como la recuperación de las características tridimensionales de una escena a partir de dos imágenes tomadas desde diferentes posiciones y orientaciones [BARN-82], [NEVA-76]. La determinación de la tercera dimensión o profundidad requiere de los siguientes pasos:

- 1. Adquisición de imágenes.
- 2. Modelado de las cámaras.
- 3. Adquisición de características.
- 4. Establecimiento de la correspondencia.
- 5. Determinación de la profundidad.
- 6. Proceso de interpolación para transformar el mapa de profundidades en una descripción 3D.

El problema fundamental reside en la localización, en cada imagen, de la proyección de un mismo punto tridimensional. Esto se conoce como el problema de la correspondencia [AGGA-81a]. diferencia en la posición, para cada imagen, de puntos características correspondientes, se conoce como disparidad [BARN-80]. disparidad es función de la posición, orientación características físicas de las cámaras. Por este motivo se hace necesario conocer y modelar [HARA-80] las transformaciones que una cámara efectúa del dominio 3D al 2D.

En orden a determinar la profundidad de cada punto, es importante conocer no sólo el modelo de cada cámara, sino el modelo entre cámaras; es decir, las posiciones y orientaciones relativas de las mismas. La estrategia seguida en estereovisión consiste, pues, en decodificar la tridimensionalidad a partir del conocimiento de las disparidades y de los modelos geométricos asociados a las cámaras.

Dos son las dificultades principales en las técnicas de visión estereoscópica:

 En primer lugar la identificación de puntos o características correspondientes en ambas imágenes (de baja resolución si comparamos con la visión humana). 2. En segundo lugar, se hace necesario establecer un compromiso entre una gran separación entre cámaras -lo que aumenta la resolución geométrica- y una separación pequeña -lo que facilita la identificación de características correspondientes-.

Para una revisión reciente de estas dos cuestiones véase [BLOS-87], [BOYE-88].

El principal problema en visión estereoscópica es encontrar características correspondientes en las imágenes izquierda y derecha. Una vez que se ha establecido dicha correspondencia, la posición 3D de esa característica se evalúa mediante un sencillo proceso de triangulación. La técnica de estereovisión presenta la ventaja de no necesitar de equipos especiales. Esta es la razón por la que el estudio de esta estrategia fué anterior al de los métodos telemétricos. En el desarrollo de la visión estereoscópica se han seguido dos aproximaciones: (1) comprender el sistema binocular humano, y (2) encontrar caminos válidos para adquirir información 3D.

1.3.3.2 ESTEREOSCOPIA FOTOMETRICA

La técnica de estereoscopía fotométrica evita algunas de las dificultades de la visión estereoscópica usando varias imágenes tomadas desde la misma posición y con la misma cámara, bajo diferentes condiciones conocidas de iluminación.

La estereoscopía fotométrica, propuesta inicialmente por Woodman [WOOD-78, 80], se basa en el hecho de que si se obtiene una pareja de imágenes del mismo objeto, pero variando la dirección de la luz incidente, será posible dibujar dos mapas de reflectancias, ya que éstos dependen, entre otros factores, de la dirección de la fuente luminosa. De esta manera la orientación de la superficie del objeto se determina localmente a través de las parejas de valores de intensidad de cada punto imagen como la intersección de líneas de brillo constante en cada mapa de reflectancias.

El mapa de reflectancias representa la relación entre la orientación de la superficie visible de un objeto y los valores de intensidad de su imagen. El diagrama formado por las líneas que unen los puntos, en el espacio gradiente (p,q), correspondientes a orientaciones de superficies que presentan el mismo brillo, se denomina mapa de reflectancias y se representa por R(p,q,x,y), siendo $p = \delta z/\delta x$, $q = \delta z/\delta y$, y z = z(x,y) la ecuación de la superficie del objeto. Bajo la aproximación de proyección ortográfica, el mapa de reflectancias es el mismo para todos los puntos (x,y) de la imagen, es decir R(p,q,x,y) se puede expresar como R(p,q).

Para un punto dado (xo,yo) del plano imagen se tienen dos valores de brillo, uno para cada tipo de iluminación, y dos mapas de reflectancias $R_1(p,q)$ y $R_2(p,q)$. Se dispone así de dos ecuaciones no lineales con dos incógnitas p y q

$$R_1(p,q) = E_1(x_0, y_0)$$

 $R_2(p,q) = E_2(x,y)$

La intersección de las líneas de isoreflectancia con valores E1 y E2 en el espacio gradiente, proporciona los valores de orientación (p,q) asociados a la superficie que proyecta en (x0,y0) (Fig. 1.8).

Puesto que el sistema de dos incognitas anterior es no lineal, puede tener más de una solución, por lo que en muchas situaciones se precisa de la información adicional suministrada por una tercera o sucesivas imágenes.

Woodham aplica esta técnica a superficies de reflexión difusa e iluminadas por una fuente puntual. Este trabajo fué ampliado por Ikeuchi [IKEU-81b] a superficies de difusión especular e iluminadas por una fuente de luz distribuida, situación ésta bastante común en entornos industriales.

Conviene indicar que aunque esta técnica no necesita de la asunción de suavidad en la superficie del objeto en estudio, los

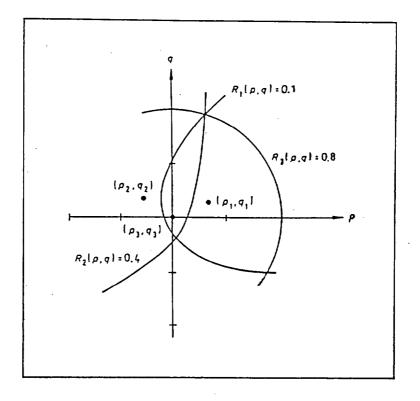


Fig. 1.8. Determinación de la orientación local de la superficie que proyecta en un punto (x0,y0) del plano imagen mediante la técnica de estereoscopía fotométrica. Para el punto en cuestión se disponen de tres medidas de brillo obtenidas bajo diferentes condiciones de iluminación. Las tres líneas de isoreflectancia intersectan en un punto cuyas coordenadas indican la normal local de la superficie.

valores de orientación obtenidos son sensibles al error cometido en las medidas de brillo. Además, puesto que todas las imágenes se obtienen desde la misma posición, se hace aquí innecesaria la determinación de la correspondencia entre imágenes como ocurre en estereovisión. Estas características y otras descritas en [WOOH-80] indican que esta técnica podría usarse en conjunción con la de estereovisión.

1.3.3.3 ANALISIS DE ESCENAS DINAMICAS

Otra aproximación general en la recuperación de la información 3D de una escena es la determinación de la estructura tridimensional a partir del estudio del movimiento. Las múltiples imágenes de un objeto en movimiento (movimiento relativo cámara-objeto) son analizadas bajo restricciones específicas para obtener descripciones explícitas de su estructura 3D. Tanto las restricciones impuestas a los objetos en la escena como el movimiento relativo cámara-objeto son importantes para la resolución del problema.

Uno de los trabajos iniciales usando múltiples vistas de un objeto para formar un modelo del mismo, fué el de Underwood y Coates [UNDE-75]. La descripción 3D de un objeto de caras planas y convexo se efectúa a partir de una secuencia de imágenes tomadas a medida que el objeto rota. La descripción se realiza en términos de las caras del objeto y sus interconexiones. El proceso de aprendizaje, definido como el cambio de una función experiencia, se lleva a cabo con las distintas vistas del objeto. A medida que se observan caras del objeto, que hasta ese momento habían quedado ocultas, éstas se etiquetan y son aprendidas como tales.

Nagel en [NAGE-83], presenta una panorámica de los distintos campos de aplicación que han dado lugar a lo que actualmente se conoce como análisis de escenas dinámicas, así como de las técnicas y estrategias empleadas en la descripción de la estructura y movimiento espacial de objetos mediante una secuencia de imágenes ordenadas en el tiempo.

En la literatura se han propuesto diferentes estrategias teniendo todas ellas en común la búsqueda de las condiciones mínimas bajo las cuales es posible determinar la unicidad en la interpretación de la estructura 3D de un conjunto de puntos así como en el movimiento que de ellos se infiere. Las condiciones bajo las cuales Ullman [ULLM-79] construye un método de reconstrucción de la estructura 3D de un objeto a partir del conocimiento de la correspondencia entre puntos característicos en diversas imágenes, se pueden resumir en el

siguiente teorema:

Dadas tres proyecciones ortogáficas distintas de cuatro puntos no coplanares de una configuración rígida, la estructura y movimiento compatibles con las tres vistas, quedan determinadas de manera única.

En 1980, Neuman [NEUM-80] estudia las interpretaciones de movimiento y profundidad compatibles con dos vistas distintas de un único punto, no estando garantizada la unicidad de la solución. En este caso la proyección empleada fué la paralela u ortográfica.

Roach y col. [ROAC-80] muestran un sistema que precisa del conocimiento de la posición y correspondencia de, al menos, cinco puntos en dos imágenes para determinar la relación espacial entre ellos. Las condiciones impuestas son la de rigidez en el objeto a reconstruir, proyección central en la adquisición de imágenes y que los cinco puntos no estén en el mismo plano. La solución se expresa en términos de un sistema de ecuaciones no lineales cuya resolución requiere de métodos iterativos. La gran sensibilidad del sistema frente a perturbaciones o errores cometidos en la localización en cada imagen de los puntos característicos impone, en apliaciones prácticas, el uso de un mayor número de puntos.

Bajo iguales consideraciones, Nagel y col. [NAGE-81a, 81b] obtienen elegantemente, y de forma compacta, un sistema no lineal de ecuaciones conteniendo los parámetros específicos de la rotación sufrida por un objeto entre dos imágenes consecutivas. Además, ponen de manifiesto que estas ecuaciones contienen como caso particular a las obtenidas por Ullman para la interpretación de los movimientos de traslación y rotación de un objeto a partir del estudio de cinco puntos en tres vistas obtenidas bajo proyección central.

Más recientemente, Tsai y Huang [TSAI-84] tras obtener las ecuaciones no lineales que relacionan las coordenadas (x,y,z) de puntos sobre la superficie de un objeto con las correspondientes (X,Y) en el plano imagen, y con los parámetros de movimiento (traslación y rotación) entre dos imágenes, suministran una demostración rigurosa

para la unicidad de solución, respondiendo a las dos cuestiones siguientes:

¿Cuántos puntos de correspondencia son suficientes para garantizar la unicidad de solución?.

¿Cómo es posible evitar los métodos iterativos para la resolución del sistema de ecuaciones no lineales?.

En este trabajo la demostración al problema de unicidad se realizó considerando proyección central en contraposición a la aproximación de proyección paralela realizada por Ullman, indicándose además que los parámetros de movimiento pueden determinarse de manera única si se conoce la correspondencia de, al menos, ocho puntos cualesquiera en dos imágenes. Este método es ligeramente diferente del de Nagel [NAGE-81a] y Longuet [LONG-81] en el sentido de que estos últimos no consideran el problema de unicidad para los parámetros de movimiento. Por otra parte, en lugar de usar métodos iterativos para la resolución del sistema de ecuaciones no lineales como en [ROAC-80], determinan los parámetros de movimiento aplicando la técnica de mínimos cuadrados a la resolución de un sistema de n ecuaciones con 8 incógnitas (n > 8). Finalmente, los parámetros anteriormente calculados se usan para determinar la estructura 3D evaluando la distancia z de separación de cada punto del objeto al plano imagen.

Como resulta evidente, la estructura de los objetos en estudio obtenida a partir de un conjunto reducido de puntos característicos de su superficie es incompleta. La información adicional para una descripción más detallada se obtiene identificando más grupos de puntos sobre el objeto en movimiento.

Los sistemas anteriores, que usan múltiples imágenes, así como otros presentes en la literatura [ASAD-83, 84], están basados en el conocimiento de la correspondencia entre puntos característicos para las diferentes imágenes. Desde este punto de vista se presupone que los objetos, ya sean rígidos o partes rígidas articuladas [WEBB-81], presentan características que son fácilmente detectables y, más

importante aún, que cada característica es identificable en la secuencia de imágenes.

No obstante, conviene indicar que el empleo de estrategias basadas en la localización de puntos característicos puede resultar problemático por diversos factores. En primer lugar la detección de tales características en la secuencia de imágenes se ve dificultada por los cambios de iluminación producidos por el movimiento relativo cámara-ordenador. Así mismo, la aparición y desaparición de puntos sobre un objeto debido a oclusiones [MART-81a, 81b], [AGGA-81b] producidas po otros objetos o autooclusiones causadas por el propio movimiento objeto-cámara es un problema frecuente al especificar la correspondencia.

Para la reconstrucción de la superficie de objetos 3D, Tsai [TSAI-83] propone un método, basado también en el uso de múltiples imágenes tomadas en el mismo instante de tiempo. La alternativa al problema de correspondencia consiste en obtener curvas monodimensionales a partir de las distintas vistas. La localización de los extremos globales de las mismas, proporciona de modo satisfactorio la posición de puntos que corresponden a la proyección de un mismo punto 3D sobre la superficie a reconstruir.

A pesar de las aproximaciones propuestas a este problema de la correspondencia, su solución se presenta difícil, siendo necesarios métodos más rápidos y eficientes. Una revisión actualizada de la problemática en los procesos implicados en la percepción del movimiento a partir de información visual, y de sus aplicaciones, se muestra en [MART-88].

Resumiendo, se puede afirmar que las técnicas de estereovisión tienen como limitación inherente el difícil problema de la correspondencia, y de alguna manera cabría pensar que las técnicas basadas en sensores activos surgieron justamente como un posible medio de evitar esa dificultad.

Aunque los métodos activos basados en luz estructurada, sensores láser, ultrasonidos, etc., son útiles en dominios fuertemente controlados tales como automatización de procesos industriales, su uso es claramente inapropiado en problemas más generales de visión por máquinas. Así por ejemplo, en la actualidad se están desarrollando sistemas de visión para robots usando imágenes adquiridas por varias cámaras [LUH-85].

1.4 RECONOCIMIENTO DE FORMAS

En las tres últimas décadas el interés en el desarrollo de métodos teóricos y experimentales en reconocimiento de formas y visión artificial ha merecido una atención considerable. Según [FU-84], y en un amplio sentido, el reconocimiento de formas trata fundamentalmente del análisis y descripción de los datos obtenidos a partir de procesos físicos o mentales. En todo proceso de reconocimiento se distinguen tres etapas:

- Alguna clase de preproceso para reducir el ruido y redundancia en los datos, de manera que se asegure una descripción eficiente y efectiva.
- La representación de cada forma en términos de un conjunto de características (numéricas o simbólicas) y de las relaciones existentes entre ellas.
- El análisis (clasificación y/o descripción) en base a la representación anterior.

La elección de un conjunto apropiado de datos evaluados sobre cada forma es primordial en orden a efectuar un buen reconocimiento. Así el reconocimiento de formas implica tanto el análisis de las características como el diseño de un sistema de reconocimiento.

Los métodos propuestos para resolver el problema expuesto pueden agruparse en dos grandes categorias: Métodos Estadísticos basados en

teoría de la decisisón y Métodos Estructurales o Sintácticos [FU-82]. Un esquema global de un sistema de reconocimiento típico desde este punto de vista, se muestra en la Fig. 1.9.

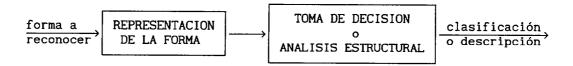


Fig. 1.9. Sistema general de reconocimiento de formas.

En la primera etapa, la mayor dificultad es la selección del esquema de representación más apropiado a cada tipo particular de problema. En la etapa final se trata de elegir una medida de similitud.

En la aproximación estadística cada forma se representa mediante un vector de N características, expresándose la medida de similitud en términos de una métrica, una función de probabilidad o una función discriminante.

En la aproximación sintáctica una forma se representa mediante una cadena, un árbol o un grafo de primitivas y sus relaciones. Aquí se establece una analogía entre la estructura de una forma y la sintaxis de un lenguaje siendo el proceso de decisión, generalmente, un proceso de análisis.

1.4.1 METODOS ESTADISTICOS

Los elementos principales en el proceso de reconocimiento de formas empleando este tipo de métodos son: extracción y selección de características, clasificación y aprendizaje (Fig. 1.10).

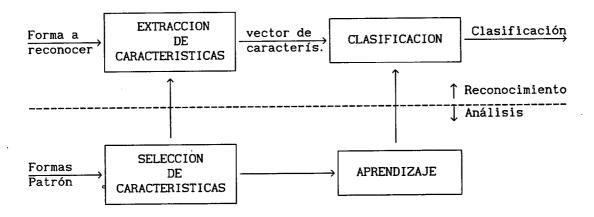


Fig. 1.10. Diagrama de bloques en reconocimiento de formas mediante métodos estadísticos.

Las características a extraer y el proceso requerido en su obtención son dependientes del dominio de aplicación ya que en diferentes problemas de reconocimiento se pueden usar características distintas. Sin embargo, los métodos de clasificación pueden ser independientes del tipo de problema a resolver. Así por ejemplo, en el reconocimiento de piezas industriales es frecuente el empleo de características geométricas (area, perímetro,...) mientras que en teledetección resulta útil el empleo de medidas multiespectrales. Otras características empleadas se basan en el análisis de texturas, extracción de contornos, etc.

1.4.1.1 EXTRACCION Y SELECCION DE CARACTERISTICAS

Generalmente los métodos de extracción y selección de características siguen alguna de las dos aproximaciones siguientes:

a) Transformación del espacio de características.

En esta aproximación el espacio original de características se transforma en uno de baja dimensión con el fin de facilitar la

representación y discriminación entre clases. Es usual utilizar criterios basados en la minimización del error cuadrático medio o criterios de entropia para la elección de la mejor transformación. La discriminación entre clases se logra maximizando la distancia entre clases o minimizando las distancias dentro de cada clase, o mediante combinación de los dos criterios simultáneamente. Transformaciones frecuentemente empleadas son las de Fourier, Haar y Wals-Hadamar entre otras. A menudo se utiliza el método de las componentes principales o desarrollos de Karhunen-Loeve para reducir la dimensión del espacio de características.

Generalmente, para una mejor separación entre clases es preferible usar transformaciones no lineales, lo que redundará en estructuras más simples en la clasificación. Sin embargo la implementación de tales transformaciones, frente a las lineales, exige un mayor coste computacional.

b) Medidas de distancia.

El objetivo perseguido en esta aproximación es la selección de un subconjunto de L características a partir de N características (L < N) de forma que la actuación del sistema de reconocimiento no se vea seriamente afectada, es decir sin un aumento considerable en la probabilidad de clasificación errónea, o más generalmente del riesgo asociado a la toma de decisiones. Debido a la dificultad de evaluar directamente la probabilidad de reconocimiento erróneo, es usual determinar criterios que sirvan de base para la selección de características. El método más común consiste en definir una medida de distancia relacionada con los límites superior y/o inferior aceptados para la probabilidad de reconocimiento erróneo. El subconjunto de características se elige de manera que la distancia así definida alcance un máximo o mínimo.

En cualquiera de las aproximaciones anteriores a) y b) la selección óptima de características es difícil de conseguir a menos que se realice una búsqueda exhaustiva.

1.4.1.2 CLASIFICACION Y APRENDIZAJE

El reconocimiento de formas mediante métodos estadísticos basados en teoría de la decisión involucra procesos de clasificación y aprendizaje. Mediante el término clasificación se hace referencia a una partición del espacio de características en regiones mutuamente excluyentes, estando asociada cada región con una clase de formas. El análisis discriminante se ha usado con éxito [ROME-73] en el proceso de clasificación.

Si las funciones densidad de probabilidad condicional son conocidas, o pueden estimarse, puede usarse un clasificador bayesiano para minimizar el riesgo de asignación errónea.

En aquellas situaciones en las que no se conozcan las funciones densidad para cada clase, será preciso hacer uso de esquemas de clasificación no paramétrica.

En la práctica, cuando se dispone de un elevado número de muestras de formas conocidas, es posible estimar o "aprender" las funciones densidad de probabilidad. En estos casos sólo será necesario aprender los parámetros de cada clase de las que componen la muestra de aprendizaje, siendo posible alcanzar clasificaciones óptimas. Por el contrario, si el número de formas de la muestra de aprendizaje es pequeño la estimación de los parámetros de las funciones densidad será pobre.

En entornos no supervisados, en los que el número de clases distintas de formas es desconocido, una aproximación práctica al problema de clasificación la constituye el análisis de agrupamientos o "clustering". El objetivo es clasificar los datos experimentales en grupos, de manera tal que las muestras de cada grupo resulten tan similares como sea posible, y que muestras de grupos distintos sean lo más distantes posibles. Esto exige diversas medidas de similitud y distancia así como criterios para agrupar las formas en el espacio de carcaterísticas [CHEN-73].

En ocasiones se da la circunstancia de poder discriminar entre formas distintas en base a una, o unas pocas características. No obstante estas características pueden no ser las mismas para cada pareja de formas¹. En tales casos es útil usar un sistema de reconocimiento estratisficado o multinivel. Los clasificadores basados en esquemas de árboles de decisión constituyen un intento de solución al problema de discriminación entre clases usando un número reducido de características en los diferentes estratos o niveles. En aplicaciones simples la elaboración del árbol puede efectuarse a mano, sin embargo el diseño óptimo de tales estructuras en problemas más complicados puede resultar computacionalmente complejo [WANG-86].

1.4.2 METODOS ESTRUCTURALES O SINTACTICOS

El reconocimiento de formas complejas requiere, generalmente, un elevado número de características por lo que los métodos estadísticos pueden resultar poco efectivos o computacionalmente inviables. Una alternativa consiste en representar una forma mediante subformas o entidades constituyentes más simples teniendo además en cuenta las relaciones entre las partes de la forma original. Este proceso puede aplicarse repetidamente a cada subforma hasta que los constituyentes últimos o primitivas puedan ser tratados fácilmente por métodos estadísticos.

Este procedimiento de descomposición de las formas en sus primitivas se conoce como aproximación sintáctica o estructural. Un diagrama de un sistema de reconocimiento basado en esta estrategia se muestra en la Fig. 1.11. En este diagrama se distinguen dos secciones: análisis y reconocimiento. El análisis incluye la selección de primitivas en función de las que se inferirá la relación o gramática estructural para describirlas y relacionarlas. En la parte de reconocimiento propiamente dicho se representa cada nueva forma en

¹ Sobre base, DOL e jemplo, es posible distinguir un águila de un perro en función de que la primera tiene alas y el segundo no. mientras que un perro y una oveja podrían distinguirse atendi endo hecho de que el perro ladra y la oveja bala.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2004

términos de sus primitivas y relaciones entre ellas, para después ser analizadas aplicando las reglas gramaticales inferidas en la sección previa a partir de formas conocidas.

En los métodos sintácticos, una forma se representa mediante una sentencia (una cadena o un árbol) en un lenguaje concreto regido por una gramática. Al lenguaje empleado se le llama lenguaje de descripción, mientras que al conjunto de reglas que gobiernan la concatenación de primitivas en dicho lenguaje se le llama gramática de la forma.

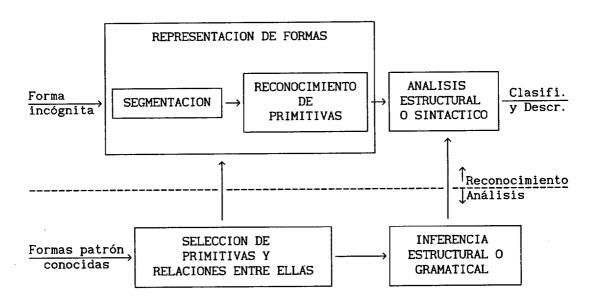


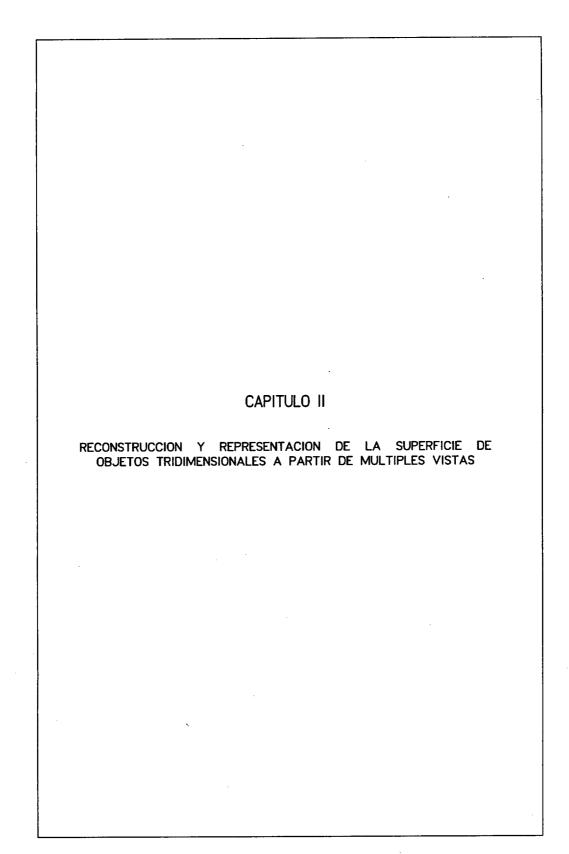
Fig. 1.11. Diagrama de bloques de un sistema de reconocimiento empleando la aproximación estructural.

Los aspectos más críticos bajo este esquema pueden resumirse en dos. El primero es la elección de las primitivas de la forma, siendo deseable que las relaciones entre ellas sean simples. El segundo problema surge al tratar con lenguajes dependientes del contexto en el análisis de formas para las que se ha empleado una descripción estructural jerárquica (estructuras de árboles y grafos). Avances considerables se han logrado al combinar atributos semánticos con los sintácticos o estructurales en la descripción de formas [FU-84].

Puesto que las primitivas de una forma son los componentes básicos de la misma, cabría esperar que su extracción fuese sencilla. Desafortunadamente este no es el caso en situaciones prácticas. Así por ejemplo, en la escritura manual los trazos son considerados como primitivas adecuadas, mientras que para el discurso hablado lo son los fonemas; sin embargo ni unos ni otros son fácilmente extraibles de manera automática. Se puede decir que en el momento actual no existe una solución general para la selección de primitivas.

Idealmente, el sistema de inferencia gramatical debería construir una gramática o reglas de descripción estructural a partir de un conjunto dado de formas. En la práctica tal tipo de máquina no está disponible excepto para problemas y entornos fuertemente restringidos. Lo usual es que el diseñador construya, bien manual o interactivamente, su propia gramática usando la información disponible a priori y su propia experiencia.

Desde el punto de vista conceptual, la técnica más simple de reconocimiento sintáctico es la de enmascaramiento, que consiste en comparar la sentencia que describe a la forma de entrada con la que describe a cada prototipo de referencia. De esta manera, la forma de entrada se asigna a aquella clase con cuyo prototipo presenta la mejor similitud. En aquellas situaciones en las que se necesite de una descripción completa de la forma, la técnica anterior no es viable, siendo necesario un análisis sintáctico más detallado.



II.1 INTRODUCCION

La recuperación de información tridimensional a partir de imágenes bidimensionales es una importante y común tarea en análisis de escenas y visión artificial. Múltiples vistas de un objeto desconocido pueden usarse para determinar su forma tridimensional. Esto es de particular interés en el análisis de escenas dinámicas [HUAN-81], consistente en el procesado de una secuencia de imágenes ordenadas en el tiempo, con el propósito de determinar no sólo las componentes más notables de la escena sino también los cambios y movimientos de objetos presentes en la misma.

Por otra parte, la aplicación tradicional de técnicas de visión en robótica ha estado casi exclusivamente orientada a la detección de eventos mediante análisis de imágenes 2D. Sin embargo, en la actualidad es de sumo interés el análisis de escenas 3D [NITZ-88]; así por ejemplo, la descripción del entorno de un robot es primordial para facilitar el movimiento de éste.

La información contenida en una imagen es variada, y dependiendo del tipo de sensor empleado, información a analizar, estrategia seguida, etc., pueden encontrarse una gran variedad de técnicas de recuperación de la estructura visible de objetos presentes en una escena [ROSE-86].

Uno de los métodos más comunes empleando múltiples vistas obtenidas mediante sensores pasivos es el de estereovisión [MARR-78b], [GRIM-81], [BARN-82]. La disparidad existente entre las proyecciones de un mismo punto 3D en las distintas vistas, se emplea para determinar la localización espacial de dicho punto. No obstante, la correspondencia entre puntos en dos o más imágenes es difícil de establecer, siendo el proceso muy sensible a los errores introducidos en la digitalización y calibrado de las cámaras [AGGA-81a], [BLOS-87], [BOYE-88].

Una vez recuperada la información tridimensional de una escena, ya sea con fines de visualización en una pantalla gráfica o para

análisis automático de imágenes, se hace necesaria una representación adecuada de la misma. Distintos esquemas se han propuesto, pero fundamentalmente éstos pueden agruparse en dos grandes categorias: representaciones basadas en la superficie У representaciones volumétricas. La diferencia entre ambos esquemas radica en el tipo de primitivas empleado. El tipo de representación elegido en cada caso dependerá del sistema de adquisición de datos, y en última instancia de la aplicación para la que se haya diseñado. Así por ejemplo, es usual el empleo de representaciones superficiales en modelado de objetos y sistemas CAD/CAM. Las representaciones volumétricas serán apropiadas en la determinación de propiedades tales como volumen de un objeto, peso, disposición interna de sus componentes, etc., siendo frecuentemente empleadas en aplicaciones médicas.

En este trabajo estamos interesados en determinar la estructura 3D de un objeto a partir de los contornos de múltiples vistas bidimensionales del mismo. La técnica empleada se basa en el método de interseccción de volúmenes usando proyección ortográfica. A diferencia de otras estrategias que también emplean múltiples vistas, este procedimiento no necesita del establecimiento de la correspondencia entre características observadas en la secuencia de imágenes. Para simplificar los desarrollos matemáticos, se ha considerado que la proyección del objeto 3D sobre cada plano imagen se efectúa mediante líneas paralelas al eje de visión de la cámara, es decir proyección ortográfica o paralela. Sin embargo esta consideración es razonable en muchas situaciones, y puede demostrarse que si la distancia cámara-objeto es relativamente mayor que la distancia focal de la cámara entonces la proyección paralela puede interpretarse como una buena aproximación de la proyección más real tipo perspectiva. La generalización de este procedimiento para otro tipo de proyecciones necesitaría previamente de la construcción de un modelo geométrico del sistema de adquisición que transforma el objeto 3D en una imagen 2D.

El procedimiento de reconstrucción tridimensional expuesto en este capítulo, fue desarrollado inicialmente por Martin y Aggarwal [MART-81a]-[MART-83] y consta de dos etapas perfectamente diferenciadas: una reconstrucción inicial a partir de dos contornos y

un posterior refinamiento de la anterior reconstrucción. La segunda etapa ha sido modificada por nosotros y convenientemente adaptada a nuestros propósitos, manteniendo inalterada la idea que lo fundamenta. Los algoritmos, y programas que los desarrollan, los hemos realizado en su totalidad a partir de los trabajos precedentes.

En el trabajo de Martin y Aggarwal el objeto 3D reconstruido se describe mediante una representación volúmica a partir de las líneas de contorno que aproximan sucesivas secciones planas paralelas del objeto real. En nuestro caso se ha optado por una representación superficial, usando como primitivas caras planas triangulares que conectan entre sí las distintas líneas de representación, y la estructura de datos asociada, resulta más apropiada para evaluar las características empleadas en el proceso de reconocimiento -momentos 3D-, y que serán definidas en el capítulo III de esta memoria. Representaciones volumétricas alternativas a la inicial de Martin y Aggarwal han sido posteriormente desarrolladas [CHIE-84], [KIM-86] empleando "octrees" paralelepípedos rectangulares como primitivas.

Obviamente, el contorno de una imagen es una información limitada acerca del objeto en observación, y por tanto también lo será el alcance del sistema presentado. Resulta claro que de los contornos de proyecciones ortográficas de un objeto, no será posible resolver cierto tipo de concavidades presentes en su superficie. En particular, aquellos puntos de la superficie del objeto para los que cualquier línea, contenida en su plano tangente y que pase por el punto en cuestión, intercepte con el objeto en algún punto de su interior, no podrán ser resueltos. Por otra parte, en algunos casos, y dependiendo de las peculiaridades del objeto observado, el número de vistas distintas necesarias para alcanzar una aproximación razonable a su superficie real puede ser elevado, siendo también un aspecto crítico la elección de la posición y orientación de cada cámara. No obstante, el tipo de formas tridimensionales que pueden ser reconstruidas mediante este procedimiento es suficientemente extenso como para justificar su uso.

El resto de este capítulo se organiza como sigue. En la sección 2 se presenta el algoritmo de reconstrucción propiamente dicho, junto a un ejemplo ilustrando la actuación del mismo. En la sección 3 se presentan, de forma resumida, distintas alternativas en la representación de formas tridimensionales. Finalmente, en la sección 4, se describe el problema de triangulación consistente en crear una representación superficial de un objeto 3D a partir de los contornos de distintas secciones planas consecutivas del mismo; las primitivas utilizadas son caras planas triangulares.

II.2 ALGORITMO DE RECONSTRUCCION

Al hablar de imagen de un objeto hacemos referencia a la proyección ortográfica del mismo sobre el plano imagen. En esta imagen se distinguen claramente dos regiones; por un lado, la silueta del objeto proyectado y por otra parte todo lo que no es silueta, también llamado fondo.

Los puntos que delimitan la frontera entre la silueta y el fondo definen el contorno de ésta [MARR-77]. Así pues, los puntos del contorno se caracterizan por poseer vecinos de dos clases:

- Vecinos que pertenecen a la silueta propiamente dicha.
- Vecinos que pertenecen al fondo.

La orientación de cada vista, o imagen, vendrá especificada por un punto y tres vectores mutuamente perpendiculares que forman un sistema coordenado a derechas, con el punto anterior como origen. El eje Y de cada sistema coordenado se considera en la dirección de la línea de visión de cada cámara, siendo ZX el plano imagen (Fig. 2.1).

La silueta de un objeto se proyecta sobre el plano imagen mediante líneas paralelas al eje Y. De esta forma, el contorno asociado a cada vista viene generado por líneas paralelas al eje de visión y que intersectan con el objeto pero solamente en puntos que pertenecen a su superficie y no al interior.

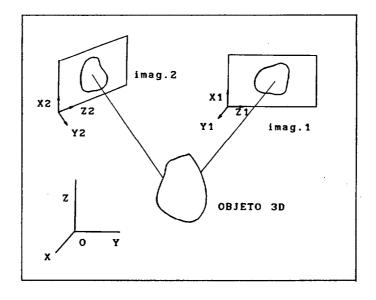


Fig. 2.1. Localización y orientación espacial de dos imágenes respecto del sistema coordenado global OXYZ.

Claramente, cada línea generadora del contorno es tangente, al menos, a un punto de la superficie del objeto. El problema que se plantea es determinar qué puntos sobre esta línea son también puntos de la superficie del objeto. Este problema se resuelve usando las restricciones impuestas por las líneas generadoras del contorno de una segunda y consecutivas imágenes.

El algoritmo de reconstrucción de la estructura tridimensional de un objeto a partir de diferentes vistas del mismo, se desarrolla en dos etapas. La primera combina la información de dos vistas con ejes de visión no paralelos para formar una reconstrucción inicial aproximada de la estructura 3D del objeto. En la segunda etapa se procesa la información de subsecuentes imágenes para refinar la reconstrucción inicial. Así, estas dos etapas analizan los distintos contornos para inicialmente construir y posteriormente refinar la representación 3D del objeto generador de esos contornos. El resultado es una superficie dentro de la cual debe de estar el objeto generador de los distintos contornos.

II.2.1 RECONSTRUCCION INICIAL

Para comprender cómo los contornos de dos imágenes, especificadas por su origen y orientación, pueden usarse para construir una primera aproximación del objeto 3D, consideremos la Fig. 2.2. Aquí se representan dos imágenes y sus contornos, cada una en su correspondiente plano imagen. Junto a estas dos vistas, se han dibujado un conjunto de líneas generadoras de contornos para cada una de ellas.

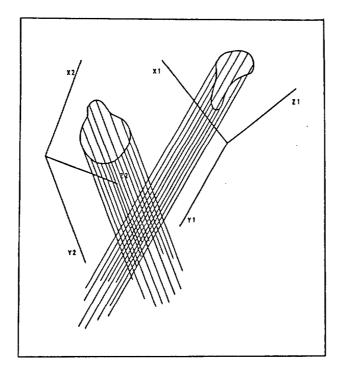


Fig. 2.2. Contornos y líneas generadoras de contornos para dos vistas distintas.

Para una vista dada, el conjunto de líneas generadoras de ese contorno definen un volumen que encierra en su interior al objeto. Evidentemente este volumen se extiende indefinidamente en la dirección del eje Y, por lo que no constituye una descripción adecuada del objeto. La segunda vista también define un volumen infinito que contiene en su interior al objeto, y si las líneas de visión para estas vistas son distintas, entonces la intersección de los dos volúmenes infinitos delimita otro volumen que circunda al objeto real generador de esos contornos (Fig. 2.3). Es este volumen resultante de la intersección el que constituye la reconstrucción inicial de la estructura 3D del objeto.

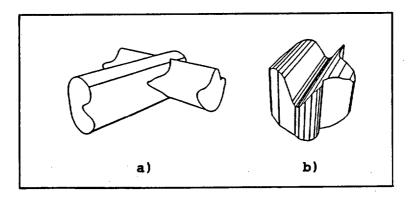


Fig. 2.3. a) Proceso de intersección basado en las siluetas de dos imágenes bajo proyección ortogonal. b) Aproximación resultante de la anterior intersección.

En la Fig. 2.2 se observa que cada línea generadora del contorno para una de las vistas, sirve para delimitar la extensión sobre la cual puntos de alguna otra línea generadora del contorno para la otra vista pueden ser también puntos de la superficie del objeto.

Obviamente, el número de líneas con las que se puede operar ha de ser finito y por tanto éstas se han de elegir de forma conveniente, como se indica en el apartado siguiente.

II.2.1.1 DETERMINACION DE LINEAS GENERADORAS DE CONTORNOS

Para clarificar el criterio de elección de las líneas generadoras de contornos, así como las mutuas restricciones que ejercen entre sí

estos dos conjuntos de líneas, fijémonos en la Fig. 2.4. De nuevo, se muestran dos contornos en sus respectivos planos imagen con las líneas de visión alineadas según los ejes Y1 e Y2 de cada cámara. Si los planos imagen (Z1,X1) y (Z2,X2) no son paralelos, deben intersectar determinando una línea que, en lo sucesivo, denominaremos "línea común". La dirección de esta línea se especifica mediante el producto vectorial de los ejes Y1 e Y2, siendo por tanto perpendicular a ellos.

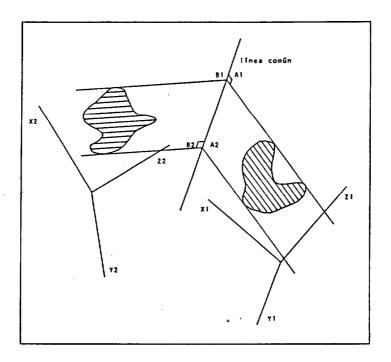


Fig. 2.4. Descripción de cada contorno mediante segmentos de rasterización perpendiculares a la línea común.

Cada uno de los contornos puede proyectarse ortogonalmente sobre la línea común. Mediante trazos continuos se representan las líneas que delimitan la proyección de cada contorno sobre la línea común.

Para distinguir entre las líneas generadoras de contornos y las líneas de proyección de cada contorno sobre la línea común, cada vez que hagamos referencia a estas últimas usaremos el término "rayo".

La propiedad fundamental que permite, a partir de estos contornos, construir una aproximación inicial a la estructura 3D del objeto es que los rayos límites para ambos contornos proyectan justamente sobre los mismos puntos de la línea común; es decir, A1 = B1, A2 = B2, siendo A1, A2 y B1, B2 las proyecciones respectivas de estos rayos para las vistas 1 y 2.

Teniendo en cuenta que la proyección de cada contorno sobre la línea común es ortogonal, se puede afirmar que dos de estos rayos (uno para cada contorno) que se encuentren en un mismo punto de la línea común, determinan un plano. Este plano es perpendicular a la línea común y contiene a las líneas generadoras de contornos correspondientes a los puntos en que ambos rayos interceptan con sus respectivos contornos.

Estas líneas generadoras de contornos, y otras igualmente obtenidas, son las que nos van a permitir construir una primera aproximación a la forma 3D del objeto.

El contorno de cada imagen se representa mediante una línea poligonal. Algoritmos eficientes para su determinación pueden encontrarse en [PAVL-80,81], [MIYA-84].

Con el fin de determinar las líneas generadoras de contornos y agruparlas en líneas coplanares, la representación inicial de cada contorno se transforma en otra representación equivalente conocida como "área rasterizada" [NEUM-79], [MART-81b]. El área delimitada por cada uno de los contornos se descompone en un conjunto de segmentos rectilíneos resultantes de la intersección de una familia de líneas rectas, paralelas y equiespaciadas, con el polígono que define cada contorno.

La dirección de estas familias de líneas rectas, o dirección de rasterización, se elige de forma que, para cada contorno, sean perpendiculares a la línea común y estén contenidas en el plano imagen ZX.

Esta nueva representación del área asociada a un contorno, comprende un conjunto ordenado de segmentos rectilíneos, cada uno de los cuales se especifica mediante las coordenadas (z,x) de sus puntos extremos, correspondiendo éstos a la intersección de cada línea de rasterización con el contorno. En la Fig. 2.4 se representan los segmentos de rasterización asociados a cada contorno, observándose también la perpendicularidad de éstos respecto a la línea común.

Los puntos extremos de cada segmento de rasterización determinan los puntos en los que se apoyan las líneas generadoras de contornos usadas para construir una aproximación inicial a la estructura 3D del objeto. Como se puede observar, las líneas generadoras de contornos elegidas de esta forma están clasificadas en función de:

- contorno al que corresponden.
- plano al que pertenecen (cada plano se asocia a una línea de rasterización o bien a un punto de la línea común).
 - segmento de rasterización a que están referidas.

II.2.1.2 INTERSECCION DE LAS LINEAS GENERADORAS DE CONTORNOS PARA DOS VISTAS DISTINTAS

Una vez especificado el criterio para elegir convenientemente las líneas generadoras de contornos, veamos como imponer las mutuas restricciones entre ambos grupos de líneas para construir la aproximación inicial a la forma 3D del objeto.

Centremos nuestra atención en un punto de la línea común (Fig. 2.5) en el que confluyan dos líneas de rasterización, una para cada uno de los contornos. Asociado a este punto hay definido un plano cuya normal es, precisamente, la dirección de la línea común.

La intersección de este plano con el objeto real determinará un corte formado por una o varias secciones disjuntas. Lo dicho para este plano es igualmente válido para todos los demás planos asociados al resto de líneas de rasterización que confluyen en alguno de los puntos

del tramo de línea común delimitado por las rayos límites de cada contorno.

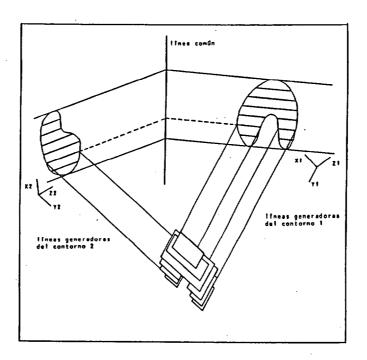


Fig. 2.5. Estructura de paralelogramos asociada a la reconstrucción inicial a partir de dos vistas distintas.

Es nuestro objetivo encontrar una aproximación poligonal a cada una de estas secciones. Esta aproximación se entiende en términos de una acotación externa en el sentido de que cada polígono circunscriba a la sección en cuestión. El conjunto de estos polígonos constituirá la reconstrucción 3D del objeto.

En la etapa inicial de reconstrucción, los polígonos anteriores serán paralelogramos. La estructura de paralelogramos para uno cualquiera de estos planos, se obtiene como sigue.

Asociadas a cada plano, existen dos líneas de rasterización, una para cada contorno. Notemos N1 al número de segmentos de rasterización para la línea del primer contorno y N2 para la línea equivalente del segundo contorno. Sobre los puntos extremos de cada uno de estos segmentos se apoyan dos líneas generadoras del contorno cuya dirección es perpendicular a cada plano imagen. La intersección de cada dos líneas de los N1 segmentos del primer contorno con cada dos líneas generadoras del contorno para los N2 segmentos de la otra vista, determina un conjunto de N1xN2 paralelogramos contenidos en ese plano. Cada una de las intersecciones delimita, sobre cada línea generadora del contorno, la extensión sobre la cual <u>pueden</u> haber puntos que, perteneciendo a esa línea, también pertenezcan a la superficie real del objeto a reconstruir.

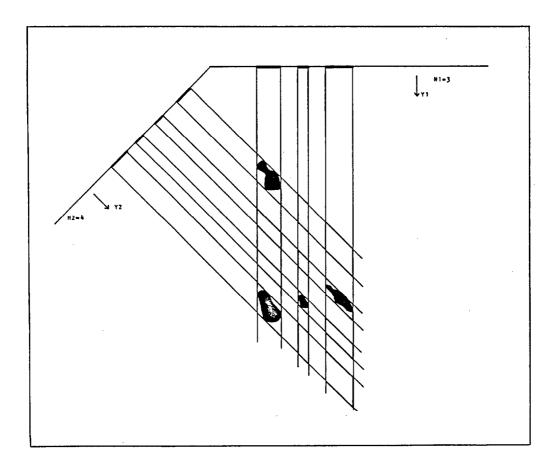


Fig. 2.6. Estructura de paralelogramos obtenida a partir de dos líneas de rasterización, una de cada contorno. Para estas líneas, el primer contorno consta de N1=3 segmentos y el segundo de N2=4 segmentos.

El número de paralelogramos generados es 12, mientras que las secciones reales del objeto (sombreadas) sólo son 4.

Conviene resaltar, que en esta primera etapa del proceso de reconstrucción, lo único que se puede afirmar es que las secciones reales del objeto, generadas por el plano en estudio, deben estar contenidas, al menos, en el interior de N de los N1xN2 paralelogramos, siendo N el máximo de los valores enteros N1 y N2 (Fig. 2.6).

Cada uno de los paralelogramos es una figura que limita la extensión sobre la que <u>puede</u> estar dispuesta la estructura tridimensional del objeto a reconstruir. Además, también se puede afirmar que la sección real del objeto debe tener, <u>al menos</u>, un punto en común con cada uno de los lados de los N paralelogramos.

A continuación se presenta un resumen del algoritmo usado para crear la reconstrucción inicial de la estructura tridimensional de un objeto.

ALGORITMO INICIAL DE RECONSTRUCCION 3D

- 1. Obtener la proyección ortográfica del objeto 3D sobre el plano imagen para la primera vista.
- 2. Determinar el número de vértices y coordenadas de los mismos para el polígono que define el contorno de la primera vista.
- 3. Obtener la proyección ortográfica del objeto 3D sobre el plano imagen para la segunda vista.
- 4. Determinar el número de vértices y coordenadas de los mismos para el polígono que define el contorno de la segunda vista.
- 5. Determinar la línea común a los planos imagen de la primera y segunda vistas.
- 6. Especificar, para cada contorno, los segmentos de rasterización con direcciones perpendiculares a la línea común.

7. Para cada dos líneas de rasterización que confluyan en un mismo punto de la línea común, determinar la estructura de paralelogramos.

De lo anteriormente expuesto es fácil deducir que la reconstrucción inicial, difícilmente describirá detalles de la superficie del objeto generador de las dos primeras vistas. Lo que se ha conseguido es acotar, mediante paralelogramos, una región del espacio en cuyo interior se encuentra localizado el objeto real.

II.2.2 ETAPA DE REFINAMIENTO

La segunda etapa del proceso de reconstrucción se completa al refinar o mejorar la reconstrucción inicial mediante la información contenida en los contornos de sucesivas vistas. El proceso de refinamiento se aplica a medida que cada nueva vista -después de las dos primeras- del objeto real es procesada.

Básicamente, el refinamiento de la reconstrucción inicial, y de sucesivas reconstrucciones, se traduce en ir creando nuevas estructuras tridimensionales que se aproximen cada vez más a la forma 3D real. Esta "mejor aproximación" se manifiesta en el hecho de que el número de aristas de los sucesivos polígonos que acotan cada sección real del objeto es una función monótona creciente con el número de vistas usadas en la etapa de refinamiento. El aumento en el número de aristas viene motivado por el hecho de que cada vez serán más los puntos de la sección real del objeto que también formen parte de las aristas de estos polígonos. Pasemos a describir los detalles de la etapa de refinamiento.

Cada vez que se adquiere una nueva vista del objeto real, debido a la proyección ortográfica de éste sobre el plano imagen, el nuevo contorno se representa mediante un polígono. Este contorno corresponde a la proyección de la silueta real del objeto que es visible desde la posición y orientación de cada nuevo plano imagen.

Como se indicó anteriormente, los elementos básicos o primitivas en la estimación de la estructura 3D son polígonos (paralelogramos en la reconstrucción inicial) contenidos en planos paralelos contiguos. El proceso de refinamiento se efectúa plano a plano, y dentro de cada plano, polígono a polígono.

El refinamiento de un polígono por la información suministrada por un nuevo contorno, consta de tres etapas sucesivas (Fig. 2.7).

A) En la primera, el polígono en cuestión se proyecta sobre el plano imagen de la nueva vista. Esta proyección también será ortográfica y según la línea de visión de la nueva imagen. La proyección del polígono se efectúa vértice a vértice, bajo la consideración, no siempre cierta, de que todos ellos son visibles desde el plano imagen de la nueva vista. En el interior del polígono proyectado debe estar contenida la proyección de la sección real del objeto asociada al polígono en estudio.

Teniendo en cuenta la posibilidad de que no toda la sección real del objeto sea visible desde la nueva vista, y que el nuevo contorno corresponde a la proyección de la silueta visible del objeto desde el nuevo plano imagen, se concluye que la proyección de la citada sección ha de estar situada en el interior del nuevo contorno.

B) Basándonos en estos argumentos, en la segunda etapa se determina la intersección del nuevo contorno con el polígono proyectado. Si la intersección fuese vacía, sería indicativo de que el polígono en estudio no contiene en su interior a parte alguna del objeto real, y por tanto debe ser descartado como parte integrante de la reconstrucción 3D. Si por el contrario la intersección fuera no vacía, el polígono o polígonos resultantes de la misma determinarían en sí mismos una mejor acotación a la sección real del objeto contenida en el polígono inicial. Si la intersección estuviese compuesta por más de un polígono, sería indicativo de que la sección real del objeto contenida en el polígono a refinar se compone de tantas subsecciones disjuntas como polígonos

resulten de la intersección.

C) Por último, y en la tercera etapa de este proceso, el polígono o polígonos resultantes de la intersección y contenidos en el plano de la nueva imagen, se vuelven a proyectar según la línea de visión, sobre el plano en el que estaba situado el polígono original sometido al refinamiento.

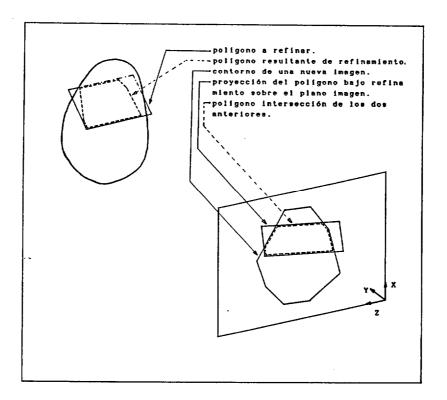


Fig. 2.7. Diagrama esquemático mostrando el proceso de refinamiento. Cada polígono aproximando una sección real del objeto 3D se proyecta sobre el plano imagen. La intersección de dicha proyección con el nuevo contorno se vuelve a proyectar sobre el plano que contiene al polígono sometido a refinamiento.

Un resumen del algoritmo empleado en la etapa de refinamiento, se muestra a continuación.

ALGORITMO DE REFINAMIENTO

- Obtener la proyección ortográfica del objeto 3D sobre el plano imagen de una nueva vista.
- Determinar el número de vértices y coordenadas de los mismos para el polígono que define el contorno de esta nueva vista.
- 3. Para cada uno de los polígonos que componen la reconstrucción 3D del objeto en estudio, actuar como sigue:
 - 3a. Proyectar ortogonalmente cada uno de sus vértices sobre el plano imagen de la nueva vista y según la dirección de su eje de visión.
 - 3b. Determinar la intersección del polígono proyectado con el contorno de la nueva imagen.
 - 3c. Proyectar cada uno de los polígonos resultantes de la intersección del punto 3b, ortogonalmente y según la dirección del eje de visión de la nueva vista, sobre el plano en el que estaba contenido el polígono en estudio.
 - 3d. Sustituir la información asociada al polígono en estudio por la relativa a los nuevos polígonos obtenidos en el punto 3c.
- Si se dispone de alguna nueva vista del objeto, volver al punto
 En caso contrario, finalizar el proceso.

Resumiendo, los datos tridimensionales de la estimación de la estructura 3D del objeto, mantienen el mismo formato a lo largo del proceso de refinamiento, siendo su contenido cada vez más cercano a los datos reales del objeto a reconstruir.

II.2.3 EJEMPLO DE RECONSTRUCCION

Para mostrar la actuación y eficacia del proceso de

reconstrucción 3D descrito anteriormente, se presenta un ejemplo en el que la superficie de un objeto conocido es reconstruida a partir de la información contenida en los contornos de varias de sus imágenes.

En la Fig. 2.8 se muestra, en perspectiva isométrica, el objeto 3D a reconstruir. La reconstrucción inicial mediante paralelogramos, junto a los contornos de las dos imagenes usadas en este proceso, se muestran en las Figs. 2.9 - 2.10.

Indiquemos que, para un objeto dado, el modelo 3D obtenido en la reconstrucción inicial depende de dos factores. En primer lugar, de cuales sean las dos vistas iniciales, y por otra parte de la separación entre los distintos planos que componen la reconstrucción.

Como se observa en la Fig. 2.10, la estructura de paralelogramos delimita pobremente la zona del espacio en la que está localizado el objeto. Obviamente, para una mejor acotación de esta zona espacial se hace necesario refinar la región definida por cada paralelogramo.

Las dos vistas iniciales determinan la dirección de la línea común, y por tanto la orientación espacial de los distintos planos o secciones que componen el modelo 3D aproximando al objeto real. Una vez especificada la posición y orientación del plano imagen para cada vista, el plano definido por las líneas de visión de las dos primeras especifica la orientación de las distintas secciones que componen el modelo 3D. La separación entre las distintas secciones paralelas la elige el usuario de manera tal que el modelo reconstruido sea descriptivo de las características más sobresalientes de la superficie del objeto a reconstruir. Este parámetro se especifica al descomponer los dos contornos iniciales en segmentos rectilíneos paralelos o segmentos de rasterización (Fig. 2.9).

Evidentemente, para una recuperación más precisa de detalles superficiales del objeto será necesario aproximar el modelo reconstruido mediante un mayor número de secciones o líneas de contorno, lo que por otra parte aumentará el tiempo de cálculo al tener que procesar un mayor número de polígonos en cada ejecución de

la etapa de refinamiento.

Para indicar cómo la información del contorno de cada nueva imagen va mejorando la reconstrucción de la estructura 3D del objeto, en las Figs. 2.11-2.14 se muestra cada nuevo contorno y la reconstrucción obtenida al aplicar el algoritmo de refinamiento.

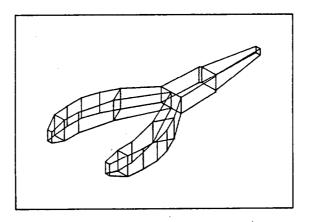


Fig. 2.8. Objeto 3D a reconstruir mediante el algoritmo descrito en la sección II.2.

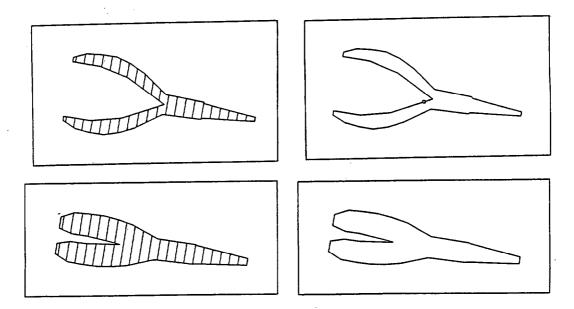


Fig. 2.9. Contornos, con líneas de rasterización, para las dos vistas usadas en el proceso de reconstrucción inicial.

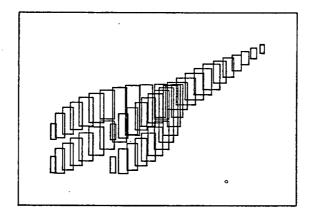
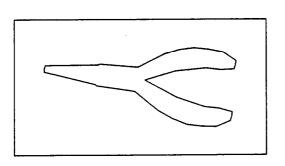


Fig. 2.10. Estructura de paralelogramos para la reconstrucción inicial del objeto de la Fig. 2.8 a partir de las vistas iniciales mostradas en la Fig. 2.9.



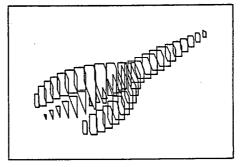
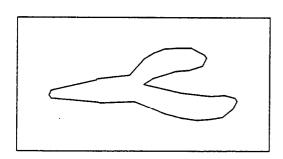


Fig. 2.11. Contorno para la tercera vista y reconstrucción obtenida al aplicar el algoritmo de refinamiento.



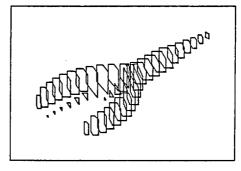
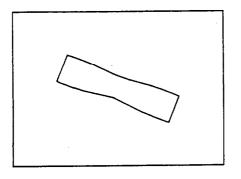


Fig. 2.12. Contorno para la cuarta vista y reconstrucción obtenida al aplicar el algoritmo de refinamiento.



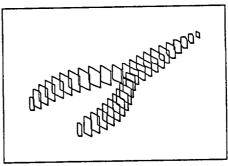
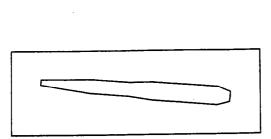


Fig. 2.13. Contorno para la quinta vista y reconstrucción obtenida al aplicar el algoritmo de refinamiento.



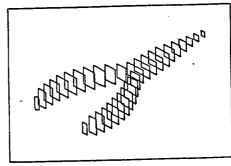


Fig. 2.14. Contorno para la sexta vista y reconstrucción obtenida al aplicar el algoritmo de refinamiento.

11.3 REPRESENTACION DE FORMAS 3D

La comprensión de estructuras tridimensionales por un ordenador requiere el empleo de modelos de objetos y algoritmos de recuperación de información tridimensional a partir de datos bidimensionales. La estructuración y representación de esta información es un tema de primordial interés y puede enfocarse desde dos puntos de vista distintos y a la vez relacionados.

En primer lugar, las representaciones comúnmente empleadas por las actuales técnicas de CAD (Computer-Aided-Design) [MIKH-78], [CARL-82] [OTA-85] están dirigidas, fundamentalmente, a conseguir refinadas visualizaciones de modelos geométricos 3D representativos de objetos reales. Este proceso debe estar precedido de una representación del modelo en el ordenador. El desarrollo del software en este campo es tal que las simulaciones alcanzan un elevado grado de realismo en lo referente a forma, color, sombreado, etc., de manera que la interpretación de estas escenas por parte de un observador es una tarea que comporta poco esfuerzo.

Por otra parte, desde el punto de vista de análisis de escenas y visión artificial por ordenador [WIST-75], [BALL-82], [FU-84], [SHIR-87], [BOYL-88], [FREE-88], [FAIR-88], el objetivo fundamental es poder descubrir, estructurar y simular el conocimiento que los humanos -y otros seres vivos- usan para así comprender e interpretar la información contenida en las imágenes.

En ambos casos se requieren procesos que permitan almacenar y representar información 3D. La representación elegida estará fuertemente condicionada por los fines y tipo de aplicación para la que haya sido diseñada.

En nuestro caso estamos interesados en la representación de formas tridimensionales con el propósito de efectuar tareas de reconocimiento.

En este contexto, Marr [MARR-78a] establece una serie de criterios y aspectos necesarios a la hora de diseñar una representación con tales fines. Siguiendo a Marr, distinguiremos entre forma, su representación y una descripción de la misma. Con el término forma haremos referencia a la geometría de la superficie física de un objeto. Una representación de esta forma será el esquema y conjunto de reglas usado para describirla. Al resultado de usar una representación determinada para describir una forma lo llamaremos descripción de la forma en esa representación.

II.3.1 CARACTERISTICAS DE LOS DISTINTOS ESQUEMAS DE REPRESENTACION

Entre los criterios a tener en cuenta al elegir un esquema de

representación y evaluar la eficiencia del mismo cabe destacar los de accesibilidad, campo de aplicación y unicidad, y estabilidad y sensibilidad.

La <u>accesibilidad</u> hace referencia a la capacidad de poder generar una descripción de una forma 3D a partir de la información proporcionada a la salida de los sensores (cámara TV, telémetros, etc.), dentro de los límites impuestos por el propio sistema de adquisición y el tiempo de proceso.

Puesto que no existe una única representación apropiada para describir todos los objetos, en el <u>campo de aplicación</u> se especifican las clases de formas para las que se ha diseñado una representación concreta. Si la representación se usa con fines de reconocimiento, será importante que la descripción de una forma sea única.

En el reconocimiento de formas sería deseable que el esquema de representación reflejara descripciones similares para formas también similares. Simultáneamente, la representación debe ser descriptiva de las diferencias existentes entre ambas formas. Estos son los criterios de estabilidad y sensibilidad respectivamente.

Los aspectos que distinguen una representación de otra son variados. Fundamentalmente pueden destacarse los siguientes:

- 1. Primitivas usadas.
- 2. Sistema de coordenadas elegido.
- 3. Operaciones que se pueden realizar sobre la representación.
- 4. Facilidad de conversión entre descripciones de una misma forma en representaciones distintas.
- 5. Organización de la información.

Para una discusión detallada de estas características véase [BADL-78], [MARR-78a], [AGGA-81c], [HEND-83]].

En el sistema de reconstrucción 3D descrito en la sección anterior los aspectos que más interesan en el esquema de

representación usado son los relativos a las primitivas y las operaciones que sobre ellas pueden efectuarse.

Las primitivas de una representación son las unidades de información más elementales usadas en la descripción de una forma. Atendiendo a este aspecto, pueden distinguirse básicamente dos amplios esquemas de representación:

- Representaciones basadas en la superficie.
- Representaciones volumétricas.

En el primer caso la descripción tridimensional de un objeto se especifica mediante la superficie que lo limita, y ésta se representa por un modelo geométrico conteniendo información acerca de los detalles del mismo. La superficie puede modelarse como una serie de puntos en el espacio tridimensional, como un conjunto de polígonos, descomponerse en zonas bien definidas analíticamente, etc., [NEUM-79], [HEND-83], [BESL-85], [QIUL-87], [FARI-88].

En las representaciones volumétricas [SRIH-80,81] se almacena información referente a la distribución espacial de esa forma. Las primitivas más comunmente usadas son cilindros, esferas, poliedros, cilindros generalizados, voxels, octrees, paralelepípedos, etc. [O'RO-79], [AGGA-81c], [KANA-83], [BESL-85], [KIM-86], [NOBO-88].

El tipo de operaciones a efectuar sobre una representación concreta puede clasificarse en dos grandes apartados:

- Operaciones de medida: área de la superficie, volumen [COOK-80,81c], normales a la superficie, etc..
- Operaciones de transformación: traslaciones, rotaciones, amplificaciones, deformaciones, etc..

II.3.2 DESCRIPCION SUPERFICIAL DE LAS RECONSTRUCCIONES 3D OBTENIDAS A PARTIR DE MULTIPLES VISTAS

En nuestro trabajo, la estructura tridimensional obtenida a

través de los contornos de múltiples proyecciones ortográficas de un objeto, la especificaremos en términos de su superficie, en contraposición a la descripción volumétrica usada por Martin y Aggarwal [MART-83] y otros autores [CHIEN-84], [KIM-86].

La razón para ello es doble. De una parte, como se indica en [WINS-75], [MARR-78a], [NISH-81], [BINF-81], la información contenida en la imagen de un objeto se debe, entre otros factores, a propiedades locales de su superficie visible. En segundo lugar porque la representación se ha elegido con fines de reconocimiento. En la etapa final del proceso de reconocimiento a toda forma se le asocia un vector de características, y en el caso que nos ocupa las características usadas son momentos 3D evaluados sobre la superficie del objeto reconstruido.

En una primera representación, la superficie del objeto 3D se describe mediante un conjunto de polígonos contenidos en planos paralelos (Fig. 2.15). Estos polígonos son los resultantes de aproximar las líneas de contorno definidas por la intersección de la superficie real del objeto con una familia de planos paralelos y equiespaciados.

La estructura de datos para esta representación estará formada por un conjunto de "planos" paralelos entre sí. Cada "plano" contiene una serie de polígonos y cada polígono se especifica por el número de vértices y las coordenadas (x,y,z) de cada uno de ellos.

La descripción de una forma dentro de este esquema, generalmente, posee un alto valor cualitativo, dada la habilidad con que un observador humano percibe la forma y estructura 3D de objetos simples [UDUP-83]. Esto se comprueba fácilmente al observar la Fig. 2.15.

Como se ha apuntado por algunos autores [CHEN-86a, 86b], la representación de una superficie mediante líneas de su contorno es poco apropiada cuando ésta se usa con propósitos de procesado automático con ordenador. En el caso que nos ocupa, y debido a la naturaleza numérica de las características extraidas sobre cada forma

tridimensional, estaríamos obligados a evaluar momentos 3D sobre las aristas de los polígonos que definen la superficie del objeto.

Para una misma forma tridimensional el número de secciones planas obtenidas en el proceso de reconstrucción es función de la separación entre secciones y de la orientación espacial de éstas, siendo estos últimos parámetros elegidos por el usuario (Fig. 2.15). Es esta la razón por la que se hace necesario, a nivel cuantitativo, pasar de la anterior representación mediante líneas de contorno, a otra que sea descriptiva de la superficie que conecta entre sí sucesivos contornos contenidos en planos paralelos y sobre la que se puedan efectuar, con más facilidad y sin ambigüedad, las operaciones necesarias para extraer las características deseadas.

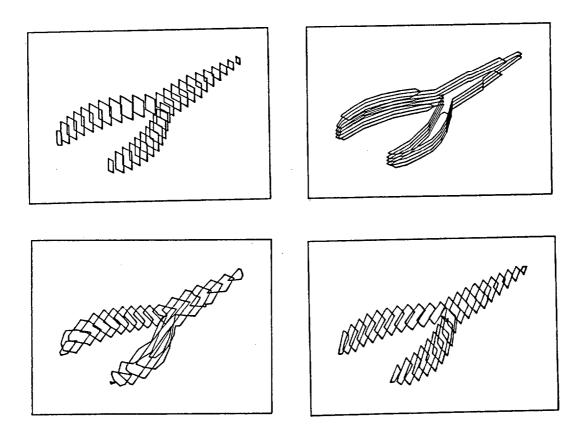


Fig. 2.15. Objeto de la Fig. 2.8, reconstruido empleando vistas distintas y con distinto espaciado entre secciones contiguas.

En nuestro caso, la solución elegida ha sido representar la superficie 3D mediante un poliedro de caras planas triangulares, correspondiendo los vértices de éste a los vértices de los polígonos que aproximan las distintas secciones del objeto. Cada triángulo conecta entre sí dos vértices de un polígono con un vértice de otro polígono situado en un plano contiguo (Fig. 2.16).

Las técnicas y algoritmos que, partiendo de la descripción de una superficie en términos de sucesivas secciones, obtienen una aproximación de la superficie real mediante triángulos, se conocen con el nombre genérico de <u>técnicas de triangulación</u> [KEPP-75], [UDUP-83], [DE FL-87]. Un resumen de las mismas se presenta en la siguiente sección.

Podemos concluir que en el proceso de reconstrucción de la estructura 3D de un objeto se han empleado dos esquemas distintos de representación usando información superficial como primitivas. El primero, en términos de polígonos que aproximan contornos planos de la superficie, es el apropiado en las etapas de reconstrucción inicial y posterior refinamiento. En esta representación la información está organizada por planos. El esquema segundo usa superficies planas triangulares como primitivas. La descripción de un objeto en esta representación se obtiene a partir de la representación anterior y una vez que se ha finalizado el proceso de reconstrucción.

La primera representación, aun presentando una estructura de datos menos compleja y requiriendo menos espacio en lo referente a lugares de memoria, no contiene explícitamente la información necesaria en las tareas de reconocimiento. Es por ello que, a nivel cuantitativo, sea preferible la segunda frente a la primera.

II.4 DESCRIPCION DE LA SUPERFICIE DE UN OBJETO MEDIANTE TRIANGULACION DE PUNTOS EN EL ESPACIO 3D

La reconstrucción de la superficie tridimensional de un objeto a

partir del conocimiento de una colección de líneas de su contorno es un problema que se presenta en diferentes campos de las ciencias aplicadas: tomografía computerizada [SLOA-81], [UDUP-83], biología [LEVI-72], [PERK-82], microscopía [SHAN-78], CAD/CAE [CHEN-86a,86b], [IDES-86], topografía, etc..

En aplicaciones de diseño por ordenador, la reconstrucción se realiza principalmente con fines de producir una visualización realista del objeto en estudio. En otras aplicaciones, la reconstrucción de la superficie está motivada por la necesidad de efectuar determinadas operaciones métricas sobre su geometría: cálculo de áreas, volúmenes, evaluación de momentos 3D, etc..

Una solución consiste en aproximar la superficie mediante un "mosaico" de superficies planas triangulares. El resultado es un poliedro de caras triangulares coincidiendo sus vértices con los de los polígonos que aproximan distintos contornos de las secciones planas del objeto.

En esta sección resumimos brevemente algunas de las aproximaciones existentes para la triangulación de la superficie de objetos tridimensionales.

II.4.1 DEFINICION DEL PROBLEMA

El problema de triangulación puede expresarse en los siguientes términos:

"Dado un conjunto de polígonos, situados sobre planos paralelos y que aproximan a las líneas de contorno de una superficie 3D, se desea obtener un conjunto de superficies triangulares eligiendo ternas de puntos de entre los vértices de

el término líneas de contorno hacemos referencia superficie intersección de una familia de planos paralelos la del objeto.

los polígonos, de forma que estos triángulos determinen una superficie "aceptable" del objeto a reconstruir."

Este es un problema complejo si se tiene en cuenta el elevado número de combinaciones que se pueden presentar [COOK-81b]. Si se considera que la superficie "aceptable" se obtiene triangulando entre polígonos contenidos en planos consecutivos, la complejidad se reduce, pero no deja de ser elevada. La validez de esta consideración estará en función de la coherencia existente en la información contenida en planos consecutivos, es decir de la similitud entre contornos de planos contiguos en lo referente a aspectos traslacionales, rotacionales y número de polígonos.

El problema se reduce, pues, al de construir una secuencia de aproximaciones parciales, cada una de las cuales conecta, mediante triángulos, los vértices de polígonos situados en planos consecutivos. En la Fig. 2.16 se muestra esquemáticamente este proceso.

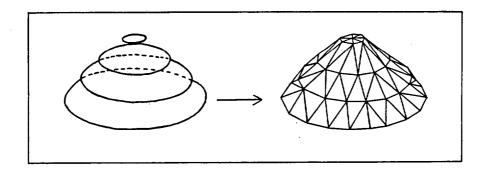


Fig. 2.16. Figura conceptual mostrando la descripción de una superficie en dos esquemas de representación distintos. a) Superficie descrita mediante líneas de su contorno. b) Superficie descrita mediante un "mosaico" de caras planas triangulares.

De aquí en adelante, y mientras no se indique lo contrario, consideraremos que la intersección de cada plano con la superficie del objeto produce una única línea de contorno. Con esta consideración, el problema de triangulación se puede reformular de la siguiente manera:

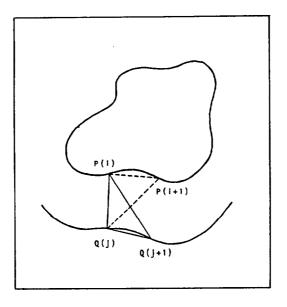


Fig. 2.17. Características de los triángulos empleados en la generación de una superficie cerrada delimitada por los contornos P y Q.

Sean dos contornos sobre planos consecutivos definidos por las secuencias de puntos P(i), $i=1,\ldots,m$ y Q(j), $j=1,\ldots,n$, siendo m y n el número de vértices para cada uno de los polígonos 2 (Fig. 2.17). Se desea construir una superficie "aceptable" entre los contornos P y Q mediante caras planas triangulares, no solapadas y de forma no necesariamente regular. Cada triángulo estará definido por tres puntos de la forma $\{P(i), P(k), Q(j)\}$ o $\{Q(i), Q(k), P(j)\}$ con las siguientes características:

- 1. Cada triángulo tiene un lado que es bien el segmento de P(k) a P(k+1) o bien el segmento de Q(k) a Q(k+1), para un valor apropiado de K.
- 2. Cada triángulo tiene dos lados cuyos puntos extremos están situados en contornos distintos; es decir, catetos que enlazan un contorno con el otro.

^{2.} Por simplicidad, y sin pérdida de generalidad, podemos considerar que ambos poligonos están orientados en el mismo sentido.

3. Si uno de los catetos del triángulo que conecta los dos contornos es el segmentos P(i) Q(j), entonces el otro cateto, con iguales caracteristicas, será P(i) Q(j+1) o P(i+1) Q(j).

Las condiciones anteriores imponen la generación de una superficie cerrada y delimitada por los contornos P y Q.

La caracterización precisa del término "aceptable" es una tarea difícil. Para clarificar este punto, comencemos indicando que el número de disposiciones distintas de triángulos que, partiendo de un segmento inicial de conexión entre los dos polígonos, pueden construirse para unos valores m y n, viene dado por [KEPP-75].

$$T(m,n) = (m+n)!/m!n!$$

Así, para polígonos con igual número de vértices, por ejemplo con $m=n=25,\ T(25,25)$ es del orden de 126 billones.

A la vista de esta expresión, se observa que el número de superficies distintas que se pueden construir entre dos contornos es suficientemente elevado, por lo que, de antemano, se descarta un análisis exhaustivo de todas las posibilidades para decidir cuál es la superficie aceptable. De aquí la necesidad de establecer un criterio para elegir una de entre todas las posibles opciones.

Dos grupos de soluciones, filosóficamente diferentes, se han propuesto. El primero, fija previamente unos criterios de optimización y reduce el problema a uno de teoría de grafos. En el segundo grupo se encuentran los algoritmos que, empleando criterios heurísticos, persiguen la determinación de una superficie que sea visualmente aceptable, cosa que no siempre está asegurada en las soluciones obtenidas mediante los criterios de optimización anteriores.

11.4.2 METODO DE KEPPEL

Keppel [KEPP-75] fue el primero en plantear una solución al

problema de elegir un conjunto óptimo de triángulos que se ajusten a una superficie tridimensional. Expresó el problema en términos de la teoría de grafos [CHRI-75]. Así, el conjunto de todas las posibles disposiciones de triángulos lo representó mediante un grafo G(V,E) donde V es el conjunto de nodos o vértices y E es un conjunto ordenado de pares de nodos llamados lados del grafo.

Cada vértice V(i,j) del grafo representa el segmento que une los puntos P(i) y Q(j). De esta forma, el grafo G(V,E) viene especificado por:

```
V = \{ V(i,j) / i = 1,2,...,m+1; j= 1,2,...,n+1 \}

E = \{ (a,b) / a = V(i,j), b = V(i+1,j) ó b= V(i,j+1) \}
```

En ambos casos la suma se elige módulo m y n para uno y otro contorno. Cada lado (a,b) del grafo representa un triángulo de vértices:

```
P(i), P(i+1), Q(j) si a = V(i,j) y b = V(i+1,j) 6
P(i), Q(j+1), Q(j) si a = V(i,j) y b = V(i,j+1)
```

El problema de triangulación queda así reducido a determinar un camino óptimo dentro del grafo G (Fig. 2.18).

Observemos que este camino ha de partir del vértice V(1,1) y finalizar en él mismo, o sea, en $V(m+1,n+1) \equiv V(1,1)$. Este camino se especificará mediante movimientos verticales y horizontales a partir de cada nodo. Dos son los únicos movimientos permitidos en el grafo: movimiento de una unidad a la derecha y movimiento de una unidad hacia abajo.

Keppel considera que el nodo de partida es conocido y lo etiqueta con los índices V(1,1). Esto equivale a elegir un triángulo inicial para la reconstrucción de la superficie; el triángulo en cuestión tiene uno de sus catetos uniendo los vértices P(1) y Q(1). La elección de estos dos vértices de entre los m y n que componen cada polígono es una tarea que ha de efectuar el usuario en función de sus necesidades.

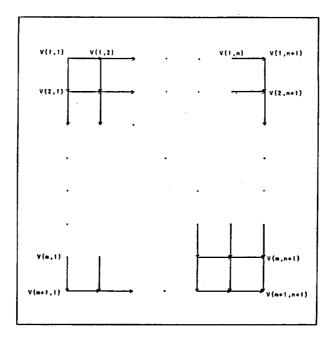


Fig. 2.18. Grafo representando todas las posibles triangulaciones en el esquema de Keppel.

Una superficie aceptable, dentro de este esquema, corresponde al camino del grafo G que hace máximo el volumen delimitado por la superficie en consideración y los dos polígonos sobre los que se apoya.

11.4.3 METODO DE FUCHS, KEDEM Y USELTON

Fuchs y col. [FUCH-77] proponen una generalización del método de Keppel que elimina el uso de criterios heurísticos, siendo así totalmente automático el proceso de determinar una superficie aceptable de acuerdo con algún criterio de optimización. El problema, igualmente se plantea en términos de teoría de grafos con la peculiaridad de tratarse ahora de un grafo toroidal, el cual se transforma convenientemente en uno plano. Cada lado conectando dos vértices vecinos del grafo representa un triángulo. A cada lado del grafo se le asigna una función de costo que coincide, en este esquema, con el área del correspondiente triángulo.

La determinación de una superficie aceptable equivale a encontrar un camino de costo mínimo dentro del grafo, siendo el costo de un camino la suma de los costos asociados a cada uno de los lados que lo componen. Teniendo en cuenta que cada camino dentro del grafo puede reescribirse de forma que empiece y acabe en un vértice de la forma V(i,1) para algún $1 \le i \le m$, el camino óptimo se determina en dos pasos. En el primero, para cada valor de i se calculan los caminos de costo mínimo, B(i), $i=1,\ldots,m$, que comienzan y acaban en el vértice V(i,1). A continuación, en el segundo paso, se elige el camino de coste mínimo de entre el conjunto $\{B(i),\ i=1,\ldots,m\}$.

El resultado de este procedimiento es la construcción de una superficie de área mínima.

La idea básica en los dos métodos anteriormente descritos ha sido formalizar el proceso de búsqueda para elegir sistemáticamente una de entre las múltiple superficies que se pueden construir sobre los contornos contenidos en dos planos paralelos. No obstante, esta forma de operar no siempre asegura que la superficie resultante, aunque óptima respecto de algún criterio, sea visualmente aceptable.

Con este fin se han desarrollado varios algoritmos (algunos de los cuales permiten la intervención del usuario experto en la superficie a reconstruir) menos complejos desde el punto de vista de cálculo y que necesitan de algún tipo de información inicial.

II.4.4 METODO DE CHRISTIANSEN Y SEDERBERG

En 1978 Christiansen y Sederberg [CHRI-78] proponen un algoritmo de triangulación para la superficie definida por dos polígonos situados en planos contiguos. Al igual que en los métodos anteriores, consideran que los polígonos están orientados de forma que se recorren en el mismo sentido. La determinación de una superficie aceptable se efectua como sigue.

A) En primer lugar, y con el fin de determinar dos vértices (uno de

cada polígono) para iniciar el proceso de triangulación, se calcula el centroide de cada polígono y se traslada uno de ellos, normalizando para inscribirlos a ambos dentro de un cuadrado unidad, de forma que coincidan los dos centroides. Se eligen dos vértices u y v tales que la distancia

$$dist[P(u),Q(v)] \le dist[P(i),Q(j)] \quad \forall \quad 1 \le i \le m, \quad 1 \le j \le n$$

De esta forma P(u) es el punto de partida para un polígono y Q(v) lo es para el otro.

B) A continuación se renumeran los vértices de ambos polígonos de forma que P(1) = P(u) y Q(1) = Q(v). A cada segmento que conecte vértices de polígonos distintos se le asocia un triángulo. El primer segmento de conexión es P(1)Q(1). El criterio para asociar un triángulo a cada segmento de conexión P(i)Q(j) es el que a continuación se indica (Fig. 2.19):

En ambos casos, y después de haber generado un triángulo, se elige como nuevo segmento de conexión el correspondiente a la menor distancia; es decir P(i)Q(j+1) en el primer caso y P(i+1)Q(j) en el segundo.

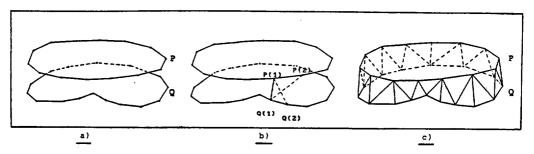


Fig. 2.19. a) Pareja de contornos a triangular. b) Selección del segmento inicial de conexión. c) Triangulación de la superficie definida por los contornos P y Q.

El proceso finaliza cuando cuando se hayan conectado todos los vértices de los polígonos P y Q, siendo el número total de triángulos generados (m+n). Aunque no explícitamente indicado por algunos autores [COOK-81b], [UDUP-83], este método se puede visualizar con un grafo de manera totalmente análoga a la usada por Keppel.

La generación de los (m+n) triángulos que componen la superficie definida por los polígonos P y Q es equivalente a encontrar un camino de (m+n) lados en el grafo de (m+1) filas y (n+1) columnas (Fig. 2.18) atendiendo a las siguientes reglas:

- 1. Los puntos de inicio y fin de este camino son los vértices del grafo V(1,1) y $V(m+1,n+1) \equiv V(1,1)$.
- 2. Los movimientos posibles a partir de un vértice V(i, j) del grafo son dos:

$$V(i,j) \longrightarrow V(i,j+1)$$
 $V(i,j) \longrightarrow V(i+1,j)$

3. Cada uno de los movimientos anteriores se efectúa a través de un lado del grafo y lleva asociado la generación de un triángulo con las características siguientes:

Mov.
$$V(i,j) \longrightarrow V(i,j+1) \iff$$
 generar triángulo $P(i)Q(j+1)Q(j)$
Mov. $V(i,j) \longrightarrow V(i+1,j) \iff$ generar triángulo $P(i)P(i+1)Q(j)$

4. La decisión del movimiento a efectuar a partir de un vértice dado V(i,j), se realiza como sigue:

Si dist[P(i),Q(j+1)] < dist[P(i+1),Q(j)]
$$\Rightarrow$$
 Mov. V(i,j) \rightarrow V(i,j+1)
Si dist[P(i),Q(j+1)] > dist[P(i+1),Q(j)] \Rightarrow Mov. V(i,j) \rightarrow V(i+1,j)

En términos de la teoría de grafos, lo anterior equivale a asignar a cada lado (a,b) del grafo una función de coste C(a,b) definida por

$$C(a,b) = \begin{cases} dist[P(i),Q(j+1)] & \text{si } b = V(i,j+1) \\ dist[P(i+1),Q(j)] & \text{si } b = V(i+1,j) \end{cases}$$

$$con a = V(i,j) y 1 \le i \le m, 1 \le j \le n$$

La determinación de una superficie aceptable según el esquema de Christiansen y Sederberg equivale a elegir un camino óptimo, o de costo mínimo, en el anterior grafo. En términos geométricos, la superficie óptima sería aquella en la que fuese mínima la longitud total de los catetos que conectan vértices situados sobre polígonos contiguos. Tal y como fue propuesto por Christiansen y Sederberg, para contornos o polígonos bastante diferentes, el algoritmo puede conducir a superficies visualmente no aceptables. En estos casos se preveía la actuación de un experto para elegir segmentos de conexión en las zonas problemáticas. Por este motivo se puede decir que este algoritmo no es completamente automático. Sin embargo, cabe indicar que se trata de un algoritmo rápido, de fácil implementación y que funciona tanto más correctamente cuanto más centrados esten los contornos y siempre que presenten una razonable similitud en tamaño y forma.

Cook y col. en 1981 [COOK-81a] proponen un algoritmo similar al objetos de forma anterior y particularmente apropiado para aproximadamente cilíndrica. Aquí, la idea básica consiste en elegir segmentos de conexión, entre uno y otro polígono, de manera tal que su orientación sea la más parecida a la del segmento que une los centroides de ambos polígonos. Partiendo de dos vértices, inicialmente elegidos con esta característica, se construye un grafo análogo al anterior. Situados en un vértice V(i,j), el movimiento hacia sus vecinos V(i,j+1) o V(i+1,j) se realiza en función de cual de los segmentos P(i)Q(j+1) o P(i+1)Q(j) esté más alineado con el segmento definido por los centroides de los dos polígonos.

Un estudio comparativo de estos algoritmos se presenta en [COOK-81b]. El único método que no necesita de información inicial es el propuesto por Fuch y col. En cuanto a complejidad, desde el punto de vista de cálculo, destaca el método de Keppel, siendo los más rápidos los de Christiansen y col. y Cook y col.

Una característica común a todos ellos es el hecho de no contemplar situaciones en las que sea necesario construir una superficie conectando más de dos polígonos situados en planos contiguos. Este es el caso en el que planos adyacentes contienen igual

número de polígonos y este número es mayor de 1, o bien contienen distinto número de polígonos. En la primera situación el problema tendría, aparentemente, una fácil solución si fuesemos capaces de obtener la correspondencia adecuada de conexión entre polígonos de uno y otro plano. Sin embargo, y de forma general, se puede afirmar que este es un problema fuertemente condicionado por la geometría del objeto en estudio.

El trabajo de Christiansen y col. contempla una situación sencilla como es el caso de conectar un único contorno de un plano con dos contenidos en el plano contiguo. En estas situaciones, se elige un punto intermedio entre los planos; seguidamente se renumeran los vértices de los dos polígonos en uno de los planos junto con el punto anterior, de manera que puedan considerarse como pertenecientes a una única línea cerrada. Finalmente se triangula de la forma usual. En un caso más general, Wang y col. [WANG-85] presentan una solución al problema de conectar distinto número de contornos en planos consecutivos. La decisión acerca de los contornos a conectar se efectúa en función del área resultante de la intersección de los distintos contornos y de la comparación de este área con la de cada uno de ellos.

No obstante, existen objetos para los cuales la superficie reconstruida usando estas aproximaciones difiere considerablemente de su superfice real. Situaciones como las anteriormente mencionadas son bastante frecuentes en el proceso de reconstrucción 3D usando múltiples vistas.

II.4.5 METODO PROPUESTO

El método de triangulación empleado en este trabajo está basado en los de Christiansen y col. y Cook y col. y, a diferencia de estos, contempla situaciones en las que la correspondencia entre contornos de planos contiguos no es necesariamente de uno a uno. Otro matiz diferenciador es el referente a la determinación de los vértices iniciales u y v para iniciar la triangulación propiamente dicha.

En una primera etapa, y dependiendo del número de polígonos que contenga cada plano o rodaja, se distingue entre las dos situaciones siguientes:

- A) Cada rodaja contiene un único polígono.
- B) Alguno, o ambos, de los planos contiguos sometidos a triangulación contiene un número de polígonos superior a la unidad.

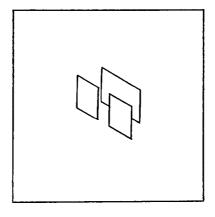
A continuación y en una segunda etapa, y con independencia de las características de cada una de las rodajas, el problema se reduce a conectar mediante triángulos cada dos polígonos -uno de cada rodaja-.

En las situaciones del tipo A), el único polígono de un plano se conecta al otro del plano contiguo. En los casos del tipo B), es decir cuando el número de polígonos en planos consecutivos sea distinto o bien, siendo igual, sea mayor que la unidad, la alternativa propuesta al problema de conexión entre polígonos de planos adyacentes es la que sigue (Fig. 2.20):

- Se determina un plano que, siendo paralelo a los anteriores, pase por el centroide de los polígonos contenidos en los mismos. Evidentemente este plano estará situado entre los dos en estudio.
- Cada polígono de cada plano se traslada hasta el plano calculado en el punto 1. Dicha traslación se efectúa según la dirección normal a la familia de planos.
- 3. Cada polígono de cada plano se triangula con su correspondiente polígono trasladado.

Resulta claro que en la triangulación del punto 3, la elección del segmento inicial de conexión entre ambos polígonos no presenta dificultad alguna, ya que los polígonos a triangular son completamente idénticos. Cada arista de un polígono y la correspondiente de su trasladado definen un rectángulo, que a su vez se descompone en dos triángulos por una de sus diagonales. En este sentido, podemos afirmar que en las situaciones del tipo B) la triangulación es inmediata,

evitándose todos los cálculos asociados a la elección del segmento inicial de conexión, y de los sucesivos que definen la familia de triángulos.



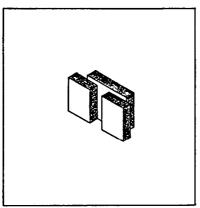


Fig. 2.20. Diagrama ilustrando la conexión entre distinto número de contornos contenidos en planos consecutivos.

La idea que subyace en los tres puntos anteriores es la de considerar que las líneas de contorno aproximando a la superfice real del objeto, y contenidas en planos consecutivos, no varían al pasar de uno cualquiera de los planos al plano intermedio. Evidentemente, esto será tanto más correcto cuanto más próximos estén los contornos a conectar. Resulta claro también, que en estos casos especiales la superficie generada entre los contornos de dos planos consecutivos no

será cerrada. No obstante, y como se pondrá de manifiesto en el capítulo IV de esta memoria, ésto no presenta mayores inconvenientes en el proceso de reconocimiento.

En los casos del tipo A), la elección del segmento inicial de conexión entre polígonos se efectúa combinando las dos ideas expresadas en los trabajos de Christiansen [CHRI-78] y Cook [COOK-81a, 81b]. Los vértices u y v que definen este segmento inicial se eligen como aquellos que, simultáneamente, muestren la mejor coherencia en lo referente a distancias de separación y orientaciones respecto del centroide.

En nuestro trabajo, y después de trasladar uno de los polígonos hasta hacer coincidir los centroides de los dos en estudio, se han calculado las distancias entre todas las posibles parejas de vértices que se pueden formar entre los m y n de los dos polígonos. Igualmente, se han determinado los vectores dirigidos desde el centroide común hasta cada vértice de cada polígono. En función de las distancias y de las diferencias en orientación entre los distintos vectores -uno de cada polígono-, se han elegido aquellos vértices que simultáneamente, presentan la menor distancia de separación y la menor diferencia en orientación.

Los detalles del algoritmo empleado en el desarrollo de este método de triangulación se muestran, con una mayor extensión, en el capítulo IV (apartado IV.4.1) de esta memoria.

En la Fig. 2.21 se muestra un ejemplo ilustrando la actuación del algoritmo propuesto. De la descripción superficial mediante líneas de contorno, indicada en la Fig. 2.14, se pasa a otro esquema de representación obtenido al triangular entre contornos de planos consecutivos usando el procedimiento anterior.

Conviene indicar que los procedimientos de triangulacón no son los únicos existentes para representar una superficie 3D definida mediante líneas de su contorno. Entre otros métodos, cabe mencionar aquellos que construyen una superfice mediante una malla de puntos enlazados entre

sí por curvas definidas paramétricamente , tanto en aplicaciones médicas [SUNG-78] como topográficas [CHEN-86a], en ingeniería [HABE-82], [ASHD-86], [QIUL-87], [FARI-88], etc.

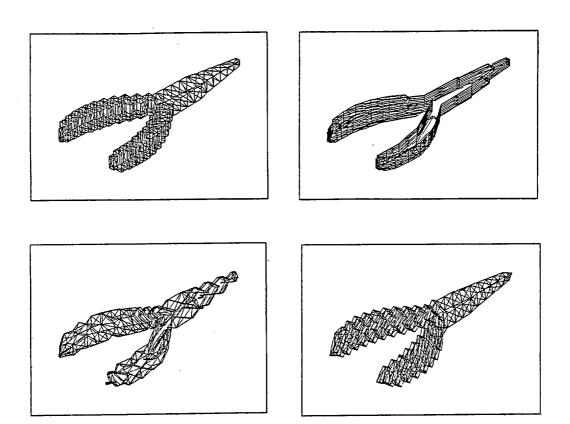
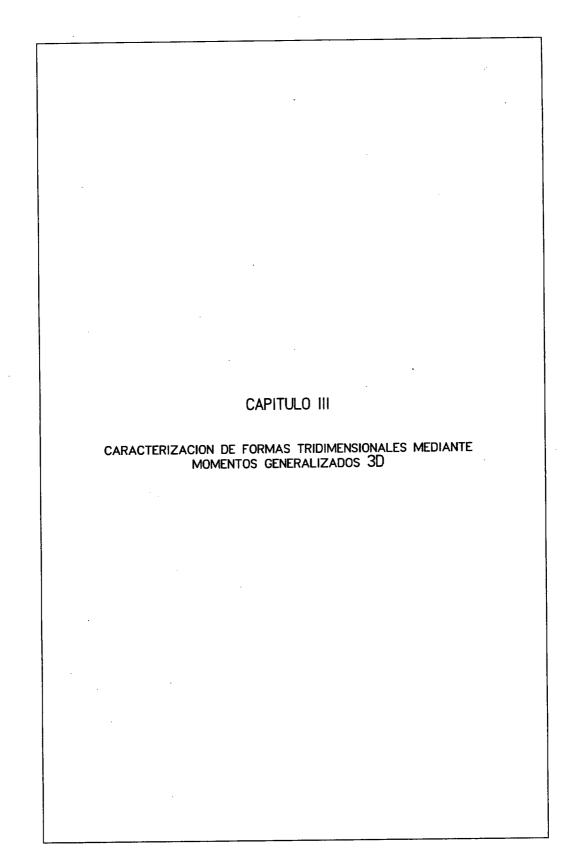


Fig. 2.21. Descripción superficial de las reconstrucciones 3D de la Fig. 2.15 aplicando el algoritmo de triangulación propuesto.

En el presente trabajo se ha optado por un método de triangulación, fundamentalmente por dos motivos. En primer lugar porque la descripción tridimensional de un objeto se realiza en términos de superficies planas triangulares bien definidas analíticamente y sobre las que se puede operar con gran facilidad. Otra razón es la relativa sencillez del algoritmo empleado y el hecho de que las descripciones obtenidas con este método están dentro de la precisión necesaria en el proceso de reconocimiento.

Notemos, finalmente, que la bondad en la reconstrucción superficial mediante las técnicas de triangulación será función del número de líneas de contorno usadas en la descripción de la superficie. En la mayoría de las aplicaciones se suele alcanzar un compromiso entre la calidad de la representación y el tiempo de proceso, ya que evidentemente, éste último es fuertemente dependiente del número de polígonos y del número de vértices de los mismos.



III.1 INTRODUCCION

Una de las etapas básicas al diseñar un sistema de reconocimiento de formas es la elección de un conjunto apropiado de características de dimensión reducida, a extraer sobre cada forma y que presenten un buen comportamiento en la etapa de clasificación.

Idealmente, las características debieran ser sensibles a la identidad de las distintas clases y no a las variaciones dentro de cada una de ellas. Sin embargo, características que no son sensibles a variaciones dentro de una misma clase, por lo general, tampoco lo son a las diferencias entre clases y viceversa.

Los momentos fueron inicialmente propuestos por HU [HU-61,62] como una solución a este problema en el reconocimiento de imágenes.

El uso de momentos puede establecerse desde una doble perspectiva. Así, invariantes obtenidos a partir de momentos forman el vector de características que mantiene inalterados los valores de sus componentes frente a traslaciones, rotaciones, amplificaciones y otras transformaciones similares [HU-62], [DUDA-77], [SADJ-80], [BOYC-83].

Por otra parte, los momentos se emplean también para normalizar una forma, es decir transformarla en una forma estándar sobre la que puedan extaerse características aptas para clasificar la forma ya normalizada.

El campo en el que la técnica de los momentos ha sido más extensamente empleada corresponde al de formas visuales bidimensionales o imágenes. Los momentos de una imagen, tal y como se definirán posteriormente, pertenecen a la categoría de características globales y de naturaleza artificial [HERN-87], [PRAT-78] ya que se obtienen tras efectuar determinadas operaciones sobre la imagen en cuestión y no aparecen directamente en la mísma como ocurre con las características naturales (brillo de una imagen, por ejemplo).

En 1962, HU obtiene un conjunto reducido de combinaciones de

momentos que permanecen invariantes frente a traslaciones, rotaciones y amplificaciones en el reconocimiento de caracteres escritos. Los trabajos de Hu se basan en la teoría de invariantes algebraicos desarrollada por Boole, Cayley y Silvester en el siglo pasado. Posteriormente DUDANI [DUDA-77], para caracterizar y reconocer objetos tridimensionales, emplea también momentos 2D calculados sobre un elevado número de imágenes de cada objeto. Maitra [MAIT-79] introduce, a partir de los invariantes de HU, otras nuevas cantidades con la peculiaridad de permanencer además invariantes frente a cambios de contraste en la imagen. Sadjadi [SADJ-80] basándose igualmente en la teoría de invariantes algebraicos, y de forma nada trivial, generaliza los resultados de HU a tres dimensiones.

El problema general de reconocimiento de imágenes es discutido por Teague [TEAG-80] dentro de una estructura más general, que además de los momentos ordinarios incluye los momentos de Zernike que gozan de unas leyes simples de transformación frente a rotaciones. La obtención de invariantes mediante esta técnica es una tarea excesivamente laboriosa y poco sistematizada; esta es la razón por la que, en el único trabajo que sobre ellos existe, se emplean únicamente 2 invariantes y que desde el punto de vista práctico presentan una pobre eficacia.

Con objeto de conseguir las necesarias invarianzas frente a transformaciones afines, [CANT-82a, 82b] propone transformar los momentos en lugar de transformar las imágenes, con lo que se evita el problema de redigitalización tras trasladar, rotar y amplificar una imagen digital.

[REDD-81] muestra que los invariantes de HU pueden obtenerse sin hacer uso de la tediosa teoría de invariantes algebraicos. Para ello propone el empleo de funciones de momentos invariantes más generales y a partir de las cuales se puedan obtener los invariantes de HU, definiendo los momentos radiales y angulares al expresar la función de distribución de intensidades en coordenadas polares en vez de rectangulares.

Más recientemente, Abu-Mostafa y col. [ABU-84, 85] introducen los momentos complejos en un intento de analizar el comportamiento analítico de los momentos invariantes como características en "pattern necagnition". Con este nuevo formalismo generalizan el criterio de normalización clásico de los ejes principales, o {N20}, a otro más amplio conocido como criterio {Npq} basado en momentos de orden superior (p+q). Posteriormente, [LUO-88], se ha puesto de manifiesto que el empleo de momentos complejos aplicados a la información suministrada por sensores táctiles, permite una identificación de objetos 3D en aquellos casos en los que no sea posible disponer de información de tipo visual.

Teh y col. [TEH-88] estudian la influencia del muestreo, digitalización y ruido sobre distintos tipos de momentos invariantes así como el problema de reconstrucción de imágenes a partir de un conjunto finito de sus momentos.

La sección 2 de este capítulo se dedica al estudio de momentos 2D, mientras que en la sección 3 se presentan los momentos 3D.

III.2 MOMENTOS BIDIMENSIONALES

En esta sección se definen varios tipos de momentos evaluados sobre formas bidimensionales y se resumen, brevemente, algunas de sus propiedades más destacadas así como su relación con los momentos ordinarios.

III.2.1 DEFINICION Y TEOREMA DE UNICIDAD

Una propiedad numérica de una imagen puede definirse como una función que transforma imágenes en valores numéricos; el número (real o complejo) obtenido a partir de una imagen f es el valor de la propiedad en cuestión. Las propiedades o características de una imagen pueden ser de naturaleza local o global dependiendo de que su valor sea o no función de solamente una pequeña zona de la imagen.

En una gran variedad de problemas de clasificación, dos o más imágenes que sólo difieren en el contraste, posición, orientación y escala o tamaño deben considerarse como pertenecientes a la misma clase. En tales situaciones, sería interesante disponer de propiedades que no dependan de los anteriores parámetros, es decir propiedades invariantes frente a las citadas transformaciones.

Puede demostrarse, [ROSE-76], que si P es una operación que transforma imágenes en números, y además es lineal, entonces para cualquier imagen f existe una función h tal que

$$\mathbb{P}(f) = \int_{-\infty}^{+\infty} h(x,y) f(x,y) dx dy$$
 (3.1)

Una gran variedad de funciones h ("templates") pueden usarse para obtener propiedades de una imagen f. Así por ejemplo, es normal el empleo de funciones simples desde el punto de vista matemático tales como polinomios, funciones de Walsh, sinusoides (en este caso los valores de las propiedades serán los coeficientes de Fourier de la imagen f). Cuando se elige un conjunto de funciones ortogonales tales como sinusoides, funciones de Walsh o polinomios ortogonales, las propiedades que se obtienen no están correlacionadas.

En el caso particular en el que h(x,y) sea el monomio x^iy^j , se obtienen las propiedades conocidas con el nombre de momentos ordinarios bidimensionales de orden (i,j) de la función f, y se denotan mediante mij

La equivalencia entre cada forma f y su representación en el espacio de momentos $\{mij\}$ viene asegurada por el teorema de unicidad, [HU-62], que expresa que una función f(x,y) está univocamente determinada por la secuencia de momentos $\{mij\}$ y recíprocamente, a cada conjunto infinito de momentos $\{mij\}$ le corresponde una, y sólo

una, función f(x,y).

A partir de ahora, y mientras no se especifique lo contrario, consideraremos que las integrales están extendidas a la región finita del plano XY en donde f(x,y) está definida.

III.2.2 MOMENTOS CENTRALES. PROPIEDADES DE LOS MOMENTOS DE BAJO ORDEN

Por orden del momento mij se entiende la suma de los enteros (i+j). El momento de orden 0 de una forma bidimensional f(x,y) viene dado por:

$$m_{oo} = \iint f(x, y) dx dy$$
 (3.3)

y representará, por ejemplo, el área de la imagen en el caso de que ésta sea binaria (dos niveles de gris). Los momentos de primer orden

$$m_{10} = \iint x f(x,y) dx dy$$
 (3.4.a)

$$m_{01} = \iint y f(x,y) dx dy$$
 (3.4.b)

junto con el momento moo determinan la localización del centroide de la imagen f en el plano XY. Las coordenadas de éste vienen dadas por

$$\overline{x} = \frac{{m \choose 10}}{{m \choose 00}} \qquad \overline{y} = \frac{{m \choose 01}}{{m \choose 00}}$$
 (3.5)

Los momentos centrales, µij, de f vienen definidos por:

$$\mu_{ij} = \iint (x - \bar{x})^i (y - \bar{y})^j f(x, y) dx dy$$
 (3.6)

y corresponden a los momentos de f evaluados tras haber trasladado el centrolde de f hasta el origen de coordenadas. De la anterior definición se deduce que los momentos centrales pueden expresarse fácilmente en términos de los momentos ordinarios de igual o inferior orden. La conexión entre ambos grupos de momentos se establece como sigue:

$$\mu_{ij} = \sum_{r=0}^{i} \sum_{s=0}^{i} {i \choose r} {j \choose s} (-1)^{(i-r+j-s)} (\overline{x})^{(i-r)} (\overline{y})^{(j-s)} m_{rs}$$
 (3.7)

Los momentos de orden 2

$$m_{20} = \iint x^2 f(x, y) dx dy$$
 $m_{11} = \iint x y f(x, y) dx dy$
 $m_{02} = \iint y^2 f(x, y) dx dy$
(3.8)

contienen información relativa al tamaño, forma y orientación de la imagen f(x,y). De hecho, cuando una forma f, centrada en el origen de coordenadas, se representa mediante momentos de segundo orden, ésta es equivalente a una elipse con igual nivel de gris en todos sus puntos (Fig. 3.1). Los tamaños de los semiejes mayor y menor, junto con la orientación del semieje mayor respecto del eje X, vienen dados por

$$a = \left(\frac{\mu_{20} + \mu_{02} + \left[(\mu_{20} - \mu_{02})^2 + 4 \mu_{11}^2\right]^{1/2}}{\mu_{00}/2}\right)^{1/2}$$
(3.9.a)

$$b = \left(\frac{\mu_{20} + \mu_{02} - [(\mu_{20} - \mu_{02})^2 + 4 \mu_{11}^2]^{1/2}}{\mu_{00}/2} \right)^{1/2}$$
(3.9.b)

$$tag 2\phi = \frac{2\mu_{11}}{\mu_{20} - \mu_{02}}$$
 $\phi = \frac{1}{2} tang^{-1} \left(\frac{2 \mu_{11}}{\mu_{20} - \mu_{02}} \right)$ (3.9.c)

Para una discusión más detallada acerca del contenido en información de los momentos de orden superior, sensibilidad al ruido,

correlación entre momentos y reconstrucción a partir de momentos, vease [TEAG-80], [ABU-84], [TEH-88].

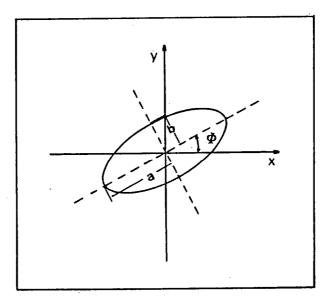


Fig. 3.1. Elipse asociada a una forma bidimensional cuando ésta se representa mediante momentos de hasta segundo orden.

III.2.3 MOMENTOS DE LEGENDRE

los polinomios de Legendre forman un conjunto completo y ortonormal de funciones en el intervalo [-1,1] sujetos a la siguiente condición de ortogonalidad

$$\int_{-1}^{1} P_{m}(x) P_{n}(x) dx = \frac{2}{2m+1} \delta_{mn}$$
 (3.10)

siendo $P_m(x)$ y $P_n(x)$ los polinomios de Legendre de grado m y n respectivamente [SPIE-70]. Dado que el tamaño de las imágenes es finito, siempre será posible elegir las unidades de longitud de forma que los momentos se calculen en la región del plano definida por $-1 \le x \le 1$, $-1 \le y \le 1$.

Los momentos ortogonales de Legendre, λ_{mn} , se definen como:

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4} \iint P_m(x) P_n(y) f(x,y) dx dy$$
 (3.11)

Teniendo en cuenta que los polinomios de Legendre pueden expresarse como

$$P_{m}(x) = \sum_{j=0}^{m} C_{mj} x^{j}$$
 (3.12)

con los coeficientes Cmj tal y como aparecen en [COUR-53], puede demostrarse que los momentos de Legendre de la función f están relacionados con los momentos ordinarios mij a través de la expresión

$$\lambda_{mn} = \frac{(2m+1)(2n+1)}{4} \sum_{i=0}^{m} \sum_{j=0}^{n} C_{mi} C_{nj} m_{ij}$$
 (3.13)

III.2.4 MOMENTOS DE ZERNIKE

Los polinomios de Zernike forman un conjunto completo de funciones ortogonales dentro del círculo unidad, y gozan de unas leyes simples de transformación frente a rotaciones. Vienen definidos de la siguiente forma

$$V_{n1}(x,y) = V_{n1}(r \text{ sen}\theta, r \cos\theta) = R_{n1}(r) e^{(il\theta)}$$
 (3.14)

siendo i el número complejo $\sqrt{-1}$, n y l enteros con n: 0,1,2,..., $\left|1\right| \leq n$, y $(n-\left|1\right|)$ par. La parte radial $R_{n1}(r)$ es un polinomio de grado n y puede expresarse como

$$R_{n1}(r) = \sum_{k=1}^{n} B_{n1k} r^{k}$$
 (3.15)

siendo Bnik los coeficientes tal y como aparecen en [BORN-75] y (n-k) par. Los polinomios de Zernike pueden obtenerse ortogonalizando los monomios x^iy^j sobre el círculo unidad $x^2 + y^2 \le 1$ [BOYC-83]. La

condición de normalización para V_{nl} se indica mediante la siguiente relación

$$\iint \left[V_{n1}(x,y) \right]^* V_{nk}(x,y) \, dx \, dy = \frac{\pi}{(n+1)} \delta_{mn} \delta_{k1}$$
 (3.16)

Cualquier función de cuadrado sumable f(x,y) puede expresarse en términos de los polinomios $V_{n1}(x,y)$ de la forma

$$f(x,y) = \sum_{n=0}^{\infty} A_{n1} V_{n1}(x,y)$$

$$|1| \le n$$
(3.17)

siendo A_{nl} los momentos ortogonales de Zernike de la función f definidos por la relación

$$A_{n1} = \frac{n+1}{\pi} \iint V_{n1}^*(x,y) f(x,y) dx dy$$
 (3.18)

donde * denota complejo conjugado.

Como se indica en [TEAG-80], los momentos de Zernike Ani pueden obtenerse a partir de los momentos centrales μ 13. La relación entre ambos es la que a continuación se indica:

$$A_{n1} = \frac{n+1}{\pi} \sum_{k=1}^{n} \sum_{j=0}^{q} \sum_{m=0}^{1} (-i)^{m} {q \choose j} {1 \choose m} B_{n1k} \mu_{k-2j-1+m, 2j+1-m}$$
(3.19)

donde q = (k-1)/2 y B_{n1k} son los coeficientes del polinomio de grado n $R_{n1}(r)$.

De la propia definición de Anı, se deduce que si la imagen f(x,y) se somete a una rotación θ en torno a su centroide, entonces los momentos finales serán

$$A_{n1} = A_{n1} e^{(-i1\theta)}$$
 (3.20)

Es decir, frente a una rotación de la imagen inicial los momentos de Zernike cambian únicamente su fase. De aquí se concluye que cualquier producto de momentos de Zernike

$$\begin{array}{ccc}
N & & & N \\
\Pi & A_{\mathbf{i}} & & \text{tal que} & \sum_{i=1}^{N} 1_{i} = 0
\end{array}$$

permanecerá invariante bajo rotaciones, siendo N el orden máximo elegido al truncar la serie (3.17) para determinar una aproximación de f.

Las características más destacadas de los momentos de Zernike son las siguientes:

- 1) En primer lugar, y al igual que los momentos de Legendre, los momentos de Zernike son independientes entre sí, hecho éste que viene garantizado por la ortogonalidad de las citadas funciones.
- 2) Presentan unas leyes muy simples de transformación frente a rotaciones.
- 3) Tanto los momentos de Zernike como los de Legendre son apropiados para reconstruir una imagen mediante un conjunto reducido de ellos.

III.2.5 MOMENTOS COMPLEJOS

Se definen los momentos complejos de orden (p,q) para una forma bidimensional f(x,y) mediante la siguiente expresión

$$C_{pq} = \iint_{-\infty}^{\infty} (x+iy)^p (x-iy)^q f(x,y) dx dy$$
 (3.21)

donde p y q son enteros no negativos e i es el complejo $\sqrt{-1}$.

Los momentos complejos, introducidos por [ABU-84,85], permiten

analizar el comportamiento de invariantes obtenidos en función de momentos, cuando éstos se usan como componentes del vector de características en "pattern recognition".

La relación entre los momentos complejos C_{pq} y los momentos ordinarios m_{ij} se obtiene al desarrollar los términos $(x+iy)^p$ y $(x-iy)^q$ que aparecen en la definición de C_{pq}

$$(x+iy) = \sum_{\substack{n_1, n_2=0\\ n_1+n_2=p}}^{p} \frac{p!}{n_1! n_2!} x^{n_1} (iy)^{n_2}$$
(3.22.a)

$$(x-iy)^{q} = \sum_{\substack{m_1, m_2=0 \\ m_1+m_2=q}}^{q} \frac{q!}{m_1! m_2!} x^{m_1} (-iy)^{m_2}$$
 (3.22.b)

Así, será posible reescribir Cpq en la forma:

$$C_{pq} = \sum_{\substack{n_1, n_2=0 \\ n_1+n_2=p}}^{p} \sum_{\substack{m_1, m_2=0 \\ n_1+n_2=p}}^{q} \frac{p! \ q!}{n_1! n_2! \ m_1! m_2!} (-1) \ i \ m_{n_1+m_1, n_2+m_2}$$
(3.23)

es decir, los momentos complejos de orden (p,q) se pueden expresar como una combinación lineal de momentos ordinarios mij, con i+j = p+q y por tanto cualquier condición que implique la presencia de momentos ordinarios puede expresarse en términos de los momentos complejos.

Al igual que los momentos de Zernike, los momentos complejos se transforman de una forma simple cuando la imagen original se somete a una rotación. El empleo de coordenadas polares facilita esta tarea. En polares, con $x = r\cos\theta$ e $y = r \sin\theta$, los momentos C_{pq} pueden escribirse como:

$$C_{pq} = \int_{0}^{2\pi} \int_{0}^{\infty} r^{(p+q+1)} e^{i(p-q)\theta} F(r,\theta) dr d\theta$$
 (3.24)

donde $F(r, \theta) = f(x, y)$.

Si la imagen f se rota un ángulo ϕ en sentido horario en torno al origen (o equivalentemente si los ejes X-Y se rotan en sentido antihorario un ángulo ϕ) entonces los nuevos momentos C'pq se relacionan con los originales Cpq mediante la expresión

$$C_{pq} = C_{pq} e^{-i(p-q)\phi}$$
 (3.25)

Esto significa que la rotación de la imagen original afecta sólo a la fase y no a la magnitud del momento complejo original C_{pq} , es decir $|C'_{pq}| = |C_{pq}|$.

III.2.6 TRANSFORMACIONES DE MOMENTOS ORDINARIOS 2D

En esta sección se muestran las relaciones que guardan entre sí los momentos ordinarios evaluados sobre una forma bidimensional f(x,y) antes y después de haberla sometido a una transformación del tipo traslación, rotación y amplificación. Cuando una forma f se somete a alguna de las operaciones anteriores, ésta experimenta un cambio y se transforma en otra f'(x,y). De manera análoga, los momentos ordinarios mij evaluados sobre f tendrán sus equivalentes en m' i j evaluados sobre f'.

Conocidas las relaciones que ligan los momentos ordinarios con los momentos resultantes después de trasladar, rotar y amplificar la imagen, será posible encontrar magnitudes, relacionadas con los momentos, que permanezcan invariantes frente a las citadas transformaciones.

III.2.6.1 TRANSFORMACION FRENTE A TRASLACIONES

Los momentos ordinarios m $\mathfrak y$ sobre la forma original f(x,y) se obtienen mediante la expresión

$$\dot{m}_{ij} = \iint x^i y^j f(x,y) dx dy$$
 (3.26)

Si la imagen f se somete a una transformación del tipo

$$x' = x + a$$

 $y' = y + b$ tal que $f(x,y) = f'(x',y')$

con a y b constantes especificando la traslación, los momentos resultantes evaluados sobre f' serán:

$$m'_{i,j} = \iint x'^{i} y'^{j} f'(x',y') dx' dy'$$
 (3.27.a)

0

$$m'_{ij} = \iint (x+a)^i (y+b)^j f(x,y) dx dy$$
 (3.27.b)

Desarrollando $(x+a)^{i}$ e $(y+b)^{j}$ en la forma

$$(x+a)^{i} = \sum_{\substack{1, n=0 \\ 1+n=i}}^{i} \frac{i!}{1! \ n!} a^{n} x^{1}$$
 (3.28.a)

$$(y+b)^{j} = \sum_{\substack{r, s=0 \\ r+s=j}}^{j} \frac{j!}{r! \ s!} \ b^{s} \ y^{r}$$
 (3.28.b)

y teniendo en cuenta las relaciones (3.26) y (3.27) se obtiene

$$m'_{ij} = \sum_{\substack{1, n=0 \ r, s=0\\1+n=i \ r+s=j}}^{i} \sum_{\substack{1! \ n! \ r! \ s!}}^{j} a^{n} b^{s} m_{1r}$$
 (3.29)

III.2.6.2 TRANSFORMACION FRENTE A ROTACIONES

Si los ejes coordenados X-Y de la imagen original se giran un ángulo θ en sentido antihorario, se obtienen unos nuevos ejes

coordenados X'-Y'. La relación entre ambos grupos de coordenadas viene dada por la transformación

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
 (3.30)

tal que f'(x',y') = f(x,y). La anterior transformación es equivalente a rotar la imagen f(x,y) un ángulo θ en sentido horario. Los momentos m'ij evaluados sobre la imagen resultante son:

$$m'_{ij} = \iint x'^{i} y'^{j} f'(x',y') dx' dy'$$

Teniendo en cuenta la transformación (3.30) y desarrollando los términos $x^{,i}$ e $y^{,j}$ según:

$$x^{i} = \sum_{r=0}^{i} {i \choose j} (x \cos \theta)^{i-r} (y \sin \theta)^{r}$$
 (3.31.a)

$$y'^{j} = \sum_{s=0}^{j} {j \choose s} (-x \operatorname{sen}\theta)^{j-s} (y \cos\theta)^{s}$$
 (3.31.b)

se obtiene

$$m'_{ij} = \sum_{r=0}^{i} \sum_{s=0}^{j} {i \choose r} {j \choose s} (-1)^{j-s} (\cos\theta)^{i-r-s} (\sin\theta)^{r+j-s} m_{i-r+j-s, r+s}$$

$$(3.32)$$

III.2.6.3 TRANSFORMACION FRENTE A AMPLIFICACIONES

Si la imagen f se somete a un cambio de escala de distinta cuantía en cada una de las cordenadas X-Y, la nueva forma resultante se caracterizará por una función f'(x',y') con

Los momentos m'_{ij} evaluados sobre f' serán

$$m'_{ij} = \iint x'^{i}y'^{j} f'(x',y') dx' dy'$$

Teniendo en cuenta la expresión de m $_{ij}$ (3.26) y que el jacobiano de la transformación (3.33) vale $\alpha\beta$, se establece la siguiente relación entre m $_{ii}$ y m $_{ij}$

$$m'_{ij} = \alpha^{i+1} \beta^{j+1} m_{ij}$$
 (3.34)

En el caso particular de amplificación isótropa, $\alpha=\beta$, la relación anterior entre momentos antes y después de transformar la imagen será:

$$m'_{ij} = \alpha^{(i+j+2)} m_{ij}$$
 (3.35)

Una traslación de una forma f en el plano imagen XY corresponde a un desplazamiento del objeto en un plano paralelo al XY, mientras que un desplazamiento del objeto a lo largo del eje óptico del sistema produce un cambio en el tamaño de la imagen. La rotación de una imagen en el plano XY va asociada a una rotación del objeto en la escena respecto de un eje perpendicular al plano imagen.

De lo anterior se deduce el claro interés en determinar características de una imagen que permanezcan invariantes frente a las transformaciones anteriormente citadas.

III.2.7 NORMALIZACION DE FORMAS BIDIMENSIONALES BASADA EN MOMENTOS 2D

En el proceso de reconocimiento, cada forma se representa mediante un vector de características de dimensión finita y reducida, dependiendo ésta, en cada caso, del tipo de aplicación para el que se ha diseñado el sistema.

El reconocimiento de formas, tanto bidimensionales como

con independencia de la posición, tridimensionales, orientación de éstas, es un tema de interés creciente por la gran cantidad de aplicaciones que de él derivan [TECH-85]. En una primera etapa del proceso, se efectúa una transformación del espacio de formas al espacio de características. El vector de características será el que permita, seguidamente y en la etapa de clasificación, asignar una forma a una de entre varias clases. De aquí que el reconocimiento de una forma dada, con independencia de su posición, orientación y lleve asociada la determinación de un vector características normalizado frente a las operaciones anteriormente citadas.

A continuación se presentan varias alternativas al proceso de normalización, teniendo todas ellas en común el empleo de momentos. Por simplicidad nos ceñiremos al caso de formas bidimensionales (2D) y en apartados posteriores se hará referencia a formas tridimensionales (3D).

Las operaciones de traslación, rotación y amplificación se representan como:

$$T: \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x+a \\ y+b \end{bmatrix}$$
 (3.36.a)

$$R : \begin{bmatrix} x', \\ y', \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix}$$
 (3.36.b)

A:
$$\begin{bmatrix} x', \\ y', \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} x', \\ y', \end{bmatrix}$$
 (3.36.c)

Una transformación que tenga en cuenta a las anteriores puede expresarse como L=A.R.T., de modo que

$$L : \begin{bmatrix} x', ', ' \\ y', ', ' \end{bmatrix} = \begin{bmatrix} c & d \\ e & f \end{bmatrix} \begin{bmatrix} x+a \\ y+b \end{bmatrix}$$
 (3.37)

En el proceso de normalización frente a la transformación L, y empleando la teoría de momentos, son posibles diversas estrategias:

- 1) Obtener determinadas combinaciones de momentos que sean invariantes frente a L (<u>Momentos Invariantes</u>).
- 2) Normalizar la forma a reconocer, es decir transformarla en una forma estándar o canónica sobre la que posteriormente se extraerán los momentos que compondrán el vector de características.
- 3) Calcular las relaciones existentes entre los momentos antes y después de aplicar las transformaciones T, R y A, normalizando en cada paso determinados momentos.

El punto 3) es el dual del 2). Mientras en 2) se transforma la imagen a reconocer, en 3) se transforman los momentos de esa imagen de manera que los momentos resultantes permanecen invariantes frente a L.

III.2.7.1 MOMENTOS INVARIANTES

Los momentos invariantes son características de una forma, calculados en términos de momentos ordinarios y que gozan de la propiedad de mantener inalterados sus valores (*invariantes*) cuando la forma se traslada, rota y/o cambia de escala.

Los momentos invariantes en 2D fueron inicialmente desarrollados por [HU-61,62] basándose en la teoría de invariantes algebraicos iniciada en el siglo XIX por Boole, Cayley y Silvester.

El empleo de momentos centrales

$$\mu_{ij} = \iint (x-\overline{x})^{i} (y-\overline{y})^{j} f(x,y) dx dy$$

en lugar de momentos ordinarios resuelve el problema de normalización frente a traslaciones. Así los momentos centrales μ 10 y μ 01 son nulos.

Frente a una amplificación del tipo (3.36.c) con $\alpha = \beta$, es decir amplificación isótropa, la normalización más extendida consiste en trabajar con las cantidades

$$\eta_{jk} = \frac{\mu_{jk}}{\mu_{00}(j+k+2)}$$
 (3.38)

Los nuevos momentos η_{ij} serán, por tanto, invariantes frente a traslaciones y amplificaciones. La normalización (3.38) equivale a imponer que el momento central de orden 0 de la imagen resultante de la amplificación A sea siempre igual a la unidad [TEAG-80].

A partir de la teoría de invariantes algebraicos [HU-62] obtiene un conjunto de invariantes absolutos frente a rotaciones. Combinando estos invariantes con los invariantes (3.38), basados en momentos centrales, HU propone 7 invariantes frente a traslaciones, amplificaciones y rotaciones en función de momentos de segundo y tercer orden. Estos son:

$$\begin{split} & I_{1} = \eta_{20} + \mu_{02} \\ & I_{2} = (\eta_{20} - \eta_{02}) + 4\eta_{11}^{2} \\ & I_{3} = (\eta_{30} - 3\eta_{12})^{2} + (3\eta_{21} - \eta_{03})^{2} \\ & I_{4} = (\eta_{30} + \eta_{12})^{2} + (\eta_{21} + \eta_{03})^{2} \\ & I_{5} = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^{2} - 3(\eta_{21} + \eta_{03})^{2} \right] + \\ & (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \left[3(\eta_{30} + \eta_{21})^{2} - (\eta_{21} + \eta_{03})^{2} \right] \\ & I_{6} = (\eta_{20} - \eta_{02}) \left[(\eta_{30} + \eta_{21})^{2} - (\eta_{21} + \eta_{03})^{2} \right] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})^{2} \\ & I_{7} = (3\eta_{21} - \eta_{03})(\eta_{03} + \eta_{12}) \left[(\eta_{30} + \eta_{12})^{2} - 3(\eta_{21} + \eta_{03})^{2} \right] - \\ & (\eta_{30} - 3\eta_{12})(\eta_{12} + \eta_{03}) \left[3(\eta_{30} + \eta_{12})^{2} - (\eta_{21} + \eta_{03})^{2} \right] \end{split}$$

I7 cambia de signo cuando la imagen se somete a una reflexión, y así

resulta útil para distinguir imágenes especulares.

[MAIT-79] transforma los invariantes anteriores de forma que permanezcan, además, invariantes frente a un cambio de contraste en f(x,y).

$$f(x,y) \xrightarrow{L} F(x'',y'') \xrightarrow{c. contraste} K F(x'',y'')$$

Los nuevos invariantes propuestos por MAITRA son:

$$\beta_{1} = \frac{I_{2}}{I_{1}} \qquad \beta_{2} = \frac{I_{2} \mu_{0}}{I_{1} I_{2}} \qquad \beta_{3} = \frac{I_{4}}{I_{3}}$$

$$\beta_{4} = \frac{I_{5}}{I_{4}} \qquad \beta_{5} = \frac{I_{6}}{I_{1} I_{4}} \qquad \beta_{6} = \frac{I_{7}}{I_{5}}$$

Los momentos de Zernike, de comportamiento simple frente a rotaciones (3.20), junto a las relaciones que los ligan a los momentos centrales (3.19), fueron empleados por [TEAG-80] para construir un conjunto de invariantes frente a L. La invarianza frente a traslaciones y amplificaciones se consigue considerando momentos centrales normalizados del tipo η_{ij} (3.38). La invarianza frente a rotaciones se determina mediante el empleo de momentos de Zernike, lo que facilita la obtención analítica de los citados invariantes, frente al método más complicado de los invariantes algebraicos de Hu.

Mientras que el número de invariantes obtenidos mediante la técnica de invariantes algebraicos es reducido, el método basado en los momentos de Zernike, y propuesto por Teague, permite obtener invariantes basados en momentos de órdenes superiores a 2 y 3.

Ambos métodos, a pesar de ser independientes de las distorsiones causadas por rotaciones y amplificaciones de la imagen original, presentan la objeción de precisar de tareas laboriosas en su construcción, y en ocasiones del uso de métodos no lineales [TEAG-80].

[BOYC-83] soslaya estos inconvenientes introduciendo los momentos rotacionales, definidos por la relación:

$$D_{n1} = \iint f(x,y) r^n e^{-il\theta} dx dy$$

con n: 0,1,2,... y l tomando valores enteros positivos o negativos. (r,θ) son coordenadas polares con x = rsen θ e y = rcos θ . Bajo rotaciones, los momentos D se transforman de igual forma que los momentos de Zernike, es decir $D_{nl} \longrightarrow D_{nl} e^{-i1} \phi$, y frente a amplificaciones del tipo r $\longrightarrow \alpha r$ se transforman según $D_{nl} \longrightarrow \alpha^{-n-2} D_{nl}$.

Boyce y col. muestran cómo, formando productos del tipo $\prod_i D_{n_i} I_i$ para los que $\sum_i I_i = 0$, se pueden obtener momentos invariantes bajo rotaciones. De igual forma, la obtención de invariantes bajo amplificaciones se lleva a cabo eligiendo cocientes de productos del tipo anterior de tal forma que las sumas $\sum_i (n_i + 2)$ para numerador y denominador sean iguales por separado.

La característica más destacable de los invariantes propuestos por Boyce es que éstos se pueden obtener hasta cualquier orden y su magnitud, por la forma en que se obtienen, no disminuye significativamente a medida que el orden aumenta.

Otra aproximación para la obtención de momentos invariantes se basa en los momentos radiales y angulares. Fueron propuestos por Reddi en 1981 siguiendo una estrategia distinta a la empleada por Hu mediante los invariantes algebraicos. Para una descripción de estos momentos invariantes y su relación con los de HU se remite al lector a la referencia [REDD-81].

III.2.7.2 NORMALIZACION DE LA IMAGEN

Este proceso consiste en transformar la función f(x,y), asociada a una imagen, en otra función F(x,y) de forma que ésta conserve toda la información relevante de la imagen original y que además satisfaga un conjunto de condiciones conocidas como criterios de normalización.

Diremos así que F(x,y) será la versión estándar o canónica de la imagen original f(x,y).

Una clase interesante de procesos de normalización se describe mediante las siguientes relaciones entre f y ${\sf F}$

$$f(x,y) = F(x^*,y^*)$$
 (3.39.a)

$$\begin{bmatrix} x_{*} \\ y^{*} \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x - \overline{x} \\ y - \overline{y} \end{bmatrix}$$
(3.39.b)

donde α , θ , \overline{x} e \overline{y} , son constantes. Este tipo de transformación, que incluye escalado isótropo, rotaciones y traslaciones, puede ampliarse para tener en cuenta otros tipos adicionales de operaciones, como por ejemplo cambios de contraste [ABU-85].

Los valores de los parámetros α , θ , \overline{x} , \overline{y} , se determinan imponiendo que la versión estándar F satisfaga un conjunto de criterios de normalización. Al decir que la normalización está basada en momentos, hacemos referencia al hecho de que los criterios de normalización son condiciones impuestas a los momentos de la imagen F.

En mecánica clásica un procedimiento de normalización frecuente, teniendo en cuenta traslaciones y rotaciones de objetos 2D, se basa en las definiciones de centroide y ejes principales. Esta misma estrategia puede aplicarse a imágenes, y como veremos más adelante, generalizarse a objetos tridimensionales [FABE-88]. Este procedimiento de normalización, también conocido como método de los ejes principales, consiste en determinar el centroide $(\overline{x},\overline{y})$ y la rotación θ que estandarizan la imagen f, normalizando momentos de hasta segundo orden.

Concretamente, las coordenadas del centroide de f vienen dadas por

$$\bar{x} = \frac{m_{10}}{m_{00}}$$
 $\bar{y} = \frac{m_{01}}{m_{00}}$

La traslación $T: \begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} x-\overline{x} \\ y-\overline{y} \end{bmatrix}$ equivale a situar el centroide de f siempre en la misma posición, el origen de coordenadas.

Una vez centrada la imagen en el origen de coordenadas, ésta se rota hasta alcanzar una orientación estándar. Una posible forma de definir esta orientación sería como sigue.

La matriz de inercia asociada a f(x,y) viene dada por

$$\begin{bmatrix} I \end{bmatrix} = \begin{bmatrix} I_{xx} & -I_{xy} \\ -I_{yx} & I_{yy} \end{bmatrix}$$
 (3.40)

que es una matriz real y simétrica, con:

$$I_{xx} = \iint y^2 f(x, y) dx dy = \mu_{02}$$

$$I_{xy} = I_{yx} = \iint x y f(x,y) dx dy = \mu_{11}$$

$$I_{yy} = \iint x^2 f(x, y) dx dy = \mu_{20}$$

Así
$$[I] = \begin{bmatrix} \mu_{02} & -\mu_{11} \\ -\mu_{11} & \mu_{20} \end{bmatrix}$$
 (3.41)

Al representar f(x,y) mediante los momentos de orden 2 μ_{20} , μ_{11} , μ_{02} , ésta es completamente equivalente a una elipse, con unos ejes y excentricidad específicos. La orientación de los ejes principales de la elipse puede definirse como el ángulo a través del cual han de rotarse los ejes del plano imagen para que se cumpla la condición μ_{11} = 0 después de haber rotado f. Este ángulo viene determinado por la transformación que diagonaliza la matriz de inercia I, [ROSE-76], y su valor es el que se indica en (3.9.c). En otras palabras, los ejes principales de inercia coinciden con los autovectores de la matriz I.

Una vez conocido θ , la orientación estándar para F la especifica el usuario. Una elección habitual consiste en alinear los ejes principales con los ejes coordenados, o viceversa.

Conviene indicar que en la obtención de θ mediante la ecuación (3.9.c) existe una ambigüedad, asociada al hecho de que cualquier rotación de los ejes principales por un múltiplo entero de $\pi/2$ conduce a otro par de ejes principales igualmente válidos. Si, como se indicó anteriormente, la orientación de la imagen está íntimamente ligada a la orientación de su correspondiente elipse, se deduce que existe ambigüedad en la determinación del ángulo θ que estandariza f(x,y).

Hu propone restricciones adicionales, como por ejemplo $\mu_{20} > \mu_{02}$, con el fin de determinar un único valor de θ según la ecuación (3.9.c). Con este mismo objetivo [TEAG-80] propone elegir el valor principal de tag⁻¹ de forma que $-\pi/2 \le \tan^{-1}\theta \le \pi/2$. No obstante, la ambigüedad no se elimina completamente. Así la condición $\mu_{20} > \mu_{02}$ reduce la ambigüedad de 4 posibilidades a 2 [ABU-85].

El valor α para la amplificación isótropa se determina, por ejemplo, imponiendo que el momento de orden cero de F, una vez centrada en el origen de coordenadas y orientada correctamente, sea una constante arbitraria B cuyo valor elige el usuario. En términos geométricos esta condición equivale a normalizar el área de la imagen al valor B.

Empleando este mismo formalismo, [ABU-85] generaliza el método de los ejes principales de manera que los criterios anteriores de normalización se refieran no a momentos ordinarios de orden 2 sino a momentos complejos de orden superior. La nueva clase de procedimientos de normalización, basada en momentos complejos de orden (p,q) con p > q, se representa por $\{N_{pq}\}$ y se define como la transformación (3.39.b) que cumple los siguientes criterios:

^{1.} C20 = B

^{2.} C10 = 0

^{3.} Cpq real y positivo

Como se observa, los dos primeros criterios son independientes de p y q, no ocurriendo lo mismo con el último. En el caso particular de p=2, q=0, se obtiene el criterio {N20} de normalización, que se reduce al clásico de los ejes principales. Esto se ve fácilmente al expresar los momentos complejos Cpq (3.23) como combinación de momentos ordinarios mpq. Así

$$C_{00} = m_{00}$$
 $C_{10} = m_{10} + im_{01}$
 $C_{20} = m_{20} - m_{02} - 2im_{11}$

Situar el centroide en el origen de coordenadas equivale a imponer que m10 = m01 = 0 y por tanto que C10 = 0. La diagonalización de la matriz de inercia, junto a la condición m20 > m02, es equivalente a decir que C20 sea real y positivo.

La característica más sobresaliente del procedimiento de normalización $\{N_{Pq}\}$ frente al clásico $\{N_{20}\}$ es que resuelve las cuestiones de ambigüedad de una manera más general y unificada, imponiendo restricciones sobre la fase de los momentos C_{Pq} .

Para una imagen f(x,y), todos los procedimientos $\{N_{Pq}\}$ proporcionan el mismo área para la imagen resultante y situan su centroide en el mismo punto. Esto resulta de la imposición de los criterios $\underline{1}$ y $\underline{2}$. Sin embargo, y en el caso general, no existirá una única imagen transformada F que satisfaga el criterio $\underline{3}$. En efecto, si llamamos $C_{Pq}(\varphi)$ al valor de C_{Pq} tras rotar los ejes de la imagen original f en sentido antihorario un ángulo φ , se cumple que

$$C_{pq}(\varphi) = C_{pq}(0) e^{-i(p-q)\varphi}$$

De aquí se observa que a medida que la imagen rota, sus momentos complejos C_{pq} mantienen inalterado su módulo $|C_{pq}|$ pero no su fase, que variará de acuerdo a la rotación φ . Por tanto, a menos que $|C_{pq}|$ sea cero, será posible encontrar algún valor de φ de manera que C_{pq} sea real y positivo. Más concretamente, se puede afirmar que existen

(p-q) valores del ángulo φ que difieren entre sí en $2\pi/(p-q)$, para los cuales C_{pq} es real y positivo. Esto se expresa diciendo que $\{N_{pq}\}$ tiene una degeneración (p-q) y por tanto la imagen estándar F no será única a menos que (p-q)=1.

De esta manera, mientras que $\{N20\}$ tiene 2 posibles versiones canónicas (una es rotación de la otra en 180°), con procedimientos $\{Npq\}$ tales que (p-q)=1 se obtiene una única versión normalizada F de la imagen original f.

III.2.7.3 NORMALIZACION DE MOMENTOS

Conceptualmente este método es similar al anterior. De igual forma, los parámetros de las transformaciones de traslación, rotación y amplificación se determinan imponiendo una serie de criterios de normalización sobre los momentos. La diferencia entre normalizar la imagen y normalizar los momentos radica en el hecho de que en el primer caso, la transformación de coordenadas (3.36) se aplica a la imagen original para estandarizarla, y en el segundo caso es a los momentos ordinarios de la forma original f a los que se aplica la citada transformación.

De esta manera no es necesario determinar la forma estandar F y calcular momentos sobre ella; por el contrario, los momentos finales usados en la clasificación se obtienen directamente, sin mas que transformar adecuadamente los momentos ordinarios evaluados sobre la forma original.

[CANT-82a, 82b] considera también el caso de amplificaciones anisótropas, es decir

$$\begin{bmatrix} x \\ y \\ y \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x - \overline{x} \\ y - \overline{y} \end{bmatrix}$$
(3.42)

Una vez determinados los parámetros $\overline{x}, \overline{y}, \theta, \alpha, \beta$, se obtienen los momentos transformados

$$\eta_{ij} = \frac{Q_{ij}}{[Q_{20}^{3j-i+2} Q_{02}^{3i-j+2}]^{1/8}}$$
 (3.43)

con

$$\eta_{ij} = \alpha^{i+1} \beta^{j+1} Q_{ij}$$
 (3.44.a)

$$Q_{ij} = \sum_{r=0}^{i} \sum_{s=0}^{j} {i \choose r} {i \choose s} (-1)^{j-s} (\cos\theta)^{i-r+s} (\sin\theta)^{r+j-s} \mu_{i-r+j-s,r+s}$$

$$(3.44.b)$$

$$\mu_{ij} = \sum_{l=0}^{i} \sum_{r=0}^{j} \binom{i}{l} \binom{j}{r} (-\overline{x})^{i-1} (-\overline{y})^{j-r} m_{1r}$$
 (3.44.c)

siendo mir momentos ordinarios evaluados sobre la imagen f. Los momentos centrales μ ij se obtienen al transformar los momentos ordinarios según la traslación de coordenadas

$$T \;:\; \begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} x - \overline{x} \\ y - \overline{y} \end{bmatrix}$$

Qij son los momentos obtenidos al aplicar la rotación R a los momentos centrales η ij.

$$R : \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

Finalmente, η_{1j} son los momentos resultantes de aplicar la transformación de escalado anisótropo A a los momentos Q.

$$A : \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}$$

Por este motivo, se puede afirmar que η_{ij} son momentos característicos de f con independencia de la posición, orientación y tamaño de ésta.

Los parámetros $\overline{x},\overline{y},\theta,\alpha,\beta$ se obtienen imponiendo los siguientes criterios de normalización

$$\mu_{10} = \mu_{01} = 0 \tag{3.45.a}$$

$$Q = 0$$
 (3.45.b)

$$\eta_{20} = \eta_{02}$$
(3.45.c)

Los criterios (3.45.a) y (3.45.b) permiten obtener (\bar{x}, \bar{y}) y θ respectivamente, mientras que (3.45.c) posibilita que los momentos finales de dos formas que únicamente difieran en el tamaño sean idénticos.

El método indicado en el anterior apartado, consistente en transformar la imagen original en otra estandarizada, presenta el inconveniente de necesitar de un proceso de redigitalización en la imagen transformada. Si en su lugar se transforman los momentos y no la imagen, esta dificultad queda eliminada. Sin embargo, y de manera general, la transformación de las características originales evaluadas sobre f no siempre será factible, dependiendo en cada caso del comportamiento de dichas características frente a las operaciones de traslación, rotación y amplificación. Las ventajas, desde un punto de vista computacional, al elegir uno u otro método dependerán en cada situación particular de las características empleadas en el reconocimiento y del esquema de representación utilizado al describir las formas.

Conviene indicar, finalmente, que en situaciones donde los momentos sean poco sensibles a la distinción entre clases, éstos no serán características adecuadas a emplear en el reconocimiento. No obstante, sería posible la <u>normalización</u> de las formas usando momentos y después, en la etapa de <u>clasificación</u>, emplear otro tipo de características.

III.3 MOMENTOS TRIDIMENSIONALES

Tradicionalmente, el reconocimiento de formas tridimensionales se abordó empleando técnicas 2D. Así por ejemplo [DUDA-77], en el reconocimiento de seis tipos diferentes de aviones, necesita extraer un vector de características de dimensión 14 para cada una de las distintas imágenes 2D que de un mismo objeto se obtienen desde diferentes posiciones; en este caso las características empleadas fueron momentos, y de cada objeto 3D se obtuvieron del orden de 500 imágenes. El manejo y tratamiento de este gran volumen de información hace preferible el empleo de otras técnicas.

En aquellas situaciones donde es factible, se trabaja directamente sobre la estructura tridimensional del objeto a reconocer. En estos casos, cada forma vendrá caracterizada por un vector de descriptores 3D.

En 1980 Sadjadi [SADJ-80] emplea, por vez primera, momentos tridimensionales en el reconocimiento de objetos con independencia de su posición, tamaño y orientación. Este trabajo es una generalización a tres dimensiones de la técnica de invariantes algebraicos, empleada por [HU-62] en la obtención de momentos invariantes bidimensionales.

Cantón y col. [CANT-84, 86] estudian la normalización de momentos 3D frente a las transformaciones de traslación, rotación y amplificación.

Recientemente, Faber y col. [FABE-88] emplean momentos tridimensionales con el fin de orientar, respecto de un estándar, imagenes médicas 3D obtenidas mediante una gamma-cámara.

Los momentos ordinarios de orden (i+j+k) de una forma 3D representada mediante una función f(x,y,z) vienen dados por la relación

$$m_{ijk} = \iiint_{-\infty} x^i y^j z^k f(x, y, z) dx dy dz$$
 (3.46)

Acceptance of the second control of the seco

De manera análoga al caso 2D, la secuencia de momentos $\{m_1j_k\}$ determina de manera unívoca la función f(x,y,z) y recíprocamente, cada forma f(x,y,z) tiene asociada una única secuencia $\{m_1j_k\}$.

A pesar del considerable esfuerzo realizado en el desarrollo de invariantes algebráicos asociados a momentos 3D, el trabajo de Sadjadi presenta el gran incoveniente de limitar a dos el número de invariantes obtenidos, lo que en la mayoría de situaciones será insuficiente para fines de reconocimiento. Los invariantes absolutos frente a traslaciones, rotaciones y amplificaciones son

$$\begin{array}{ccc}
J_{1\mu}^{2} & & \Delta_{2\mu} \\
\hline
J_{2\mu} & & J_{1\mu}^{3}
\end{array}$$

con

$$J_{1\mu} = \mu_{200} + \mu_{020} + \mu_{002}$$

$$J_{2\mu} = \mu_{020}\mu_{002} - \mu_{011}^2 + \mu_{200}\mu_{002} - \mu_{101}^2 + \mu_{200}\mu_{020} - \mu_{110}^2$$

$$\Delta_{2\mu} = \left| \begin{array}{ccc} \mu_{200} & \mu_{110} & \mu_{101} \\ \mu_{110} & \mu_{020} & \mu_{011} \\ \mu_{101} & \mu_{011} & \mu_{002} \end{array} \right|$$

y $\boldsymbol{\mu}_{i\;jk}$ momentos centrales definidos por

$$\mu_{ijk} = \iiint_{-\infty}^{+\infty} (x-\overline{x})^i (y-\overline{y})^j (z-\overline{z})^k f(x,y,z) dx dy dz$$

$$\bar{x} = \frac{m_{100}}{m_{000}}$$
 $\bar{y} = \frac{m_{010}}{m_{000}}$ $\bar{z} = \frac{m_{001}}{m_{000}}$

III.3.1 MOMENTOS 3D EVALUADOS MEDIANTE INTEGRALES DE SUPERFICIE

Como se indicó en el capítulo II de esta memoria, la

representación 3D de los objetos en estudio se efectúa en términos de su superficie. La superficie externa del objeto se aproxima mediante un poliedro de caras triangulares, estando cada una de ellas especificada por las coordenadas (x,y,z) de sus vértices. Puesto que la superficie es la parte visible de un objeto, y ésta lo describe sin ambigüedad en la mayoría de las situaciones, resulta conveniente evaluar el vector de características sobre dicha superficie.

Los momentos ordinarios de orden i+j+k evaluados sobre la superficie de la forma 3D se definen por la siguiente relación

$$m_{ijk} = \iint_{S} x^{i} y^{j} [z(x,y)]^{k} f(x,y,z) ds$$
 (3.47)

siendo S la superficie poliédrica que representa a la forma 3D y z(x,y) la relación que liga entre sí las coordenadas (x,y,z) de cada una de las caras triangulares que componen S. La función f(x,y,z) representa la función de distribución de intensidades asociada a cada punto (x,y,z) en donde está definida la forma 3D.

Puesto que nuestro interés va dirigido al reconocimiento de formas con independencia de su posición, tamaño y orientación pero sin tener en cuenta cualidades tales como el color, brillo, etc, resulta claro que son las propiedades geométricas de cada forma las que merecen nuestra principal atención.

Teniendo en cuenta lo anterior, la función f(x,y,z) se ha elegido de manera que tome un valor constante sobre la superficie S. Por simplicidad, la constante se hace igual a la unidad.

$$f(x,y,z) = \begin{cases} 1 & (x,y,z) \in S \\ 0 & \text{en el resto} \end{cases}$$

Con estas consideraciones, la expresión (3.47) para mijk se reduce a

$$m_{ijk} = \iint_{S} x^{i} y^{j} [z(x,y)]^{k} ds$$
 (3.48)

La superficie poliédrica se expresa como $S = \{T_1, T_2, \dots, T_N\}$, donde T_n representa cada una de las N caras triangulares de S. De esta forma la ecuación anterior se transforma en

$$m_{ijk} = \sum_{n=1}^{N} \iint_{T_n} x^i y^j [z(x,y)]^k ds = \sum_{n=1}^{N} (m_{ijk})_n$$
 (3.49)

lo que equivale a decir que el momento mijk se obtiene como suma, para todos los triángulos Tn, de los momentos del mismo orden evaluados sobre cada cara del poliedro S.

III.3.1.1 EVALUACION DE MOMENTOS ORDINARIOS SOBRE UNA SUPERFICIE TRIANGULAR

De aquí en adelante, mientras no se indique lo contrario y por simplicidad en la notación, mediante el término mijk nos referiremos al momento de orden i+j+k evaluado sobre la superficie triangular de cada cara del objeto 3D en estudio.

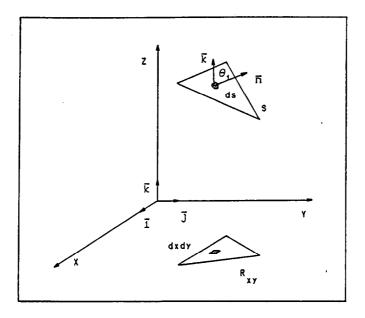


Fig. 3.2. Evaluación de m sobre la superficie triangular S. ijk

Sea S la superficie de uno cualquiera de los triángulos, tal y como se indica en la Fig. 3.2. El momento mijk sobre S será

$$m_{ijk} = \iint_{S} x^{i} y^{j} z^{k} ds$$
 (3.50)

Las variables x,y,z sobre S no son independientes sino que están ligadas por la ecuación del plano Π que contiene a S.

$$\Pi : ax + by + cz = 0 \longrightarrow \overrightarrow{n} = (a, b, c)$$
(3.51)

La expresión (3.50) puede reescribirse como

$$m_{ijk} = \frac{1}{\cos\theta_1} \iint_{R} x^i y^j [z(x,y)]^k dx dy$$
 (3.52.a)

$$m_{ijk} = \frac{1}{\cos \theta_2} \iint_{R_{xz}} x^i z^k [y(x,z)]^j dx dz$$
 (3.52.b)

$$m_{ijk} = \frac{1}{\cos\theta_3} \iint_{R_{VZ}} y^j z^k [x(y,z)]^i dy dz$$
 (3.52.c)

Siendo Rxy, Rxz, y Ryz las regiones de proyección de la superficie S sobre los planos XY, XZ e YZ respectivamente. Los ángulos θ_1 , θ_2 y θ_3 vienen definidos por las relaciones

$$\theta_1: \quad \cos\theta_1 = \left|\cos(\vec{n}, \vec{k})\right| \; ; \quad \cos\theta_1 = \frac{\left|\vec{n} \cdot \vec{k}\right|}{\left|\vec{n}\right|} = \frac{\left|n_z\right|}{\left|\vec{n}\right|} \tag{3.53.a}$$

$$\theta z: \quad \cos \theta z = \left| \cos \left(\overrightarrow{n}, \overrightarrow{j} \right) \right| ; \quad \cos \theta z = \frac{\left| \overrightarrow{n} \cdot \overrightarrow{j} \right|}{\left| \overrightarrow{n} \right|} = \frac{\left| n_y \right|}{\left| \overrightarrow{n} \right|}$$
 (3.53.b)

$$\theta_3: \quad \cos\theta_3 = \left|\cos(\vec{n}, \vec{1})\right| ; \quad \cos\theta_3 = \frac{\left|\vec{n} \cdot \vec{1}\right|}{\left|\vec{n}\right|} = \frac{\left|n_x\right|}{\left|\vec{n}\right|}$$
(3.53.c)

y corresponden al menor de los ángulos determinados por la dirección normal al plano de S, \vec{n} , con cada una de las direcciones \vec{k} , \vec{j} e \vec{i} respectivamente. De esta forma, el elemento de superficie ds será tal que su proyección según las direcciones \vec{k} , \vec{j} e \vec{i} coincidirá con dxdy, dxdz y dydz respectivamente. Es decir

 $ds cos\theta1 = dxdy$; $ds cos\theta2 = dxdz$; $ds cos\theta3 = dydz$

La elección de una de las expresiones (3.52) para evaluar mijk se hará en función de cual sea la orientación de la superficie S, o equivalentemente de su normal, en el espacio. Así por ejemplo, si la superficie S está contenida en un plano perpendicular al XY, o lo que es lo mismo si \vec{n} y \vec{k} son perpendiculares, la proyección del triángulo en cuestión sobre el plano XY degenerará en un segmento rectilíneo; por este motivo no será aplicable la expresión (3.52.a) para el cálculo de mijk, siendo necesario el empleo de (3.52.b) o (3.52.c). En cualquiera de los casos, se puede afirmar que (3.52.a) (3.52.b) y (3.52.c) son equivalentes y además siempre será factible evaluar mijk según alguna de estas expresiones dependiendo de que $c \neq 0$, $b \neq 0$ ó a $\neq 0$ respectivamente.

La relaciones (3.52), expresadas en función de los parámetros a,b,c y d de la ecuación del plano Π , quedan como sigue:

$$m_{ijk} = \frac{1}{\cos \theta i} \iint_{R_{xy}} x^i y^j \left[-\frac{ax + by + d}{c} \right]^k dx dy$$
 (3.54.a)

$$m_{ijk} = \frac{1}{\cos \theta_2} \iint_{R} x^i z^k \left[-\frac{ax + cz + d}{b} \right]^j dx dz$$
 (3.54.b)

$$m_{ijk} = \frac{1}{\cos \theta 3} \iint_{R} y^{j} z^{k} \left[-\frac{by + cz + d}{a} \right]^{i} dy dz$$
 (3.54.c)

Generalmente, la proyección de la superficie triangular S sobre los planos XY, XZ e YZ será otro triángulo, siendo éste último el que

determine los límites de integración en las expresiones (3.54).

Sea el triángulo de la Fig. 3.3 el correspondiente a los dominios de integración Rxy, Rxz, o Ryz sobre los planos XY, XZ o YZ. Mediante (u,v) se indican, genéricamente, las coordenadas de cada uno de estos planos.

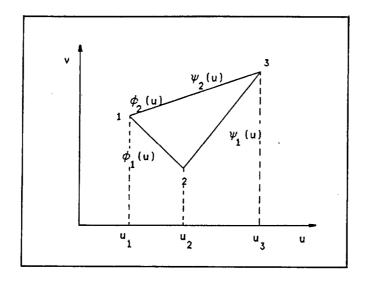


Fig. 3.3. Dominio de integración para la evaluación de mijk sobre la superficie S mediante las expresiones (3.50) y (3.54)

Para cada valor de la coordenada u comprendido en alguno de los intervalos [u1,u2] ó [u2,u3], los límites de variación para la otra coordenada v vienen definidos por las ecuaciones de dos segmentos rectilíneos asociados a los lados del triángulo. Así

$$\forall u \in [u_1, u_2] \qquad \phi_1(u) \leq v \leq \phi_2(u)$$

$$\forall u \in [u2, u3]$$
 $\psi_1(u) \le v \le \psi_2(u)$

siendo ϕ_1 , ϕ_2 , ψ_1 , y ψ_2 las ecuaciones de las rectas definidas por los vértices $1\longleftrightarrow 2$, $1\longleftrightarrow 3$, $2\longleftrightarrow 3$, y $1\longleftrightarrow 3$ respectivamente.

$$\phi_{1}(u) = A_{1}u + B_{1}$$

$$\phi_{2}(u) = A_{2}u + B_{2}$$

$$\psi_{1}(u) = C_{1}u + D_{1}$$

$$\psi_{2}(u) = C_{2}u + D_{2}$$

Las expresiones (3.54) pueden reducirse a una sola, dada por

$$m_{ijk} = \int_{u1}^{u2} du \int_{\phi_1}^{\phi_2} F(v, u) dv + \int_{u2}^{u3} du \int_{\psi_1}^{\psi_2} F(v, u) dv$$
 (3.55)

con

$$F(v,u) = \frac{1}{\cos\theta} u^{I1} v^{I2} \left(-\frac{T_2 u + T_3 v + T_4}{T_1} \right)^{I_3}$$
 (3.56)

y teniendo en cuenta la siguiente tabla de asignaciones

plano proy.	u	v	I1	12	13	Т1	Т2	13	T4	θ
ху	x	у	1	j	k	С	a	ъ	d	θ1
xz	×	z	1	k	j	ь	a	С	d	0 2
yz	у	z	ţ	k	i	a	ь	С	d	0 3

TABLA 3.I. Cuadro de asignaciones para el cálculo de m según las expresiones (3.50) y (3.54).

III.3.2 TRANSFORMACIONES DE MOMENTOS ORDINARIOS

En esta sección se establecen las relaciones que guardan entre sí los momentos ordinarios mijk evaluados sobre una forma tridimensional y los momentos resultantes después de aplicar una transformación del tipo traslación, rotación y amplificación.

III.3.2.1 TRANSFORMACION FRENTE A TRASLACIONES

Frente a una traslación dada por las relaciones

$$x' = x + a$$

 $y' = y + b$
 $z' = z + c$ (3.57)

los momentos mijk de una forma 3D se transforman en m'ijk, siendo a,b,c los parámetros que definen la traslación y m'ijk los momentos evaluados sobre la nueva forma resultante de dicha transformación.

Sobre cada una de las caras triangulares, S, que componen la superficie de un objeto, los momentos mijk vienen definidos por:

$$m_{ijk} = \frac{1}{\cos\theta} \iint_{R_{xy}} x^i y^j \left[-\frac{Ax + By + D}{C} \right]^k dx dy$$
 (3.58)

con

$$z(x,y) = -\frac{Ax + By + D}{C}$$

la ecuación del plano definido por la superficie S, y Rxy la región de proyección de S sobre el plano XY (Fig. 3.4).

La superfice S' resultante de trasladar S según (3.57) será otro triángulo con igual orientación que S. La ecuación del plano definido por S' será

$$z' = z + c = -\frac{Ax + By + D}{C} + c$$

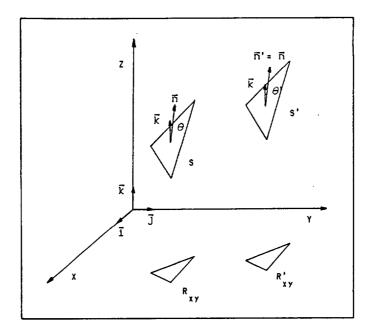


Fig. 3.4. Traslación aplicada a la superficie triangular S. Los momentos mijk, evaluados sobre S, se transforman en m'ijk, evaluados sobre S'.

La superfice S' resultante de trasladar S según (3.57) será otro triángulo con igual orientación que S. La ecuación del plano definido por S' será

$$z' = z + c = -\frac{Ax + By + D}{C} + c$$

Los momentos m'ıjk, evaluados sobre S', se expresan como

$$m'_{ijk} = \frac{1}{\cos\theta'} \iint_{R'_{xy}} x'^{i} y'^{j} [z'(x',y')]^{k} dx' dy'$$

Puesto que S y S' tienen la misma orientación, $\vec{n} = \vec{n}' = (A, B, C)$, se deduce que $\cos\theta = \cos\theta'$. De esta manera

$$m'_{ijk} = \frac{1}{\cos\theta} \iint_{R_{xy}} (x+a)^{i} (y+b)^{j} \left[-\frac{Ax + By + D}{C} \right]^{k} dx dy$$
 (3.59)

Expresando:

$$(x+a)^{i} = \sum_{1=0}^{i} {i \choose 1} x^{1} a^{(i-1)}$$

$$(y+b)^{j} = \sum_{m=0}^{j} {j \choose m} y^{m} b^{(j-m)}$$

$$\left[-\frac{Ax + By + D}{C} + C \right]^{k} = \sum_{n=0}^{k} {k \choose n} \left[-\frac{Ax + By + D}{C} \right]^{n} c^{(k-n)}$$

sustituyendo en (3.59) y teniendo en cuenta (3.58) se obtiene

$$m'_{ijk} = \sum_{l=0}^{i} \sum_{m=0}^{j} \sum_{p=0}^{k} {i \choose l} {j \choose m} {k \choose n} a^{(l-1)} b^{(j-m)} c^{(k-n)} m_{lmn}$$
(3.60)

Esta expresión nos indica que los momentos m'ijk sobre S' se obtienen mediante combinación lineal de momentos mima de igual o inferior orden evaluados sobre el triángulo original S. Además, los coeficientes de dicha combinación dependen exclusivamente de la traslación (a,b,c) y no de la ecuación de la superficie triangular S. Por este motivo, la ecuación (3.60), obtenida anteriormente para cada una de las caras triangulares, será igualmente válida cuando los momentos mima y m'ijk se evaluen sobre toda la superficie poliédrica del objeto en consideración.

III.3.2.2 TRANSFORMACION FRENTE A AMPLIFICACIONES

La transformación

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \longrightarrow \begin{cases} x' = \alpha x \\ y' = \beta y \\ z' = \gamma z \end{cases}$$
(3.61)

define un cambio de escala sobre los ejes coordenados X, Y y Z. Bajo la misma, los momentos mija de una forma tridimensional se transforman

en m'ijk siendo éstos los momentos evaluados sobre el objeto resultante de aplicar (3.61) a la forma original. Frente a esta transformación, cada cara triangular de la superficie del objeto se transforma en un nuevo triángulo de distinta orientación al original y situado en una posición diferente (Fig. 3.5)

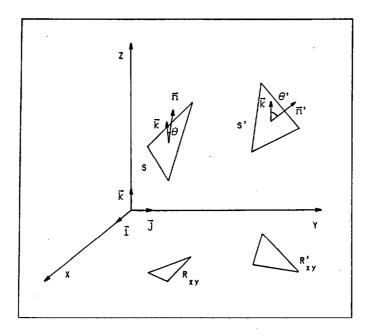


Fig. 3.5. Frente a una transformación de escalado, la superficie S se transforma en otra S' con una orientación n' \neq n .

Si

$$z = -\frac{Ax + By + D}{C}$$
 \acute{o} $Ax + By + Cz = 0$ (3.62)

representa al plano definido por una de las caras S del objeto original, el plano resultante de la transformación (3.61) sobre S vendrá dado por

$$z' = \gamma z$$
 \acute{o} $\frac{\gamma}{\alpha} Ax' + \frac{\gamma}{\beta} By' + Cz' + \gamma D = 0$ (3.63)

siendo \vec{n} = (A, B, C) y \vec{n} = ($\gamma A/\alpha$, $\gamma B/\beta$, C) las direcciones

perpendiculares a cada una de las caras S y S'.

Admitamos, por simplicidad, que tanto el triángulo S como el transformado S' proyectan sobre el plano XY sin dar lugar a un segmento rectilíneo como resultado de la proyección (esto no implica pérdida de generalidad, pues de no ser así siempre será posible la proyección sobre alguno de los planos XY o XZ).

los momentos mijk y m'ijk, evaluados sobre S y S' respectivamente, vienen dados por:

$$m_{ijk} = \frac{1}{\cos\theta} \iint_{R_{xV}} x^i y^j \left[-\frac{Ax + By + D}{C} \right]^k dx dy$$
 (3.64)

$$m'_{ijk} = \frac{1}{\cos\theta'} \iint_{R'_{xy}} x'^{i} y'^{j} \left[-\gamma \frac{A/\alpha x' + B/\beta y' + D}{C} \right]^{k} dx' dy'$$
 (3.65)

siendo Rxy y R'xy los dominios de proyección de S y S' sobre el plano XY. Los ángulos θ y θ ' se caracterizan por

$$\cos\theta = \frac{|\vec{n} \cdot \vec{k}|}{|\vec{n}| |\vec{k}|} = \frac{|C|}{\sqrt{A^2 + B^2 + C^2}}$$
(3.66)

$$\cos \theta' = \frac{|\vec{n}' \cdot \vec{k}|}{|\vec{n}'| |\vec{k}|} = \frac{|C|}{\sqrt{\frac{\gamma^2}{\alpha^2} A^2 + \frac{\gamma^2}{\beta^2} B^2 + C^2}}$$
(3.67)

Teniendo en cuenta (3.61) y las expresiones (3.63) y (3.64), la relación que liga entre sí los momentos mijk y m'ijk, definidos en (3.64) y (3.65), se establece como sigue

$$m'_{ijk} = \alpha^{(i+1)} \beta^{(j+1)} \gamma^k \frac{\cos\theta}{\cos\theta}, m_{ijk}$$
(3.68)

De (3.66) y (3.67) se observa que el cociente $\cos\theta/\cos\theta$, depende

de los parámetros α, β, γ de la transformación así como de la orientación (A,B,C) de la superficie original S. Puesto que este cociente, en general, no será el mismo para todas las caras triangulares de la superficie del objeto, se deduce que la relación (3.68), válida para una superfice específica S, no será igualmente válida cuando mijk y m'ijk se evaluen sobre toda la superficie del objeto. En este último caso tendriamos

$$\mathbf{m'}_{ijk} = \alpha^{(i+1)} \beta^{(j+1)} \gamma^{k} \left[\sum_{t=1}^{N} \left(\frac{\cos \theta}{\cos \theta'} \mathbf{m'}_{ijk} \right)_{t} \right]$$
(3.69)

siendo ahora m'ijk momentos evaluados sobre la superficie completa del objeto transformado según (3.61), y los términos con subíndice t los correspondientes evaluados sobre cada uno de los N triángulos del objeto original.

En el caso particular en que la transformación de escalado (3.61) a que se somete la forma original fuese isótropa, es decir $\alpha=\beta=\gamma$, la relación entre m'ijk y mijk, tanto si se evaluan sobre un único triángulo o sobre el total de los que componen la superficie poliédrica del objeto, sería

$$m'_{ijk} = \alpha^{(i+j+k+2)} m_{ijk}$$
(3.70)

Esta expresión se obtiene a partir de (3.68) y (3.69) haciendo $\alpha=\beta=\gamma$ y teniendo en cuenta que ahora $\cos\theta/\cos\theta'=1$ para todos los triángulos, ya que los vectores \vec{n} y \vec{n}' asociados a S y S' respectivamente, son idénticos.

III.3.2.3 TRANSFORMACION FRENTE A ROTACIONES

Sea la transformación $R(\varphi, \vartheta, \psi) = R(\varphi) R(\vartheta) R(\psi)$ definida por tres rotaciones consecutivas en torno a los ejes X,Y,Z respectivamente con:

$$R (\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{bmatrix}; \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R (\psi) \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
(3.71.a)

$$R (\vartheta) = \begin{bmatrix} \cos\vartheta & 0 & -\operatorname{sen}\vartheta \\ 0 & 1 & 0 \\ \operatorname{sen}\vartheta & 0 & \cos\vartheta \end{bmatrix}; \begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix} = R (\vartheta) \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$
(3.71.b)

$$R (\varphi) = \begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}; \begin{bmatrix} x''' \\ y''' \\ z''' \end{bmatrix} = R (\varphi) \begin{bmatrix} x'' \\ y'' \\ z'' \end{bmatrix}$$
(3.71.c)

Bajo la rotación $R(\varphi,\vartheta,\psi)$, los momentos mijk evaluados sobre un triángulo S se transforman en m''ijk, evaluados sobre el triángulo s''' resultante de rotar S según R. La relación entre mijk y m'''ijk puede establecerse en un proceso de tres pasos, cada uno asociado a las rotaciones simples $R(\psi)$, $R(\vartheta)$ y $R(\varphi)$.

La rotación R (ψ) transforma a cada triángulo S en otro S' de distinta orientación al primero y situado en una posición diferente (Fig. 3.6).

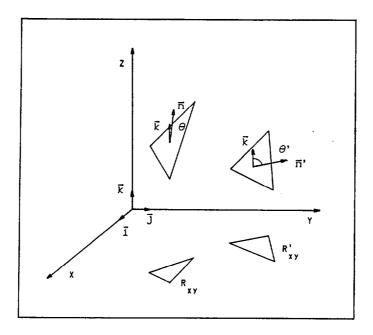


Fig. 3.6. Superficie S y su transformada S' según la rotación $R(\varphi,\vartheta,\psi)$

Admitiendo que los triángulos S y S' proyectan en sendos triángulos Rxy y R'xy sobre el plano XY, será posible escribir las siguientes ecuaciones para los planos definidos por S y S'

S:
$$Ax + By + Cz + D = 0$$
 \acute{o} $z = -\frac{Ax + By + D}{C}$ (3.72.a)

$$S': z' = -\frac{Ax' + (Csen\psi + Bcos\psi)y' + D}{(Ccos\psi - Bsen\psi)}$$
(3.72.b)

Los momentos mijk y m'ijk evaluados sobre una y otra superficie son:

$$m_{ijk} = \frac{1}{\cos\theta} \iint_{R_{XV}} x^i y^j \left(-\frac{Ax + By + D}{C} \right)^k dx dy$$
 (3.73)

$$m'_{ijk} = \frac{1}{\cos\theta'} \iint_{R'_{xy}} x'^{i} y'^{j} [z'(x',y')]^{k} dx' dy'$$
 (3.74)

con

$$\cos\theta = \frac{|\vec{n} \cdot \vec{k}|}{|\vec{n}|}$$
 $y \cos\theta' = \frac{|\vec{n}' \cdot \vec{k}|}{|\vec{n}'|}$

siendo \vec{n} = (A, B, C) y \vec{n} ' = (A, Csen ψ +Bcos ψ , Ccos ψ -Bsen ψ) los vectores normales a S y S' de forma tal que \vec{n} ' = R (ψ) \vec{n} , es decir $|\vec{n}| = |\vec{n}'|$. Así

$$\cos\theta = \frac{|C|}{|\vec{n}|} \qquad \cos\theta' = \frac{|C\cos\psi - B\sin\psi|}{|\vec{n}|}$$

La relación (3.71.a) permite escribir x'^{i} , y'^{j} , $y^{z'^{k}}$ de la siguiente forma:

$$x^{i} = x^{i}$$
 (3.75.a)

$$y'^{J} = \left[\left(-\frac{Ax + By + D}{C} \right) \operatorname{sen}\psi + y \cos\psi \right]^{J}$$
 (3.75.b)

$$z^{,k} = \left[\left(-\frac{Ax + By + D}{C} \right) \cos \psi - y \sin \psi \right]^{k}$$
 (3.75.c)

Desarrollando las expresiones (3.75.b) y (3.75.c) de la forma

$$y'^{j} = \sum_{\substack{j1, j2=0\\ j1+j2=j}}^{j} \frac{j!}{j!! j2!} \left(-\frac{Ax + By + D}{C} \right)^{j1} (sen\psi)^{j1} (cos\psi)^{j2} y^{j2}$$
(3.76.a)

$$z^{,k} = \sum_{\substack{k1, k2=0 \\ k1+k2=k}}^{k} \frac{k!}{k1! \ k2!} (-1)^{k2} \left(-\frac{Ax + By + D}{C} \right)^{k1} (\cos\psi)^{k1} (\sin\psi)^{k2} y^{k2}$$
(3.76.b)

sustituyendo (3.75.a) y (3.76) en (3.74) y teniendo en cuenta que el jacobiano de la transformación (3.71.a) es tal que

$$\left| J\left(\frac{x'y'}{xy}\right) \right| \frac{\cos\theta}{\cos\theta'} = 1$$

puede establecerse la siguiente relación entre (3.73) y (3.74)

De aquí se observa que frente a la transformación $R_{\mathbf{x}}(\psi)$, los momentos resultantes m'ijk se obtienen como combinación lineal de momentos miln con $0 \le 1 \le j+k$, $0 \le n \le j+k$; es decir los momentos m'ijk, de orden (i+j+k), se obtienen en función de momentos originales miln de hasta orden (i+2(j+k)).

La expresión (3.77), que relaciona m'ijk y min cuando estos se evaluan sobre un único triángulo S y su transformado S', resulta igualmente válida cuando m'ijk y min representen los momentos evaluados sobre el conjunto de triángulos que definen la superficie

completa de un objeto y de su transformado según (3.71.a). Esto es debido a que los coeficientes $(-1)^{k2}(\text{sen}\psi)^{j1+k2}(\cos\psi)^{j2+k1}$ sólo dependen de la rotación $R_x(\psi)$ y no de las peculiaridades de cada triángulo particular, siendo, por tanto, los mismos para todas las caras del poliedro.

Siguiendo un proceso análogo al anterior, puede demostrarse que la relación entre los momentos antes y después de aplicar las rotaciones elementales R (ϑ) y R (φ) viene dada por:

$$m''_{ijk} = \sum_{\substack{11, i2=0\\11+i2=i}}^{1} \frac{i!}{i1! \, i2!} \, (-1)^{12} \sum_{\substack{k1, k2=0\\k1+k2=k}}^{k} \frac{k!}{k1! \, k2!} \, (\cos\vartheta)^{11+k2} \, (\sin\vartheta)^{12+k1} \, .$$

$$(3.78)$$

Combinando las expresiones (3.77) (3.78) y (3.79) se observa que los momentos m''ijk, resultantes de la transformación $R(\varphi,\vartheta,\psi)$, se obtienen en función de momentos originales mima con

$$0 \le 1 \le i+k+j$$

 $0 \le m \le 2(i+j) + k$
 $0 \le n \le 2(i+j) + k$

III.3.3 NORMALIZACION DE FORMAS TRIDIMENSIONALES MEDIANTE MOMENTOS 3D

Los momentos 3D, definidos en (3.48), se emplearán tanto en las tareas de normalización como en las propias de clasificación como componentes del vector de características.

En este apartado se describe cómo los momentos 3D conducen a la

obtención de una forma canónica o estándar para cada objeto tridimensional en estudio. Una vez dispongamos de las formas canónicas de los objetos que componen las distintas clases, se estará en condiciones de extraer un mismo conjunto de características sobre cada una con el fin de distinguir entre las distintas clases.

La estrategia seguida en el proceso de normalización consiste en transformar la forma original 3D en otra, de manera que ésta última satisfaga una serie de criterios. En nuestro caso, los criterios serán condiciones impuestas a los momentos 3D evaluados sobre la superficie de cada objeto.

En términos geométricos, la forma estándar se define como aquella en la que cada objeto 3D tiene su centroide en una misma posición, está igualmente orientada y además con un tamaño específico. En este sentido, las características extraidas sobre cada objeto estándar serán independientes de la posición, orientación y tamaño.

La transformación de cualquier forma tridimensional en su versión canónica final se consigue mediante la aplicación sucesiva de una traslación T, una rotación R y una amplificación A definidas como sigue:

$$T : \begin{bmatrix} x \\ y \\ z \end{bmatrix} \longrightarrow \begin{bmatrix} x + T \\ y + T^{X} \\ z + T^{Y} \end{bmatrix}$$
 (3.81.a)

 $R : R_z(\varphi)R_y(\vartheta)R_x(\psi)$ con

$$R_{X}(\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{bmatrix}$$
 (3.81.b)

$$R_{\mathbf{y}}(\vartheta) = \begin{bmatrix} \cos\vartheta & 0 & -\operatorname{sen}\vartheta \\ 0 & 1 & 0 \\ \operatorname{sen}\vartheta & 0 & \cos\vartheta \end{bmatrix}$$
(3.81.c)

$$R_{Z}(\varphi) = \begin{bmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
 (3.81.d)

$$A: \left[\begin{array}{ccc} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{array} \right] \tag{3.81.e}$$

La determinación de los coeficientes T_x , T_y , T_z , ψ , ϑ , φ y α , se lleva a cabo imponiendo una serie de criterios de normalización sobre la forma estándar resultante.

La expresión (3.80) permite la determinación por separado de T, R y A de manera tal que las formas obtenidas después de cada transformación sigan manteniéndose normalizadas frente a las transformaciones anteriores, es decir:

III.3.3.1 ESTANDAR DE POSICION

El centroide de una forma 3D se define como el punto de coordenadas $(\overline{x}, \overline{y}, \overline{z})$ respecto del referencial absoluto OXYZ con vectores unitarios $\{\vec{i}, \vec{j}, \vec{k}\}$.

$$\overline{x} = \frac{m_{100}}{m_{000}}$$
 $\overline{y} = \frac{m_{010}}{m_{000}}$ $\overline{z} = \frac{m_{001}}{m_{000}}$

siendo m100, m010, y m001 momentos ordinarios de orden 1 y m000 el momento de orden 0 evaluados sobre la superficie del objeto en estudio.

La elección más frecuente para el estándar de posición consiste en situar el centroide de todas las formas 3D en el origen de coordenadas. En términos de momentos 3D la anterior elección equivale a imponer que los momentos mioo, moio y mooi sean nulos sobre la forma canónica. Este criterio se satisface sin mas que aplicar una traslación $\vec{T}=(T_x,T_y,T_z)$ a cada objeto de manera que situe su centroide en el origen de coordenadas. Los valores T_x,T_y,T_z vienen dados por:

$$T_{x} = -\frac{m_{100}}{m_{000}}$$
 $T_{y} = -\frac{m_{010}}{m_{000}}$ $T_{z} = -\frac{m_{001}}{m_{000}}$ (3.82)

siendo mijk los momentos ordinarios evaluados sobre la forma original. Esquemáticamente esta transformación T se indica como

$$T \;:\; \left[\begin{array}{c} x \\ y \\ z \end{array}\right] \longrightarrow \left[\begin{array}{c} x - \overline{\underline{x}} \\ y - \overline{\underline{y}} \\ z - \overline{z} \end{array}\right]$$

Los momentos obtenidos tras situar el centroide en el origen de cordenadas son los momentos centrales μ_{1jk}

$$\mu_{1jk} = \iint_{S} (x - \overline{x})^{1} (y - \overline{y})^{j} [z(x, y) - \overline{z}]^{k} ds$$
 (3.83)

III.3.3.2 ESTANDAR DE ORIENTACION

Como indica el teorema de unicidad [SADJ-80], cada objeto o forma tridimensional queda univocamente definida si se especifica una serie infinita de momentos {mijk}. Sin embargo, si una forma 3D se representa únicamente mediante momentos de hasta orden 2, ésta es

completamente equivalente a un elipsoide. La generalización a 3D del método de los ejes principales nos va a permitir orientar todos los objetos de igual manera. Cada usuario definirá el estándar de orientación -el mismo para todos los objetos- indicando, una vez que ha situado el centroide en el origen de coordenadas, cual es la orientación que ha de poseer el elipsoide de inercia asociado al objeto en estudio.

Es usual elegir el estándar de orientación de manera que los ejes principales se alineen con los ejes coordenados OX, OY y OZ, aunque conviene indicar que esta elección no es la única.

Una vez se ha situado el centroide en el origen de coordenadas mediante la traslación (3.82), será preciso determinar la rotación que alinea los ejes principales del objeto 3D con los ejes del referencial OXYZ.

III.3.3.2.1 ELIPSOIDE DE INERCIA. EJES PRINCIPALES DE INERCIA

Como es de todos conocido, a cada sólido se le asigna una matriz de inercia [GOLD-72], [ARFK-70], [FABE-88] dada por

$$I = \begin{bmatrix} I_{xx} & -I_{xy} & -I_{xz} \\ -I_{xx} & I_{xy} & yz \\ -I_{yx} & yy & yz \\ -I_{zx} & -I_{zz} & I_{zz} \end{bmatrix}$$
(3.84)

Se trata de una matriz real y simétrica. En nuestro caso, y teniendo en cuenta que el centroide se ha situado en el origen de coordenadas, los momentos de inercia Ixx, Iyy, Izz, así como los productos de inercia Ixy, Izx e Iyz se expresan como

$$I_{xx} = \iint_{S} (y^2 + z^2) ds = \mu_{020} + \mu_{002}$$

$$I_{yy} = \iint_{S} (x^2 + z^2) ds = \mu_{200} + \mu_{002}$$

$$I_{zz} = \iint_{S} (x^2 + y^2) ds = \mu_{200} + \mu_{020}$$

$$I_{xy} = I_{yx} = \iint_{S} xy \, ds = \mu_{110}$$

$$I_{xz} = I_{zx} = \iint_{C} xz \, ds = \mu_{101}$$

$$I_{yz} = I_{zy} = \iint_{S} yz \, ds = \mu_{011}$$

donde S es el dominio de integración y representa a la superficie del objeto en estudio y (x,y,z) son las coordenadas de los puntos de dicha superficie. Así la matriz I puede escribirse como

$$I = \begin{bmatrix} \mu_{020} + \mu_{002} & -\mu_{110} & -\mu_{101} \\ -\mu_{110} & \mu_{200} + \mu_{002} & -\mu_{011} \\ -\mu_{101} & -\mu_{011} & \mu_{200} + \mu_{020} \end{bmatrix}$$
(3.85)

El momento de inercia de un cuerpo respecto de un eje cualquiera que pase por el origen, tal como OL en la Fig. 3.7, se define como [BEER-83]

$$I_{OL} = \int p^2 dm = \int (\vec{\lambda} \times \vec{r})^2 dm \qquad (3.86)$$

En nuestro caso la región de integración se reduce a la superfcie S del objeto y por tanto IoL se expresa como

$$I_{OL} = \iint_{S} (\vec{\lambda} \times \vec{r})^{2} ds$$
 (3.87)

siendo $\vec{\lambda}$ el vector unitario asociado al eje OL y \vec{r} = (x,y,z) el vector de posición de cada punto de la superficie S.

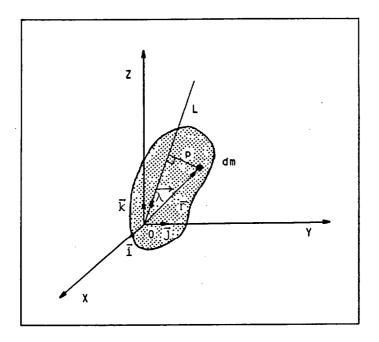


Fig. 3.7. Momento de inercia de un cuerpo respecto de OL. La distancia del elemento de masa dm al eje OL se indica mediante p.

Desarrollando el producto vectorial $\vec{\lambda} \times \vec{r}$ y teniendo en cuenta que las componentes λ_x , λ_y , λ_z coinciden con los cosenos directores de OL, se tiene

$$I_{OL} = I_{xx} \frac{\lambda^{2}}{x} + I_{yy} \frac{\lambda^{2}}{y} + I_{zz} \frac{\lambda^{2}}{z} - 2I_{xy} \frac{\lambda}{x} \frac{\lambda}{y} - 2I_{yz} \frac{\lambda}{y} \frac{\lambda}{z} - 2I_{xz} \frac{\lambda}{x} \frac{\lambda}{z}$$
(3.88)

De la definición (3.86) se observa que IoL es siempre distinto de cero.

Admitamos que el momento de inercia del objeto en estudio se ha calculado para un gran número de ejes OL pasando todos ellos por el origen O, y que sobre cada eje OL se representa un punto P a una distancia OP = $1/\sqrt{IoL}$. El lugar geométrico de los puntos P así definidos determina una superficie.

Teniendo en cuenta que $(OP)\lambda_x = x$, $(OP)\lambda_y = y$, $(OP)\lambda_z = z$, se obtiene, a partir de (3.88), la ecuación de dicha superfice

$$I_{xx}^2 + I_{yy}^2 + I_{zz}^2 - 2I_{xy}xy - 2I_{yz}yz - 2I_{xz}zz = 1$$
 (3.89)

donde x,y,z son ahora las coordenadas rectangulares de los puntos P. Esta ecuación es la de una cuádrica y más concretamente corresponde a un elipsoide centrado en el origen de coordenadas, también conocido por elipsoide de inercia.

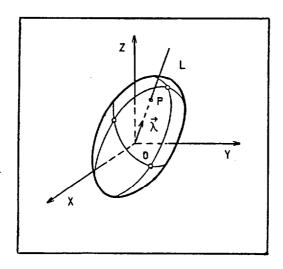


Fig.3.8. Elipsoide de inercia asociado al objeto de la figura 3.7

Ahora bien, si los ejes coordenados, o el sólido, se giran respecto del punto O, entonces los momentos y productos de inercia de (3.89) también cambiarán. Siempre será factible girar los ejes coordenados, o el cuerpo en estudio, de forma que en (3.89) desaparezcan los términos Ixy, Ixz e Iyz. En esta situación la ecuación (3.89) viene dada por

$$I_{xx}^{'}x^{2} + I_{yy}^{'}y^{2} + I_{zz}^{'}z^{2} = 1$$
 (3.90)

que corresponde a la de un elipsoide centrado en el origen de coordenadas y con tamaños de semiejes dados por

$$a = \frac{1}{\sqrt{\frac{I'}{xx}}}$$
 $b = \frac{1}{\sqrt{\frac{I'}{yy}}}$ $c = \frac{1}{\sqrt{\frac{I'}{zz}}}$ (3.91)

siendo I'xx, I'yy I'zz los nuevos valores de los momentos principales de inercia después de haber rotado el elipsoide. A las direcciones del objeto coincidentes con los ejes cartesianos se las conoce como *ejes* principales de inercia.

En términos de la matriz I, la rotación del elipsoide hasta alinear sus ejes principales con los ejes coordenados, o equivalentemente la rotación que aplicada a los ejes del sistema coordenado $\{\vec{1}, \vec{j}, \vec{k}\}$ los alinea con los ejes principales, viene especificada por la transformación ortogonal S que diagonaliza a la matriz I, es decir

$$S^{T} I S = D (3.92)$$

siendo S^T la traspuesta de S y D una matriz diagonal

$$S = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} \qquad D = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$$

Los vectores columna de la matriz S vienen especificados por los autovectores de I y determinan un nuevo sistema coordenado $\{\vec{u}_1,\vec{u}_2,\vec{u}_3\}$ respecto del cual la matriz de inercia será diagonal y adoptará la expresión indicada por D.

$$I \vec{u}_1 = \lambda_1 \vec{u}_1$$

$$I \vec{u}_2 = \lambda_2 \vec{u}_2$$

$$I \vec{u}_3 = \lambda_3 \vec{u}_3$$

$$S = \begin{bmatrix} (\vec{u}_1) & (\vec{u}_2) & (\vec{u}_3) \end{bmatrix}$$

Los vectores ortonormales $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, o autovectores de I, indican la orientación de los ejes principales de inercia del objeto en estudio respecto del referencial $\{\vec{i}, \vec{j}, \vec{k}\}$ (Fig. 3.9). Los momentos (μ 020 + μ 002), (μ 200 + μ 002) y (μ 200 + μ 020) evaluados respecto del sistema

 $\{\vec{u}_1,\vec{u}_2,\vec{u}_3\}$, que son los elementos de la diagonal principal de D, coinciden con los autovalores λ de I y se conocen como momentos principales de inercia.

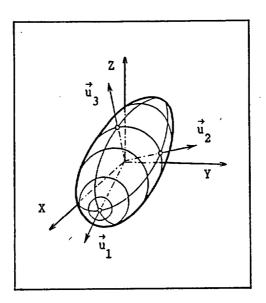


Fig. 3.9. Orientación de los ejes principales de inercia $\{\overrightarrow{u1}, \overrightarrow{u2}, \overrightarrow{u3}\}$ respecto del referencial $\{\overrightarrow{1}, \overrightarrow{j}, \overrightarrow{k}\}$.

III.3.3.2.2 ORIENTACION DE FORMAS 3D

Con el fin de caracterizar la orientación espacial de una forma 3D será necesario asignarle previamente un único referencial, preferiblemente centrado en el propio objeto, lo que facilitará, además, las tareas de reconocimiento [MARR-78a]. La obtención de dicho referencial puede llevarse a cabo en función de los conceptos, anteriormente definidos, de elipsoide y ejes principales de inercia.

Cuando un objeto se representa mediante momentos de hasta orden 2, éste es completamente equivalente a un elipsoide con una orientación y tamaño de semiejes específicos.

Una vez determinados, y clasificados en orden creciente, los

autovalores $\{\lambda_i / \lambda_1 \le \lambda_2 \le \lambda_3\}$ de la matriz de inercia I (3.85) del objeto, tras haber situado su centroide en el origen de coordenadas, se puede afirmar que los autovectores correspondientes $\{\vec{u}_i\}$ forman un sistema ortonormal de vectores de manera que \vec{u}_1 indica la dirección del eje mayor del elipsoide, \vec{u}_3 la del eje menor y \vec{u}_2 la del eje intermedio, tal y como se muestra en la Tabla 3.II.

autovalor	autovector	eje del elipsoide	momento princ. inercia
λ1	√ u1	mayor	menor
λ2	ù 2	intermedio	intermedio
λз	, úз	menor	mayor

TABLA 3.II. Autovalores y autovectores de la matriz de inercia I y su relación con el elipsolde de inercia.

Conviene resaltar que los vectores propios $\{\vec{u}_i\}$ simplemente definen una dirección en el espacio, la asociada a los ejes del elipsoide, pero no un sentido; es decir, los vectores \vec{u}_i y $-\vec{u}_i$ son igualmente válidos como autovectores asociados al autovalor λ_i de la matriz I (3.85).

$$I \vec{u}_i = \lambda \vec{u}_i$$

$$I (-\vec{u}_i) = \lambda (-\vec{u}_i)$$

Así, a cada objeto pueden asignarsele 8 ternas distintas de vectores unitarios y ortogonales $\{\pm \ \overrightarrow{u}_1, \pm \ \overrightarrow{u}_2, \pm \ \overrightarrow{u}_3\}$ diferenciadas entre sí sólo por los sentidos de dichos vectores.

A cada objeto corresponde un único elipsoide de inercia, no siendo cierta la afirmación inversa; es decir, un mismo elipsoide puede corresponder, o estar asociado, a varios objetos distintos e incluso a distintas disposiciones espaciales de un mismo objeto.

De lo anterior se deduce que los autovectores $\{\overrightarrow{u}_i\}$, por sí solos,

no permiten orientar de manera única al sólido al cual están asociados. Para ello será necesario reducir a un único referencial las 8 opciones $\{\pm \ \dot{u}_1, \ \pm \ \dot{u}_2, \ \pm \ \dot{u}_3\}$. En términos de la ecuación (3.92) podriamos decir que la determinación de la matriz de rotación S que aplicada a una forma 3D centrada en el origen de coordenadas la transforma en otra tal que su matriz de inercia sea diagonal, no es única.

Los autovectores $\{\vec{u}_i\}$ aunque son ortonormales, no siempre definen un referencial a derechas, es decir no se puede afirmar que

$$\overrightarrow{u}_1 \times \overrightarrow{u}_2 = \overrightarrow{u}_3$$

$$\overrightarrow{u}_2 \times \overrightarrow{u}_3 = \overrightarrow{u}_1$$

$$\overrightarrow{u}_3 \times \overrightarrow{u}_1 = \overrightarrow{u}_2$$

Lo único que se puede afirmar es que de las 8 posibles ternas $\{\pm \vec{u}_1, \pm \vec{u}_2, \pm \vec{u}_3\}$, 4 definen triedros a derechas y las 4 ternas restantes no. Así, si se exige que el referencial centrado en el objeto sea a derechas, se reducen de 8 a 4 el número de orientaciones ambiguas que puede presentar un objeto al imponer como criterio de normalización frente a rotaciones el hecho de que la matriz de inercia sea diagonal, es decir

$$Q_{101} = 0$$
 $Q_{001} = 0$
 $Q_{110} = 0$
(3.93)

donde Qijk son los momentos evaluados sobre la superficie del objeto tras centrarlo en el origen de coordenadas y someterlo a la rotación S.

Admitiendo que $\{\vec{a}1, \vec{a}2, \vec{a}3\}$ representa uno de los referenciales a derechas asociado al objeto, las 4 orientaciones ambiguas de éste corresponden a giros de π radianes en torno de los ejes $\vec{a}1$, $\vec{a}2$ y $\vec{a}3$ respectivamente, como se indica en la Fig. 3.10.

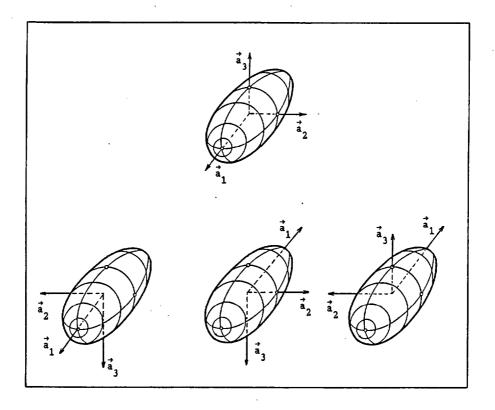


Fig. 3.10. 4 disposiciones diferentes de un mismo referencial $\{\vec{a1}, \vec{a2}, \vec{a3}\}$ a derechas y solidario al objeto, de manera que su elipsoide de inercia permanece inalterado. Las 4 disposiciones corresponden a rotaciones de π radianes en torno de los ejes $\vec{a1}$, $\vec{a2}$ y $\vec{a3}$ coincidentes con las direcciones principales de inercia.

Cuando un objeto 3D asociado al primer referencial se somete a las rotaciones indicadas en la Fig. 3.10, la disposición espacial del mismo cambia, no ocurriendo lo mismo con la forma de su elipsoide asociado, que permanece invariante frente a las citadas rotaciones; es decir, existen 4 disposiciones diferentes de un mismo objeto, de manera que su elipsoide asociado sea indistinguible.

Este hecho sugiere que, en lo que a orientación respecta, la representación de un objeto mediante momentos de segundo orden es insuficiente por no conducir a una única orientación. Esto podría subsanarse estableciendo criterios de normalización basados en momentos de orden superior a dos, o bien manteniendo los criterios

(3.93) junto con criterios heurísticos adicionales.

La primera solución podría conducir a la generalización a 3D de los criterios $\{N_{Pq}\}$ presentados por [ABU-85] para el caso bidimensional.

En nuestro trabajo se ha optado por la segunda vía, es decir imponer una serie de criterios que, junto al criterio de diagonalización de la matriz de inercia, permitan reducir a una única las orientaciones ambiguas que puede presentar cualquier objeto cuando se orienta en términos de su elipsoide de inercia.

La determinación de un único referencial ortonormal y a derechas $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ de entre los 4 triedros, también a derechas, $\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$, $\{\vec{a}_1, -\vec{a}_2, -\vec{a}_3\}$, $\{-\vec{a}_1, \vec{a}_2, -\vec{a}_3\}$, $\{-\vec{a}_1, -\vec{a}_2, \vec{a}_3\}$ alineados con los ejes principales de inercia, se efectua imponiendo las siguientes restricciones [GALV-89]:

- A) El sentido del autovector \overrightarrow{v}_1 , asociado al menor autovalor λ_1 de I, ha de ser tal que esté dirigido desde el centroide hasta el punto del objeto más distante del centroide según la <u>dirección</u> \overrightarrow{v}_1 .
- B) El sentido del autovector \vec{v}_2 , asociado al autovalor λz de I, ha de ser tal que esté dirigido desde el centroide hasta el punto del objeto más distante del centroide según la dirección \vec{v}_2 .
- C) El autovector \vec{v}_3 , asociado al menor de los autovalores λ_3 , se elige de forma que $\vec{v}_1 \times \vec{v}_2 = \vec{v}_3$.

El referencial $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ así elegido se caracteriza por formar un triedro a derechas, con \vec{v}_1 , \vec{v}_2 y \vec{v}_3 autovectores de la matriz de inercia (3.85) del objeto en cuestión, asociados a los autovalores $\lambda_1 \le \lambda_2 \le \lambda_3$ respectivamente, siendo además el único referencial centrado en el objeto que satisface los criterios A), B) y C).

La unicidad está asegurada salvo en aquellas situaciones de

elevada simetría del objeto (por ejemplo en el caso de esferas, cilindros,...) o en los casos de degeneración de los autovalores de I.

A continuación se muestra un resumen del algoritmo empleado en la determinación de \vec{v}_1 y \vec{v}_2 .

- 1. Determinar la línea que pasa por el centroide del objeto y tiene dirección \vec{u}_1 (\vec{u}_2).
- 2. Calcular la intersección de la línea anterior con la superficie del objeto.
- De los puntos intersección calculados en 2 elegir el más distante del centroide.
- 4. Hacer $\vec{v}_1 = \pm \vec{u}_1$ ($\vec{v}_2 = \pm \vec{u}_2$) dependiendo de que el vector dirigido desde el centroide hasta el punto calculado en 3 tenga o no el mismo sentido de \vec{u}_1 (\vec{u}_2)

Notemos que en la determinación de \vec{v}_1 y \vec{v}_2 se presupone que la línea recta especificada en el apartado 1, intersecta con la superficie del objeto. Esta intersección sólo queda garantizada en aquellos casos en los que el centroide sea interior al objeto, situación ésta bastante usual en una gran variedad de formas tales como piezas industriales, por ejemplo. No obstante en situaciones en las que el centroide no sea interior al objeto también será posible determinar \vec{v}_1 y \vec{v}_2 atendiendo al proceso seguido en la reconstrucción de la superfice del objeto a partir de distintas vistas del mismo.

Una vez se ha asignado un único referencial $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ a cada objeto 3D centrado en el origen de coordenadas, sólo resta rotarlo hasta conseguir el estándar de orientación.

La orientación canónica para cada objeto tridimensional viene definida por la rotación R que transforma el triedro $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ en el triedro $\{\vec{k}, \vec{l}, \vec{j}\}$ de forma que alinee \vec{v}_1 con \vec{k} , \vec{v}_2 con \vec{l} y \vec{v}_3 con \vec{j} (Fig. 3.11).

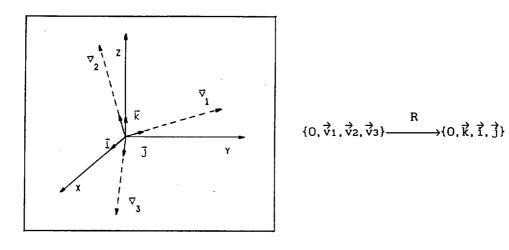


Fig. 3.11. Rotación R que alinea los ejes principales de inercia $\{\vec{v}1, \vec{v}2, \vec{v}3\}$ con el referencial absoluto $\{\vec{k}, \vec{1}, \vec{j}\}$.

Dicha transformación puede especificarse como combinación de tres rotaciones sucesivas $R_x(\psi)$, $R_y(\vartheta)$, $R_z(\varphi)$ en torno de los ejes X,Y,Z respectivamente

$$R (\varphi, \vartheta, \psi) = R_{z}(\varphi) R_{y}(\vartheta) R_{x}(\psi)$$
(3.94)

siendo $R_z(\varphi)$, $R_y(\vartheta)$ y $R_x(\psi)$ matrices de rotación cuya forma es similar a las dadas en (3.81.b)-(3.81.d). Los elementos de esta matriz se expresan en función de los ángulos de giro ψ, ϑ, φ , y éstos a su vez se determinan en función de las coordenadas de \vec{v}_1 , \vec{v}_2 y \vec{v}_3 respecto del referencial $\{\vec{1}, \vec{j}, \vec{k}\}$. Como se muestra en el apéndice A la transformación R viene dada por:

$$R(\psi,\vartheta,\varphi) = \begin{bmatrix} \cos\varphi\cos\vartheta & \cos\varphi\sin\psi\sin\vartheta + \sin\varphi\cos\psi & -\cos\varphi\cos\psi\sin\vartheta + \sin\varphi\sin\psi \\ -\sin\varphi\cos\vartheta & -\sin\varphi\sin\vartheta + \cos\varphi\cos\psi & \sin\varphi\cos\psi\sin\vartheta + \cos\varphi\sin\psi \\ \\ \sin\vartheta & -\sin\varphi\cos\vartheta & \cos\psi\cos\vartheta \end{bmatrix}$$

Los ángulos ψ , ϑ , y φ se caracterizan por

$$\cos \psi = \frac{v_{1z}}{v} \qquad \cos \vartheta = v \qquad \cos \varphi = v_{2Tx}$$

$$\sin \psi = -\frac{v_{1y}}{v} \qquad \sin \vartheta = v_{1x} \qquad \sin \varphi = v_{2Ty}$$
(3.95)

con

$$v = \sqrt{v_{1y}^{2} + v_{1z}^{2}}$$

$$\vec{v}_{1} = v_{1x} \vec{1} + v_{1y} \vec{j} + v_{1z} \vec{k}$$

$$\vec{v}_{2} = v_{2x} \vec{1} + v_{2y} \vec{j} + v_{2z} \vec{k}$$

$$\vec{v}_{3} = v_{3x} \vec{1} + v_{3y} \vec{j} + v_{3z} \vec{k}$$

$$\vec{v}_{2T} = v_{2Tx} \vec{1} + v_{2Ty} \vec{j} + v_{2Tz} \vec{k}$$

$$\vec{v}_{2T} = R_{v}(\vartheta) R_{x}(\psi) \vec{v}_{2}$$

III.3.3.3 ESTANDAR DE TAMAÑO

Cuando objetos geométricamente similares pero de diferentes tamaños quieran ser reconocidos como iguales, es decir asignados a la misma clase, la forma canónica asociada a cada objeto habrá de estar normalizada no sólo frente a traslaciones y rotaciones sino que además ha de poseer un tamaño estándar.

Una vez que la forma 3D se ha centrado en el origen de coordenadas aplicando la traslación T (3.81.a), (3.82), y se ha rotado hasta alinear sus ejes principales con los ejes coordenados mediante la transformación R (3.94), el tamaño estándar vendrá definido por la transformación de escalado isótropo (3.81.e) tal que la versión canónica resultante de aplicar A al estándar de posición y orientación satisfaga un conjunto de condiciones de normalización.

Los criterios impuestos sobre la versión estándar final no necesariamente han de ser únicos. La diferencia entre ellos estriba en la estrategia seguida en la determinación de un único valor α que los

satisfaga.

Dos posibles criterios de normalización frente a cambios de escala, y de una fácil interpretación geométrica, son los siguientes.

CRITERIO 1

El tamaño estándar se define de manera que el momento de orden 0 evaluado sobre la forma canónica final tome un valor constante C previamente establecido.

$$\eta_{000} = C \tag{3.96}$$

La relación entre η_1 j $_k$ y los momentos Q_1 j $_k$, obtenidos sobre el estándar de posición y orientación, viene dada por

$$\eta_{ijk} = \alpha^{(i+j+k+2)} Q_{ijk}$$
(3.97)

siendo η_{ijk} momentos evaluados sobre el estándar final, y por tanto normalizados frente a traslaciones, rotaciones y cambios de escala. El parámetro α que satisface la condición de normalización (3.96) será

$$\alpha = \left(\frac{\eta_{000}}{Q_{000}}\right)^{1/2} = \left(\frac{C}{Q_{000}}\right)^{1/2}$$
 (3.98.a)

o en función de momentos ordinarios mijk sobre la forma original

$$\alpha = \left(\frac{C}{m_{000}}\right)^{1/2} \tag{3.98.b}$$

ya que mooo = μ ooo = Qooo

En términos geométricos, el valor de α especificado en (3.98) equivale a imponer que todas las formas en su versión estándar final tengan el mismo valor C para el área de su superficie.

CRITERIO 2

En este caso el parámetro α se elige de manera que todas las formas en su estándar final se caractericen por tener un mismo valor L para el semieje mayor del elipsoide de inercia asociado.

Una vez la forma tridimensional está en su estándar de posición y orientación, la matriz de inercia

$$\mathbf{I} = \begin{bmatrix} \mathbf{Q}_{020} + \mathbf{Q}_{002} & -\mathbf{Q}_{110} & -\mathbf{Q}_{101} \\ -\mathbf{Q}_{110} & \mathbf{Q}_{200} + \mathbf{Q}_{002} & -\mathbf{Q}_{011} \\ -\mathbf{Q}_{101} & -\mathbf{Q}_{011} & \mathbf{Q}_{200} + \mathbf{Q}_{020} \end{bmatrix}$$

es diagonal, siendo sus valores propios los momentos principales de inercia de manera tal que

$$(Q_{200} + Q_{020}) \le (Q_{020} + Q_{002}) \le (Q_{200} + Q_{002})$$

es decir, el menor autovalor (Q200 + Q020) coincide con el momento principal de inercia en torno del eje z. Además, de (3.91) se observa que el tamaño del semieje mayor, asociado a este autovalor, viene dado por

$$\frac{1}{\sqrt{(Q_{200} + Q_{020})}}$$

Cuando a esta forma se la somete a una transformación del tipo (3.81.e), la matriz de inercia de la forma resultante sigue siendo diagonal, teniendo ahora por expresión

$$\left[\begin{array}{ccccc} \eta_{020} + \eta_{002} & 0 & 0 \\ 0 & \eta_{200} + \eta_{002} & 0 \\ 0 & 0 & \eta_{200} + \eta_{020} \end{array}\right]$$

con

$$(\eta_{020} + \eta_{002}) = \alpha^4 (Q_{020} + Q_{002})$$

$$(\eta_{200} + \eta_{002}) = \alpha^4 (Q_{200} + Q_{002})$$

$$(\eta_{200} + \eta_{020}) = \alpha^4 (Q_{200} + Q_{020})$$

como se sigue de (3.97). Esto es, la forma resultante sigue estando normalizada frente a traslaciones y rotaciones con la única salvedad de que ahora ha cambiado el tamaño de sus semiejes en la misma proporción. Concretamente, el nuevo valor para el semieje mayor viene dado por

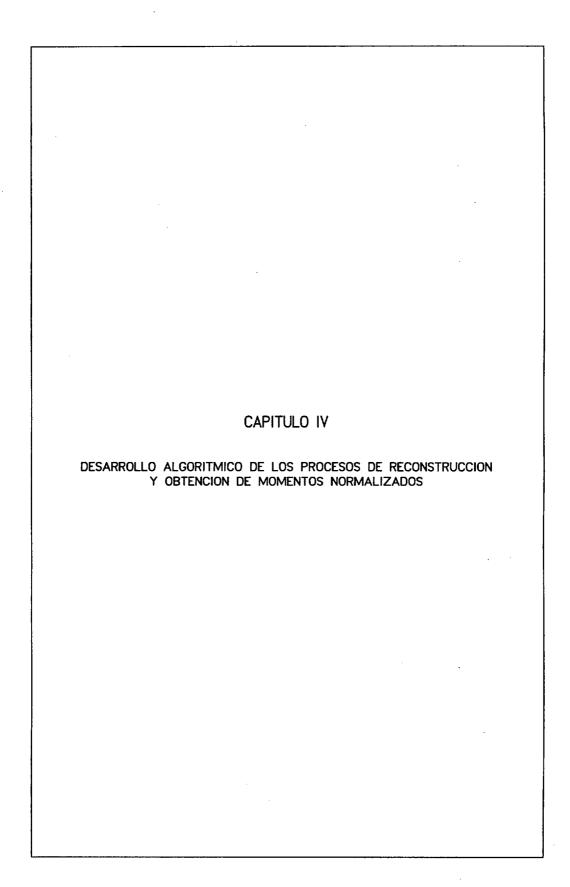
sem. mayor =
$$\frac{1}{\sqrt{(\eta_{200} + \eta_{020})}} = \frac{1}{\alpha^2 \sqrt{(Q_{200} + Q_{020})}}$$

El parámetro de escalado α que normaliza a L el tamaño del semieje mayor para todos los objetos en su versión estándar final será

$$\alpha = \left[\frac{1}{L \sqrt{(Q_{200} + Q_{020})}} \right]^{1/2}$$
 (3.99)

Conviene indicar, finalmente, que cuando la forma está en su estándar de posición y se la somete a la rotación R que la normaliza frente a rotaciones, la forma resultante, además de alinear sus ejes principales con los ejes coordenados, sigue manteniendo su centroide en el origen de coordenadas. Igualmente, si al estándar de posición y orientación se le somete a la amplificación A que normaliza su tamaño, se sigue cumpliendo que la forma resultante mantiene su centroide en el origen de coordenadas sin haber alterado su orientación.

De esta manera se ha mostrado un procedimiento conducente a la obtención de la versión canónica para cualquier forma 3D. Los momentos de cualquier orden que sobre ella se calculen, serán independientes de la posición, orientación y tamaño de la forma original. Esto supone una mejora de sustancial importancia respecto de los 2 invariantes 3D obtenidos por Sadjadi a partir de la compleja y tediosa teoría de invariantes algebraicos.



IV.1 PROYECCION PARALELA. COMPARACION CON OTRAS TRANSFORMACIONES DE PROYECCION

Tareas computacionales relativas a la visión deben explicar cuestiones tales como: ¿Cómo se perciben correctamente propiedades tridimensionales de objetos a partir de proyecciones bidimensionales?. Puesto que existe una pérdida de información al proyectar una escena tridimensional sobre una superficie bidimensional, la interpretación de imágenes, posiblemente afectadas de ruido, incompletas, ambiguas, etc., requerirá de algún modelo cognitivo computacional.

Entre las propiedades espaciales que de un objeto percibimos destacan, fundamentalmente, el <u>tamaño</u>, <u>la forma</u> y <u>la profundidad</u>. Mientras el tamaño y forma se definen en relación al propio objeto, la profundidad de éste se define en relación al observador. El tamaño de un objeto usualmente se determina mediante el empleo de una métrica euclídea: longitud, área, volumen. Por otra parte, la definición de la forma de un objeto es una tarea más difícil; no obstante, con este término se hace referencia a la disposición espacial de los contornos y superficies que componen el objeto.

En una gran cantidad de situaciones, la forma de un objeto puede describirse adecuadamente mediante líneas rectas y planos. Estos descriptores geométricos tan simples son comunes en conteniendo objetos manufacturados por el hombre y quizás menos en naturales. Si se es capaz de desarrollar métodos computacionales eficientes para la percepción de líneas y planos, probablemente éstos puedan generalizarse para poder tratar formas más complejas.

En la recuperación de formas 3D a partir de proyeciones 2D, es esencial disponer de algún modelo de la transformación proyectiva $3D\longrightarrow 2D$. Dos son las más usuales: proyección paralela u ortográfica y proyección central o perspectiva (Fig. 4.1).

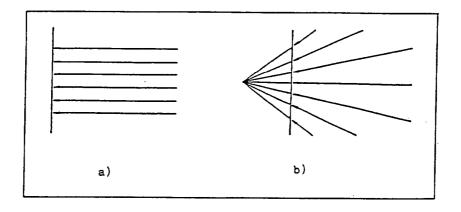
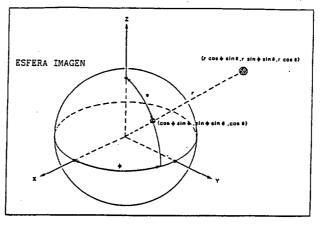


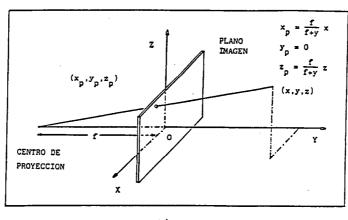
Fig. 4.1. a) Proyección Paralela. b) Proyección Central.

En proyección paralela, la imagen se forma mediante rayos paralelos, usualmente perpendiculares al plano imagen. En proyección central, la imagen se forma mediante rayos que convergen en un punto llamado punto focal. Conviene indicar que la proyección central es el modelo apropiado para la visión en humanos y mediante cámaras [BARN-83]. No obstante, otros modelos de proyección han sido propuestos. En [IKEU-84] se presenta un modelo de proyección esférica con características similares a la proyección central, siendo una esfera el lugar donde proyectan los rayos para formar la imagen en vez de un plano como en proyección central. En la Fig. 4.2 se indican las características de estos tres modelos de proyección.

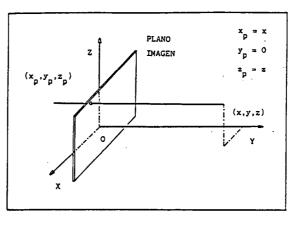
El parámetro más importante que distingue la proyección central de la paralela es el ángulo de visión, definido como el ángulo entre dos rayos con máxima separación angular. El empleo de proyección ortográfica equivale a considerar ángulo de visión nulo, mientras que en proyección central y esférica éste es distinto de cero. Localmente, proyecciones perspectiva У esférica son aproximadamente ortográficas puesto que en este caso la separación angular entre rayos próximos es pequeña. Sin embargo, cuando se requiere la consideración de la imagen en su conjunto, las características de perspectiva resultan importantes. Si la distancia focal es grande en comparación con las dimensiones lineales de la imagen, el ángulo de visión será pequeño y la aproximación de proyección paralela será razonable.



a)



b)



c)

Fig. 4.2. a) PROYECCION ESFERICA: el observador está situado en el centro de la esfera. Cualquier punto se proyecta sobre la superficie esférica a través de su centro. b) PROYECCION CENTRAL: cualquier punto (x,y,z) proyecta en el punto de coordenadas (xp,0,zp) del plano imagen. c) PROYECCION PARALELA: cualquier punto (x,y,z) proyecta en el punto de coordenadas (xp,0,zp) del plano imagen.

Además, puede demostrarse [WANG-87] que la proyección ortográfica constituye una buena aproximación de la proyección central en situaciones en las que la distancia cámara-objeto sea relativamente mayor que la distancia focal de la cámara.

Otro factor que distingue las transformaciones perspectiva y esférica de la paralela, incluso cuando el ángulo de visión sea pequeño, es el efecto de distancia o razón entre las distancias al plano imagen que presentan distintos objetos de una misma escena. Bajo los modelos de proyección central y esférica, el área proyectada de un objeto varía inversamente con la distancia del objeto al punto focal. Sin embargo, bajo proyección ortográfica el tamaño de un objeto en la imagen es independiente de la separación objeto-plano imagen debido, justamente, a la consideración de distancia focal infinita.

En las transformaciones de proyección paralela y esférica, la línea de visión es perpendicular a la superficie sensible o de formación de imagen (plano, esfera). De esta manera, ninguna de ellas muestra distorsiones en la imagen, mientras que bajo proyección central la forma del objeto proyectado variará con la localización espacial del objeto, siendo, además, función del ángulo formado por la línea de visión y el plano imagen. El requerimiento de línea de visión perpendicular a la superficie de formación de imágenes se satisface en la visión humana por la necesidad de mover los ojos y/o cabeza para examinar objetos desde diferentes direcciones. Sin embargo, los sistemas usuales de formación de imágenes fabricados por el hombre tienen un área plana como superficie sensible, siendo ésta la principal razón para el empleo de proyección central frente a la esférica.

Existen también diferencias cualitativas entre los modelos de proyección central y paralela. Asi, por ejemplo, la interpretación de imágenes ambiguas (ejemplo del cubo de Necker, [BARN-83]) presenta menos dificultades en proyección central que en ortográfica o paralela. Además, los puntos y líneas de fuga no se encuentran en proyección paralela, siendo característicos en proyección central.

Por último, los modelos de proyección central y esférica comportan un tratamiento matemático considerablemente más complejo que en proyección paralela [HARA-80]. Este es el motivo por el que muchos sistemas de visión artificial computerizados se desarrollan, inicialmente, bajo la consideración de proyección ortográfica o paralela. Así, por ejemplo, Horn [HORN-75] usó esta aproximación en el análisis de "Shape From Shading". Otros trabajos en los que se considera proyección ortográfica pueden encontrarse en [IKEU-81b], [WOOD-80], [MART-81,82a,82b,83], [KANA-83], [WANG-85,87].

En nuestro trabajo se ha optado por el modelo de proyección paralela debido principalmente a la sencillez matemática que conlleva, sobre todo en la implementación del algoritmo de reconstrucción 3D, frente a la más laboriosa tarea que comportaría el empleo de proyección central. Aún tratándose de una aproximación, no siempre válida, su uso queda, además, justificado al trabajar con objetos tridimensionales simulados mediante modelos poliédricos y no con datos reales.

IV.2 CONSTRUCCION DE MODELOS PARA OBJETOS TRIDIMENSIONALES

Como se mencionó anteriormente, en nuestro trabajo no se han empleado datos correspondientes a escenas reales. La razón fundamental radica en la no disponibilidad de un sistema de adquisición y procesado de imágenes en el momento de iniciar el trabajo. Ello motivó el uso de objetos sintéticos modelados y simulados sobre un ordenador. Análogamente, el proceso de obtención de imágenes, así como la etapa de reconstrucción 3D a partir de las citadas vistas 2D, fueron simulados en su totalidad sobre un ordenador VAX 11/780. Como contrapartida, el empleo de datos sintéticos frente a datos reales ha permitido trabajar bajo unas condiciones perfectamente controladas, evitando así problemas adicionales derivados de las condiciones de iluminación, ruido, calibración de cámaras, segmentación, etc. En esta situación ha sido posible evaluar de manera detallada la bondad y eficiencia del método de reconstrucción 3D utilizado.

mediante configuraciones objetos pueden modelarse Los tridimensionales de primitivas basadas en superficies o volúmenes con unos atributos y relaciones específicas entre ellas. Dichas primitivas y sus relaciones admiten, bien una descripción geométrica o una descripción simbólica más abstracta. Los modelos geométricos son más apropiados para describir objetos específicos, particularmente los manufacturados por el hombre y que presenten estructuras regulares. Los modelos simbólicos, representados formalmente mediante un grafo cuyos nodos son las primitivas 3D y cuyos arcos son los atributos y relaciones entre primitivas, son más apropiados para naturales, mejor definidos en términos de características genéricas que en función de su forma precisa.

La propia naturaleza del proceso de reconocimiento será función del modelo empleado para describir los objetos. Desafortunadamente, y con frecuencia, las representaciones apropiadas para una aplicación no son computacionalmente aptas para otros procesos [BALL-82].

En nuestro caso se ha optado por un modelo geométrico cuyas primitivas son superficies planas. Estos modelos son frecuentemente utilizados en visión artificial ya que la superficie es la parte visible de un objeto y ésta lo describe sin ambigüedad en la mayoria de las situaciones [REQU-80]. Cada modelo 3D consiste en un par ordenado (P,R) donde $P = \{P_1, P_2, \ldots, P_N\}$ es un conjunto de vértices en el espacio tridimensional y R es una relación, R c $P \times P$, que contiene todos los pares de vértices en P que están conectados por una arista.

El banco de objetos empleados como base de trabajo en la reconstrucción 3D y posterior reconocimiento, se compone de las 13 piezas tridimensionales mostradas en la Fig. 4.3 y cuyas formas presentan distintos grados de complejidad. Estos objetos se han elegido sólidos rígidos y sin agujeros. El tratar con objetos rígidos viene motivado por ser los más frecuentemente encontrados en entornos industriales. La no consideración de objetos con agujeros no implica pérdida de generalidad en el proceso de reconstrucción 3D presentado; por el contrario, presenta la simplificación de no necesitar de detección de contornos interiores de las imágenes que de ellos se obtengan.

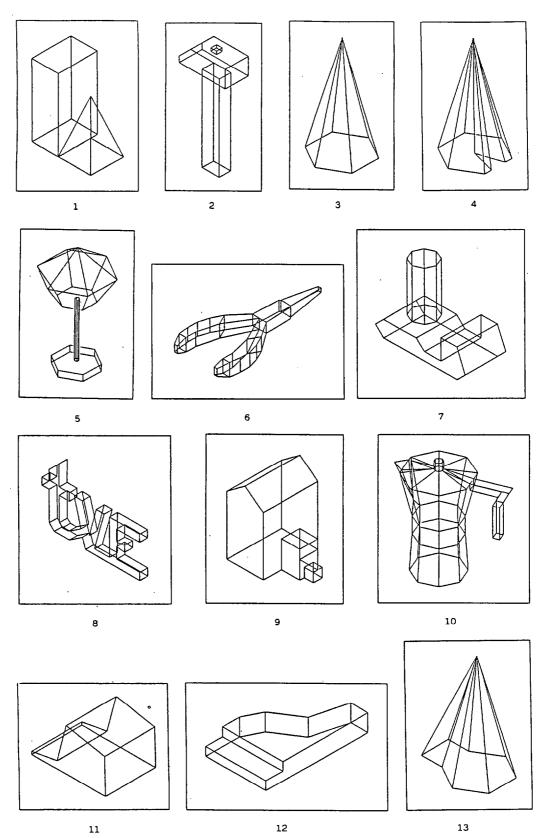


Fig. 4.3. Banco de objetos empleado en el proceso de reconstrucción 3D y posterior reconocimiento.

IV.2.1 MODELADO GEOMETRICO

En la descripción de cada uno de los objetos indicados en la Fig. 4.3 se han empleado dos modelos íntimamente relacionados entre sí. Cada objeto se modela mediante una superficie poliédrica de caras planas (preferentemente triangulares, aunque no siempre ha sido así). El citado modelo poliédrico se obtiene a partir de uno de alambres (wireframe) en el que se distinguen claramante dos tipos de datos. De un lado los referentes a la información geométrica y por otro lado los relativos a la información topológica. La geometría del modelo implica las dimensiones y localización espacial de cada componente, mientras que la topología indica las conexiones entre los distintos elementos.

Los elementos constituyentes en el modelo de alambres son los vértices y aristas que lo definen, mientras que en el modelo poliédrico, además de los anteriores, hay que indicar las caras que componen la superficie. La representación mediante un modelo de alambres no especifica explícitamente las caras de su superficie. Esto se lleva a cabo asignando superficies planas a cada conjunto de aristas que definan un camino cerrado, de manera que el objeto descrito por el modelo de alambres (vértices + aristas) esté limitado por dicha estructura.

La información relativa a cada uno de los modelos representativos de los objetos indicados en la Fig. 4.3 se ha almacenado en un fichero. A modo de ejemplo, en la Fig. 4.4. se muestra el correspondiente al objeto número 12.

En primer lugar aparecen 5 registros con datos identificativos referentes al nombre del objeto, organización y contenido de los registros siguientes.

IV.2.1.1 INFORMACION TOPOLOGICA

A continuación, y en un bloque de registros de número variable dependiendo del número de vértices del modelo, se presenta la

información topológica del objeto 3D. Esta información hace referencia al número de vértices y conexiones entre ellos que definen las aristas.

```
'OBJETO 3D : PIANO'
'El número de vértices NV de este objeto es 18. Datos en registro 6'
'Matriz INDICE(I:NV) en registro 7'
'Matriz CONEXION(I, J) I=1, NV J=1, INDICE(I) escrita por filas en reg. 8:25'
'Coordenadas 3D X,Y,Z de cada vértice en registros 26:43. Centroide en 44'
18
18*3
7 2 8
1 3 9
2 4 10
3 5 11
4 6 16
5 7 15
6 1 14
14 9 1
8 10 2
9 11 3
10 12 4
11 13 17
12 14 18
13 8 7
18 16 6
15 17 5
16 18 12
15 17 13
1.0 6.0 4.0
5.0 5.0 4.0
7.0 2.0 4.0
9.0 1.0 4.0
11.0 1.0 4.0
11.0 8.0 4.0
1.0 8.0 4.0
1.0 6.0 2.0
5.0 5.0 2.0
7.0 2.0 2.0
9.0 1.0 2.0
13.0 1.0 2.0
13.0 8.0 2.0
1.0 8.0 2.0
11.0 8.0 3.0
11.0 1.0 3.0
13.0 1.0 3.0
13.0 8.0 3.0
                                     2.913945
7.966008
                5. 175625
```

Fig. 4.4. Fichero que caracteriza al objeto nº. 12 de la Fig. 4.3.

Cada vértice se etiqueta asignándole un número de manera que al

hacer referencia a un vértice concreto, éste se representará por su número correspondiente. En el registro 6 del fichero de datos siempre se indica el número de vértices NV. Además de los vértices será preciso definir las aristas que los conectan entre sí, y que conforman el modelo de alambres. Cada arista vendrá especificada por los dos vértices finales que la caracterizan. Esta información está contenida en los arrays INDICE y CONEXION de la siguiente forma. En el registro 7 se almacena el array INDICE de dimensión 1xNV. El elemento INDICE(I) es un entero que indica el número de vértices que conectan con el vértice I mediante una arista, es decir especifica el número de aristas que confluyen en el vértice I. En los NV registros siguientes se almacena, por filas, el contenido de CONEXION. El significado de cada elemento de esta array es el siguiente: el vértice CONEXION(I, J) define, junto con el vértice I, una de las aristas del modelo 3D. Esta arista corresponde a la conexión J-ésima del vértice I.

IV.2.1.2 INFORMACION GEOMETRICA

A continuación, y en los NV registros siguientes al array CONEXION, se indica la información geométrica del modelo. En cada uno de los NV registros se especifica, para cada vértice, su posición espacial mediante las coordenadas (x,y,z) respecto del sistema de referencia global de coordenadas centrado en el origen OXYZ. Dicho sistema de referencia será el que se emplee también para especificar la orientación y tamaño de cada objeto 3D, así como la posición y orientación de cada plano imagen sobre el que se obtengan las distintas vistas bidimensionales del objeto en cuestión.

Finalmente, y en el último registro del fichero, se indican las coordenadas del centroide del objeto. Estas coordenadas se obtienen, no en función de las aristas del modelo de alambres sino a partir de las caras definidas por estas aristas, y que conforman la superficie poliédrica del objeto. En la evaluación de dichas coordenadas (Xc, Yc, Zc) se le asigna un peso unidad a cada punto de la superficie poliédrica y cero en el resto.

IV.3 DESARROLLLO DEL PROCESO DE RECONSTRUCCION 3D

En la Fig. 4.5 se muestra un esquema global del proceso de reconstrucción tridimensional, a partir de múltiples vistas, descrito en el capítulo II de esta memoria. Dicho diagrama corresponde a los algoritmos de reconstrucción inicial y posterior refinamiento allí descritos. Es el objetivo de los siguientes apartados presentar el desarrollo, así como las características más notables, de los distintos programas aquí empleados. Estos son:

- PROYECCION
- CONTORNO
- RASTERIZAR
- PARALELOGRAMOS
- REFINAMIENTO

Una vez se ha construido el modelo de alambres y se conoce la lista de vértices, aristas y caras planas que componen el poliedro representativo del objeto 3D a reconstruir, se obtiene mediante simulación, y aplicando el programa PROYECCION, una colección de vistas desde posiciones distintas; dichas vistas se almacenan en los ficheros de nombre genérico vistan.imag, con n = 1,2,... Seguidamente se extraen los contornos de las siluetas definidas por las distintas vistas; esto se consigue después de aplicar el programa CONTORNO. Los polígonos representativos de cada contorno se almacenan en los ficheros del tipo <u>vistan.contor</u>, n = 1,2,3... A partir de las dos primeras vistas se determina la intersección de los dos planos imagen. Mediante el programa RASTERIZAR se obtiene una descripción alternativa de los contornos <u>vistal.contor</u> y <u>vista2.contor</u>; la nueva descripción corresponde al resultado de evaluar la intersección de sendas máscaras de líneas rectas paralelas- y contenidas en cada plano imagen- con el correspondiente contorno. Para cada plano imagen la dirección de las líneas de rasterización es perpendicular a la línea común a ambos planos. Cada imagen rasterizada se almacena en los ficheros vista1.ras y <u>vista2.ras</u> respectivamente.

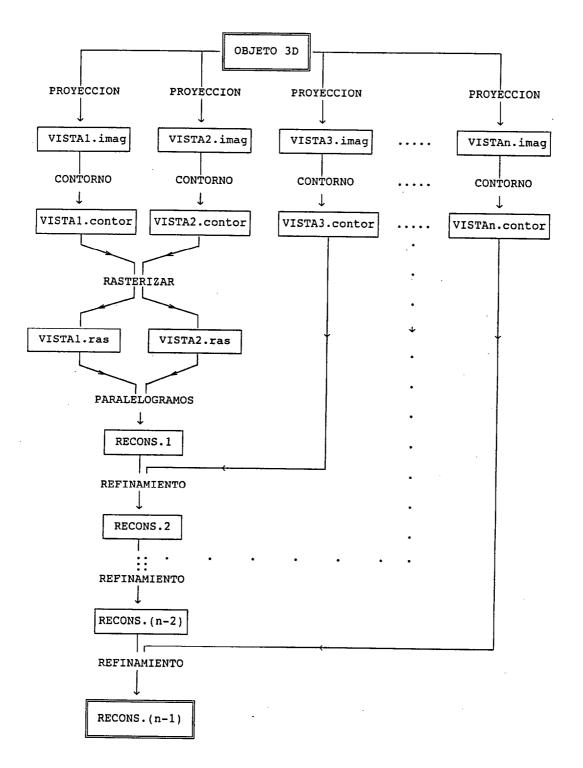


Fig. 4.5. Diagrama global del proceso de reconstrucción 3D a partir de múltiples vistas.

El conjunto de segmentos de rasterización que describen a las dos primeras vistas se combinan adecuadamente en el programa PARALELOGRAMOS para obtener, finalmente, la primera reconstrucción tridimensional del objeto en estudio. Dicha reconstrucción se compone de un conjunto de paralelogramos que en su interior encierran al objeto 3D de partida. Puesto que la reconstrucción inicial recons.1, por lo general, constituye una aproximación burda del objeto 3D, se precisa de la información adicional contenida en los contornos de nuevas vistas, vista3.contor, vista4.contor, ..., vistan.contor, para ir acotando con más exactitud la zona espacial ocupada por el objeto en reconstrucción. Esta tarea se lleva a cabo mediante repetidas aplicaciones del programa REFINAMIENTO.

IV.3.1 PROCESO DE ADQUISICION DE VISTAS DE OBJETOS 3D

En este punto se describe la aplicación del proceso de proyección para la obtención del conjunto de vistas 2D que posteriormente se emplearán en la reconstrucción 3D.

El proceso de adquisición de las vistas por las cámaras ha sido simulado, usando como objetos de prueba los de la Fig. 4.3. Un aspecto interesante es el posicionamiento y orientación de las cámaras. Teniendo en cuenta que se trabaja bajo la consideración de proyección paralela, cada cámara puede sustituirse por su plano imagen.

De cada objeto se han obtenido 10 vistas desde distintas posiciones y orientaciones. Estos parámetros se especifican indicando un punto del plano imagen y la dirección normal al mismo. Para cada cámara las orientaciones se han elegido de forma que no existan dos suficientemente similares, con el fin de poder adquirir información lo más variada y menos redundante posible. En la Tabla 4. I se muestran los vectores representativos de las mencionadas orientaciones, así como un punto de cada plano, ambos expresados en la base $\{\vec{1}, \vec{j}, \vec{k}\}$ del sistema global de referencia OXYZ.

VISTA	ORIENTACION	POSICION
1	(1, 0, 0)	(100, 0, 0)
2	(0, 1, 0)	(0, 100, 0)
3	(0, 0, 1)	(0, 0, 100)
4	(0.889, -0.249, -0.383)	(889, -249, -383)
5	(-0.501, -0.079, 0.862)	(-501, -78, 862)
6	(-0.452, 0.634, 0.628)	(-452, 634, 628)
7	(0.652, 0.757, -0.047)	(652, 757, -47)
8 .	(0.956, -0.042, -0.290)	(956, -42, -290)
9	(0.703, -0.146, 0.696)	(703, -146, 696)
10	(-0.691, -0.434, -0.578)	(-691, -434, -578)

TABLA 4.I. Disposición espacial de los planos imagen para las 10 cámaras. Cada plano se especifica mediante su orientación (vector normal)

Para las tres primeras vistas, los planos de cada cámara se toman perpendiculares a los vectores ortonormales \vec{i} , \vec{j} , \vec{k} , siendo , por tanto, perpendiculares entre sí. Las 7 direcciones restantes se han elegido aleatoriamente, con la única condición de no determinar un ángulo menor de 14° entre dos cualesquiera de ellas. En la Fig. 4.6 se muestra la disposición espacial del conjunto de las cámaras en el referencial absoluto OXYZ.

Los objetos en estudio se consideran en reposo, tratándose por tanto de escenas estáticas, lo que simplifica el problema de reconstrucción al no tener que hacer intervenir la variable tiempo como en análisis de escenas dinámicas. Así, el conjunto multicámara podrá ser sustituido por una única cámara colocada en posiciones distintas en diferentes instantes de tiempo.

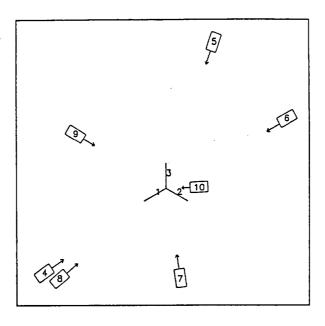


Fig. 4.6. Orientación y localización espacial del conjunto de cámaras respecto del sistema global de referencia.

Al programa PROYECCION se le proporciona, como datos de entrada, la siguiente información:

- -Fichero conteniendo la posición y orientación de cada plano imagen.
- -Fichero con datos topológicos y geométricos del objeto 3D en observación.

A la salida, PROYECCION genera un fichero en el que se almacenan los datos de posición y orientación de la cámara, así como la misma información topológica del objeto 3D. Este fichero contiene, además, las coordenadas de cada vértice del objeto pero proyectadas sobre el plano imagen. Dichas coordenadas no se expresan en función de la base $\{\vec{i}, \vec{j}, \vec{k}\}$ sino en términos de una nueva base $\{\vec{u}, \vec{v}\}$ que define los ejes horizontal y vertical del plano imagen. Esta base $\{\vec{u}, \vec{v}\}$ se determina durante la ejecución de PROYECCION y en función de las coordenadas 3D de cada vertice proyectado, de manera que todos estos puntos presenten coordenadas positivas en la base $\{\vec{u}, \vec{v}\}$, como

corresponde a una cámara real.

De manera resumida, el proceso PROYECCION se ejecuta en los siguientes pasos:

- Leer información del objeto 3D a proyectar (lista de vértices + aristas).
- 2. Situar la cámara en la posición y orientación deseadas.
- 3. Proyectar cada uno de los vértices del modelo poliédrico 3D perpendicularmente al plano imagen (proyección ortográfica).
- 4. Determinar el punto origen y vectores $\{\vec{u}, \vec{v}\}$ que definen la base del plano imagen (ejes horizontal y vertical del plano de la cámara).
- 5. Expresar cada punto proyectado en la base $\{\vec{u}, \vec{v}\}\$ (Coordenadas en el plano imagen).

6. Almacenar:

- Datos identificativos e informativos del fichero resultante
- Posición y orientación de la cámara.
- Número de vértices proyectados.
- Información topológica de conexión entre vértices.
- Cordenadas de cada punto proyectado en la base del plano imagen.

7. Finalizar.

En la Fig. 4.7 se muestra la proyección del objeto nº. 12 después de aplicar el programa PROYECCION para obtener la vista nº. 4. La información topológica es la misma en los ficheros de entrada y salida, debido a que el proceso de proyección, ya sea paralela o central, mantiene conectados en 2D a los mismos vértices que ya lo estaban en 3D.

En el programa PROYECCION se hace uso de la subrutina INTER_PLANO_RECTA para determinar la intersección del plano de la cámara con cada una de las líneas que, partiendo de cada vértice, se dirigen perpendicularmente del objeto a la cámara.

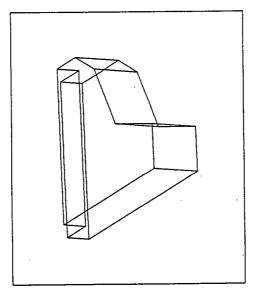


Fig. 4.7. Proyección paralela del modelo de alambres del objeto n° . 12 sobre el plano imagen para la vista n° . 4.

IV.3.2 EXTRACCION DE CONTORNOS

El conjunto de vértices que definen cada objeto 3D se ha transformado, mediante el proceso de proyección ortográfica, en otro conjunto de puntos contenidos en el plano de la cámara. Dichos puntos se conectan entre sí mediante segmentos rectilíneos, que son la proyección de las aristas que en 3D conectaban los correspondientes vértices del modelo poliédrico del objeto.

Como se apuntó en el capítulo II, el método de reconstrucción tridimensional empleado se basa, exclusivamente, en los contornos de las distintas vistas 2D sin tener en cuenta otro tipo de información que la imagen pueda contener. Es el objetivo de este apartado, una vez que se dispone de las distintas vistas obtenidas al simular el proceso de adquisición por una cámara, extraer y representar de manera adecuada los contornos de las citadas proyecciones 2D.

En el caso de imágenes reales, la extracción de contornos podría efectuarse en dos etapas consecutivas. La silueta de una imagen se obtiene tras un simple proceso de umbralizado de los valores de intensidad de cada pixel; la imagen binaria así obtenida se somete,

seguidamente, a un análisis de componentes conectadas [ROSE-76], [PAVL-82], [MART-83].

En nuestro caso, el contorno de cada vista vendrá definido por el polígono más externo que determinan los distintos segmentos rectilíneos en la proyección 2D (Fig. 4.7). El contorno vendrá así especificado por una lista ordenada de vértices, de manera que cada vértice determine con el siguiente un segmento rectilíneo, a excepción del último vértice que conecta con el primero.

A diferencia del caso real, en donde es posible la extracción de contornos interiores de la silueta, en el procedimiento de simulación desarrollado no se contempla la determinación de posibles contornos interiores. Esto se debe, básicamente, al tipo de modelo de alambres 3D empleado en la construcción de objetos tridimensionales. No obstante, en aquellas vistas donde hubiese tales contornos interiores, la no determinación de ellos no afecta grandemente al proceso de reconstrucción 3D, ya que dicha deficiencia se ve paliada por la información que conllevan los contornos de restantes vistas.

La extracción del contorno de cada vista se lleva a cabo mediante el programa CONTORNO. Conocidas las coordenadas de cada vértice proyectado sobre el plano imagen, el trazado del contorno se efectúa en sentido antihorario a partir de uno de los vértices del polígono a determinar. Como vértice de partida se elige aquel que presente mayor valor en la coordenada vertical del plano imagen.

La estrategia seguida en la extracción del contorno externo para cada vista 2D puede entenderse fácilmente, y de manera gráfica, con la ayuda del siguiente símil (Fig. 4.7): "Los segmentos rectilíneos conectando vértices en la proyección bidimensional ortográfica, pueden representar carreteras de una red vial. El contorno más externo correspondería al recorrido de un móvil que, partiendo del vértice superior (el de mayor coordenada vertical) y moviéndose en sentido antihorario, cada vez que alcanzara un cruce de carreteras eligiera siempre la primera carretera situada a su derecha hasta retornar al punto inicial de partida".

superior (el de mayor coordenada vertical) y moviéndose en sentido antihorario, cada vez que alcanzara un cruce de carreteras eligiera siempre la primera carretera situada a su derecha hasta retornar al punto inicial de partida".

Al programa CONTORNO se le suministran como datos de entrada los contenidos en el fichero de salida del programa PROYECCION. A la salida, el programa CONTORNO genera un fichero conteniendo, igualmente, la información correspondiente a la orientación de la vista especificada por un punto y tres vectores mutuamente perpendiculares que forman un sistema coordenado a derechas con el punto dado como origen. Las direcciones de estos vectores indican los ejes horizontal y vertical de la cámara así como la perpendicular al plano imagen o línea de visión. Además de estos datos, el fichero de salida contiene la lista de vértices, y coordenadas de los mismos, que conforman el polígono-contorno de la vista en cuestión.

De forma simplificada, el procedimiento CONTORNO puede resumirse en los siguientes pasos.

IV.3.2.1 ALGORITMO DE EXTRACCION DE CONTORNOS

- Leer la información correspondiente a los vértices, y conexión entre ellos, de la vista 2D en estudio, así como la posición y orientación espacial de ésta.
- 2. Determinación del vértice inicial a partir del cual se inicia el trazado del contorno. Como vertice inicial VI se elige aquel que presente el mayor valor en la coordenada vertical. Las coordenadas horizontal y vertical de VI se almacenan en (ZPARTIDA, XPARTIDA) y se escriben en el fichero de salida.
- 3. Determinar todos los segmentos rectilíneos definidos por VI y el resto de vértices que con él conectan. De estos segmentos, elegir aquel que forme el menor ángulo, medido en sentido antihorario, con la dirección positiva del eje horizontal de la cámara. Asignar a VF el número de vértice final asociado a dicho segmento.
- 4. Calcular la intersección del segmento rectilíneo VI—VF con el resto de segmentos rectilíneos, proyección en 2D de las aristas 3D del modelo poliédrico, siempre que dicha

intersección no tenga lugar en el vértice VI.

Cada punto intersección está asociado, además de al segmento $VI \longrightarrow VF$, a algún otro segmento caracterizado por sus vértices extremos I y K.

Para cada segmento(s) asociado(s) a cada punto intersección, determinar el vértice I ó K situado a la derecha del segmento VI—VF cuando éste se recorre desde VI hasta VF.

5. Analizar todas las intersecciones del segmento VI——VF obtenidas en el punto 4.

 \underline{SI} existe algún(os) punto(s) intersección cuyo segmento(s) asociados tenga(n) alguno de sus vértices extremos, I ó K, a la derecha del segmento $VI \longrightarrow VF$, $\underline{ENTONCES}$:

- Elegir de entre ellos el más próximo al vértice VI.
- Almacenar las coordenadas de dicho punto en (CIZ,CIX).
- Asignar a VD el número de vértice, I ó K, situado a la derecha del segmento VI—VF.

CASO CONTRARIO

De todos los segmentos con uno de sus extremos, I ó K, intersectando con el vértice VF, elegir aquel que forme el menor ángulo, medido en sentido antihorario, con la dirección positiva del eje horizontal. A continuación:

- Almacenar las coordenadas del vértice VF en (CIZ, CIX).
- Asignar a VD el número de vértice, I ó K, que no haya interceptado con el vértice VF.
- 6. El segmento rectilíneo definido por las coordenadas del vértice VI y las del punto intersección (CIZ,CIX) obtenido en el punto 5, será una arista del polígono-contorno.

<u>SI</u> las coordenadas (CIZ, CIX) son iguales a las del vértice inicial de partida (ZPARTIDA, XPARTIDA) <u>ENTONCES</u>:

- Finalizar el procedimiento CONTORNO (pasar al punto 9)

CASO CONTRARIO

- Almacenar las coordenadas (CIZ, CIX) en el fichero de salida.
- 7. Elección de un nuevo segmento rectilíneo VI→VF.
 - Como vértice inicial VI se toma el punto intersección (CIZ, CIX) obtenido en el punto 5.
 - Como vértice final VF se elige aquel cuyo número se había asignado a VD en el punto 5.

4

8. Repetir los pasos 4-7.

9. Finalizar.

En la implementación del anterior algoritmo ha sido necesario el desarrollo de las siguientes subrutinas:

INTER_SEGMENT_SEGMENT_2D

Evalúa la intersección, en 2D, de dos segmentos rectilíneos.

INTER_RECTA_SEGMENT_2D

Evalúa la intersección, en 2D, de una línea recta y un segmento rectilíneo.

INTER_RECTAS_2D

Evalúa la intersección, en 2D, de dos líneas rectas.

DERECHA

Determina la posición de un punto respecto de una recta, en 2D, cuando ésta se recorre en un sentido previamente especificado.

ACTUAL

Subrutina que actualiza, a medida que se van obteniendo todas las intersecciones con un segmento dado $VI \rightarrow VF$, la intersección adecuada de acuerdo con el criterio descrito en el punto 5 del procedimiento CONTORNO.

ANGULO

Subrutina que evalúa el menor de los ángulos que entre sí determinan dos direcciones específicas en 2D.

IGUAL

Subrutina que afirma o niega la igualdad de dos puntos en 2D dentro de una cota de error previamente especificado.

A modo de ejemplo, en la Fig. 4.8 se presenta el contorno de la vista 2D de la Fig. 4.7.

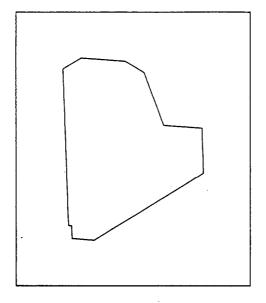


Fig. 4.8. Contorno para la vista de la Fig. 4.7.

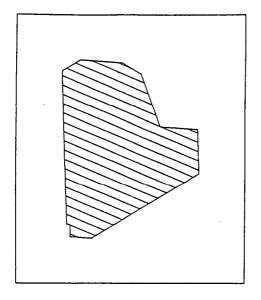
IV.3.3 RASTERIZACION DE LAS DOS PRIMERAS VISTAS

Una vez que se han determinado los contornos de todas las vistas, y como paso previo a la obtención de una primera reconstrucción 3D del objeto generador de las mismas, se necesita de una nueva representación para los contornos de las dos primeras, conocida como área rasterizada [NEUM-79], [MART-81b], [PAVL-82].

En esta nueva representación, en lugar de describir una imagen mediante la lista ordenada de vértices que definen su contorno, la silueta de cada vista se visualiza mediante un conjunto ordenado de segmentos rectilíneos equiespaciados y paralelos entre sí. La Fig. 4.9 muestra una misma silueta usando dos direcciones distintas para los segmentos de rasterización.

Cada vista rasterizada puede almacenarse mediante una simple estructura de datos compuesta de un conjunto ordenado de líneas de

rasterización que interceptan la silueta; cada línea se compone de un conjunto de segmentos rectilíneos, especificados por una lista ordenada de coordenadas de sus puntos extremos, correspondientes a la intersección de la línea con el contorno.



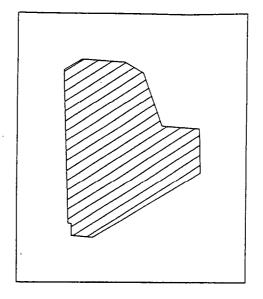


Fig. 4.9. Vista 4 del objeto n°. 12 con segmentos de rasterización según dos direcciones diferentes.

De acuerdo a lo expuesto en el capítulo II, son las dos primeras vistas las que especifican mutuamente las direcciones de rasterización en cada imagen. La intersección de ambos planos imagen determina una línea recta conocida como línea común. En cada vista, la dirección de rasterización se toma de manera que los segmentos de rasterización sean perpendiculares a la línea común. Así, y como paso previo a la rasterización, será preciso expresar la posición y orientación de la línea común en las bases 2D que definen los ejes horizontal y vertical de cada plano imagen (Apéndice B).

Esquemáticamente este proceso puede indicarse como sigue

PROCEDIMIENTO DE RASTERIZACION

- Leer la información relativa al contorno de las dos primeras vistas:
 - Posición y orientación de cada plano imagen.
 - Número y coordenadas de los vértices de cada contorno.
- 2. Determinación de la línea común mediante la intersección de los dos planos imagen.
- 3. Expresar un punto y el vector director de la línea común en términos de:
 - -base $\{\vec{u}_1, \vec{v}_1\}$ que define los ejes de la primera cámara. -base $\{\vec{u}_2, \vec{v}_2\}$ que define los ejes de la segunda cámara.
- 4. RASTERIZAR la primera vista.
- 5. RASTERIZAR la segunda vista.
- 6. Finalizar.

El núcleo fundamental de este procedimiento lo constituye la propia subrutina de rasterización, RASTERIZAR, llamada en los puntos 4 y 5. Dicha subrutina es la encargada de generar las dos máscaras de líneas de rasterización y efectuar la intersección de ellas con ambos contornos, para posteriormente decidir qué conjunto de los segmentos rectilíneos resultantes de la intersección son o no interiores a cada contorno y forman parte, por tanto, del área rasterizada.

IV.3.3.1 CONSTRUCCION DE LAS LINEAS DE RASTERIZACION

Las dos máscaras de líneas de rasterización, además de estar contenidas una en cada vista y ser perpendiculares a la línea común, han de confluir en los mismos puntos de la línea común, de manera que se pueda hablar de líneas de rasterización asociadas a puntos concretos de la línea común. Así, será posible ordenar estas líneas, por ejemplo, en función de las coordenadas de los puntos intersección con la línea común.

El primer punto sobre la línea común en el que confluyen dos líneas de rasterización se elige arbitrariamente, con la única condición de asegurar una intersección no nula de ambas líneas con sus respectivos contornos. Para la primera imagen, la línea de rasterización se elige de forma que pase por el punto intermedio de la primera arista del contorno. Esta línea cortará a la línea común en un punto determinado. Las líneas restantes se construyen a partir de la primera, estando asociadas a los distintos puntos que, sobre la línea común y en los dos sentidos de ésta, se obtienen a intervalos fijos contados desde el punto inicial.

El parámetro EQUI, que especifica la separación entre los distintos puntos de la línea común en los que confluyen líneas de rasterización, es de especial interés en todo el proceso reconstrucción posterior. El valor de EQUI especificará la separación entre líneas de rasterización y por tanto entre las distintas "roda jas" que conformen la reconstrucción 3D del objeto. Evidentemente, a menor valor de EQUI mayor será el número de segmentos de rasterización empleados para una misma vista, con lo que se consigue una mejor descripción de su silueta al estar los distintos segmentos más próximos entre sí. En el límite, cuando EQUI se hace tender a cero, se conseguiría una descripción continua, y no discreta, de la silueta limitada por cada contorno. Sin embargo, una disminución en el valor de EQUI se traduciría en un aumento considerable en el tiempo de cálculo para el procedimiento de rasterización descrito en este apartado y, más importante aún, para las posteriores y sucesivas ejecuciones del procedimiento empleado en la etapa de refinamiento.

En nuestro trabajo, y con los objetos mostrados en la Fig. 4.3, el valor de EQUI se ha elegido en función de los tamaños de cada uno de ellos, de manera que al RASTERIZAR, cada imagen venga representada por algunas decenas de líneas. De este modo se consigue que el número de líneas no sea excesivamente elevado y que a la vez sea suficiente para poder retener y describir pequeños cambios y detalles presentes en cada contorno. Por esta razón, la recuperación de detalles en la superficie de cada objeto reconstruido vendrá limitada, en última instancia, por la separación -elegida por el usuario- entre líneas de

rasterización.

IV.3.3.2 OBTENCION DE LOS SEGMENTOS DE RASTERIZACION

Una vez que se conoce, para cada contorno, la posición y orientación de la máscara de líneas de rasterización, se procede a determinar la intersección de éstas con cada contorno. En la intersección de una de estas líneas con todas y cada una de las aristas que definen el contorno, la primera queda descompuesta en un conjunto finito de segmentos rectilíneos (Fig. 4.10).

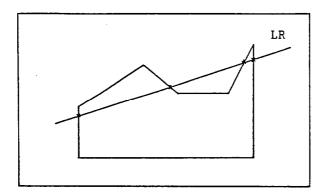


Fig. 4.10. Intersección de una línea de rasterización LR con el contorno de una vista.

IV.3.3.3 CRITERIOS DE INTERIORIDAD

La ordenación según valores crecientes de la coordenada horizontal, o según valores crecientes de la vertical si lo anterior no fuese posible, permite la especificación correcta de cada uno de los segmentos de rasterización. Sólo queda por decidir cuales, de entre los anteriores segmentos, son interiores al contorno.

En el caso más frecuente en el que la línea de rasterización, LR, sólo presente intersecciones en puntos de las aristas del contorno

distintos a los vértices, el criterio de selección de segmentos interiores es simple.

CRITERIO I

Empezando por el primero, van siendo etiquetados como interiores todos los segmentos cuyo número de orden sea impar, es decir el 1°, 3°, 5°,... Los segmentos pares son, evidentemente, exteriores al contorno.

Este criterio habrá de sustituirse por otro, cuando en la intersección de LR con el contorno se presente alguna -o ambas- de las situaciones siguientes (Fig. 4.11):

- a) Que alguno(s) de los puntos intersección coincida con algún(os) vértice(s) del contorno (Fig. 4.11.a)
- b) Que alguno(s) de los segmentos rectilíneos resultantes de la intersección -entre LR y el contorno- sea precisamente alguna(s) de las aristas del contorno.

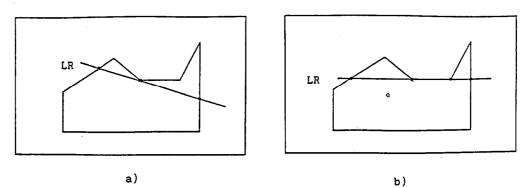


Fig. 4.11. Esquema ilustrativo de las situaciones conflictivas a) y b) para el criterio I de interioridad.

En los casos a) y b) se observa que la aplicación del CRITERIO I no siempre conduce a resultados satisfactorios. Es por ello que se necesita de otro análisis alternativo en estas situaciones peculiares y, por otro lado, poco frecuentes.

Si existe algún segmento que corresponda al caso b) entonces directamente será etiquetado como interior al contorno. Para las situaciones restantes, es decir del tipo a), a cada segmento se le somete a un test de interioridad. Teniendo en cuenta que si un punto cualquiera de alguno de los posibles segmentos de rasterización, y distinto de sus extremos, es interior al contorno entonces el segmento en cuestión también lo es, se puede establecer el siguiente criterio de interioridad.

CRITERIO II

Para cada posible segmento de rasterización actuar como sigue:

- Aleatoriamente se elige un punto de él que sea distinto de sus extremos.
- 2. De manera igualmente aleatoria se elige una dirección tal que no esté alineada con el eje vertical de la imagen ni con la dirección de la línea de rasterización LR.
- 3. Con la recta definida por el punto y dirección aleatorios obtenidos en 1 y 2 se determina la intersección entre esta línea y el contorno.
- 4. Repetir los pasos 1-3 hasta que todos los puntos de intersección obtenidos en 3 no se ajusten a las circunstancias de los apartados a) y b).
- 5. Contar el número de intersecciones obtenidas en 3 que estén situadas a izquierda y derecha del punto elegido en 1.
 - 5.a. \underline{SI} el número de intersecciones a la izquierda es impar $\underline{ENTONCES}$ el segmento en estudio es interior al contorno.
 - 5.b. \underline{SI} el número de intersecciones a la izquierda es cero y el número de intersecciones a la derecha es impar $\underline{ENTONCES}$ el segmento en estudio es interior al contorno.
 - En el resto de los casos el segmento en estudio será exterior al contorno.

6. Finalizar.

Resumiendo, podría decirse que la subrutina RASTERIZAR transforma la descripción de una vista 2D, en términos de su contorno, en otra

representación en la que el interior de la imagen se especifica por la intersección de un conjunto de líneas, paralelas y equiespaciadas, con el contorno de la imagen.

Los datos de entrada más destacables en esta subrutina están constituidos por:

- Información completa de posición, orientación y coordenadas del contorno de cada vista.
- Posición de un punto y dirección de la línea común.
- Distancia de separación entre líneas de rasterización (EQUI).

A la salida, la subrutina RASTERIZAR genera un fichero conteniendo las coordenadas de cada segmento de rasterización. Estas coordenadas están ordenadas por valores crecientes de la coordenada horizontal del plano imagen, o por valores crecientes de la coordenada vertical si lo anterior no fuera posible. Cada conjunto de segmentos asociados a una línea están agrupados en un único registro. Asociado a cada línea se indica el número de orden de ésta a partir de la línea LR = 0. Los índices asociados a cada línea pueden ser positivos, negativos o cero, correspondiendo los positivos y negativos a líneas de rasterización situadas a un lado u otro de la línea inicial (LR=0). Igualmente se indica el punto de la línea común por el que pasa cada línea de rasterización, así como el número de segmentos de que consta.

IV.3.3.4 ALGORITMO DE RASTERIZACION

A continuación se presenta un esquema del algoritmo empleado en la subrutina RASTERIZAR.

 Leer la información relativa al contorno a rasterizar, así como la posición y dirección de la línea común y la distancia de separación entre líneas (EQUI).

- 2. Elección de la posición para la línea inicial de rasterización (LR = 0). Esta se elige de manera que, siendo perpendicular a la línea común y estando contenida en el plano imagen de la cámara, pase por el punto intermedio de la primera arista del contorno.
- 3. Elegir, arbitrariamente, uno de los dos posibles sentidos de movimiento a lo largo de la línea común.
- 4. Determinar la interseccción de la línea LR con todos los lados del polígono-contorno. Ordenar los puntos de intersección, así obtenidos, por valores crecientes de la coordenada horizontal de la cámara o, a igualdad de éstos, por valores crecientes de la coordenada vertical. Cada punto intersección determinará con el siguiente un segmento rectilíneo.
- 5. \underline{SI} la línea de rasterización LR no intercepta con el contorno, $\underline{ENTONCES}$ invertir el sentido de movimiento adoptado en el punto 3.

<u>SI</u> no es la primera vez que se invierte dicho sentido <u>ENTONCES</u> finalizar la subrutina (pasar al punto 9)

- 6. Estudiar la interioridad, respecto del contorno en estudio, para cada uno de los segmentos rectilíneos obtenidos en el punto 4. Los segmentos rectilíneos, asociados a LR y ordenados por valores crecientes de las coordenadas horizontal o vertical, que sean interiores al contorno se almacenan en el fichero de salida como segmentos de rasterización. Los segmentos restantes son descartados.
- 7. Elegir una nueva línea de rasterización, paralela a la anterior y situada a una distancia EQUI de la misma según el sentido de movimiento vigente sobre la línea común.
- 8. Volver al punto 4.
- 9. Finalizar.

En la implementación de los algoritmos descritos en esta sección se ha hecho uso de las subrutinas adicionales, ya mencionadas en la anterior sección, INTER_RECTAS_2D, IGUAL, así como de la subrutina INTER_PLANOS que, como su nombre indica, determina la línea de intersección entre dos planos (caso de que exista).

IV.3.4 OBTENCION DE LA ESTRUCTURA DE PARALELOGRAMOS ASOCIADA A LA RECONSTRUCCION 3D INICIAL

Siguiendo el esquema de la Fig. 4.5, y según lo descrito en el capítulo II, el siguiente paso consiste en obtener una primera reconstrucción 3D del objeto a partir, únicamente, de la información contenida en las imágenes, ya rasterizadas, de las dos primeras vistas. Dentro del proceso de simulación, esta etapa la resuelve el programa PARALELOGRAMOS del que a continuación indicaremos las características más notables.

El objetivo de dicho programa es acotar espacialmente una región en cuyo interior esté situado el objeto en reconstrucción. Esta zona espacial se decribe mediante un conjunto de paralelogramos contenidos en una familia de planos paralelos adyacentes. El número de planos es igual al número de líneas de rasterización empleadas en la descripción de cada vista mediante el programa RASTERIZACION. Evidentemente, la separación entre planos dependerá del valor que el usuario elija para el parámetro EQUI presentado en la anterior sección.

La orientación de la familia de planos se elige de manera que sus normales estén alineadas con la línea común de las dos vistas iniciales, estando especificada la posición de cada uno de ellos por el correspondiente punto, sobre la línea común, en el que confluyan dos líneas de rasterización -una para cada imagen-.

La descripción de la reconstrucción inicial de cada objeto se realiza, así, en términos de un conjunto de paralelogramos que, posiblemente, contengan en su interior a secciones del objeto real obtenidas al efectuar la intersección del propio objeto tridimensional con la mencionada familia de planos.

El número de paralelogramos asociados a cada plano viene dado por N1xN2, siendo N1 y N2 el número de segmentos de rasterización en cada vista para la línea de rasterización asociada a este plano.

Conviene indicar que no necesariamente todos los N1xN2

paralelogramos han de contener en su interior a secciones reales del objeto. Como se justificó en el capítulo 2, lo que se puede afirmar es que, de los N1xN2 paralelogramos, al menos existen N (N = max (N1, N2)) que circunscriben a la sección real del objeto, pudiendo estar el resto vacios.

Los datos de entrada al programa PARALELOGRAMOS están constituidos por la información completa referida a cada imagen rasterizada. Esto es:

- Posición y orientación de cada vista.
- Conjunto de segmentos rectilíneos asociados a cada línea de rasterización, junto al índice que identifica a ésta última.
- Coordenadas del punto de la línea común en donde intercepta cada línea de rasterización.

A la salida, el programa PARALELOGRAMOS genera un fichero con la información de cada uno de los paralelogramos creados. Esta información se almacena por planos consecutivos, indicando, para cada plano, el número de paralelogramos que contiene y las coordenadas de los vértices que lo definen.

De manera resumida, el programa PARALELOGRAMOS se desarrolla de acuerdo a los siguientes puntos.

ALGORITMO DE GENERACION DE PARALELOGRAMOS

- Para cada punto de la línea común en el que intercepten líneas de rasterización para las vistas 1 y 2, actuar como sigue:
 - 1.1 Leer el número de segmentos de rasterización (N1) y coordenadas de los mismos para la vista 1.
 - 1.2 Leer el número de segmentos de rasterización (N2) y coordenadas de los mismos para la vista 2.
 - 1.3 Determinar las ecuaciones de las parejas de líneas rectas que, pasando por los extremos de cada uno de los N1 segmentos, estén alineados con la línea de visión de la cámara 1.

- 1.4 Determinar las ecuaciones de las parejas de líneas rectas que, pasando por los extremos de cada uno de los N2 segmentos, estén alineados con la línea de visión de la cámara 2.
- 2. Para cada uno de los N1 segmentos de rasterización de la vista 1, actuar como sigue:
 - 2.1 Para cada uno de los N2 segmentos de rasterización de la vista 2, actuar como sigue:
 - 2.1.a Evaluar la intersección de las dos parejas de líneas rectas, (Una asociada a cada segmento de una vista distinta). Los cuatro puntos resultantes de la intersección definen un paralelogramo contenido en un plano perpendicular a la línea común y que pasa por el punto de ésta que se está estudiando.
 - 2.1.b Almacenar las coordenadas del paralelogramo anterior en el fichero de salida y junto a las de los paralelogramos ya obtenidos y asociados al mismo plano indicado en 2.1.a.

3. Finalizar.

Un aspecto interesante en la implementación del anterior algoritmo, y que no aparece explícitamente recogido en él, es el siguiente. Los ficheros generados a la salida de RASTERIZACION, y que son los que PARALELOGRAMOS necesita como entrada, contienen la información de rasterización para cada vista. Esta información, como se comentó en la sección anterior, se especifica por líneas de rasterización y para cada línea se indica el número de segmentos de rasterización y coordenadas de los mismos. Sin embargo, y debido al propio desarrollo del algoritmo RASTERIZACION, puede -y de hecho ocurre con frecuencia- que en el fichero de salida, la información escrita secuencialmente no corresponda a líneas de rasterización contiguas.

Es por ello que en el programa PARALELOGRAMOS, y después de haber analizado cada línea de rasterización, sea necesario un reordenamiento de los distintos planos que componen la reconstrucción inicial. Dicho ordenamiento se efectúa, no por el orden en que han sido generados, sino según valores crecientes o decrecientes de alguna de las

cordenadas X, Y o Z del punto de la línea común asociado a cada plano. Conviene indicar, también, que las intersecciones de parejas de rectas del punto 2.1.a se llevan a cabo de manera que el orden en el que se especifican los distintos vértices del paralelogramo resultante sea tal que, al pasar de un vértice al siguiente, el sentido en el que se recorren todos los paralelogramos sea el mismo.

Otra cuestión a tener en cuenta en la implementación del programa PARALELOGRAMOS, como en los anteriores PROYECCION, CONTORNO y RASTERIZAR, es el hecho de tener que expresar un mismo punto espacial en función, tanto de las coordenadas (x,y,z) del sistema global de referencia como en términos de las coordenadas que lo localizan dentro de la imagen (coordenadas horizontal y vertical). Para una aclaración de estas relaciones, véase el apéndice B.

Las subrutinas llamadas desde dentro desde PARALELOGRAMOS son las ya mencionadas ANGULO e INTER_RECTAS_2D.

IV.3.5 REFINAMIENTO DE LA RECONSTRUCCION 3D

El último paso a efectuar en la reconstrucción de un objeto a partir de múltiples vistas consiste en mejorar la reconstruccioón inicial usando para ello, y de manera secuencial, los contornos de nuevas vistas. En la primera reconstrucción a partir de dos vistas, cada sección real del objeto se aproxima mediante uno o varios paralelogramos que lo acotan externamente. Como se muestra en la Fig. 4.12, la diferencia entre la silueta de una sección real del objeto y un paralelogramo que lo circunscriba, permite afirmar que la aproximación de ésta mediante paralelogramos, no siempre será aceptable y por tanto será susceptile de ser mejorada.

La mejora o refinamiento de la estructura de paralelogramos que constituye la reconstrucción 3D inicial se entiende en los dos sentidos siguientes:

A) En primer lugar en conseguir la aproximación de cada sección del

objeto mediante una línea poligonal cerrada que se ajuste externamente a dicha sección, de manera que llegue a ser descriptiva de detalles y cambios en su forma.

B) Además, como se indicó en la sección previa, cabe la posibilidad de que no todos los paralelogramos que conforman una "rodaja" de la reconstrucción inicial contengan en su interior a parte alguna de la sección real del objeto. Sería, por tanto, deseable la detección y posterior eliminación de tales paralelogramos (y polígonos en las reconstrucciones sucesivas) como parte integrante de la reconstrucción 3D del objeto.

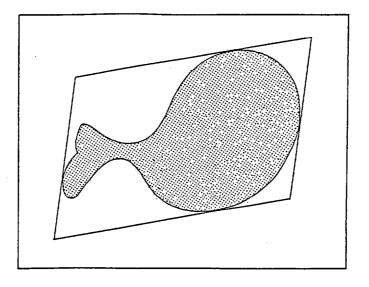


Fig. 4.12. Sección real de un objeto y un paralelogramo que la circunscribe.

Los dos objetivos anteriores se alcanzan contrastando los polígonos (paralelogramos en la reconstrucción inicial) que aproximan a las distintas secciones del objeto, con la información contenida en los contornos de nuevas imágenes. Este contraste y adaptación de información se realiza en los siguientes términos.

Puesto que el contorno de cada imagen de un objeto engloba a la silueta visible del mísmo desde la posición y orientación de la

cámara, la proyección de cualquier sección real del objeto sobre el plano de la cámara debería estar contenida dentro del contorno de dicha imagen. En el proceso de refinamiento, si en lugar de proyectar la sección real del objeto -que por otra parte no se conoce- se proyecta el polígono que externamente lo acota, se puede afirmar que las partes de este polígono proyectado que no intercepten con el contorno de la imagen podrán se descartadas como partes integrantes de la reconstrucción 3D. Sólo aquel polígono que hubiese dado lugar a la proyección del polígono resultante de la intersección entre el polígono inicial proyectado y el contorno de la imagen sería susceptible de formar parte de la reconstrucción 3D.

Evidentemente, si la intersección entre el contorno y uno de los polígonos proyectados es el vacío, esto indica que el polígono en cuestión, que hasta ese momento se consideraba parte integrante de la reconstrucción 3D, no contiene en su interior a parte alguna de la sección real del objeto y, por tanto, puede ser eliminado.

Dentro del proceso global de simulación empleado en este trabajo, los objetivos de mejora en cada reconstrucción 3D se logran mediante el programa de nombre REFINAMIENTO (Fig. 4.5), del que seguidamente pasamos a presentar un resumen y las características más destacables.

IV.3.5.1 PROCEDIMIENTO REFINAMIENTO

El programa REFINAMIENTO necesita a la entrada los datos referentes a la reconstrucción 3D que se desea refinar, junto a la información completa relativa al contorno de cada nueva imagen empleada para ello (Fig. 4.5). La primera vez que se ejecuta dicho programa, la reconstrucción 3D está compuesta de una serie de paralelogramos contenidos en planos paralelos consecutivos. En posteriores ejecuciones, cada uno de estos paralelogramos se irá sustituyendo por distintos polígonos que delimiten, con más precisión, las secciones reales del objeto. En algunos casos, tales polígonos desaparecen a medida que se ejecuta el programa REFINAMIENTO.

Después de cada ejecución se genera un fichero de formato análogo al de entrada, conteniendo igualmente la información de la reconstrución 3D, pero mejorada. Recordemos que cada reconstrucción está estructurada en rodajas contiguas y que para cada rodaja se especifican el número de polígonos que la componen y el número de vértices y coordenadas de los mismos.

El proceso de refinamiento se lleva a cabo rodaja a rodaja y dentro de cada rodaja polígono a polígono. La información, ya mejorada, para cada rodaja se almacena en el fichero de salida y se repite el proceso hasta completar todas las rodajas que componen la reconstrucción 3D. Un resumen del algoritmo empleado se indica a continuación.

ALGORITMO DE REFINAMIENTO

- 1. Leer la información relativa al contorno de la imagen empleada en el proceso de refinamiento:
 - Posición y orientación del plano imagen de la cámara.
 - Número de vértices que componen el contorno y coordenadas de los mismos.
- 2. Para cada una de las rodajas que componen la reconstrucción 3D del objeto a refinar, actuar como sigue:
 - 2.1 Leer la información relativa a la rodaja en estudio:
 - Indice asociado a este plano (rodaja).
 - Número de polígonos de que consta (NP).
 - Número de vértices de cada polígono y coordenadas de los mismos.
 - 2.2 Para cada uno de los NP polígonos, actuar como sigue:
 - 2.2.a Proyectar cada uno de sus vértices sobre el plano imagen y según la línea de visión de ésta.
 - 2.2.b Determinar la intersección del polígono proyectado con el contorno de la imagen empleada en el refinamiento.

- 2.2.c Proyectar, según la línea de visión del plano imagen y sobre el plano que contenía a la rodaja sometida a refinamiento, el polígono o polígonos resultantes de la intersección indicada en 2.2.b.
- 2.2.d Almacenar la nueva información asociada a la rodaja sometida a refinamiento en el fichero de salida.

3. Finalizar.

Conviene indicar que no todas las imágenes serán útiles en el proceso de refinamiento. La validez de cada nueva imagen estará en función de cuales hayan sido las dos primeras empleadas en la reconstrucción inicial y cual sea la orientación de cada una de las siguientes respecto de la línea común a las dos primeras. Como se ha apuntado ya, la orientación espacial de las distintas rodajas es tal que estarán contenidas en planos perpendiculares a la línea común a las dos primeras vistas. Para cada nueva vista que se emplee en REFINAMIENTO será preciso proyectar, ortogonalmente y sobre su plano imagen, los polígonos de las distintas rodajas. Para aquellas vistas cuya línea de visión sea perpendicular a la línea común de las dos primeras, la proyección de cada polígono de la reconstrucción degenerará en un segmento rectilíneo sobre el plano imagen. De ahí que la intersección entre el contorno y cada uno de estos segmentos rectilíneos no suministre una información directamente utilizable y que contribuya a refinar cada polígono.

Esta ha sido la razón por la cual en la implementación del programa REFINAMIENTO, antes de comenzar el proceso y una vez conocida la posición y orientación de cada nueva vista, se detecta la adecuación de ésta última. A efectos prácticos, siempre que la alineación entre la familia de planos (rodajas) y la línea de visión de cada vista no sea menor de 0.02 radianes (\simeq 1.15°), ésta última ha sido considerada como apta para el proceso.

Nuevamente, en este programa, se hace necesaria la conversión de coordenadas de un mismo punto de una base 2D a otra 3D y viceversa. Esto se requiere en el punto 2.2.a como paso previo a la intersección

del punto 2.2.b, ya que aquí se exige que los dos polígonos (el contorno y el polígono proyectado) estén expresados en la misma base 2D definida por los ejes horizontal y vertical del plano de la cámara.

En la implementación de este programa se requiere de las subrutinas INTER_PLANO_RECTA, ANGULO_3D e INTER_POLIGONOS. La subrutina ANGULO_3D evalúa el ángulo, entre 0 y π radianes, que determinan dos vectores en el espacio 3D.

La etapa fundamental en el proceso de refinamiento está constituida por la intersección entre dos polígonos, proceso éste que se efectúa repetidamente mediante llamadas a la subrutina INTER POLIGONOS en el punto 2.2.b.

IV.3.5.2 INTERSECCION ENTRE DOS POLIGONOS

El objetivo de la subrutina INTER_POLIGONOS es la determinación y especificación correctas de la intersección entre dos polígonos cualesquiera (cóncavos o convexos) cuando de éstos se conoce el número y orden de sus vértices, además de las coordenadas de los mismos. Dicha intersección se lleva a cabo en 2D y bajo la consideración de que ambos polígonos están contenidos en el mismo plano.

Este problema ha recibido una atención considerable, por la gran cantidad de aplicaciones que lo involucran, en áreas como *Gráficos por Ordenador* [NEUM-79], [FOLE-84], *Procesado de Imágenes* [PAVL-82] y *Geometría Computacional* [TOUS-85] entre otras.

En la literatura [SUTH-74], [CHAZ-80] existen algoritmos dirigidos a la resolución de problemas particulares, como por ejemplo la intersección entre polígonos convexos. En el caso de polígonos de uno cualquiera de los tipos, una estrategia bastante extendida consiste en descomponer, previamente, los polígonos cóncavos en un conjunto de polígonos convexos y seguidamente aplicar los algoritmos desarrollados para la intersección entre polígonos convexos.

Un algoritmo que contemple el caso general de la intersección para obtener, como resultado, nuevos polígonos, debe tener en cuenta una gran variedad de situaciones tales como las indicadas en la Fig. 4.13.

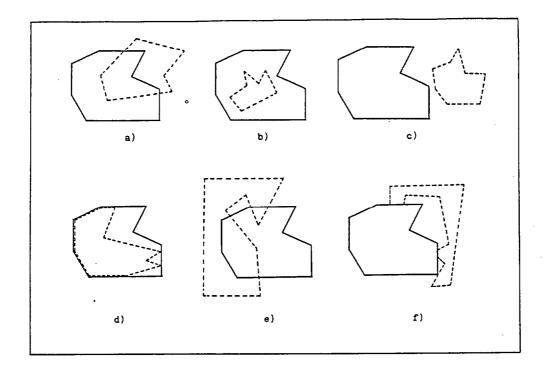


Fig. 4.13. Ejemplos de intersección de polígonos.

En el caso más general, el proceso de intersección llevará aparejada la obtención de uno o más polígonos (Fig. 4.13.a y Fig. 4.13.e) o incluso ninguno cuando se trate de intersección vacía (Fig. 4.13.c). En algunas situaciones puede que el resultado sea una colección de segmentos rectilíneos (Fig. 4.13.f). Así, algunas de las aristas de los polígonos en estudio deberán ser descartadas, otras nuevas deberán ser incluidas o incluso algunas serán divididas en varios segmentos. De ahí que sea preciso la elección de una estrategia organizada para poder tratar todas estas situaciones.

Cada polígono se especifica mediante una serie de vértices V1, V2,...VN, de manera que cada uno determine con el siguiente una

arista, a excepción del último VN que conecta con el primero V1.

IV.3.5.2.1 CARACTERISTICAS DEL ALGORITMO DESARROLLADO

El algoritmo empleado en la subrutina INTER_POLIGONOS "recorre" ordenadamente cada uno de los polígonos, arista a arista, y decide para cada una de ellas qué partes son o no interiores al otro polígono, es decir pertenecientes a la intersección entre ambos. Finalizados los recorridos por ambos polígonos, se dispone de una serie de NA segmentos pertenecientes a la intersección. A cada segmento se hace referencia mediante un número entero que indica el orden en que fué obtenido. Cada segmento se especifica por las coordenadas de sus extremos inicial y final.

Un aspecto interesante de esta subrutina es el relativo al sentido de movimiento a través de cada polígono. Se exige que ambos se recorran en el mismo sentido (horario u antihorario). Entre otras razones, este hecho viene motivado por la necesidad de especificar correctamente y distinguir entre los extremos inicial y final de cada segmento. La distinción entre ambos se realizará en función de cual sea el primero en detectarse según el sentido de movimiento elegido.

IV.3.5.2.1.1 CONEXIONES ENTRE SEGMENTOS DE LA INTERSECCION

Además del número identificativo de cada segmento y de las coordenadas de sus puntos extremos, será preciso establecer las posibles relaciones de conexión que entre sí guardan dichos segmentos. Así, a medida que se va obteniendo cada nuevo segmento de la intersección, se estudia la posible conexión de sus extremos inicial y final con los extremos, que aún estén libres, de los segmentos previamente obtenidos. Dichas conexiones se ajustan a los siguientes criterios:

1. El segundo extremo de un segmento habrá de conectar con el extremo inicial de otro segmento distinto y, viceversa, el primer extremo

de un segmento deberá conectar con el segundo de otro segmento.

 Nunca podrán conectar entre sí dos extremos iniciales o dos finales, ni tampoco un extremo de un segmento con otro de otro segmento que ya estuviese conectado.

Al finalizar los recorridos por uno y otro polígono, dichas conexiones se indican mediante un puntero NEXT asociado al extremo final de cada segmento. Para cada segmento I, $1 \le I \le NA$, el contenido del puntero NEXT(I) puede tomar los valores $0 \le J$, $1 \le J \le NA$, con $J \ne I$. Si NEXT(I) = J nos indica que el extremo final del segmento I conecta con el extremo inicial del segmento número J. Si NEXT(I) fuese igual a 0 (NEXT(I) = 0) se entiende que su extremo final está libre, es decir no conecta con ningún otro segmento. Es más, puede existir alguno o algunos segmentos I para los cuales su extremo inicial esté libre, es decir para los que no exista ningún otro segmento J que conecte su extremo final con el inicial de I, $\{\neg \exists J, 1 \le J \le NA / NEXT(J) = I\}$.

El estudio de las citadas conexiones constituye otra de las razones para exigir el mismo sentido de recorrido por los dos polígonos sometidos a intersección.

IV.3.5.2.1.2 ESPECIFICACION DE LOS POLIGONOS RESULTANTES DE LA INTERSECCION

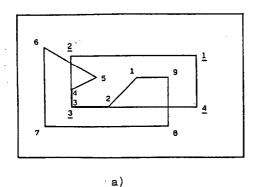
Con esta información queda aún por especificar cuántos y cuales son los polígonos que conforman la intersección de los dos en estudio, olvidándonos de los posibles segmentos que, perteneciendo a la intersección de ambos, no forman parte de ninguno de los polígonos resultantes de la misma. Previo a este último paso, se detecta si la intersección corresponde a alguno de los siguientes casos particulares:

a) La intersección es alguno de los dos polígonos en estudio. Este caso corresponde a la situación en la que los dos polígonos son

coincidentes, o bien uno es completamente interior al otro (Fig. 4.13.b, Fig. 4.13.d).

- b) La intersección es el vacio. Esta situación se detecta al no obtener ningún segmento de un polígono que sea interior al otro y viceversa (Fig. 4.13.c).
- c) La intersección está compuesta, únicamente, por un conjunto de segmentos rectilíneos, no necesariamente conectados (Fig. 4.13.f).

En el resto de situaciones la intersección estará compuesta por uno o más polígonos, posiblemente conectados entre sí por segmentos rectilíneos (Fig. 4.13, Fig. 4.14).



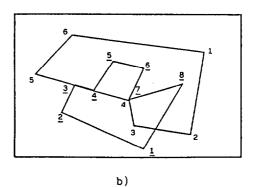


Fig. 4.14. Casos en que la intersección consta de uno o más polígonos conectados por segmentos rectilíneos (a), o bien de segmentos rectilíneos que no forman parte de ninguno de los polígonos intersección (b). Cada número representa el número de vértice de cada polígono (números subrayados para polígono 1 y números normales para polígono 2).

Teniendo en cuenta los fines expuestos en el proceso de refinamiento, es decir la sustitución de cada polígono de la reconstrucción 3D por otro que acote con más exactitud a cada sección real del objeto 3D, o en el peor de los casos la eliminación de algunos de estos polígonos, queda claro que el objetivo fundamental en la intersección es la detección de, únicamente, los polígonos que la

conforman y no de los segmentos rectilíneos que no pertenezcan a alguno de estos polígonos resultantes de la intersección de los dos en estudio (segmentos 2 y 3 de Fig. 4.15.a y segmento 1 en la Fig. 4.15.b).

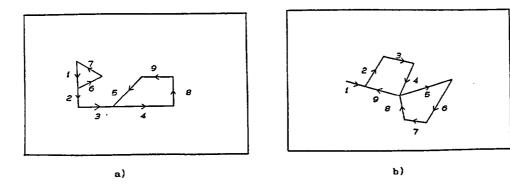


Fig. 4.15. Segmentos, con su número de orden, obtenidos al efectuar las intersecciones de los polígonos de las Fig. 4.14.a-b. La flecha sobre cada segmento indica cuales son sus extremos inicial y

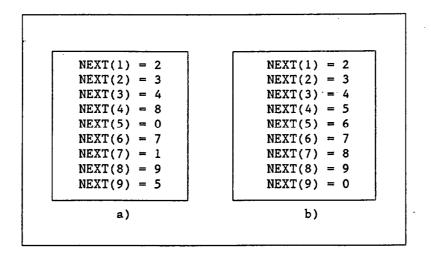


TABLA 4.II. Punteros NEXT asociados a los segmentos de la Fig. 4.15 obtenidos tras el recorrido por los polígonos de la Fig. 4.14 y según los criterios 1 y 2 del apartado IV.3.5.2.1.1. a) Los correspondientes a la Fig. 4.15.a. b) Idem para la Fig. 4.15.b.

La especificación correcta del número de polígonos de que consta

la intersección, así como del número de vértices y coordenadas de los mismos, se lleva a cabo examinando, y ampliando si procede, la información contenida en el puntero NEXT de cada segmento.

LLegados a este punto, el puntero NEXT habrá de modificarse para tener en cuenta todas las posibles conexiones del extremo final de un segmento con todos los extremos iniciales de todos los segmentos restantes. Esto viene motivado por el hecho de que el extremo final de un segmento puede conectar no sólo con el extremo inicial de un único segmento sino con los extremos iniciales de más de uno, o bien porque en el extremo inicial de un segmento pueden conectar su extremo final más de un segmento. Así, la Tabla 4.II obtenida al finalizar los recorridos por los dos polígonos en intersección de las Fig. 4.14.a-b, y ajustándose a los criterios 1 y 2 expuestos en este apartado, debería sustituirse por la Tabla 4.III.

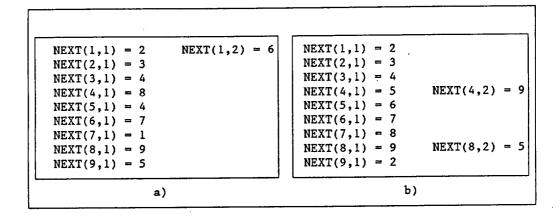


TABLA 4.III. Punteros NEXT de la TABLA 4.II pero ampliados al tener en cuenta todas las posibles conexiones -de extremos finales con iniciales- para los segmentos de las Fig. 4.15. a-b.

Todo polígono viene definido por una línea quebrada que determina un contorno cerrado. Así, la estrategia empleada para la especificación correcta de los polígonos que componen la intersección consiste en ir analizando, uno a uno, todos los segmentos de la intersección y determinando los caminos cerrados que partiendo de cada segmento permiten volver hasta él.

Aquellos segmentos de partida para los que no sea posible establecer un camino cerrado serán descartados y por tanto no formarán parte de la intersección entre los dos polígonos en estudio. El segmento hacia el que se efectúa el movimiento a partir de uno dado, viene especificado por las posibilidades contenidas en el puntero NEXT ampliado.

Los datos de entrada a la subrutina INTER_POLIGONOS los constituyen la información completa -número de vértices y coordenadas ordenadas de los mismos- de cada uno de los polígonos en estudio, así como el error con el que aceptamos la igualdad de dos números. A la salida, esta subrutina proporciona el número de polígonos que componen la intersección, el número de aristas para cada uno de ellos y las coordenadas de sus puntos extremos.

IV.3.5.2.2 ALGORITMO INTER_POLIGONOS

RECORRIDO POR EL PRIMER POLICONO

- 1. Para cada una de las aristas del polígono 1 actuar como sigue:
 - 1.1 Determinar la intersección de la arista del polígono 1 con todas y cada una de las aristas del polígono 2.

La arista del polígono 1 queda, así, descompuesta en una serie de segmentos rectilíneos. Ordenar estos segmentos por cercanía al vértice inicial de la arista en estudio del primer polígono. Si el número de intersecciones obtenidas fuese cero, la arista quedaría descompuesta en un único segmento rectilíneo: ella misma.

- 1.2 Determinar la interioridad, respecto del polígono 2, para cada uno de los segmentos rectilíneos obtenidos en 1.1.
- 1.3 Para cada uno de los segmentos del polígono 1 etiquetados como interiores al segundo polígono, estudiar la posible conexión de su extremo inicial con el extremo final del segmento previamente etiquetado

como interior. Si se trata del último segmento de la última arista del polígono 1 estudiar, además, la posible conexión de su extremo final con el inicial del primer segmento del polígono 1 etiquetado como interior al polígono 2.

<u>DECIDIR SI LA INTERSECCION ENTRE LOS DOS POLIGONOS LA CONSTITUYE EL PROPIO POLIGONO 1</u>

2. SI todos los segmentos rectilíneos en que quedó descompuesto el polígono 1 son interiores al polígono 2, es decir se han etiquetado como pertenecientes a la intersección entre ambos, ENTONCES el polígono 1 constituye, él mismo, el polígono intersección.

Finalizar el proceso (pasar al punto 7).

RECORRER LOS DOS POLIGONOS EN EL MISMO SENTIDO

3. Determinar si el orden de los vértices del polígono 2 es tal que cuando se pasa de uno al siguiente, el sentido de recorrido es igual al del polígono 1. Si no es así, entonces reordenar los vértices del segundo polígono de forma que ambos se recorran en el mismo sentido.

RECORRIDO POR EL SEGUNDO POLIGONO

- 4. Para cada una de las aristas del polígono 2, actuar como sigue:
 - 4.1 Determinar la intersección de la arista del polígono 2 con todas y cada una de las aristas del polígono 1.

La arista del polígono 2 queda, así, descompuesta en una serie de segmentos rectilíneos. Ordenar estos segmentos por cercanía al vértice inicial de la arista en estudio del polígono 2. Si el número de intersecciones obtenidas es cero, la arista quedaría descompuesta en un único segmento: la propia arista.

- 4.2 Determinar la interioridad, respecto del polígono 1, para cada uno de los segmentos rectilíneos obtenidos en 4.1.
- 4.3 Para cada uno de los segmentos etiquetados como interiores en 4.2, estudiar las posibles conexiones que sus extremos pueden presentar con otros segmentos también etiquetados como pertenecientes a la intersección y obtenidos hasta este momento. Las distintas posibilidades son:
 - El extremo primero, o inicial, de cada uno de estos segmentos <u>podría</u> conectar con alguno de los extremos finales de los segmentos del polígono 1 que pertenezcan al(los) polígono(s) intersección.

- El primer extremo de cada segmento del polígono 2 podría conectar con el extremo final del anterior segmento del polígono 2 que hubiese sido etiquetado como perteneciente a la intersección.
- El extremo final de cada segmento del polígono 2 podría conectar con alguno de los extremos iniciales de los segmentos del polígono 1 etiquetados como pertenecientes a la intersección entre ambos. Si se trata del último segmento de la última arista del polígono 2, entonces éste podría conectar, además, su segundo extremo con el extremo inicial del primero de los segmentos del segundo polígono etiquetado como parte de la intersección.
- 6. <u>DECIDIR SI LA INTERSECCION ENTRE POLIGONOS ES EL POLIGONO SEGUNDO O EL VACIO O BIEN DEGENERA EN UN CONJUNTO DE SEGMENTOS RECTILINEOS</u>
 - SI todos los segmentos rectilíneos en que quedó descompuesto el polígono 2 son interiores al polígono 1, es decir se han etiquetado como pertenecientes a la intersección entre ambos, <u>ENTONCES</u> el polígono 2 es completamente interior al polígono 1 y, por tanto, él mismo constituye el polígono intersección. Finalizar el procedimiento (pasar al punto 7).
 - SI al finalizar los puntos 1 y 4 se detecta que todas las aristas del polígono 1 son exteriores al polígono 2 y que todas las aristas del polígono 2 son exteriores al polígono 1, ENTONCES se puede afirmar que ambos polígonos no interceptan. Finalizar el procedimiento (pasar al punto 7).
 - <u>SI</u> al finalizar los puntos 1 y 4 se encuentra, para cada polígono, que el número de segmentos rectilíneos interiores al otro polígono es menor que el número total de segmentos rectilíneos en que quedaron descompuestas sus aristas, y además en el recorrido por el polígono 2 no se obtuvo ningún nuevo segmento perteneciente a la intersección que no hubiera sido ya obtenido al finalizar el recorrido por el polígono 1, <u>ENTONCES</u> se puede afirmar que la intersección degenera en un conjunto de segmentos rectilíneos no necesariamente conectados entre sí.

Finalizar el procedimiento (pasar al punto 7).

6. ESPECIFICACION DEL (DE LOS) POLIGONO(S) INTERSECCION

6.1 Para cada segmento de los NA obtenidos hasta este momento y pertenecientes a la intersección de los polígonos en estudio, examinar si, además de las conexiones de sus extremos especificadas hasta ahora, son posibles otras conexiones de sus extremos con los de cualquier otro segmento de los NA pero siempre

atendiendo al criterio:

- Segundo extremo de un segmento ha de conectar con el primero de otro segmento distinto.
- Nunca podrán conectar entre sí dos extremos iniciales o dos finales.

Actualizar el puntero NEXT teniendo en cuenta todas las posibles conexiones (antiguas y nuevas).

- 6.2 Descartar todos aquellos segmentos que no presenten conexión en alguno de sus extremos.
 - Actualizar el número -comenzando por 1- y conexiones de los segmentos restantes (NA).
- 6.3 De entre todos los posibles recorridos cerrados que se puedan construir de longitud máxima NA segmentos y partiendo del primer segmento según el puntero de conexiones NEXT, elegir el primero que se encuentre.
- 6.4 En caso de ejecución afirmativa para el punto 6.3, actuar como sigue:
 - 6.4.a Las aristas que definen dicho camino cerrado forman uno de los polígonos resultantes de la intersección. Almacenar el número de vértices y coordenadas de los mismos para este nuevo polígono.
 - 6.4.b <u>SI</u> el número de segmentos(aristas) restantes que aún no forman parte de alguno de los polígonos intersección es cero, <u>ENTONCES</u> finalizar el proceso de intersección (pasar al punto 7).
 - 6.4.c Actualizar el número -comenzando por 1- y conexiones de los restantes segmentos (NA) que aún no formen parte de alguno de los polígonos intersección, y volver a iniciar el punto 6.2.
- 6.5 SI la ejecución del punto 6.3 no alcanza el éxito en la obtención de algún polígono (camino cerrado), ENTONCES se puede afirmar que el segmento de partida elegido en 6.3 no forma parte de ninguno de los polígonos resultantes de la intersección de los dos en estudio. Actuar como sigue:
 - 6.5.a Eliminar este primer segmento de partida, así como sus conexiones, como parte integrante de la intersección entre los polígonos 1 y 2.
 - 6.5.b Actualizar el número -comenzando por 1- y conexiones de los segmentos restantes (NA).

6.5.c Volver a iniciar el punto 6.2.

7. Finalizar el procedimiento de intersección.

En la implementación de la subrutina INTER_POLIGONOS se efectúan sucesivas llamadas a otras subrutinas específicamente diseñadas como complementarias de la principal. Estas son:

IN_SEG_SEG_POL

Efectúa la intersección de dos segmentos rectilíneos dados. Esta subrutina se diseñó con unas características peculiares que la diferencian en su desarrollo de la ya mencionada INTER_SEGMENT_ZED.

INTERIOR

Determina si un segmento rectilíneo en 2D, especificado por las coordenadas de sus extremos, es o no interior a un polígono dado.

NEWARISTA

Determina si una arista en 2D, especificada por las coordenadas de sus extremos, es igual o no a alguna de las ya existentes.

CONEXION

Determina si un segmento conecta alguno de sus extremos con alguno de los extremos libres de alguno de los segmentos ya existentes.

PUNTERO

Actualiza el puntero de conexiones a medida que se obtiene cada nuevo segmento de la intersección.

SENTIDO_MOVIMIENTO

Determina la igualdad o no de los sentidos de movimiento, especificados por el orden y coordenadas de los vértices, para dos polígonos.

REORDENAMIENTO

Reordena y renumera todos los vértices de un polígono invirtiendo su sentido de movimiento. Esta operación sólo se realiza cuando el polígono en cuestión se esté recorriendo en sentido opuesto a otro de referencia.

SUAVIZAR

Examina cada polígono detectando la proximidad de cada dos vértices consecutivos. Aquellos vértices muy próximos son sustituidos por uno solo a mitad de camino entre ambos.

IGUAL

IV.4 ETAPA DE NORMALIZACION

Una vez se ha finalizado el proceso de reconstrucción 3D a partir de múltiples vistas, queda por efectuar la etapa de reconocimiento. Teniendo en cuenta que se está trabajando con objetos tridimensionales y que, en el caso más general, éstos pueden aparecer en una escena con tamaños, posiciones y orientaciones diferentes, será necesario, como paso previo al reconocimiento, normalizar las reconstrucciones 3D que de diferentes objetos se efectuen con el fin de poder compararlas entre sí, y por tanto distinguirlas y asignarlas a clases o categorias diferentes.

Esquemáticamente, los procesos de reconstrucción y normalización previos al reconocimiento se indican en el diagrama de la Fig. 4.16.

El proceso de normalización, a partir de la reconstrucción 3D en términos de "rodajas", se realiza en dos etapas sucesivas mediante los programas TRIANGULAR Y STANDARD.

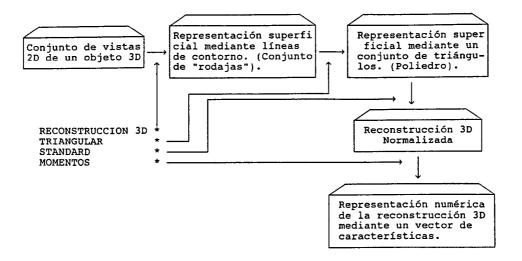


Fig. 4.16. Diagrama global para los procesos de reconstrucción 3D y posterior normalización.

IV.4.1 TRIANGULACION ENTRE RODAJAS CONSECUTIVAS

El proceso de reconstrucción 3D expuesto en las secciones previas conduce a una descripción de la superficie del objeto en términos de un conjunto de líneas de contorno paralelas y equiespaciadas, aproximadas mediante polígonos. La orientación de dicha familia de líneas es función, únicamente, de la orientación relativa de las dos primeras vistas empleadas en la reconstrucción. Quiere ello decir que si para un mismo objeto 3D se emplean otras dos vistas diferentes, el número de líneas de contorno y orientación espacial de las mismas sería distinto.

Como el proceso de normalización está basado en el empleo de características numéricas, si éstas se evaluasen sobre las líneas de contorno ("rodajas") se obtendrían tantas versiones normalizadas como grupos de dos vistas iniciales fuesen posibles. Para evitar este problema y obtener una única versión normalizada, será preciso que las características numéricas empleadas sean independientes de las vistas iniciales -y restantes- que conducen a la reconstrucción 3D.

Al observar la Fig. 4.17 (o igualmente la Fig. 2.15) se concluye que las diferentes familias de "rodajas", aún con orientaciones distintas y distinto número de líneas, definen la misma superficie del objeto y por tanto el mismo volumen. Es esta la razón por la cual las características numéricas usadas en este trabajo -momentos 3D- se han evaluado sobre la superficie de la reconstrucción y no sobre las líneas de contorno que la caracterizan.

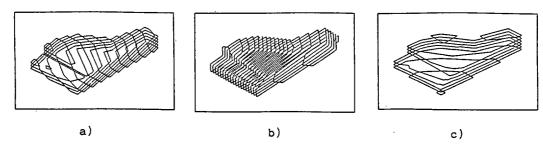


Fig. 4.17. Reconstrucciones del objeto n°. 12 con el mismo conjunto de vistas 2D. La diferencia entre a), b) y c) radica en el empleo de diferentes vistas iniciales en el proceso de reconstrucción.

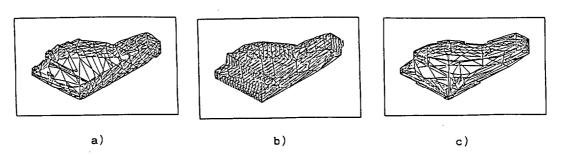


Fig. 4.18. Reconstrucciones 3D de la Fig. 4.17. después de triangular entre rodajas consecutivas.

Se necesita, por tanto, pasar de un esquema de representación superficial en términos de líneas de contorno a otro, también superficial, que no sea función del orden en que se combinen las vistas 2D en la etapa de reconstrucción 3D. En este segundo esquema la

superfice del objeto reconstruido se aproxima mediante un poliedro de caras planas triangulares según se indicó en el apartado II.4 de esta memoria.

En la práctica, la "no dependencia" anterior, con el orden de combinación de vistas 2D, se sustituye por una dependencia poco estricta o una dependencia pobre, debido a que las superfices obtenidas triangulando un conjunto de rodajas u otro pueden diferir ligeramente entre sí (Fig. 4.18).

IV.4.1.1 CARACTERISITCAS DEL PROCESO DE TRIANGULACION

El proceso de triangulación descrito en el capítulo II (apartado II.4) aproxima la superfice real del objeto que se apoya en los polígonos de dos rodajas consecutivas, mediante un conjunto de triángulos. Dichos triángulos se eligen de manera que dos de sus vértices pertenezcan a una rodaja y el tercero a la rodaja siguiente.

Este proceso se lleva a cabo mediante el programa TRIANGULAR cuyas características más notables destacamos a continuación. Los datos de entrada a este programa los constituyen el fichero obtenido a la salida de REFINAMIENTO, en el que se indican el número de polígonos, vértices y coordenadas de los mismos para cada una de las rodajas. A la salida, el programa TRIANGULAR genera un fichero que contiene, de manera ordenada, el número de triángulos y coordenadas de sus vértices para los triángulos obtenidos entre cada dos rodajas consecutivas.

Un aspecto interesante de este programa, y que lo distingue de otros presentes en la literatura [FUCH-77], [CHRI-78], [COOK-81a, 81b], es la posibilidad de triangular entre rodajas consecutivas en las que alguna -o ambas- presente un número de polígonos superior a la unidad. En estas situaciones, la mayor dificultad se presenta al decidir entre qué polígonos de una y otra rodaja se efectúa la triangulación. En el caso en que ambas rodajas consten de un único polígono cada una, la elección no presenta dudas.

La gran cantidad de esfuerzos que actualmente se dedican a este problema, [O'ROU-81], [TOUS-82], [BOIS-84], [DE FL-87], nos permite afirmar que aún no está completamente resuelto y por tanto no se dispone de una solución que funcione de manera fiable en toda la gran variedad de situaciones que se puedan presentar.

El algoritmo empleado en el programa TRIANGULAR distingue entre los dos grupos de situaciones siguientes:

- A) Las dos rodajas a triangular contienen un único polígono cada una.
- B) Las dos rodajas a triangular no están en la situación A), es decir o bien contienen distinto número de polígonos o bien, teniendo igual número de polígonos, éste es mayor que 1.

En el caso A) el polígono de una rodaja se triangula con el polígono de la otra. Cuando ocurra la situación B) se traslada cada polígono de cada rodaja hasta un plano intermedio entre ambas. A continuación se triangula cada polígono con su correspondiente trasladado.

Como se observa, el problema se ha reducido a triangular entre cada dos polígonos. La diferencia entre las situaciones A) y B) estriba en el hecho de que en A) se obtiene una superficie cerrada entre las dos rodajas, mientras que en B) esto no puede afirmarse. Además, puesto que uno de los polígonos -el trasladado- no constituye ninguna acotación real a la sección del objeto, sino que se considera como tal siguiendo un proceso de interpolación, el conjunto de triángulos así obtenidos aproxima a la superfice real del objeto pero en forma menos precisa que en las situaciones del tipo A). En la actualidad estamos trabajando en la búsqueda de estrategias válidas para conseguir que dicha superficie sea cerrada.

Cabe mencionar que las situaciones del tipo B) se presentan con cierta frecuencia, tanto mayor cuanto más compleja sea la forma del objeto en estudio y más cambios presente ésta última. Los resultados por nosotros obtenidos al trabajar con objetos conocidos y simulados,

nos han mostrado que la forma heurística de proceder en el caso B) conduce, en la mayoría de los ejemplos estudiados, a buenas aproximaciones de su superficie. Evidentemente, la superficie reconstruida se ajustará más finamente a la superficie real cuanto menor sea la separación entre rodajas, variable que controla el usuario mediante el parámetro EQUI.

Otro aspecto interesante a tener en cuenta es el hecho de recorrer en el mismo sentido todos los polígonos que componen las distintas rodajas. Esto viene impuesto por la forma en que se han generado los paralelogramos de la reconstrucción inicial y el haber exigido que los polígonos sometidos a intersección en REFINAMIENTO se recorran en igual sentido.

IV.4.1.2 PROCEDIMIENTO TRIANGULAR

El proceso de triangular se ejecuta, para cada dos rodajas consecutivas, comenzando por la primera y finalizando cuando se alcance la última.

Una vez finalizada esta tarea, se procede a descomponer en triángulos las superficies definidas por los polígonos que componen la primera y última rodajas. Para distinguir entre las situaciones A) y B) anteriormente comentadas se emplea la variable lógica ESPECIAL, pudiendo tomar los valores VERDADERO o FALSO. Cuando a ESPECIAL se le asigna el valor VERDADERO se entiende que se está en la situación B) y en caso contrario en la situación A).

El algoritmo desarrollado y empleado en el programa TRIANGULAR es el que a continuación se presenta:

- Leer la información relativa a las dos primeras rodajas de la reconstrucción 3D del objeto en estudio:
 - Indice de cada rodaja.
 - Número de polígonos de que consta.
 - Número de vértices de cada polígono y coordenadas de los mismos.

- Comparar el número de polígonos (NP2) de que consta la segunda rodaja con el número de polígonos (NP1) de la primera.
 - 2.1 <u>SI</u> ambas rodajas presentan el mismo número de polígonos y este número es igual a 1 (NP2 = NP1 = 1) <u>ENTONCES</u>:
 - 2.1.a Asignar el valor FALSO a la variable lógica ESPECIAL.
 - 2.1.b Ejecutar la subrutina TRIANGULACION para los dos polígonos de las dos rodajas.
 - 2.1.c Almacenar, en el fichero de salida, la información relativa al número de triángulos obtenidos y coordenadas de sus vértices.
 - 2.2 SI no se da la situación indicada en 2.1, es decir si alguna -o ambas- de las rodajas presenta un número de polígonos distinto de 1, ENTONCES:
 - 2.2.a Asignar el valor VERDADERO a la variable lógica ESPECIAL.
 - 2.2.b Elegir un punto intermedio (PC) situado entre los planos definidos por las dos rodajas en estudio. El punto seleccionado es el centroide de ambas rodajas.
 - 2.2.c Trasladar cada polígono de cada rodaja, según la dirección perpendicular a éstas, hasta el plano que pasa por el punto PC y es paralelo al de las rodajas.
 - 2.2.d Para cada polígono de cada rodaja y su correspondiente polígono trasladado, según se indica en 2.2.c, actuar como sigue:
 - 2.2.d.1 Ejecutar la subrutina TRIANGULACION.
 - 2.2.d.2 Almacenar, en el fichero de salida, la información relativa al número de triángulos obtenidos y coordenadas de sus vértices.
- 3. <u>SI</u> la segunda rodaja de las estudiadas es la última de las que componen la estructura 3D del objeto en estudio, <u>ENTONCES</u> pasar al punto 6.
- 4. Considerar ahora como rodaja primera la que hasta este momento había sido la segunda.
- 5. Leer la información correspondiente a una nueva rodaja considerando a ésta como la rodaja segunda y volver al punto 2.

6. Finalizar el procedimiento.

Las subrutinas empleadas en el programa TRIANGULAR son:

TRIANGULACION

Comentada en el siguiente apartado.

CENTROIDE

Determina la posición del centroide de un polígono atendiendo exclusivamente al "peso" de cada una de sus aristas.

CDM

Determina el centroide de un conjunto de puntos cuyos "pesos" son conocidos.

IV.4.1.3 TRIANGULACION ENTRE DOS POLIGONOS

El bloque fundamental del anterior procedimiento TRIANGULAR lo constituye la etapa de triangulación entre dos polígonos, uno de cada rodaja. Esta tarea se ejecuta repetidamente mediante sucesivas llamadas a la subrutina TRIANGULACION.

Los datos de entrada a esta subrutina están constituidos por el número y coordenadas de los vértices de cada polígono, así como por la variable ESPECIAL ya comentada. A la salida, esta subrutina proporciona el número de triángulos generados y las coordenadas de sus vértices.

Teniendo en cuenta lo ya dicho en el capítulo II (apartado II.4) referente al elevado número de triangulaciones posibles entre dos polígonos diferentes, parece clara la necesidad de emplear algún método que reduzca drásticamente este número de posibilidades conduciendo a un conjunto de triángulos que constituyan una aproximación aceptable de la superficie.

la estrategia aquí empleada consta, básicamente, de dos pasos:

- En primer lugar se elige, siguiendo algún criterio, un segmento inicial de conexión entre ambos polígonos.
- 2. A continuación se van generando, a partir de este segmento inicial, un conjunto de triángulos hasta alcanzar nuevamente el segmento de partida. La generación de cada triángulo obedecerá, lógicamente, a un conjunto de reglas.

IV.4.1.3.1 ELECCION DEL SEGMENTO INICIAL DE CONEXION ENTRE POLIGONOS

El segmento de partida se determina siguiendo las ideas expuestas por [CHRI-78] y [COOK-81a,81b].

Christiansen y col. transforman los polígonos a triangular circunscribiéndolos a ambos en un cuadrado de lado unidad; seguidamente se eligen aquellos dos vértices -uno de cada polígonoque tras la transformación estén separados por la menor de las distancias. En los trabajos de Cook y col., los vértices que definen el segmento de partida son aquellos para los que los vectores dirigidos desde sus respectivos centroides hasta los vértices en cuestión, presentan la menor diferencia en orientación.

En nuestro trabajo se ha desarrollado un nuevo método que combina las dos ideas expresadas anteriormente. Por un lado se calculan los centroides de los dos polígonos y se traslada uno de ellos sobre el otro de manera que ambos centroides coincidan. En esta situación, y para todas las posibles parejas de vértices de uno y otro polígono, se calculan las distancias entre vértices y los ángulos formados por los vectores dirigidos desde el centroide común hasta cada vértice. A continuación se establecen dos listas ordenadas y distintas de vértices; una según valores crecientes de distancias y otra según valores crecientes de ángulos.

Examinando ambas listas simultáneamente, empezando por el primer lugar y aumentando en una unidad a cada paso, se selecciona aquella pareja de vértices que por primera vez sea detectada en ambas listas.

Estos vértices -uno de cada polígono- son los que determinan el segmento inicial para comenzar la triangulación.

Al igual que los criterios de selección de Christiansen y col. y Cook y col., el criterio expuesto en este apartado no es óptimo en ninguno de los sentidos propuestos en [KEPP-75] y [FUCH-77]. Se trata, sin embargo, de un criterio heurístico que conduce a un algoritmo más rápido que los de Keppel y col. y Fuchs y col., y que en nuestro trabajo ha proporcionado unos excelentes resultados.

IV.4.1.3.2 GENERACION DE TRIANGULOS

La generación de los (m + n) triángulos, a partir del segmento inicial, se lleva a cabo de acuerdo a la regla de la menor distancia entre vértices consecutivos de distintos polígonos, expuesta en el apartado II.2.4 de esta memoria. Aquí, m y n representan el número de vértices de cada polígono.

IV.4.1.3.3 ALGORITMO DE TRIANGULACION

En este apartado se presenta, resumidamente, el algoritmo empleado en la subrutina TRIANGULACION.

- SI la variable ESPECIAL toma el valor VERDADERO (caso de triangulación entre un polígono y él mismo trasladado hasta el plano intermedio a las rodajas en estudio), <u>ENTONCES</u> pasar al punto 6.
- 2. Elección del segmento inicial para comenzar la triangulación. Este segmento habrá de enlazar entre sí dos vértices, uno de cada polígono en estudio (polígonos 1 y 2).
 - 2.1 Determinar el centroide para cada uno de los polígonos 1 y 2.
 - 2.2 Determinar el vector traslación que va del centroide del segundo polígono hasta el centroide del primero.
 - 2.3 Trasladar el segundo polígono, según el vector obtenido en 2.2, de manera que coincidan los centroides de los

polígonos 1 y 2 trasladado.

- 2.4 Para cada vértice del polígono primero, actuar como sigue:
 - 2.4.1 Para cada vértice del polígono segundo trasladado sobre el primero, actuar como sigue:
 - 2.4.1.a Determinar la distancia que separa a ambos vértices.
 - 2.4.1.b Determinar el ángulo que entre sí forman los vectores dirigidos desde el centroide común hasta cada uno de los dos vértices.
- 2.5 Ordenar las parejas de vértices de los polígonos 1 y 2 por valores crecientes de las distancias obtenidas en 2.4.1.a, y según valores crecientes de los ángulos obtenidos en 2.4.1.b.
- 2.6 Como segmento de partida se elige aquel definido por la pareja de vértices (ISPAN, JSPAN) -uno de cada polígono 1 y 2- que mejor lugar ocupe simultaneamente en la ordenación según distancias y en la ordenación según ángulos del punto 2.5.
- 3. Renumerar los polígonos sometidos a triangulación tomando ahora como vértices iniciales, etiquetados con el número 1, los que antes ocupaban las posiciones ISPAN y JSPAN en los polígonos primero y segundo respectivamente.
- 4. Generar (m+n) triángulos entre los polígonos 1 y 2 aplicando el criterio de la menor distancia de Christiansen y Sederberg, tomando como segmento inicial el definido por los vértices primeros de cada polígono ya renumerado.
- 5. Pasar al punto 7.
- 6. A este punto se accede en el caso especial de triangular un polígono (polígono 1) con él mismo después de aplicarle una traslación (polígono 2). Cada arista del polígono 1 y la correspondiente de su trasladado definen un rectángulo que a su vez queda dividido en dos triángulos por una de sus diagonales.
 - Obtener los citados triángulos, asociados a todas las aristas de los polígonos 1 y 2.
- 7. Finalizar la subrutina.

En el desarrollo de la subrutina TRIANGULACION se efectúan, a su vez, llamadas a las subrutinas complementarias siguientes:

RENUMERAR

Renumera los vértices de un polígono manteniendo el mismo sentido de movimiento a través de sus vértices.

CENTROIDE

ANGULO_3D

1V.4.2 OBTENCION DE LA VERSION NORMALIZADA PARA LA RECONSTRUCCION 3D

Una vez que la reconstrucción 3D se describe en una forma apropiada para poder evaluar momentos tridimensionales sobre su superficie, estamos en condiciones de transformar dicha reconstrucción en su forma canónica o normalizada. Esta etapa se efectúa mediante la aplicación del programa STANDARD del que seguidamente nos disponemos a indicar sus peculiaridades más notables.

El objetivo de STANDARD consiste en la determinación de los parámetros de las transformaciones de traslación, rotación y escalado descritas en la sección III.3.3 y apartados que la desarrollan.

La normalización frente a traslaciones y rotaciones será necesaria en todos los casos; sin embargo la estandarización de tamaños sólo habrá que efectuarla en aquellas situaciones en las que se quieran reconocer como idénticos aquellos objetos del mismo tipo pero con diferentes dimensiones.

Las versiones normalizadas para todos los objetos 3D en estudio se caracterizarán por estar situadas siempre en la misma posición, con igual orientación y, en su caso, con igual tamaño. Los valores característicos de posición, orientación y tamaño los especifica cada ususario de acuerdo a sus exigencias.

IV.4.2.1 CARACTERISTICAS DEL PROCESO DE NORMALIZACION

La normalización frente a traslaciones se consigue situando el centroide de la reconstrucción 3D en el origen de coordenadas del referencial absoluto OXYZ. En la práctica esta sería la última etapa a realizar, una vez se haya determinado la versión estándar respecto a orientación y tamaño. Si la posición estándar se eligiera en una localización distinta a la (0, 0, 0), la traslación hasta la posición canónica se efectuaría en dos etapas. En primer lugar se determina el centroide del objeto reconstruido, y éste se traslada hasta centrarlo en el origen de coordenadas. En esta situación, la reconstrucción 3D se orienta según el estándar establecido y se le asigna el tamaño adecuado (si procede normalizar los tamaños). A continuación sólo resta trasladar el centroide desde el origen de coordenadas hasta la posición deseada por el usuario.

De acuerdo a lo expuesto en el apartado III.3.3.2, la orientación canónica para cada reconstrucción 3D se alcanza aplicando una rotación adecuada después de haberla centrado en el origen de coordendas (Véase el apéndice A). La forma de operar consta de dos pasos consecutivos:

- 1. Determinar un único referencial ortonormal y a derechas centrado en cada reconstrucción 3D y solidario a ella.
- Orientar el mencionado referencial, y por tanto la reconstrucción
 según el estándar previamente establecido.

La orientación estándar o canónica la hemos elegido generalizando a 3D el método clásico de los ejes principales para 2D. Básicamente, se trata de orientar una forma tridimensional de igual manera que su elipsoide de inercia asociado; la razón para ello estriba en el hecho de que la transformación experimentada por un objeto al rotarlo es idéntica a la que experimenta su elipsoide de inercia. En términos del mencionado elipsoide, el estándar de orientación se define como sigue:

- Eje mayor del elipsoide de inercia alineado con el vector unitario \vec{k}

- Eje menor del elipsoide de inercia alineado con el vector unitario
- Eje intermedio del elipsoide alineado con el vector unitario 1.

No obstante, existen 8 orientaciones distintas de un objeto 3D centrado origen de coordenadas que satisfacen en e1 especificaciones anteriores. La eliminación de esta ambigüedad [GALV-89] se lleva a cabo imponiendo una serie de restricciones sobre los vectores propios de la matriz de inercia en su relación con la superficie de la reconstrucción 3D. Las citadas restricciones, descritas en el apartado III.3.3.2.2, van dirigidas a determinar un único referencial $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ de entre los autovectores $\{\pm \vec{u}_1, \pm \vec{u}_2, \vec{v}_3\}$ $\pm \vec{u}$ 3}, ya ordenados por valores crecientes de sus correspondientes autovalores $\{\lambda_1 \le \lambda_2 \le \lambda_3\}$ de la matriz de inercia. La elección de v_1 y v2 de entre ±u1 y ±u2 respectivamente, se efectúa en función de la posición que ocupe el punto del objeto más distante del centroide y según las <u>direcciones</u> de los vectores ún y úz respectivamente.

En la determinación de estos puntos sobre la superficie del objeto 3D reconstruido, será preciso efectuar la intersección de cada uno de los triángulos que componen la reconstrucción con las rectas asociadas a los vectores un y uz (direcciones definidas por los ejes mayor e intermedio del elipsoide de inercia). Si en el proceso de TRIANGULAR entre las distintas rodajas de la reconstrucción 3D se presenta el caso B) del apartado IV.4.1.1, es decir la triangulación de un polígono con él mismo pero trasladado, el conjunto de triángulos que componen la superficie del objeto reconstruido no definirán, generalmente, una superficie cerrada. En estos casos no siempre queda asegurada la intersección de las direcciones principales de inercia con la superficie reconstruida del objeto. Una posible solución, y sobre la que estamos trabajando actualmente, consistiría en el desarrollo de algoritmos más completos para el problema triangulación.

En nuestro trabajo hemos optado por detectar aquellas situaciones en las que no se produzca intersección alguna y volver a iniciar el proceso de reconstrucción, y posterior normalización, alterando el valor del parámetro EQUI hasta evitar la situación conflictiva anteriormente mencionada.

En lo concerniente a la normalización de los tamaños de los objetos reconstruidos, dos han sido los criterios empleados y ya descritos en el apartado III.3.3.3. Estos hacen referencia a caraterísticas geométricas para establecer el tamaño estándar de un objeto cuando sea requerida dicha normalización.

El primer criterio de normalización exige la igualdad de área para la superficie de todas las reconstrucciones en su versión estándar, mientras que el segundo criterio exige la igualdad de longitud de los ejes mayores para los elipsoides de inercia asociados a las reconstrucciones 3D ya normalizadas. Indiquemos que estos criterios son de aplicación independiente, es decir se elige uno u otro pero nunca simultáneamente. Como valores estándar para el área y tamaño del eje mayor en los dos criterios, se eligió el valor de 100 unidades en cada caso.

Conviene destacar que la obtención de los parámetros de normalización frente a traslaciones, rotaciones y cambios de escala, se ha efectuado exclusivamente en función de momentos 3D de hasta segundo orden con una fácil interpretación geométrica.

Tanto los datos de entrada, como los proporcionados a la salida del programa STANDARD, poseen el mismo formato: para cada dos rodajas consecutivas se especifica el número de triángulos que en ellas se apoyan así como las coordenadas de sus vértices. Adicionalmente, y a la salida, STANDARD genera un fichero conteniendo los parámetros asociados a las transformaciones de rotación, escalado isótropo y traslación que normalizan cada reconstrucción 3D.

IV.4.2.2 ALGORITMO STANDARD

1. Leer la información relativa al conjunto de triángulos que conforman la superficie de la reconstrucción 3D en estudio:

- Número de triángulos.
- Coordenadas de sus vértices.
- 2. Determinar la posición espacial, respecto del referencial OXYZ, para el centroide de la reconstrucción 3D.
- 3. Trasladar la reconstrucción 3D de forma que su centroide se situe en el origen de coordenadas.
- 4. Determinar la matriz de inercia asociada a la reconstrucción 3D ya centrada en el origen. Actuar como sigue:
 - 4.1 Diagonalizar esta matriz obteniendo sus autovalores y autovectores asociados.
 - 4.2 Ordenar, de menor a mayor, los valores propios anteriores $\{\lambda_1 \le \lambda_2 \le \lambda_3\}$ así como sus correspondientes vectores propios $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$.
- 5. De entre las 8 posibles opciones de autovectores $\{\pm \vec{u}_1, \pm \vec{u}_2, \pm \vec{u}_3\}$, alineados según las direcciones principales de inercia de la reconstrucción 3D, elegir un único referencial a derechas $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ atendiendo a los criterios expuestos en el apartado 3.3.3.2.2 de esta memoria.
- 6. Rotar la reconstrucción centrada en el origen de coordenadas de manera que alinee el referencial $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ solidario a ella con el referencial $\{\vec{k}, \vec{i}, \vec{j}\}$ del sistema global de referencia.
- 7. \underline{SI} se desea normalizar el tamaño de la reconstrucción 3D $\underline{ENTONCES}$:
 - 7.1 Determinar el parámetro de amplificación isótropa de acuerdo con el criterio elegido para normalización de tamaños.
 - 7.2 Aplicar la transformación de escalado isótropo a la reconstrucción 3D centrada en el origen y orientada según 6, empleando para ello el parámetro obtenido en 7.1.
- 8. Trasladar la reconstrucción 3D, que ya posee tamaño y orientación estandar de manera que situe su centroide en la posición estandar especificada por el usuario.
- 9. Finalizar.

En las Fig. 4.19.a - 4.19.i se ilustran, con un ejemplo, las 8 posibles disposiciones que de un objeto se pueden obtener imponiendo el criterio clásico de normalización basado en los ejes principales de inercia. En la Tabla 4.IV se muestran los autovalores y autovectores

obtenidos en la diagonalización de la matriz de inercia asociada al objeto de la Fig. 4.19.a después de centrarlo en el origen de coordenadas. La imposición de los criterios heurísticos adicionales para la obtención de un único referencial a derechas solidario al objeto, conduce a la orientación estándar mostrada en la Fig. 4.19.i. La matriz de rotación que aplicada al objeto de la Fig. 4.19.a, después de centrarlo en el origen, lo reorienta según se indica en la Fig. 4.19.i, se muestra en la Tabla 4.IV.

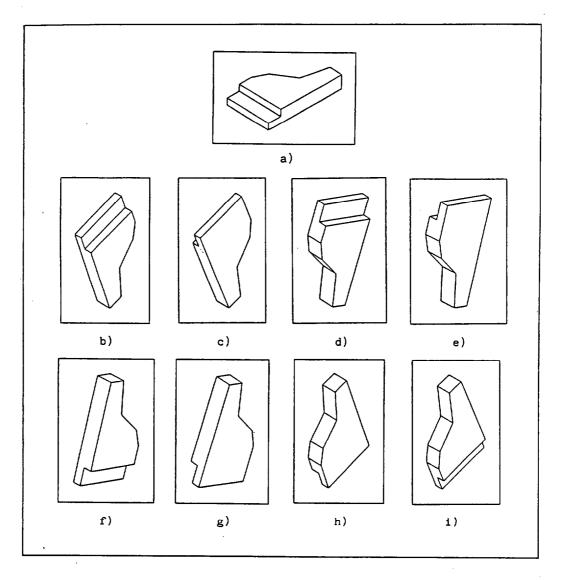


Fig. 4.19. 8 orientaciones posibles para el objeto n^* . 12 según el método clásico de los ejes principales de inercia.

AUTOVALORES	VECTORES PROPIOS	MATRIZ DE ROTACION
$\lambda_1 = 828.7$	$u_1 = \pm (.9392,3420,0316)$	[34149396 .0246]
$\lambda_2 = 2491.3$	$u_2 = \pm (.3413, .9396,0246)$.0382 .0123 .9992
$\lambda_3 = 3072.1$	$u_3 = \pm (.0382, .0123, .9992)$	[9392 .3420 .0316]

TABLA 4.IV. Autovalores, autovectores y matriz de rotación que transforma el objeto de la Fig. 4.19.a en la versión estándar de la Fig. 4.19.i

En la implementación del programa STANDARD, según el algoritmo anteriormente descrito, se llama directamente a las siguientes subrutinas:

PUNTO_POLIGONO

Se decide si un punto es o no interior a un polígono dado. Si el punto en cuestión forma parte de alguna de las aristas del polígono, también será considerado como interior. Se entiende que el punto y el polígono están contenidos en un mismo plano y que sus coordenadas se expresan en una base 2D.

ALINEAR_EJES

Obtiene los elementos de la matriz de rotación que alinea el sistema ortonormal a derechas $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ con los vectores unitarios $\{\vec{k}, \vec{i}, \vec{j}\}$ respectivamente. La matriz de rotación se expresa como combinación sucesiva de tres rotaciones, una en torno de cada eje OX, OY, OZ, del sistema global de referencia.

IGUAL

INTER_PLANO_RECTA

CENTROIDE

MOMENTOS_OR

Subrutina que determina el momento ordinario 3D mijk sobre la superficie de un triángulo. Esta subrutina se comenta con más detalle en la siguiente sección.

F02ABF

Subrutina que calcula los autovalores y autovectores de una matriz real y simétrica. Esta subrutina se ha tomado de la librería NAG [NAG-82].

IV.5 OBTENCION DE CARACTERISTICAS: MOMENTOS 3D NORMALIZADOS

Los momentos 3D mijk, definidos en (3.47), se han empleado en dos tareas distintas. Por un lado, como herramienta en el proceso de estandarización de cada forma tridimensional. Desde otro punto de vista los momentos 3D evaluados sobre la superficie son las características que usaremos en el proceso de asignación de un objeto desconocido a una de entre varias clases conocidas a priori. Cabe indicar que cualquiera de estos dos procesos, normalización y reconocimiento, podría haberse efectuado empleando otro tipo adecuado de características y otro tipo de estrategia de normalización.

La obtención de momentos 3D se lleva a cabo, de acuerdo a lo expuesto en el apartado III.3.1, mediante el programa MOMENTOS. Los momentos ordinarios mija se evalúan sobre cualquier forma tridimensional siempre que su superficie venga descrita por una colección de superficies planas triangulares. En este contexto, trabajar con superficies triangulares presenta la ventaja de un fácil tratamiento analítico, frente a la dificultad inherente en la superficie de objetos naturales.

La estrategia seguida en la obtención de momentos obedece a la aditividad de estas características, indicada en la ecuación (3.49). Al estar expresadas las cantidades mijk mediante integrales dobles evaluadas sobre la superficie, y pudiendo aproximar ésta última por una colección de triángulos, los momentos se obtienen como suma de momentos individuales del mismo orden y evaluados sobre cada uno de los triángulos.

Los datos de entrada al programa MOMENTOS están constituidos por el número de triángulos que aproximan la superficie y las coordenadas de sus vértices. Esta información está organizada por grupos de triángulos asociados a la triangulación entre cada dos rodajas, y corresponde a cualquiera de los ficheros proporcionados tras la ejecución de los programas TRIANGULAR o STANDARD. Además, MOMENTOS requiere a la entrada un número (1, 2 ó 3) representativo de alguna de las siguientes opciones:

- Opción 1 Evaluar un único momento mijk especificado por los subíndices i, j, k.
- Opción 2 Calcular todos los momentos, hasta un orden dado, comenzando por el momento de orden cero mono.
- Opción 3 Calcular solamente los momentos de un orden dado.

Recordemos que por orden de un momento mijk se entiende la suma de los enteros (i + j + k). A la salida, este programa suministra una lista ordenada de los momentos solicitados.

Recordemos, también, que en la evaluación de mijk mediante alguna de las expresiones (3.54) era necesario determinar, previamente, un dominio de integración sobre los planos XY, XZ o YZ, correspondiente a la proyección de cada triángulo (con la condición de que dicha proyección no degenerase en un segmento rectilíneo). La determinación del dominio adecuado se efectúa mediante una sencilla operación implicando los productos escalares del vector normal a cada triángulo y los vectores unitarios \vec{k} , \vec{j} e \vec{i} perpendiculares a los planos XY, XZ e YZ respectivamente.

IV.5.1 PROCEDIMIENTO MOMENTOS

A continuación se muestra un resumen de los distintos pasos seguidos en la ejecución del programa MOMENTOS.

1. Lectura de la información relativa a:

- Número de triángulos NT.
- Coordenadas de sus vértices.
- Opción para el cálculo de momentos.

- 2. Para cada uno de los NT triángulos que componen la superficie del objeto en estudio, actuar como sigue:
 - 2.1 Determinar el plano XY ó XZ ó YZ sobre el que puede proyectarse, perpendicularmente, dicho triángulo de manera que su proyección no degenere en un segmento rectilíneo.
 - 2.2 Evaluar el momento deseados sobre el triángulo en estudio mediante llamada a la subrutina MOMENTOS_OR.
- 3. Para cada grupo de subíndices i, j, k, representativos de los momentos a calcular, sumar los momentos evaluados sobre cada uno de los triángulos en el punto 2.2
- 4. Finalizar.

El bloque fundamental en el procedimiento anterior lo constituye la subrutina MOMENTOS_OR, la única que es llamada directamente desde MOMENTOS, y que a continuación describimos.

IV.5.2 SUBRUTINA MOMENTOS_OR

La subrutina MOMENTOS_OR, diseñada para la evaluación de mijk sobre un único triángulo, se ejecuta, para cada grupo de subíndices i,j,k, tantas veces como número de triángulos tenga la superficie tridimensional en estudio.

De acuerdo a lo expuesto en el apartado III.3.1.1, se dispone de una expresión de validez general para la obtención de mijk. La relación (3.55), que por comodidad aquí repetimos, junto a (3.56) y el cuadro de asignaciones que allí le sigue, viene expresada directamente en términos de:

- Coeficientes de la ecuación normal del plano que contiene al triángulo en estudio.
- Indices i, j,k que caracterizan al momento mijk que se desea evaluar.
- El coseno del menor de los ángulos que determinan las dos

direcciones siguientes: normal al triángulo y normal al plano de proyección.

$$m_{ijk} = \int_{u_1}^{u_2} du \int_{u_1}^{\phi_2(u)} F(v, u) dv + \int_{u_2}^{3} du \int_{u_1}^{2} F(v, u) dv$$
 (3.55)

Los datos de entrada para la subrutina MOMENTOS_OR son los siguientes:

- Coordenadas de los vértices del triángulo en estudio.
- Plano (XY ó XZ ó YZ) de proyección adecuado para dicho triángulo.
- Indices del momento a calcular.

Con esta información, y después de proyectar el triángulo sobre el plano correspondiente, la subrutina determina los límites para cada una de las integrales dobles que aparecen en (3.55). Dichos límites son los que especifican los dominios de integración que coinciden, justamente, con la superficie triangular ya proyectada.

La obtención de los parámetros u1, u2, u3, así como de las funciones lineales $\phi1$, $\phi2$, $\psi1$ y $\psi2$ se lleva a cabo como se indica en el apartado III.3.1.1.

Finalmente, las integrales dobles de (3.55) se evalúan mediante la llamada a la subrutina DO1DAF de la librería NAG [NAG-82]. De manera resumida, esta subrutina procede a evaluar una integral definida de la forma:

$$I = \int_{a}^{b} \int_{1}^{2} f(x, y) dx dy$$

donde a y b son constantes y $\Gamma_1(y)$ y $\Gamma_2(y)$ son funciones de la variable y. El procedimiento seguido consta de sucesivas aplicaciones del método descrito por Patterson en [PATT-68] hasta alcanzar un valor dentro de la precisión especificada por el usuario $(10^{-6}$ en nuestro caso).

A la salida, MOMENTOS_OR proporciona el valor de mijk siempre que el proceso en DO1DAF converja. Notemos la ausencia de problemas de convergencia en todo nuestro trabajo en relación con DO1DAF.

IV.6 CONSIDERACIONES PRACTICAS EN LAS ETAPAS DE RECONSTRUCCION, NORMALIZACION Y OBTENCION DE CARACTERISTICAS

En este apartado se presentan algunos datos y valores numéricos de parámetros de interés en los programas encargados de la reconstrucción y posterior normalización de objetos tridimensionales.

En nuestro trabajo se ha empleado un conjunto de 13 objetos poliédricos 3D (Fig. 4.3). Todos ellos se han situado en el primer octante del referencial absoluto $\{\vec{i}, \vec{j}, \vec{k}\}$ de manera que las coordenadas de sus vértices sean todas positivas. El número de vértices varía entre 7 y 74, correspondiendo estos valores a los objetos 3 y 8 respectivamente.

IV.6.1 PROGRAMA PROYECCION

Con el fin de disponer de reconstrucciones de un mismo objeto, pero situado en posiciones y con orientaciones diferentes, a cada uno se le somete a una serie de 11 transformaciones. Estas constan de una traslación y una rotación efectuadas de acuerdo a los tres pasos siguientes:

- 1. Trasladar el centroide hasta el origen de coordenadas.
- 2. Aplicar una rotación en torno de un eje que pase por el origen.
- Trasladar nuevamente el centroide hasta situarlo en la posición deseada.

Los parámetros característicos de las 11 transformaciones, y que hacen referencia a la posición final del centroide así como al eje y ángulo de giro, aparecen en la Tabla 4.V. La transformación inicial,

etiquetada con el número O, equivale a mantener inalterado a cada objeto de la Fig. 4.3.

NUMERO DE TRANSFOR.	POSICION FINAL DEL CENTROIDE	EJE DE GIRO	ANGULO DE GIRO
0	la propia de cada objeto		
1	(0, 0, 0)	(1, 0, 0)	90.0
2	(0, 0, 0)	(0, 1, 0)	90.0
3	(0, 0, 0)	(0, 0, 1)	90.0
4	(-2, 26, 23)	(0.493, 0.458, 0740)	40.3
5	(29, -19, -3)	(0.839, -0.352, -0.414)	114.2
6	(-1, 34, 35)	(-0.391, 0.920, -0.035)	223.1
7	(-10, -16, 21)	(-0.266, 0.891, -0.367)	322.0
8	(-10, 20, -10)	(-0.179, -0.578, 0.796)	60.2
9	(25, -15, -13)	(-0.203, -0.320, 0.925)	155.3
10	(-5, 24, 35)	(0.877, -0.146, -0.457)	180.5

TABLA 4.V. Transformaciones de traslación y rotación aplicadas a cada uno de los objetos de la Fig. 4.3.

Para cada versión de cada objeto se obtienen 10 vistas bidimensionales distintas, mediante el programa PROYECCION, correspondientes a 10 posiciones y orientaciones diferentes del plano de la cámara. Los valores para estos parámetros se indican en la Tabla 4.I, siendo los mismos para todos los objetos. Notemos que las orientaciones relativas entre las líneas de visión de cada dos cámaras oscilan entre 14° y 145° aproximadamente.

IV.6.2 PROGRAMA CONTORNO

En la ejecución del programa CONTORNO, así como en todos aquellos en los que sea necesario decidir la igualdad o no de dos puntos en 2 ó 3D, se precisa del parámetro ERROR. El valor asignado a ERROR ha sido

 10^{-6} , aceptando, de esta manera, que dos puntos (vértices de segmentos rectilíneos) son iguales cuando sus coordenadas respectivas difieran en una cantidad inferior a 10^{-6} .

Teniendo en cuenta que se trabaja con 13 objetos, que a cada uno se aplican 11 transformaciones distintas, y que de cada objeto transformado se toman 10 vistas diferentes, el número total de imágenes 2D, así como sus correspondientes contornos, es del orden de 10³.

IV.6.3 PROGRAMAS RASTERIZAR Y PARALELOGRAMOS

Como se comentó en el apartado IV.3.3, referente a la descripción del contorno de una imagen mediante líneas de rasterización paralelas, el parámetro EQUI, que define la separación entre líneas, juega un papel importante. Los valores de EQUI empleados para cada uno de los objetos en estudio son los mostrados en la Tabla 4.VI. La razón para utilizar distintos valores de EQUI reside en la diferencia que presentan las dimensiones características de unos y otros objetos.

Recordemos que las vistas empleadas en la ejecución de RASTERIZACION son las dos primeras obtenidas para cada objeto, estando los ejes de visión de ambas cámaras alineados con los ejes OX y OY respectivamente.

Ciertos valores para EQUI pueden conducir a errores en la ejecución del programa REFINAMIENTO. Cabe la posibilidad de que alguna de las líneas de rasterización extremas, en alguna de las dos vistas, quede reducida a un único punto como resultado de la intersección del contorno con la familia de líneas paralelas. La primera reconstrucción 3D, mediante un conjunto de paralelogramos, se obtiene al ejecutar el programa PARALELOGRAMOS sobre las vistas primera y segunda ya rasterizadas. De esta manera, es posible que alguno -o varios- de los paralelogramos degenere en un único punto o en un segmento rectilíneo. Estas circunstancias se detectan en PARALELOGRAMOS, informando al usuario acerca de la conveniencia de volver a iniciar el proceso de

RASTERIZACION con otro valor de EQUI ligeramente distinto al inicialmente utilizado.

OBJETO	PARAMETRO EQUI
1	0.3
2	0.3
3	0.63
4	0.63
5	0.47
6	0.47
7	0.47
8	0.23
9	0.37
10	0.63
11	0.37
12	0.37
13	0.37

TABLA 4.VI. Valores del parámetro EQUI en la ejecución del programa RASTERIZAR.

IV.6.4 PROGRAMA REFINAMIENTO

Como ya se ha indicado, en cada ejecución de REFINAMIENTO la estructura 3D de la reconstrucción se ve mejorada. Para cada una de las transformaciones aplicadas a los objetos de partida, el programa REFINAMIENTO se ejecuta 8 veces consecutivas empleando para ello las 8 vistas 2D distintas a las 2 primeras. No obstante, se ha podido constatar que en la mayoría de los casos sería suficiente con el empleo de un menor número de vistas. Evidentemente, no existe una regla de validez general para especificar el número óptimo de vistas

que proporcionen un buen compromiso entre el tiempo de cálculo y la bondad de la reconstrucción.

Parece claro que dicho número sea función directa de la complejidad del objeto a reconstruir. Además, las orientaciones relativas de las distintas vistas entre sí y de éstas con el objeto a reconstruir, constituyen otro de los aspectos a tener en cuenta. En efecto, este factor determinará el contenido de información redundante en las distintas vistas. A medida que las vistas presenten menos similitud la convergencia en el proceso de REFINAMIENTO será más rápida. Esto se ha tenido en cuenta al elegir las orientaciones de las 10 cámaras mostradas en la Tabla 4.I.

IV.6.5 PROGRAMA TRIANGULAR

Uno de los programas más costosos en lo referente al tiempo de cálculo es TRIANGULAR. Dos son los factores más decisivos que sobre ello influyen. De una parte el número de vértices de que consta cada polígono, y por otro lado el número de rodajas -dependiente de EQUIque conformen la reconstrucción.

Recordemos que el número de triángulos generados entre dos polígonos de m y n vértices es (m + n). Además la selección, para cada dos polígonos, del segmento inicial de conexión lleva aparejado el cálculo de distancias y orientaciones entre cada dos de los m y n vértices. Puesto que la triangulación se determina entre cada dos rodajas consecutivas, el proceso anterior se repetirá tantas veces como rodajas tenga la reconstrucción, siendo el número de éstas función inversa del parámetro EQUI.

A modo ilustrativo, indicar que el número de triángulos que componen la superficie reconstruida para los 13 objetos en su transformación n°. O, oscila entre 442 para el objeto n°. 3 y 1572 para el objeto n°. 10.

IV.6.6 PROGRAMA STANDARD

El reconocimiento de objetos con independencia de la posición, orientación y tamaño que posean, llevaría asociada la reconstrucción de diferentes versiones de un mismo objeto y posterior normalización de las mismas. La obtención de tales versiones se efectuaría aplicando las correspondientes operaciones de traslación, rotación y cambio de escala al objeto real de partida.

No obstante, en ocasiones, puede ser de interés considerar como distintos aquellos objetos que, aún presentando igual forma externa, difieran en sus dimensiones. En este caso sólo se tendría que normalizar frente a traslaciones y rotaciones. Por este motivo, inicialmente se obtuvieron 11 versiones distintas -en lo referente a posición y orientación- de cada objeto.

En aquellos casos en los que, además, se desee normalizar frente a cambios de escala no ha sido necesario volver a obtener otras versiones de diferentes tamaños. Las mismas versiones anteriores, diferenciadas sólo en posición y orientación, se han empleado para normalizar frente a cambios isótropos de tamaño. La razón para ello estriba en que si se amplifica un objeto para después reconstruirlo, cambiando en igual proporción el valor de EQUI, entonces la reconstrucción 3D de la versión amplificada es idéntica a la amplificación de la reconstrucción de la versión inicial (Fig. 4.20).

La posición estándar para los centroides de todas las reconstrucciones ya normalizadas se eligió en el punto de coordenadas (0, 0, 0). El criterio de normalización frente a cambios de escala fué el de la igualdad en el área de todas las versiones normalizadas. El valor estándar de área se tomó de 100 unidades. El estándar de orientación es el descrito en el apartado III.3.3.2.2.

A modo ilustrativo, en la Fig. 4.21 se presentan las reconstrucciones 3D, en términos de líneas de contorno, para los objetos de la Fig. 4.3 en su transformación n°. 0, con la posición, orientación y tamaños allí indicados. La comparación de ambas figuras

(4.3 y 4.21) nos permite afirmar que el método de reconstrucción empleado proporciona unos resultados excelentes.

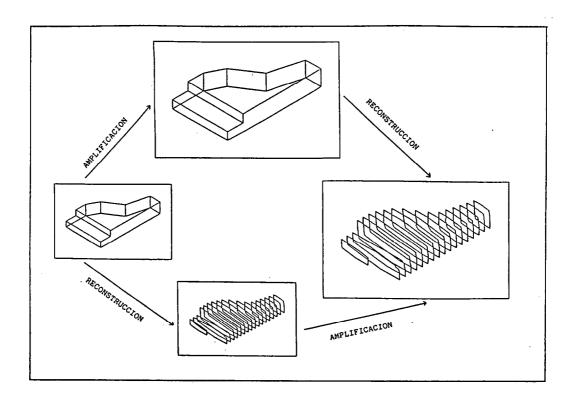


Fig. 4.20. Ejemplo ilustrando la conmutatividad entre los procesos de reconstrucción 3D y amplificación, siempre que el parámetro EQUI se amplifique en la misma proporción.

La superficie de las anteriores reconstrucciones, pero descrita mediante una colección de triángulos al aplicar el programa TRIANGULAR, se muestra en la Fig. 4.22.

En la Fig. 4.23 aparecen los mismos objetos de partida de la Fig. 4.3 pero en su versión normalizada de posición, orientación y tamaño. Después de aplicar el programa STANDARD a cada una de las reconstrucciones de la Fig. 4.22 se obtienen sus correspondientes versiones canónicas o normalizadas tal y como se indican en la Fig. 4.24.

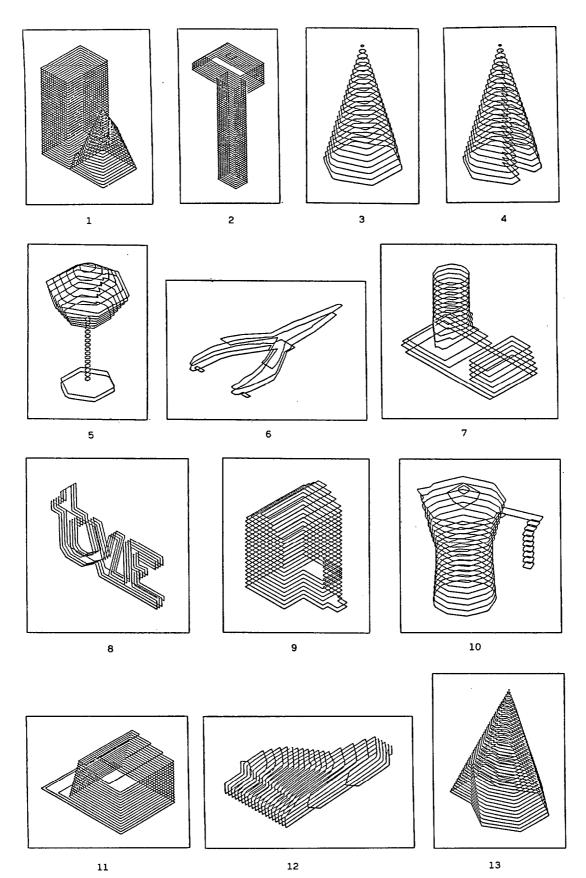


Fig. 4.21. Reconstrucciones 3D para los objetos de la Fig. 4.3.

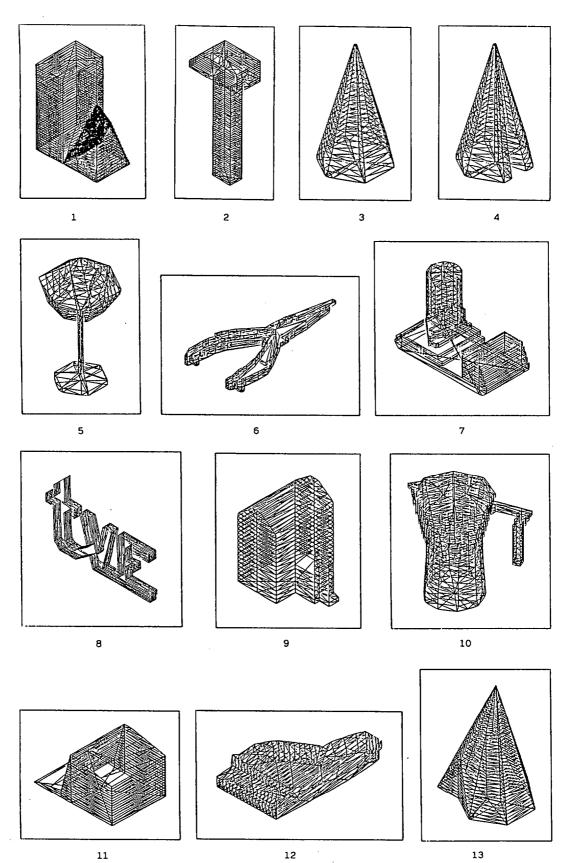


Fig. 4.22. Resultados obtenidos al aplicar el programa TRIANGULAR a los objetos de la Fig. 4.21.

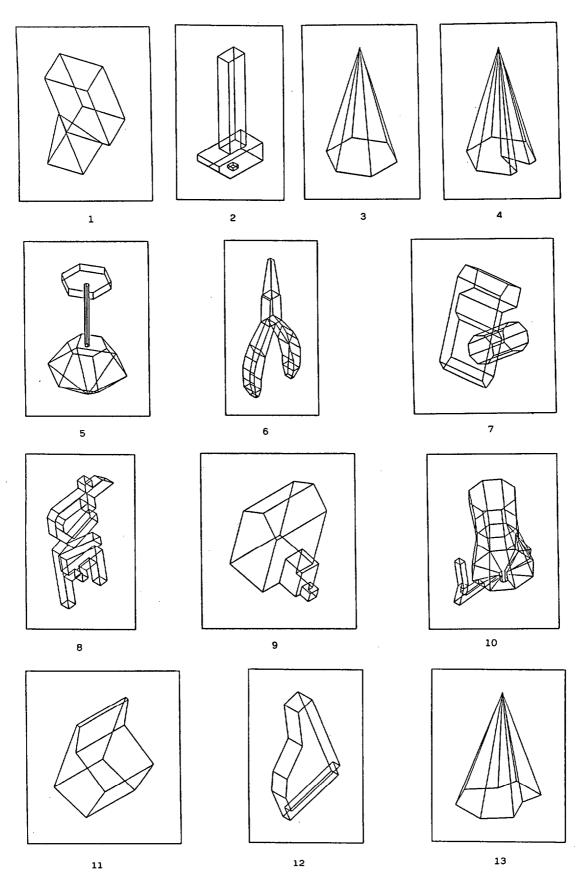


Fig. 4.23. Versión canónica o estándar para cada uno de los objetos de la Fig. 4.3.

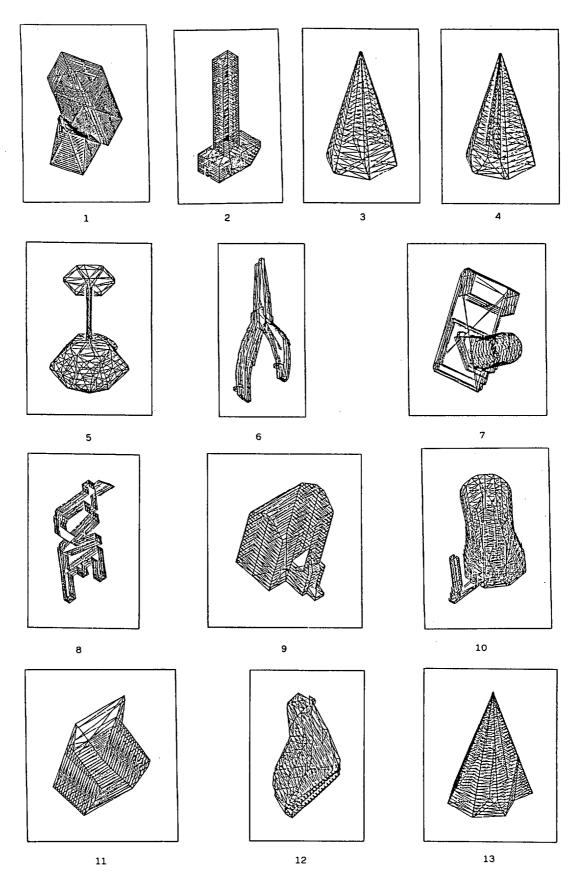


Fig. 4.24. Versión estándar para cada una de las reconstrucciones de la Fig. 4.22.

La eficacia del procedimiento de normalización empleado en el programa STANDARD se puede poner de manifiesto al contrastar las reconstrucciones 3D normalizadas, para las distintas versiones de un objeto, con el objeto real de partida en su versión normalizada. La Tabla 4.VII muestra el número de normalizaciones correctas e incorrectas obtenidas para cada objeto en estudio.

OBJETO	NOR. CORRECTAS	NOR. INCORRECTAS
1	8	3
2	8	3
3	11	0
4	. 6	5
5	11	0
6	10	1
7	11	0
8	5	6
9	5	6
10	9 .	2
11	7	4
12	9	2
13	11	0

TABLA 4.VII. Resultados obtenidos al determinar la forma canónica o normalizada para cada una de las 11 versiones que, de un mismo objeto, se disponen.

La causa más frecuente de normalización errónea es una orientación incorrecta en la versión normalizada, y ello ocasionado fundamentalmente por una triangulación desacertada entre rodajas consecutivas. Las triangulaciones problemáticas ocurren, generalmente, entre rodajas con poca similitud en lo referente al número, tamaño y disposición espacial de polígonos dentro de cada una.

Las razones previas indican dos direcciones futuras de trabajo:

- A) De una parte, investigar nuevas estrategias de triangulación, o mejora de las existentes, que permitan una buena descripción de superficies 3D a partir del conocimiento de líneas de su contorno.
- B) Por otro lado, y habida cuenta de los excelentes resultados obtenidos al representar las superficies de las reconstrucciones mediante una familia de líneas de su contorno, cabría la posibilidad de estudiar nuevas estrategias de normalización basadas en este esquema de representación.

A modo ilustrativo, en la Tabla 4.VIII se indica el orden de magnitud para el tiempo de CPU sobre el ordenador VAX 11/780 en la ejecución de los programas de reconstrucción y normalización para el objeto 12 en su transformación n°. O y empleando como vistas iniciales para la reconstrucción las 1 y 5.

PROGRAMA	N° EJECUCIONES	TIEMPO CPU (en s)
PROYECCION	10	9.7
CONTORNO	10	3.3
RASTERIZAR	1	1.4
PARALELOGRAMOS	1	6.8
REFINAMIENTO	8	89.9
TRIANGULAR	1	36.4
STANDARD	1	26.5
MOMENTOS	1	514.5

TABLA 4.VIII. Tiempos de ejecución para los programas de reconstrucción, normalización y obtención de características para el objeto n°. 12 en su transformación n°. 0.

IV.6.7 PROGRAMA MOMENTOS

En la obtención de momentos 3D sobre las versiones normalizadas de cada reconstrucción 3D, se han evaluado momentos de hasta orden 5. El tiempo de ejecución del programa MOMENTOS es directamente proporcional al número de triángulos que componen la superficie reconstruida. Por otra parte, y al ejecutarse para cada triángulo, será función de la rapidez con que se alcance la convergencia en la evaluación de las integrales dobles (3.55) mediante la subrutina DO1DAF de la librería NAG.

El número total de momentos evaluados sobre cada objeto 3D es 56. No todos ellos son independientes entre sí, y además algunos momentos toman siempre los mismos valores con independencia del objeto de que se trate. Esta característica viene motivada por el hecho de haber utilizado momentos de hasta orden 2 para transformar un objeto en su versión estándar.

Más concretamente, si sólo normalizamos frente a traslaciones y rotaciones, los momentos mioo, moio, mooi, que especifican la posición final del centroide (mioo/mooo, moio/mooo, mooi/mooo), así como los elementos no diagonales de la matriz de inercia (ec. 3.85, 3.60) moii, mioi, mio están prefijados para todos los objetos¹.

```
m100 = 0 \times m000 = 0

m010 = 0 \times m000 = 0

m001 = 0 \times m000 = 0

m011 = m101 = m110 = 0
```

De esta forma, de los 56 momentos evaluados podemos eliminar m100, m010, m011, m101, m110 ya que estos no permiten distinguir unos objetos de otros.

Si se normaliza frente a traslaciones, rotaciones y cambios de

¹ Se entiende que estos momentos son los evaluados sobre la versión ya estandarizada.

escala, podemos eliminar los momentos empleados en dichas normalizaciones. Eligiendo como criterio de normalización frente a cambios de escala la igualdad en el área (100 unidades) de todas las versiones estándar, se tiene:

mooo = 100

m100 = m010 = m001 = 0

mo11 = m101 = m110 = 0

Así, de los 56 momentos obtenidos sólo emplearemos 49 en el proceso de reconocimiento.

IV.6.8 MATERIAL Y UTILIDADES

Mencionemos que todos los programas y subrutinas empleadas en este trabajo han sido desarrollados por nosotros mismos, a excepción de las subrutinas DO1DAF y FO2ABF para la evaluación de integrales dobles y obtención de valores y vectores propios, respectivamente, tomadas de la librería NAG. Para análisis discriminante se ha utilizado el programa MAHAL3, tomado del texto de Romeder [ROME-73] y adaptado a nuestras necesidades. Todos los programas han sido implementados en VAX-11 FORTRAN y se encuentran a la disposición de cualquier interesado.

Los medios materiales empleados son los siguientes:

- Ordenador VAX 11/780.
- Terminal gráfico TEKTRONIX 4115B con una resolución de 1024 x 1280.
- Terminal TELEVIDEO 9220 con capacidad de emulación del terminal TEKTRONIX 4115B y con una resolución de 640 x 240.
- Plotter TEKTRONIX 4663.

Para la visualización y trazado de los objetos, vistas, contornos y reconstrucciones 3D se ha hecho uso del paquete de subrutinas gráficas para 2 y 3 dimensiones PGPLOT [PGPL-87].

En la Fig. 4.25 se muestra un esquema del sistema utilizado en la simulación del proceso global de reconstrucción y reconocimiento de objetos 3D.

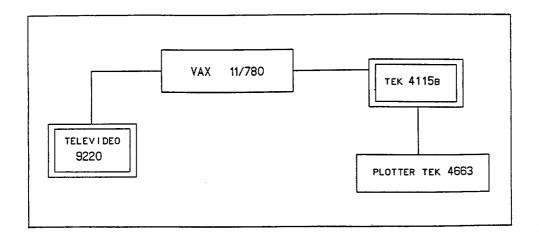
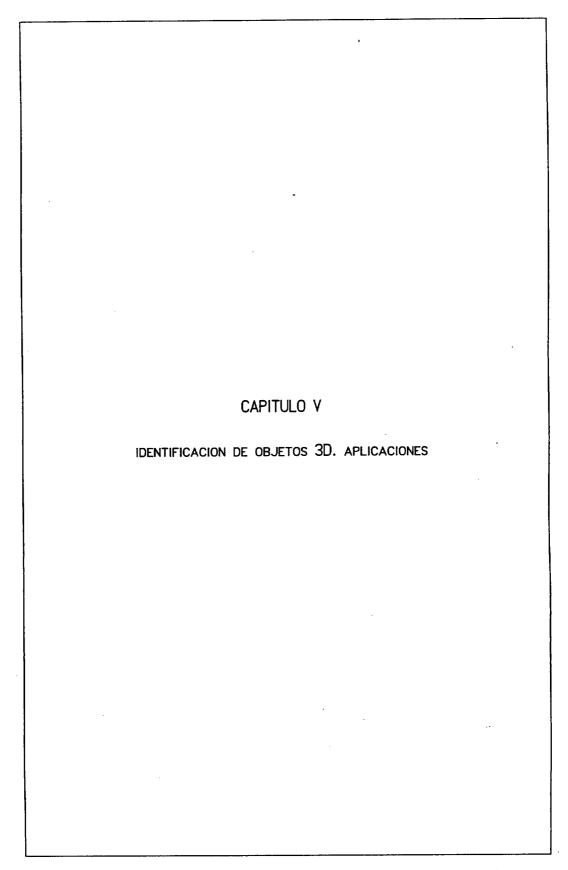


Fig. 4.25. Esquema del sistema disponible en el desarrollo de este trabajo.



V.1 INTRODUCCION

Reconstruida cada forma tridimensional a partir de vistas 2D, y normalizada de manera conveniente para que las características que de ella se extraigan sean invariantes frente a traslaciones, rotaciones y cambios de escala, sólo resta pasar a la etapa de reconocimiento.

Más concretamente, el problema estudiado en este trabajo ha sido la identificación de objetos 3D, estando cada uno de ellos representado mediante un vector de características en el espacio \mathbb{R}^p , donde p representa el número de momentos 3D evaluados sobre cada objeto.

La identificación se consigue al asignar cada objeto desconocido a alguna de las clases o grupos conocidos a priori. Una vez se ha establecido la existencia de varias clases de objetos, los métodos de discriminación [ROME-73] proporcionan una herramienta útil para caracterizar, mediante un número reducido de variables (momentos 3D), la pertenencia a un grupo determinado. Dichos métodos están sujetos a la siguiente hipótesis de base: "Se dispone de un conjunto de datos relativos a individuos (objetos 3D) pertenecientes a dos o más grupos -clases o categorias- definidos a priori de forma precisa". Es decir, no se cuestiona la existencia de clases de objetos 3D sino que se desea encontrar, de entre las variables disponibles, aquellas que permitan decidir satisfactoriamente la pertenencia de un objeto desconocido a alguna de las categorias existentes.

Nuestro criterio de asignación consiste en comparar entre sí las distancias de cada nuevo individuo a los centros de cada clase (distancias medidas con una métrica apropiada).

En la elección de las características que mejor discriminan entre clases, y por tanto comportan un menor riesgo de asignación errónea, se ha seguido un método de paso a paso. Esta técnica consiste en seleccionar, de entre las características empleadas, la mejor variable al primer paso; a continuación las dos mejores; seguidamente las tres mejores, y así sucesivamente.

V.2 DISCRIMINACION: DEFINICION

Sea un conjunto de individuos repartidos en varias clases o grupos definidos a priori. Sea, también, un conjunto de variables o características medidas sobre cada individuo. El problema general de discriminación puede plantearse en los términos siguientes:

Las variables anteriores, ¿permiten diferenciar o discriminar entre las clases a priori definidas?.

Dos grandes categorias de problemas pueden plantearse:

- A) Discriminación descriptiva.
- B) Discriminación desde un punto de vista de toma de decisiones.

V.2.1 DISCRIMINACION DESDE UN PUNTO DE VISTA DESCRIPTIVO

En este caso se desea mostrar el poder discriminante de las variables medidas sobre cada individuo, conocidas las diferentes clases y unas muestras de individuos para cada una de ellas. Por un lado se pretende asegurar una buena separación entre clases, y por otro lado poder comparar el poder discriminante de cada una de las variables empleadas en la separación.

Este concepto de discriminación comporta dos aspectos distintos. De una parte la existencia de una separación entre clases con poco solapamiento entre ellas. Por otra parte la exigencia de una separación estadísticamente significativa.

V.2.2 IDENTIFICACION

En este caso se trata de tomar una decisión. Cada nuevo individuo sobre el que se han medido un conjunto de características debe ser asignado a una de las clases definidas a priori, bajo la suposición de que pertenece a alguna de ellas. Se trata, por tanto, de identificar

la categoria a la que pertenece en función de su vector de características.

Se entiende así que la identificación, o toma de decisiones, es una etapa que, necesariamente, sucede a la primera de discriminación descriptiva.

En el primer caso se trata de responder a la pregunta: ¿COMO agrupar un conjunto de individuos, de los que se conocen ciertos atributos, en clases lo más distintas posibles?.

En el segundo caso se trata de afectar un individuo, del que se conocen ciertos atributos, a una de entre varias clases. La pregunta a la que se debe responder es: ¿DONDE se asigna el individuo ?.

V.3 METODO DE ASIGNACION

Pasemos ahora a describir el método de asignación empleado. Este método, en el caso particular de 2 clases, se reduce a utilizar el método clásico de la función lineal discriminante introducida por Fisher en 1936 [FISH-36]. Para cada individuo a identificar se evalúa la distancia que lo separa del centro de cada una de las clases. La asignación se efectuará a aquel grupo con el que presente una mayor similitud, o sea a aquel cuya distancia, según la métrica elegida, sea la menor.

V.3.1 NOTACION

Para caracterizar las distintas clases y poder definir los criterios de discriminación entre ellas, se necesitará disponer de muestras de individuos para cada clase y haber evaluado un mismo conjunto de características sobre cada uno de ellos. Será preciso, después de haber definido estas funciones discriminantes, poder evaluar su validez. Un indicador, para este último aspecto, lo constituye el porcentaje de clasificación correcta.

© Universidad de Las Palmas de Gran Canaria. Biblioteca Digital, 2004

La notación empleada es la siguiente:

Se considera un conjunto X de N individuos x sobre los cuales se han medido P variables.

X está dividido en K clases, siendo Y el conjunto de clases y.

Se tiene así:

x : individuos $x \in X$ y : clases $y \in Y$

A cada uno de los puntos x del espacio \mathbb{R}^p se le asigna un peso mx = 1/N. Si notamos Ny al número de individuos de la clase y, el peso asignado a cada una de ellas será my = Ny/N.

Los centros de gravedad, o valores medios, \bar{x} de X e \bar{y} de Y, son coincidentes y los notaremos por el vector g.

 x_i : i-ésima componente del individuo (vector) x

yi : i-ésima componente del vector medio del grupo y

gi : i-ésima componente del vector medio del conjunto de N individuos.

con i = 1, 2, P.

$$g = \overline{x} = \overline{y} = \sum \{ m_x x / x \in X \} = \sum \{ m_y y / y \in Y \}$$

Para el problema de identificación o asignación se dispondrá de individuos desconocidos o anónimos. A éstos se les designará por la letra "a", correspondiéndoles también un vector en \mathbb{R}^p de coordenadas (a1, a2,...ap).

V.3.2 DEFINICION DEL CRITERIO

Dado un nuevo individuo anónimo "a", se desea saber la clase a la que pertenece. Se supone, efectivamente, que ha de pertenecer a alguna de las clases.

Establecida una partición de \mathbb{R}^P en K regiones correspondientes a

las K clases, se calcula la distancia, mediante la métrica definida por T^{-1} , que separa al individuo "a" del centro de la clase y. Esta distancia viene dada por:

$$d(a, y) = (a - y)' T^{-1} (a - y)$$
 (5.1)

siendo:

(a-y) : vector columna definido por la diferencia entre los correspondientes vectores columna "a" e "y".

(a-y)': Matriz traspuesta de (a - y). Se trata de un vector fila.

T: Matriz de covarianza total de orden PxP. Sus elementos diagonales son las varianzas relativas a cada una de las P variables; los términos no diagonales son las varianzas relativas a cada pareja de variables.

$$T = [tij] t_{ij} = \frac{1}{N} \sum \{ (x_i - g_i) (x_j - g_j) / x \in X \} (5.2)$$

$$i, j : 1, 2, P$$

En esta expresión, la suma está extendida a todos los individuos x que componen la muestra de partida.

Después de evaluar (5.1) para cada grupo y \in Y, se decide asignar el individuo anónimo "a" a la clase yo tal que

$$d(a, yo) = min \{ d(a, y) / y \in Y \}$$
 (5.3)

La probabilidad de que las distancias de un punto a dos grupos distintos sean iguales es nula si las variables medidas toman un conjunto no discreto de valores.

Las cantidades d(a,y) definidas en (5.1) son funciones cuadráticas de "a"; sin embargo, todas ellas tienen en común el término a' $T^{-1}a$. Así, en lugar de comparar entre sí las funciones cuadráticas (5.1), será posible comparar las siguientes funciones

lineales de "a", relativas a cada grupo y:

$$V_{v}(a) = d(a,y) - a'T^{-1}a = -2y'T^{-1}a + y'T^{-1}y$$
 (5.4)

o más simplificadamente

$$V_{V}(a) = y' T^{-1}(y - 2a)$$
 (5.5)

Con estas consideraciones, el criterio de asignación definido en (5.3) puede expresarse como: El individuo anónimo "a" se asigna a la clase yo tal que

$$V_{y0}(a) = \min \{ V_{y}(a) / y \in Y \}$$
 (5.6)

donde Vy(a) está definida por (5.5)

V.4 PROCEDIMIENTOS DE PASO A PASO

El fin último perseguido en discriminación es la obtención de métodos que permitan alcanzar las menores tasas de asignaciones erróneas. Las condiciones óptimas para ello corresponderían a situaciones en las que:

- a) Las diferentes clases estuvieran bastante separadas entre sí (gran varianza entre-clases).
- b) Dentro de cada clase, los diferentes individuos que la componen presentasen una gran similitud o un elevado grado de agregación (pequeña varianza dentro de cada clase).

De manera general, y habida cuenta de la gran diversidad que pueden presentar las características medidas, resulta evidente que un número insuficiente de éstas podría conducir, en ocasiones, a resultados insatisfactorios [CHEN-73]. Por otro lado, trabajar con vectores de características de muy elevada dimensión puede resultar

impracticable. Se debe establecer, por tanto, un compromiso entre ambas situaciones que permita alcanzar un grado de confianza adecuado en la actuación del proceso de identificación.

Se trata pués de encontrar el subconjunto de características que aseguren la mejor discriminación. Un procedimiento eficiente, aunque subóptimo, es la técnica de paso a paso. Esta consiste en selecionar, mediante algún criterio, el mejor conjunto de variables de manera secuencial, añadiendo a cada paso una variable suplementaria al subconjunto obtenido en el paso anterior.

La búsqueda óptima consistiría en examinar, a cada paso q, todas las posibles combinaciones de q variables, lo que resulta computacionalmente prohibitivo cuando el número P de características medidas es elevado. En este caso se necesitaría evaluar el criterio de selección 2^P veces, mientras que con la técnica de paso a paso este criterio precisa ser ejecutado un máximo de P(P+1)/2 veces.

V.4.1 CRITERIO DE SELECCION DE VARIABLES

La elección de una nueva variable a cada paso puede efectuarse atendiendo a criterios diversos [ROME-73], [CHEN-73]. Un posible criterio está basado en el porcentaje de clasificación correcta para los individuos de la muestra inicial conocida. El procedimiento de paso a paso correspondiente a este criterio evalúa a cada paso, y para cada variable, el porcentaje de asignación correcta a partir de las distancias de cada individuo a cada grupo, sabiendo el grupo al que realmente pertenece. De las variables examinadas se elige aquella asociada al mayor porcentaje de clasificación correcta. No obstante, y como se apunta en [ROME-73] este criterio exige un análisis preciso ya que su utilización puede inducir a resultados erróneos.

El criterio empleado en nuestro trabajo consiste en maximizar la traza de la matriz $(T^{-1}B)$ como se indica a continuación.

V.4.2 MAXIMIZAR TRAZA (T⁻¹B)

La selección de una nueva variable a cada paso, mediante este criterio no necesita necesariamente de la definición de un procedimiento de clasificación. Para el caso anterior -criterio del porcentaje de clasificación correcta- el método de asignación era el de la mínima distancia.

T es la matriz de covarianza total definida en (5.2), de manera tal que la forma cuadrática uTu' es definida positiva definiendo sobre \mathbb{R}^P una estructura euclídea cuyo producto escalar asociado viene especificado por la matriz inversa T^{-1} . Así la distancia entre dos puntos de \mathbb{R}^P , W1 y W2 será

$$d(W_1, W_2) = (W_1 - W_2)' T^{-1}(W_1 - W_2)$$
(5.7)

B es la matriz de covarianza entre-clases definida por

$$B = [b_{ij}] b_{ij} = \sum \left\{ \frac{N_y}{N} (y_i - g_i) (y_j - g_j) / y \in Y \right\}$$

$$i, j : 1, 2, ..., P$$
(5.8)

Notemos que en esta expresión la suma está extendida sobre todas las clases y del conjunto Y.

Determinaremos, a cada paso, el conjunto de variables que hacen máxima la inercia de la nube Y^2 , calculada con la métrica T^{-1} , respecto de su centro de gravedad (el de la nube, $\overline{y} = \overline{x} = g$). Concretamente, en el paso q se elige como mejor subconjunto de q variables aquel que proporcione el valor máximo para la expresión escalar

¹ U representa un vector fila en R^p.

Mediante el término "nube" hacemos referencia al conjunto de puntos de R definidos por los centros de gravedad de cada clase y e Y.

$$\sum \left\{ \begin{array}{c} N \\ \overline{N} \end{array} (y - \overline{y}), \ T_{q}^{-1} \ (y - \overline{y}) / y \in Y \end{array} \right\}$$
 (5.9)

donde la suma está extendida a todos los vectores "y" representativos de las distintas clases pero considerando sólo q componentes (las q-1 primeras son las correspondientes a las q-1 variables ya seleccionadas en pasos anteriores, y la última componente es la variable a ensayar en este paso q). La matriz Tq designa a la matriz de orden p deducida de T al considerar sólo los términos relativos a las q variables.

Teniendo en cuenta que (5.9) es un escalar, éste puede ser considerado igual a la traza de la correspondiente matriz reducida a un único elemento (él mismo). Además, recordando que traza (AB) = traza (BA), para A y B matrices de órdenes mxn y nxm respectivamente, y considerando la linealidad del operador traza, podemos escribir la relación (5.9) como:

Traza
$$\left\{ T_{q}^{-1} \left[\sum \frac{Ny}{N} (y-\overline{y}) (y-\overline{y})' \right] \right\}$$
 (5.10)

Igualmente, si se considera a Bq como la matriz obtenida a partir de la matriz B de covarianza entre-clases (5.8) teniendo en cuenta sólo los elementos relativos a las q variables, (5.10) puede expresarse como:

$$Traza \left(\begin{array}{cc} T_q^{-1} & B_q \end{array} \right) \tag{5.11}$$

Así, el criterio de selección de la variable número q, en el paso q, puede indicarse mediante

Maximizar Traza
$$\left(\begin{array}{cc} T_q^{-1} & B_q \end{array}\right)$$
 (5.12)

Es decir, de entre las [P-(q-1)] variables que aún no han sido

seleccionadas se elige aquella que hace máximo el escalar Traza $(T_q^{-1}B_q)$. Este criterio equivale, como veremos a continuación, a hacer máxima la distancia D^2 de Mahalanobis.

V.4.3 CASO DE DOS GRUPOS

En el caso en que los individuos de la muestra de partida estén distribuidos únicamente en 2 grupos con efectivos N1 y N2 respectivamente, podemos ver que el criterio de paso a paso propuesto en (5.12) está íntimamente relacionado con la distancia D^2 de Mahalanobis. En esta situación se tiene:

 \overline{y} : centro de gravedad de ambas clases; $\overline{y} = \frac{N1y_1 + N2y_2}{N}$

y1 : vector medio de la clase y1y2 : vector medio de la clase y2

La matriz de covarianza, Bq, entre las clases 1 y 2 será:

$$B_{q} = \frac{N1}{N} (y_{1} - \overline{y}) (y_{1} - \overline{y})' + \frac{N2}{N} (y_{2} - \overline{y}) (y_{2} - \overline{y})'$$

$$\downarrow \qquad \qquad \downarrow \qquad \qquad \qquad \qquad \downarrow \qquad \qquad \qquad \downarrow \qquad \qquad \downarrow$$

Operando sobre esta expresión, la podemos transformar en

$$B_{q} = \frac{N1N2}{N^{2}} (y_{1} - y_{2}) (y_{1} - y_{2})'$$
 (5.14)

Introduciendo (5.14) en (5.12), el criterio de paso a paso se reduce a:

Maximizar Traza
$$\left\{ \frac{N1N2}{N^2} (y_1 - y_2), T_q^{-1} (y_1 - y_2) \right\}$$
 (5.15)

En esta última relación el escalar contenido en el corchete, salvo el factor $N1N2/N^2$, es precisamente la distancia D^2 de Mahalanobis entre las clases y1 e y2

$$D^{2} = (y_{1} - y_{2})' T_{q}^{-1} (y_{1} - y_{2})$$
 (5.16)

Más concretamente, D^2 mide la distancia entre los centros de ambas clases mediante la métrica definida por T_q^{-1} .

V.4.4 RESUMEN DEL CRITERIO MAX TRAZA (Ta1Ba)

La cantidad determinada en (5.12) constituye una generalización de D^2 de Mahalanobis al caso de más de dos grupos o clases, y puede ser considerada como un índice de separación entre varios grupos en el espacio \mathbb{R}^p .

Para el método de paso a paso la forma de proceder es simple:

- Inicialmente se dispone de las matrices T y B calculadas para las P variables y sobre el conjunto de individuos conocidos a priori.
- En el primer paso, q = 1, se calcula, para cada variable de las P disponibles, la cantidad Traza (Ti¹B1), que se reduce al cociente entre los términos diagonales de B y T correspondientes a la variable considerada. De todas las variables se elige aquella que proporcione el valor máximo de las trazas calculadas.
- En el segundo paso, q = 2, se utilizan matrices T_2 y B_2 , de orden 2, relativas a la variable anteriormente seleccionada y a una nueva variable añadida. Para cada variable añadida se calcula el escalar Traza $(T_2^{-1}B_2)$. Como segunda variable se selecciona aquella

que proporcione la traza máxima.

- Este proceso se vuelve a repetir hasta tener en cuenta, si así se desea, la totalidad de las P variables.

Resulta claro que mediante el criterio de maximizar Traza $(T_q^{-1}B_q)$ no se dispone de un test natural para finalizar el proceso de paso a paso. En efecto, la cantidad Traza $(T_q^{-1}B_q)$ podrá crecer de un paso al siguiente sin que por ello la discriminación se vea mejorada.

Sin embargo en el criterio del porcentaje de clasificación correcta se puede establecer un umbral para dicho porcentaje de tal forma que una vez alcanzado se ponga fin al proceso de selección de nuevas variables. Esta es la razón por la que dicho porcentaje se emplea como uno de los indicadores para establecer el final del proceso mediante el criterio basado en la Traza $(T_q^{-1}B_q)$. Otra alternativa consiste en prefijar, de antemano, el número de pasos a ejecutar.

Indiquemos, por último, que el criterio definido en (5.12) equivale a hacer máxima la suma de valores propios, λ_i , de la matriz $T_q^{-1}B_q$, ya que como se demuestra en [ROME-73]

Traza
$$(T_q^{-1}B_q) = \sum_{i=1}^q \lambda_i$$

V.5 DESCRIPCION DE MAHAL3

La asignación de objetos 3D desconocidos a alguna de las clases de objetos previamente definidas, se lleva a cabo mediante el programa MAHAL3. Dicho programa ha sido tomado de la librería expuesta en [ROME-73] y adaptado a nuestras características particulares.

Se trata de un programa de discriminación lineal entre 2 ó más grupos. Contempla la posibilidad de introducir 3 conjuntos distintos

de individuos: conjunto de aprendizaje (Training-Set), individuos de prueba (Test-Set) e individuos anónimos.

V.5.1 MUESTRA DE APRENDIZAJE

Con el fin de definir criterios discriminantes es preciso disponer de una muestra de individuos cuyo prupo de pertenencia es a priori conocido. Tales criterios discriminantes se ajustarán de forma que los diferentes grupos que conforman esta muestra estén lo más separados posible en el espacio \mathbb{R}^P , asumiendo que de cada individuo se conocen los valores de un mismo conjunto de P variables o características.

V.5.2 MUESTRA DE PRUEBA

La muestra de individuos que componen este conjunto no sirve mas que para <u>verificar a posteriori</u> que la separación o discriminación entre clases es satisfactoria. Igualmente, se considera conocido el grupo de pertenencia para cada uno de estos individuos. El grado de separación entre clases vendrá determinado por el porcentaje de asignación correcta, evaluado al contabilizar el número de individuos que, habiendo sido asignados a una clase pertenecen realmente a ella.

En la práctica, la construcción de una muestra de prueba sólo será factible cuando el número de individuos sea suficientemente elevado. Es usual elegir entre un 20-30 % de los individuos disponibles para construir el Test-Set, mientras que los restantes conforman la muestra de aprendizaje. En MAHAL3 no es preceptivo disponer de la muestra de prueba.

V.5.3 INDIVIDUOS ANONIMOS

Además de las dos muestras anteriores, se dispondrá también de individuos anónimos siendo éstos, precisamente, a los que se quiere

identificar como pertenecientes a alguna de las clases a priori conocidas. La asignación se efectuará en función de los criterios discriminantes previamente determinados sobre la muestra de aprendizaje.

V.5.4 DESARROLLO DEL PROGRAMA

El criterio de paso a paso utilizado en MAHAL3 es el consistente en maximizar la Traza $(T_q^{-1}B_q)$ y, como se ha mencionado, constituye una generalización de D^2 de Mahalanobis. Conviene indicar que la determinación de la mejor variable, a cada paso, se efectúa a partir de los individuos de la muestra de aprendizaje.

Mientras que el método de paso a paso se emplea para conseguir la mejor separación entre clases, el criterio de asignación basado en la mínima distancia definida en (5.3), es el utilizado para clasificar los individuos de aprendizaje, prueba y anónimos. Recordemos que en lugar de comparar las distancias (5.1) que separan a cada individuo del centro de las distintas clases, se pueden comparar las funciones lineales correspondientes definidas en (5.4), y que justifican el calificativo "lineal" dado a este método de discriminación.

A cada paso se imprime la tabla de clasificaciones y los porcentajes de asignaciones correctas relativas a la muestra de aprendizaje y a la de prueba (si la hay).

En cualquier caso, habrá que indicar a MAHAL3 el número máximo de pasos que se desean ejecutar (en su defecto el proceso de paso a paso se repetirá hasta agotar todas las variables disponibles). MAHAL3 finalizará en el último paso, salvo que antes el porcentaje de asignación correcta alcance el 100 % en la muestra de aprendizaje, en cuyo caso el proceso finaliza.

Además de las tablas anteriores de clasificación, a cada paso se imprime sistemáticamente la lista de variables seleccionadas, en número igual al número del paso.

Los resultados anteriores no permiten identificar a los individuos mal clasificados ni, tampoco, apreciar cuantitativamente el motivo de esta asignación errónea. Esta es la razón por la cual se imprimen resultados adicionales para aquellos pasos en los que se detecte un máximo relativo en el porcentaje de clasificación correcta para la muestra de aprendizaje. Esencialmente, estos resultados se refieren a las asignaciones y distancias a las diferentes clases para cada individuo de las muestras de aprendizaje y prueba, así como para los individuos anónimos.

Por último, indicar que el empleo de MAHAL3 está sujeto a la restricción $N \ge P+1$, en lo referente al número de individuos N y al número de características, P, medidas sobre cada uno.

V.6 CONSIDERACIONES PRACTICAS EN LA IDENTIFICACION DE FORMAS 3D

El banco de objetos utilizado en las aplicaciones de identificación es el de la Fig. 4.3. Consta de 13 objetos y de cada uno de ellos se tienen 11 versiones situadas y orientadas de manera diferente. Se dispone por tanto de 143 formas 3D reconstruidas agrupadas en 13 clases o categorias distintas.

Cada forma se representa mediante un vector de características cuya dimensión varía en función del número de momentos empleados para describirla. El número máximo de momentos evaluados sobre cada forma 3D es 49 y corresponde a la obtención de los 56 momentos, de orden 0 hasta 5, después de eliminar los 7 momentos utilizados en la normalización frente a traslaciones, rotaciones y cambios de escala.

Con objeto de evaluar la eficacia de los momentos como descriptores de formas 3D y a la vista de la Tabla 4.VII, que muestra el número de normalizaciones correctas e incorrectas para las 11 versiones de cada objeto, éstos se han dividido en dos tipos: objetos con un número elevado de normalizaciones correctas (objetos 1, 2, 3, 5, 6, 7, 10, 11, 12 y 13) y objetos que presentan un número similar de normalizaciones erróneas y correctas (objetos 4, 8 y 9).

Teniendo en cuenta los resultados obtenidos en la reconstrucción 3D (rodajas), triangulación (poliedro) y normalización, estamos en condiciones de afirmar que son las dos últimas etapas (programas TRIANGULAR y STANDARD) las que mayor influencia muestran en la obtención de la versión canónica correcta. La causa más frecuente de normalizaciones incorrectas ha sido una orientación inadecuada de la forma 3D respecto de su patrón estándar (compárense, por ejemplo, las Fig. 4.21 2) y Fig. 4.23 2)). En estos casos se ha podido constatar que este hecho viene asociado a la ocurrencia de alguna -o ambas- de las situaciones siguientes:

- Obtención de triangulaciones visualmente inaceptables en la ejecución del programa TRIANGULAR.
- 2. Determinación incorrecta de un único referencial $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ de entre las 8 opciones $\{\pm \vec{u}_1, \pm \vec{u}_2, \pm \vec{u}_3\}$ asociados a las direcciones principales de inercia durante la ejecución de STANDARD.

En el primer caso las triangulaciones obtenidas no siempre se ajustan a la superficie real del objeto y ello está originado por la existencia de rodajas consecutivas con diferente número y disposición espacial de los polígonos que las componen.

En el caso 2, y unido al hecho de que STANDARD trabaja sobre los resultados obtenidos con el programa TRIANGULAR, cabe además la opción de que en algunas ocasiones el centroide de la forma reconstruida esté situado a distancias similares respecto de su superficie según alguno de los dos sentidos de las direcciones principales de inercia. Evidentemente, en estas situaciones la probabilidad de seleccionar uno u otro sentido es muy parecida, pudiendo conducir a una elección desacertada del referencial $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

Conviene resaltar que las anteriores situaciones no han sido las más frecuentemente encontradas y que, incluso cuando han estado presentes, las reconstrucciones 3D en términos de rodajas obtenidas para estos objetos constituyen una aproximación excelente de su

superficie real.

De las 11 reconstrucciones normalizadas que componen cada grupo de objetos, 7 se han elegido como componentes de la muestra de aprendizaje, y las 4 restantes para construir la muestra de prueba.

En la ejecución de MAHAL3 se ha desechado el empleo de individuos anónimos (formas 3D). En su lugar, el conjunto de prueba (Test-Set) cumple la misma función, permitiéndonos además conocer cuales de estas formas 3D son clasificadas correctamente y cuales no.

V.6.1 PRUEBAS REALIZADAS Y DISCUSION DE LOS RESULTADOS OBTENIDOS

En este apartado se presentan y comentan algunas de las pruebas más significativas realizadas en identificación de formas 3D empleando momentos 3D normalizados. Estas son:

- PRUEBA 1 Identificación, mediante MAHAL3, de formas 3D correspondientes a los objetos 1, 2, 3, 5, 6, 7, 10, 11, 12 y 13.
- PRUEBA 2 Ejecución de MAHAL3 para las tres categorias de objetos 4, 8 y 9.
- <u>PRUEBA</u> 3 Ejecución de MAHAL3 para identificar distintas versiones de los objetos 3, 4 y 13.
- PRUEBA 4 Ejecución de MAHAL3 para el caso de considerar un número reducido de grupos de objetos que no presentan problemas en la obtención de las versiones normalizadas (objetos 3, 5, 7 y 13).

En cada una de las etapas indicadas, MAHAL3 se ejecutó 7 veces considerando distintos grupos de características (momentos 3D normalizados) para describir las formas 3D. Esto equivale a representar cada forma 3D en el espacio de características mediante

vectores de diferente dimensión, lo que permitirá establecer los grupos de momentos que proporcionan mayores tasas de reconocimineto y por tanto una mejor distinción entre las distintas categorias de objetos 3D. En la Tabla 5.I se muestra el número y tipo de momentos empleados en cada ejecución de MAHAL3.

EJECUCION	MOMENTOS 3D	N° de MOMENTOS
1	de orden 2	3
2	de orden 3	10
3	de orden 4	15
4	de orden 5	21
5	de hasta orden 3	13
6	de hasta orden 4	28
7	de hasta orden 5	49

TABLA 5.I. Momentos que describen cada forma 3D en las distintas ejecuciones de MAHAL3.

Como se indicó en V.5.4, la finalización de MAHAL3 se produce cuando la tasa de asignación correcta de los individuos de la muestra de aprendizaje alcance el 100 %. Si esto no llegase a ocurrir, se ha previsto la finalización de MAHAL3 estableciendo un máximo de 10 ejecuciones del procedimiento de paso a paso para la selección de variables basado en maximizar Traza $(T^{-1}B)$.

V.6.1.1 PRUEBA 1

En las Tablas 5.II-5.VIII se muestran los resultados obtenidos tras las 7 ejecuciones de MAHAL3 para los objetos 1, 2, 3, 5, 6, 7, 10, 11, 12 y 13. Cada uno de los 10 grupos de objetos consta de 11

individuos (11 versiones distintas). Así la muestra de aprendizaje está formada por 70 formas 3D mientras que las 40 restantes se eligen como muestra de prueba.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 002	55.7	72.5
2	m 020	74.3	95.0
3	m 200	85.7	92.5

TABLA 5.II. Ejecución de MAHAL3 con momentos de orden 2.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 021	62.9	67.5
2	m 003	92.9	97.5
3	m 201	95.7	100
4	m 120	95.7	90.0
5	m 210	92.9	82.5
6	m300	92. 9	85.0
7	т 030	95.7	85.0
8	m 111	91.4	85.0
9	m 012	94.3	82.5
10	m 102	94.3	80.0

TABLA 5.III. Ejecución de MAHAL3 con momentos de orden 3.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 022	44.3	55.0
2	m 040	84.3	95.0
3	m 121	92.9	90.0
4	m 004	92.9	90.0
5	m 400	98.6	95.0
6	m 220	95.7	92.5
7	m 202	97.1	92.5
8	m 130	97.1	95.0
9	m 112	97.1	95.0
10	m 211	98.6	95.0

TABLA 5.IV. Ejecución de MAHAL3 con momentos de orden 4.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m O 4 1	68.6	75.0
2	m 221	90.0	100
3	m 023	90.0	97.5
4	m 140	87.1	87.5
5	M 401	88.6	87.5
6	m 005	91.4	90.0
7	m 410	91.4	87.5
8	m 203	91.4	87.5
9	m 500	97.1	92.5
10	m 212	95.7	92.5

TABLA 5.V. Ejecución de MAHAL3 con momentos de orden 5.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 021	62.9	67.5
2	m 002	94.3	100
3	m 201	98.6	100
4	m 120	95.7	87.5
5	m 200	94.3	90.0
6	m 020	95.7	92.5
7	m 210	98.6	92.5
8	т 300	98.6	92.5
9	m 030	100	92.5

TABLA 5.VI. Ejecución de MAHAL3 con momentos de hasta orden 3.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	^m 021	62.9	67.5
2	m 002	94.3	100
3	m 201	98.6	100
4	m 022	94.3	. 100
5	m 121	94.3	95.0
6	m 200	95.7	95.0
7	m 220	98.6	95.0
8	m 040	98.6	95.0
9	m 120	98.6	95.0
10	m 130	97.1	95.0

TABLA 5.VII. Ejecución de MAHAL3 con momentos de hasta orden 4.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m o21	62.9	67.5
2	m 002	94.3	100
3	m 201	98.6	100
4	m ozz	94.3	100
5	m 121	94.3	95.0
6	m 200	95.7	95.0
7	m 220	98.6	95.0
8	m 041	98.6	95.0
9	m 140	98.6	95.0
10	m 122	98.6	95.0

TABLA 5. VIII. Ejecución de MAHAL3 con momentos de hasta orden 5.

Puesto que las 10 categorias de objetos elegidos corresponden a aquellos que presentan pocas normalizaciones incorrectas, se observa que en cualquiera de las ejecuciones de MAHAL3 al describir cada forma 3D mediante distinto número de momentos y de órdenes distintos, se obtienen unas tasas de reconocimiento elevadas. Estas oscilan entre un 85.7% y un 100% para las formas que componen la muestra de aprendizaje, y entre un 95% y un 100% para la muestra de prueba. Al representar cada forma 3D mediante momentos de un único orden, la mayor tasa de asignación correcta en la muestra de prueba se obtiene al emplear momentos de orden 4 (98.6%). Podemos observar también que en la descripción mediante momentos de orden 3, el uso de 9 de ellos proporciona un 100% de separación correcta entre los individuos de la muestra de prueba (Tabla 5.VI).

A la vista de los resultados precedentes parece razonable limitarse a las funciones discriminantes definidas con la ayuda de los 4 ó 5 primeros pasos, ya que el añadir más momentos a estos primeros

no siempre mejora los resultados de la clasificación para los dos conjuntos de formas 3D (prueba y aprendizaje). Notemos que las tasas inferiores de clasificación correcta se obtienen al emplear únicamente los 3 momentos de orden 2 (mooz, mozo, mzoo), mientras que el tener en cuenta momentos de órdenes 4 y 5 además de los 15 momentos de hasta orden 3 no mejora los resultados obtenidos con estos últimos.

En lo relativo a los tiempos de CPU sobre VAX 11/780, indicar que oscilan entre 17 y 30 segundos aproximadamente cuando se ejecuta MAHAL3.

Por tanto podríamos decir que la información proporcionada por los momentos de órdenes 2 y 3, y empleando sólo 4 ó 5 de ellos, proporciona para este banco de objetos una excelente discriminación.

V.6.1.2 PRUEBA 2

Aquí se aplica MAHAL3 a las 33 formas 3D normalizadas asociadas a los objetos 4, 8 y 9. Los tres tipos de objetos tienen en común el hecho de que el programa STANDARD, cuando se aplica a cada una de las 11 versiones 3D reconstruidas para cada objeto, no conduce a una única versión canónica.

La muestra de aprendizaje se compone de 21 formas 3D (7 asociadas a cada objeto) mientras que las 12 restantes se han empleado como muestra de prueba para comprobar, a posteriori, la validez de la separación entre clases conseguida con la muestra de aprendizaje. Los resultados obtenidos en las 7 ejecuciones de MAHAL3 al clasificar con momentos de distintos órdenes, se pueden resumir en las Tablas 5. IX-5. XV.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 020	100	91.7

TABLA 5. IX. Ejecución de MAHAL3 con momentos de orden 2.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	м 003	90.5	83.3
2	m 102	90.5	91.7
3	m 021	90.5	91.7
4	m 120	95.2	100
5	m 003	95.2	100
6	m 012	100	100

TABLA 5.X. Ejecución de MAHAL3 con momentos de orden 3.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 220	100	100

TABLA 5.XI. Ejecución de MAHAL3 con momentos de orden 4.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 221	95.2	100
2	m 401	85.7	75.0
3	M 041	95.2	91.7
4	m 320	100	91.7

TABLA 5.XII. Ejecución de MAHAL3 con momentos de orden 5.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 020	100	91.7

TABLA 5.XIII. Ejecución de MAHAL3 con momentos de hasta orden 3.

		PORCENTAJE DE ASIGNACION CORRECTA		
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA	
1	m 020	. 100	91.7	

TABLA 5.XIV. Ejecución de MAHAL3 con momentos de hasta orden 4.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 020	100	91.7

TABLA 5.XV. Ejecución de MAHAL3 con momentos de hasta orden 5.

Se observa aquí que, incluso existiendo normalizaciones no todo lo correctas que cabría esperar, en 5 de las ejecuciones se alcanza una clasificación correcta del 100% empleando una única característica. Esto equivale a establecer una nítida separación, sin solapamientos, en el espacio de características al representar las 21 formas 3D de aprendizaje, asociadas a los tres tipos de objetos, mediante un sólo momento 3D normalizado.

De manera análoga al caso anterior, en el que se estudiaba la discriminación entre 10 objetos, podemos afirmar aquí que el número de pasos necesarios para finalizar MAHAL3 con 100% de clasificación correcta oscila entre 4 y 6 pasos (cuando se precise más de 1 paso).

La bondad de estos resultados, aún tratándose de formas 3D obtenidas tras una pobre actuación de la etapa de normalización, puede explicarse atendiendo al hecho de tratar sólo con tres categorias de objetos y que presentan superficies bien diferenciadas entre sí (téngase en cuenta que un reparto aleatorio de cada una de las 21 formas en tres grupos distintos conduciría a un porcentaje de clasificación correcta del 33%).

De las Tablas 5.XIII-5.XV se deduce que al considerar momentos de hasta órdenes 4 y 5 se obtienen las mismas tasas de clasificación correcta que al emplear momentos de hasta orden 3, y que además es el momento mozo el que proporciona, por sí mismo, una clasificación del 100% para la muestra de aprendizaje y un 91.7% para los individuos, supuestamente incógnitas, que componen el conjunto de prueba. Esto indica que, en lo referente a discriminación entre las tres categorias de objetos en estudio, la información contenida en los momentos de órdenes 3, 4 y 5 no añade nada nuevo a la proporcionada por los momentos de orden 2.

Los tiempos de ejecución para MAHAL3 oscilan, para esta prueba, entre 3 y 7 segundos aproximadamente.

V.6.1.3 PRUEBA 3

En este caso se presenta la actuación de MAHAL3 al discriminar entre tres categorias de objetos 3D con la peculiaridad de poseer formas externas con gran similitud. Se trata concretamente de las reconstrucciones 3D obtenidas a partir de las 11 transformaciones aplicadas a cada uno de los objetos 3, 4 y 13 de la Fig. 4.3. El objeto 3 es una pirámide de base exagonal, correspondiendo los objetos 4 y 13 a variantes obtenidas a partir del 3 eliminando o añadiendo partes nuevas.

Mientras que las reconstrucciones para los objetos 3 y 13 muestran su versión canónica con orientación similar a la de la versión normalizada de los objetos de partida (Fig. 4.23 3) y 13)), las versiones normalizadas asociadas al objeto 4 difieren en posiciones giradas 180° respecto del eje vertical OZ (6 versiones de acuerdo a la normalización de la Fig. 4.23 4) y 4 normalizaciones giradas 180° respecto de las anteriores); (Véase Tabla 4.VIII).

Un resumen de los resultados obtenidos en las 7 ejecuciones de MAHAL3 se muestra en las Tablas 5.XVI-5.XXII.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 200	90.5	91.7
2	m oso	85.7	66.7
3	m 002	85.7	66.7

TABLA 5.XVI. Ejecución de MAHAL3 con momentos de orden 2.

		PORCENTAJE DE ASIGNACION CORRECTA		
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA	
1	m 102	66.7	66.7	
2	m 021	90.5	83.3	
3	m 120	95.2	91.7	
4	т 300	95.2	91.7	
5	m 201	100	91.7	

TABLA 5.XVII. Ejecución de MAHAL3 con momentos de orden 3. .

		PORCENTAJE DE ASIGNACION CORRECTA		
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA	
1	m 400	95.2	91.7	
2	M 040	85.7	83.3	
3	m 301	95.2	83.3	
4	m 022	100	91.7	

TABLA 5.XVIII. Ejecución de MAHAL3 con momentos de orden 4.

		PORCENTAJE DE ASIGNACION CORRECTA		
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA	
1	m 302	66.7	66.7	
2	m 221	95.2	91.7	
3	m 320	95.2	91.7	
4	M 014	95.2	91.7	
5	m 023	100	91.7	

TABLA 5.XIX. Ejecución de MAHAL3 con momentos de orden 5.

		PORCENTAJE DE ASIGNACION CORRECTA		
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA	
1	m 200	90.5	91.7	
2	m 020	85.7	66.7	
3	m 300	90.5	66.7	
4	m 021	90.5	91.7	
5	m 102	100	91.7	

TABLA 5.XX. Ejecución de MAHAL3 con momentos de hasta orden 3.

		PORCENTAJE DE ASIGNACION CORRECTA		
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA	
1	m 400	95.2	91.7	
2	m 020	85.7	66.7	
3	m 301	95.2	75.0	
4	m 021	100	91.7	

TABLA 5.XXI. Ejecución de MAHAL3 con momentos de hasta orden 4.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 400	95.2	91.7
2	m 041	90.5	83.3
3	M 301	100	91.7

TABLA 5.XXII. Ejecución de MAHAL3 con momentos de hasta orden 5.

A la vista de estos resultados podemos decir que, en el peor de los casos, se necesitan 5 pasos para lograr una clasificación correcta del 100% en la muestra de aprendizaje al emplear momentos de orden 3, o de orden 5 o de hasta orden 3 (Tablas 5.XVII, 5.XIX y 5.XX). En el resto de situaciones se alcanzan idénticos resultados con un menor número de pasos, a excepción de la primera ejecución de MAHAL3 al emplear sólo los 3 momentos de orden 2 en donde se obtienen tasas de asignación correcta, en los conjuntos de prueba y aprendizaje, sensiblemente inferiores.

Estos resultados están de acuerdo a lo expuesto en el apartado III. 3.3.2 si se tiene en cuenta que al describir una forma 3D exclusivamente mediante momentos de hasta orden 2 ésta es equivalente a su elipsoide de inercia. Así, puesto que los tres objetos en estudio tienen superficies muy similares es razonable que sus elipsoides de inercia también lo sean y por tanto que la primera ejecución de MAHAL3 (Tabla 5.XVI) proporcione las menores tasas de clasificación correcta. Al representar las formas mediante momentos de orden superior se tienen en cuenta más detalles de los cambios presentes en su superficie. De aquí que los mejores resultados, para estas tres categorias de objetos, se obtengan al emplear momentos de hasta órdenes 4 y 5 (Tablas 5.XXI-5.XXII).

Si se compara la actuación de MAHAL3 sobre los objetos 4, 8 y 9 al emplear momentos de hasta órdenes 3, 4 y 5 (Tablas 5.XIII-5.XV) con los mismos resultados al discriminar entre los objetos 3, 4 y 13 (Tablas 5.XX-5.XXII) se observa que en el primer caso, al tratarse de objetos de formas externas normalizadas bien diferenciadas, sólo se necesita un único momento de orden 2 (mozo) para distinguirlos, mientras que para los objetos 3, 4 y 13, al presentar gran similitud, son los momentos de órdenes 4 y 5 los que establecen la mejor discriminación, siendo necesarios entre 4 y 5 de ellos.

Indiquemos también que para la muestra de prueba (formas 3D a identificar) se obtienen tasas de clasificación correcta inferiores a las obtenidas en las dos pruebas anteriores de discriminación entre 10 y 3 objetos respectivamente. De cualquier manera, los resultados son siempre superiores al 91%. De los resultados complementarios proporcionados por MAHAL3 puede observarse que ello es debido, básicamente, a que versiones 3D del objeto 4 son asignadas a la categoria definida por el objeto 3, dado el parecido entre ambos y a la actuación defectuosa de las normalizaciones efectuadas sobre las 11 versiones del objeto 4.

Los tiempos de CPU en las 7 ejecuciones de MAHAL3 varian entre 6 y 9 segundos aproximadamente.

A título ilustrativo, en la Fig. 5.1 se muestra la separación entre clases obtenida en la séptima ejecución de MAHAL3 al emplear momentos de hasta orden 5 (Tabla 5.XXII). Se representan las distintas formas de la muestra de aprendizaje según los dos primeros factores discriminantes obtenidos al realizar el análisis factorial discriminante en el paso número 3. De esta gráfica puede observarse la mayor cercanía entre las versiones de los objetos 3 y 4.

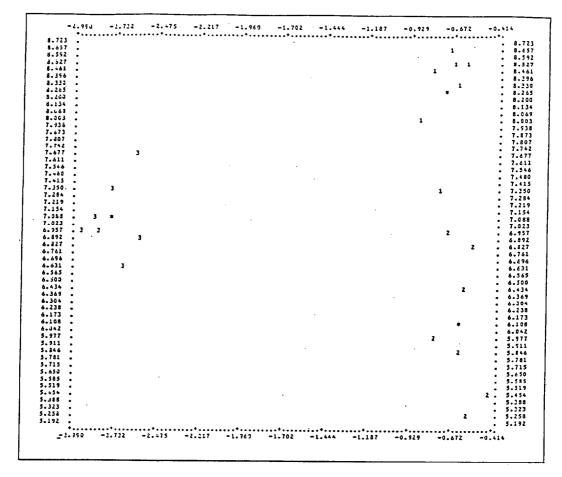


Fig. 5.1. Separación entre los tres grupos de objetos 3, 4 y 13. Mediante un * se indica la media de cada grupo. El código empleado es el siguiente:

- 1: versiones del objeto 3
- 2: versiones del objeto 4
- 3: versiones del objeto 13

V.6.1.4 PRUEBA4

El conjunto de aprendizaje está constituido por las 28 versiones reconstruidas y normalizadas obtenidas a partir de los 4 objetos 3, 5, 7 y 13 (7 reconstrucciones de cada objeto). El conjunto de prueba lo forman las 16 versiones restantes, también normalizadas. La diferencia con la prueba 1 reside en el hecho de considerar aquí un menor número

de clases al haber seleccionada sólo 4 objetos distintos, tratándose además de aquellos que no han mostrado ninguna anomalía en la etapa de normalización (Tabla 4.VII).

Los porcentajes de asignaciones correctas para los individuos de la muestra de prueba y aprendizaje en las 7 ejecuciones de MAHAL3 son los indicados en las Tablas 5.XXIII-5.XXVI.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 002	60.7	75.0
2	m 200	85.7	93.8
3	m 020	85.7	93.8

TABLA 5.XXIII. Ejecución de MAHAL3 con momentos de orden 2.

		PORCENTAJE DE ASIGNACION CORRECTA		
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA	
1	m 003	100	100	

TABLA 5.XXIV. Ejecución de MAHAL3 con momentos de orden 3, de hasta orden 4 y de hasta orden 5. En las 4 ejecuciones se obtienen los mismos resultados.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m 004	89.3	100
2	m 301	96.4	100
3	m 400	96.4	100
4	m 121	96.4	100
5	m 040	100	100

TABLA 5.XXV. Ejecución de MAHAL3 con momentos de orden 4.

		PORCENTAJE DE ASIGNACION CORRECTA	
PASO	MOMENTOS	MUESTRA DE APRENDIZ.	MUESTRA DE PRUEBA
1	m O 4 1	96.4	100
2	m 023	96.4	100
3	m 140	100	100

TABLA 5.XXVI. Ejecución de MAHAL3 con momentos de orden 5.

Como puede observarse, al emplear momentos de orden 3, de hasta orden3, de hasta orden 4 o de hasta orden 5 como componentes del vector de características se obtienen idénticos resultados (Tabla 5.XXIV). En estas 4 ejecuciones de MAHAL3, con una única variable (moo3) se alcanza un 100% de clasificación correcta en las muestras de aprendizaje y prueba. Nótese que bajo estas mismas circunstancias pero al discriminar entre 10 objetos distintos (PRUEBA1) que incluyen a los 4 en estudio, sólo se obtuvo el 100% de clasificación correcta al emplear momentos de hasta orden 3, siendo necesarios 9 de ellos (Tabla 5.VI).

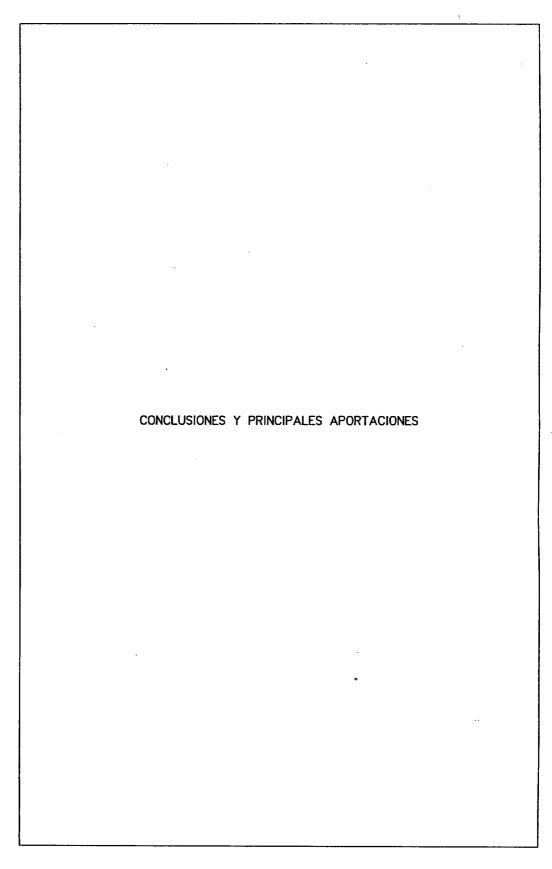
Sin embargo, la descripción mediante momentos de orden 2 conduce, en ambas pruebas, a resultados similares aún al considerar en la PRUEBA4 un menor número de objetos. En esta prueba las asignaciones erróneas (cuando se producen) tienen lugar, fundamentalmente, entre los objetos 3 y 13. La razón para ello reside en el gran parecido que muestran los citados objetos.

Obsérvese también que la ejecución de MAHAL3 finaliza en los pasos 5 y 3 al emplear momentos de órdenes 4 y 5 respectivamente (Tablas 5.XXV-5.XXVI) ya que se alcanza, en ambos casos, un 100% de clasificiación correcta para los individuos de prueba y aprendizaje. En las mismas ejecuciones de MAHAL3 en la PRUEBA1, la finalización se producía al ejecutar el último paso de los 10 permitidos sin alcanzar, en ningún caso, una tasa del 100%.

De estos resultados, y para el banco de objetos empleado en este trabajo, podemos indicar que las mejores tasas de reconocimiento se consiguen al utilizar momentos de hasta orden 3 ó superiores, poniéndose además de manifiesto que un aumento en el número de clases exige el empleo de un mayor número de momentos 3D para lograr tasas similares de clasificación correcta.

Un análisis comparativo de los resultados obtenidos en esta PRUEBA4 con los proporcionados en las pruebas 2 y 3 empleando un número similar de objetos, nos permite concluir que en aquellas situaciones de discriminación entre objetos con superficies similares, son los momentos de orden superior (hasta 4 ó 5) los que conducen a las mejores tasas de discriminación. Además, cuando los objetos en estudio no muestran triangulaciones o normalizaciones erróneas se alcanzan las menores tasas de asignación errónea al clasificar un objeto como miembro de una clase distinta a la que realmente pertenece.

Conviene puntualizar, por último, que los resultados obtenidos en el proceso de discriminación lineal mediante el empleo de MAHAL3, son función de las formas 3D que componen el banco de objetos estudiados. En este sentido, los resultados obtenidos en lo referente al número de pasos necesarios, y por tanto de coeficientes descriptores (momentos 3D), no serían extrapolables a otros bancos de objetos (más numerosos o de formas más complejas) mas que en sentido cualitativo.



CONCLUSIONES Y PRINCIPALES APORTACIONES

Esta Tesis representa un estudio acerca de las cuestiones implicadas en el reconocimiento de objetos tridimensionales, empleando como herramienta principal los momentos 3D. Los momentos 3D se han propuesto como una solución, desde una doble perspectiva:

- En la normalización de las formas tridimensionales a reconocer, mediante el empleo de momentos de hasta segundo orden.
- En la caracterización de formas 3D, como descriptores globales apropiados para clasificar la forma ya normalizada.

La manera de extraer y transformar la información contenida en una imagen, afecta de modo crucial a la eficiencia de los sistemas de visión artificial. Dicha eficiencia puede evaluarse al juzgar la actuación de las tres grandes fases siguientes:

- Recuperación de información 3D a través de las medidas efectuadas sobre los objetos.
- Transformación de los datos medidos en representaciones del propio objeto.
- Uso de las anteriores representaciones en tareas de reconocimiento e interpretación.

En función de lo anterior, las aportaciones más destacadas de nuestro trabajo se resumen en los siguientes puntos.

1. Se ha modificado y adaptado a nuestros fines un procedimiento basado en la técnica de intersección de volúmenes que, analizando los contornos de múltiples vistas de un objeto, permite reconstruir la superficie de éste. En la adquisición de las distintas vistas se ha empleado proyección ortográfica debido a la sencillez que comporta frente a otro tipo de proyecciones, sin que por ello se vea reducida la potencia del método de reconstrucción. A diferencia de otras estrategias que también emplean múltiples vistas, este procedimiento no necesita del establecimiento de la correspondencia entre características observadas en la secuencia de imágenes. El tipo de

objetos que pueden ser reconstruidos mediante este procedimiento es suficientemente amplio como para justificar su uso.

- 2. En el proceso de reconstrucción se distinguen dos etapas perfectamente diferenciadas: una reconstrucción inicial, y un posterior refinamiento de la anterior reconstrucción. Se ha modificado el anterior proceso, reemplazando la estructura de datos original, asociada a una descripción volúmica del objeto, por otra que es descriptiva de la superficie y más apropiada a nuestros fines. Ello ha comportado la reestructuración de la etapa de refinamiento, siendo necesaria la generación de un procedimiento de intersección entre polígonos, problema éste que presenta un gran interés en geometría computacional.
- 3. La descripción resultante, para la estructura 3D reconstruida, se compone de una familia de líneas de contorno, contenidas en planos paralelos consecutivos y equiespaciados, que acotan externamente a la superficie real del objeto. La estructura de datos para este esquema de representación se ha organizado por planos contiguos. Para cada plano se indica el número de contornos que lo componen, así como las coordenadas de los puntos que los definen. Los atributos más destacados en este esquema de representación son los siguientes:
 - Facilitar la actualización de la estructura de datos, necesaria en la etapa de refinamiento.
 - Ser descriptiva de detalles de la superficie reconstruida.
 - Permitir la transformación a otros esquemas de representación, ya sean superficiales o volúmicos.
- 4. La investigación en el campo de representación de formas tridimensionales es muy activa, concluyéndose que no existe una representación universal que se adapte a todos los problemas. Las características de un esquema concreto vienen determinadas, bien por la técnica de adquisición de datos o, en última instancia, por el tipo de aplicación para la que haya sido diseñada. Desde este punto de vista, se ha desarrollado una representación poliédrica para facilitar

la extracción de características invariantes de la forma 3D, con independencia de cuales fueron las vistas empleadas en su reconstrucción y el orden en que fueron procesadas. La descripción de la superfice de objetos mediante caras planas triangulares presenta la ventaja, frente a otro tipo de representaciones superficiales, de comportar un fácil tratamiento analítico en la normalización y obtención de características.

5. El esquema de representación indicado en el punto 3 se transforma en otro, también superficial, cuyas primitivas son superficies elementales triangulares. En este sentido, cabe mencionar lo siguiente:

Se ha descrito el problema general de triangulación, junto a una metodología que clasifica las soluciones existentes en óptimas y heurísticas. Se ha presentado, y desarrollado, un nuevo procedimiento heurístico de triangulación entre contornos de secciones contiguas que, a diferencia de los métodos existentes, incorpora una solución en el caso de que las secciones a conectar presenten distinto número de contornos. Al igual que otros métodos heurísticos, el nuestro sólo requiere (m+n) pasos para triangular entre dos contornos de m y n vértices respectivamente. Las aplicaciones efectuadas sobre el banco de objetos estudiado en esta memoria, constatan la bondad del método en la generación de superficies visualmente aceptables.

6. Se han definido los momentos 3D de una forma tridimensional, mediante integrales dobles evaluadas sobre superficie, investigándose las reglas de transformación de los momentos cuando la forma se somete a traslaciones, rotaciones y cambios de escala. En este sentido, las expresiones obtenidas ponen de manifiesto la dificultad en las transformaciones de rotación, al depender los momentos resultantes, no sólo de los parámetros de la rotación sino también de la orientación de cada una de las superficies elementales triangulares que componen la superficie reconstruida. Por este motivo, se ha optado por la normalización de cada forma 3D en lugar de normalizar directamente los momentos sobre ella evaluados.

- 7. En relación con la normalización de las formas 3D, se ha propuesto una estrategia que, a diferencia de otras existentes en 2D, no intenta evaluar directamente momentos invariantes, sino determinar los coeficientes de la transformación afín que relaciona la disposición espacial de un objeto cualquiera, con alguna posición, orientación y tamaño estándar previamente definidos. Las versiones normalizadas o canónicas para los objetos 3D se caracterizan por estar situadas siempre en la misma posición, con igual orientación y, en su caso, con igual tamaño. La transformación apropiada se obtiene al imponer una serie de criterios de normalización, de una fácil interpretación geométrica, sobre los momentos 3D de hasta segundo orden para la versión estándar.
- 8. Se ha investigado el método clásico de los ejes principales en relación con la orientación de formas 3D. Se concluye, y los resultados así lo confirman, que la diagonalización de la matriz de inercia conduce a ocho orientaciones diferentes para una misma forma 3D, constituyendo esta situación de ambigüedad un importante problema en el reconocimiento de objetos 3D. En conjunción con el método de los ejes principales, se han propuesto, y esto también es innovador, una serie de restricciones adicionales sobre los vectores propios de la matriz de inercia en su relación con la superficie reconstruida. Las citadas restricciones permiten obtener un único referencial ligado a cada objeto, de entre las ocho posibles opciones proporcionadas por los autovectores de su matriz de inercia. Los resultados obtenidos muestran que este nuevo procedimiento de normalización puede usarse con éxito en la determinación de la orientación estándar para una gran variedad de objetos 3D.
- 9. La técnica de análisis discriminante, en el reconocimiento de formas 3D con independencia de su posición, tamaño y orientación, ha proporcionado unos resultados altamente satisfactorios (en torno al 95%) al caracterizar cada forma normalizada mediante momentos 3D. De los resultados obtenidos, y para el banco de objetos estudiado, parece razonable limitarse a las funciones discriminantes definidas con la ayuda de los cuatro o cinco primeros pasos, ya que al añadir más momentos a estos primeros no siempre se mejoran los resultados de clasificación.

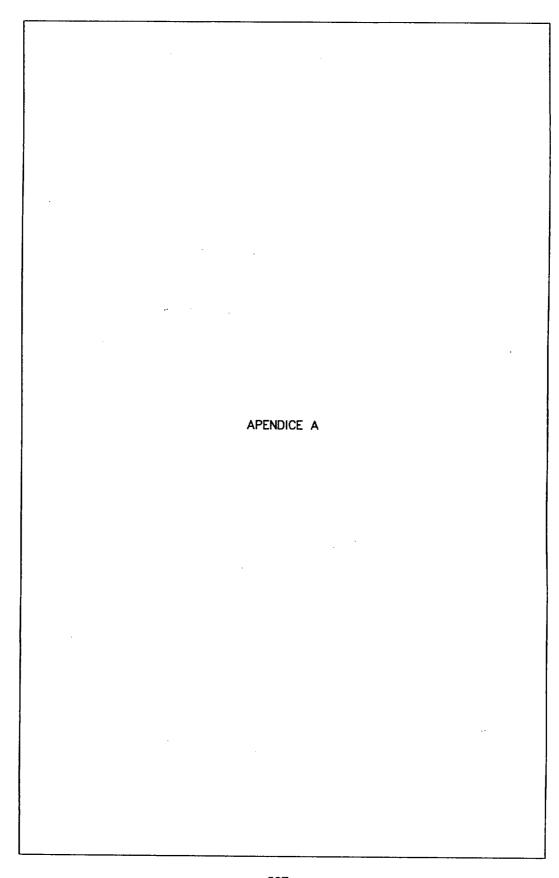
10. En general, las mejores tasas de clasificación se han obtenido al utilizar momentos de hasta orden 3 o superiores, poniéndose de manifiesto que un aumento en el número de clases exige el empleo de un mayor número de momentos para lograr tasas similares de clasificación correcta. Al discriminar entre objetos cuya geometría es similar, son los momentos de orden superior -de hasta órdenes 4 ó 5- los que conducen a las mejores tasas de discriminación, ya que de esta manera se tienen en cuenta pequeños detalles y cambios presentes en sus superficies. Por el contrario, para objetos de geometría bien diferenciada sus elipsoides de inercia respectivos también presentarán notables diferencias, siendo ésta la razón por la cual con momentos de bajo orden -de hasta orden 3- se obtienen elevadas tasas de asignación correcta.

AREAS DE FUTURAS INVESTIGACIONES

Directamente relacionadas con el trabajo expuesto en esta memoria, existen una serie de cuestiones interesantes que necesitan de consideraciones adicionales, sirviendo como punto de partida para investigaciones futuras. De entre ellas destacamos las siguientes:

- a) Habida cuenta de los excelentes resultados obtenidos al representar las superficies de las reconstrucciones mediante una familia de líneas de su contorno, resulta de sumo interés la investigación de estrategias de normalización basadas en este esquema de representación.
- b) La propia naturaleza del método de triangulación propuesto, conduce a la generación de una superfice compuesta de elementos planos triangulares que aproximan a la superficie real del objeto. Además, cuando se triangula entre rodajas contiguas con distinto número de contornos, la superficie generada, aunque visualmente aceptable, no es cerrada, lo que añade un problema adicional en la etapa de normalización. Actualmente estamos trabajando en la mejora del método de triangulación propuesto, con el fin de eliminar la mencionada dificultad.

- c) Aún habiendo proporcionado unos excelentes resultados en los casos estudiados, la técnica de normalización frente a rotaciones no está libre de inconvenientes. La unicidad en la obtención de la rotación que transforma cualquier objeto 3D en su versión estándar, no queda garantizada en el caso de degeneración en los valores propios de la matriz de inercia. Esta situación está asociada a la existencia de simetrías en torno a alguno de los ejes principales del objeto. Además, si el centroide del objeto no es interior a su superficie, la solución propuesta no siempre funcionará correctamente. Queda así abierta una línea de trabajo conducente al desarrollo de criterios de normalización que, basados en la imposición de restricciones sobre momentos de orden superior a dos, eliminen las ambigüedades inherentes al normalizar sólo con momentos de bajo orden.
- d) Los tiempos de cómputo en los procesos de triangulación y obtención de características son función directa del número de triángulos de que consta la superficie 3D reconstruida. En reconstrucciones de objetos con formas geométricas complejas, el número de triángulos requeridos es elevado, razón por la que las tareas anteriores, y fundamentalmente la evaluación de momentos 3D, resultan computacionalmente costosas. La naturaleza aditiva de los momentos, así como la independencia de los procesos de triangulación entre cada dos parejas de contornos 3D, resultan idóneas para su implementación en estructuras de tipo concurrente. Queda así abierta una vía dirigida a la disminución en el tiempo de proceso mediante el empleo de sistemas multiprocesadores.
- e) El proceso de convertir varias vistas bidimensionales de un objeto en una representación tridimensional completa de la superficie de éste, es una tarea que, aunque habitualmente realizada por algunos seres vivos, está aún lejos de ser resuelta en su totalidad de forma automática. Ello requerirá la utilización conjunta de toda la información contenida en las imágenes, y la suministrada por otro tipo de sensores. Un trabajo futuro que consideramos de interés, sería la cooperación de los métodos aquí empleados con otras estrategias de recuperación de información 3D.



ROTACIONES EN TRES DIMENSIONES

En tres dimensiones las traslaciones y transformaciones de escalado se expresan de la misma forma que en el caso bidimensional. No obstante, en las rotaciones es obligada la elección de un eje. Las expresiones para transformar las coordenadas de un punto tras someterlo a una rotación en torno de un eje arbitrario son bastante más complicadas; si embargo es posible presentarlas de una manera concisa mediante la ayuda del algebra vectorial.

Introduzcamos a continuación dos operaciones vectoriales adicionales que, por no ser tan extensamente conocidas como los productos escalar y vectorial, nos son menos familiares:

Definición:

El producto vectorial, $\overrightarrow{A} \times \overrightarrow{v}$, de una matriz A y un vector \overrightarrow{v} (vector columna) es otra matriz cuyas columnas se obtienen como el producto vectorial de las columnas de A con el vector \overrightarrow{v} .

<u>Definición:</u>

Una diada es el producto de un vector columna \vec{v} con un vector fila \vec{v} , y se representa como \vec{v} , siendo \vec{v} el traspuesto del vector columna \vec{v} .

A.1 ROTACION EN TORNO DE UN EJE QUE PASA POR EL ORIGEN

Sea un punto genérico P cuya posición respecto del sistema cartesiano OXYZ se especifica por las coordenadas (x,y,z), o bien por el punto extremo del vector \vec{b} (Fig. A.1). Sea un eje que pasa por el origen de coordenadas y viene definido por los ángulos ϕ_1 , ϕ_2 , ϕ_3 que forma con los ejes coordenados OX, OY y OZ respectivamente. La orientación de este eje puede especificarse mediante el vector unitario \vec{a} o vector columna

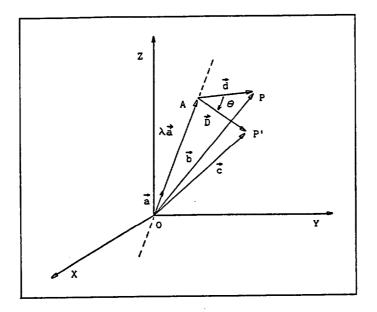


Fig. A.1. Rotación de un punto P en torno de un eje que pasa por el origen de coordenadas.

La rotación de valor θ en torno del eje definido por \vec{a} , y según el sentido contrario al indicado por la regla de la mano derecha, puede representarse por la matriz $R(\theta)$ que transforma las coordenadas de P en P'. Dicha matriz viene dada por [PAVL-82]:

$$R(\theta) = \overrightarrow{aa}' + \cos\theta(I - \overrightarrow{aa}') + \sin\theta(I \times \overrightarrow{a})$$
 (A.2)

siendo :

I : matriz identidad de orden 3x3.

aa': diada representada por la matriz resultante del producto de matrices

$$\begin{pmatrix} \cos\phi_1 \\ \cos\phi_2 \\ \cos\phi_3 \end{pmatrix} \begin{pmatrix} \cos\phi_1 & \cos\phi_2 & \cos\phi_3 \end{pmatrix} = \begin{pmatrix} \cos^2\phi_1 & \cos\phi_1\cos\phi_2 & \cos\phi_1\cos\phi_3 \\ \cos\phi_2\cos\phi_1 & \cos^2\phi_2 & \cos\phi_2\cos\phi_3 \\ \cos\phi_3\cos\phi_1 & \cos\phi_3\cos\phi_2 & \cos^2\phi_3 \end{pmatrix}$$

 $\vec{l} \times \vec{a}$: Matriz igual al producto vectorial de la matriz I y el vector columna \vec{a}

$$\overrightarrow{I \times a} = \begin{pmatrix} 0 & -\cos\phi_3 & \cos\phi_2 \\ \cos\phi_3 & 0 & -\cos\phi_1 \\ -\cos\phi_2 & \cos\phi_1 & 0 \end{pmatrix}$$

La validez de (A.2) se obtiene al comprobar que, efectivamente, después de aplicar la rotación $R(\theta)$ al vector \vec{b} se obtiene el vector \vec{c} que define al punto P' transformado de P.

Siempre es posible expresar \vec{b} como suma de dos vectores, uno $\lambda \vec{a}$ paralelo al eje de giro y otro, \vec{d} , normal a él:

$$\vec{b} = \lambda \vec{a} + \vec{d} \tag{A.3}$$

con un valor apropiado para la constante λ .

Calculando el producto $R(\theta)\vec{b}$ y usando las relaciones (A.2) y (A.3) se obtiene :

$$R \vec{b} = \lambda \vec{a} + \lambda \cos\theta \vec{d} + \sin\theta (\vec{l} \times \vec{a}) \vec{d}$$
 (A.4)

y desarrollando Ixa podemos escribir

$$R \vec{b} = \lambda \vec{a} + \cos\theta \vec{d} + \sin\theta \vec{d} \times \vec{a}$$
 (A.5)

Comprobemos, finalmente, que el vector $\cos\theta d + \sin\theta d \times d$ de (A.5) corresponde al vector \vec{D} de la Fig. A.1, de forma tal que $\vec{c} = \lambda \vec{a} + \vec{D}$. $\vec{d} \times \vec{a}$ es un vector de módulo d, perpendicular al plano definido por \vec{d} y \vec{a} y cuyo sentido es el propio de la rotación θ . por tanto es mutuamente perpendicular a \vec{a} y \vec{d} . Además, puesto que \vec{d} es perpendicular a \vec{a} , podemos afirmar que el vector $\cos\theta \vec{d} + \sin\theta \vec{d} \times \vec{a}$ será otro vector perpendicular al eje de giro y contenido en el plano definido por \vec{d} al rotar en torno de \vec{a} .

La Fig. A.2 muestra un esquema del proceso de transformación de P en P' visto desde el eje de giro y mirando hacia el origen de coordenadas desde el extremo del vector a.

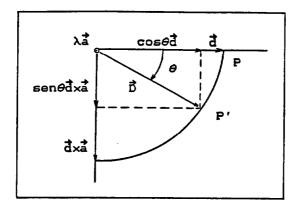


Fig. A.2. Plano de rotación perpendicular al eje de giro mostrando la relación entre los vectores \overrightarrow{d} , $\overrightarrow{d} \times \overrightarrow{a}$ y \overrightarrow{D} .

De esta figura se observa que el punto P', o su vector asociado \vec{c} , viene definido por la suma de los vectores $\lambda \vec{a}$ y \vec{D}

$$\vec{c} = \lambda \vec{a} + \vec{D}$$

$$\vec{D} = \cos\theta \vec{d} + \sin\theta \vec{d} \times \vec{a}$$
(A. 6)

Comparando (A.5) y (A.6) se obtiene que $R(\theta)\vec{b} = \vec{c}$, poniéndose de manifiesto la validez de (A.2). Desarrollando (A.2), la matriz de rotación $R(\theta)$ puede expresarse en función del ángulo de giro θ y de la orientación ϕ 1, ϕ 2, ϕ 3 del eje de giro. En forma compacta, los elementos Rij de esta matriz se obtienen como sigue:

$$R_{ij} = \begin{cases} \cos\phi_i \cos\phi_j (1 - \cos\theta) - \cos\phi_k \sin\theta (-1)^{i+j} & \text{Si i < j} \\ \cos^2\phi_i (1 - \cos\theta) + \cos\theta & \text{Si i = j} \\ \cos\phi_i \cos\phi_j (1 - \cos\theta) + \cos\phi_k \sin\theta (-1)^{i+j} & \text{Si i > j} \end{cases}$$

$$\cot 1 \le i, j, k \le 3 \quad y \quad k \ne i, j$$

A modo de ejemplo, y con la ayuda de (A.7), se pueden obtener fácilmente las matrices de rotación en torno de los ejes OX, OY, OZ para unos ángulos de giro 01, 02 y 03 respectivamente. Estas son:

$$R_{\mathbf{x}}(\theta_1) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & -\sin\theta_1 & \cos\theta_1 \end{bmatrix}$$
 (A.8.a)

$$R_{y}(\theta_{2}) = \begin{bmatrix} \cos\theta_{2} & 0 & -\sin\theta_{2} \\ 0 & 1 & 0 \\ \sin\theta_{2} & 0 & \cos\theta_{2} \end{bmatrix}$$
 (A.8.b)

$$R_{z}(\theta 3) = \begin{bmatrix} \cos \theta 3 & \sin \theta 3 & 0 \\ -\sin \theta 3 & \cos \theta 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
(A.8.c)

A.2 ROTACION QUE ALINEA DOS SISTEMAS DE REFERENCIA A DERECHAS Y CENTRADOS EN EL ORIGEN

A.2.1 Planteamiento del problema

Sea el sistema de ejes cartesianos OXYZ caracterizado por el triedro ortonormal $\{\vec{1}, \vec{j}, \vec{k}\}$ que lo define, y centrado en el origen de coordenadas. Sea otro referencial cualquiera, también a derechas, definido por el sistema ortonormal $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ centrado en el origen de coordenadas (Fig. A.3).

Conocidas las expresiones de \vec{v}_1 , \vec{v}_2 y \vec{v}_3 respecto de la base $\{\vec{i},\ \vec{j},\ \vec{k}\}$

$$\vec{v}_1 = v_{1x} \vec{i} + v_{1y} \vec{j} + v_{1z} \vec{k}$$

$$\vec{v}_2 = v_{2x} \vec{i} + v_{2y} \vec{j} + v_{2z} \vec{k}$$

$$\vec{v}_3 = v_{3x} \vec{i} + v_{3y} \vec{j} + v_{3z} \vec{k}$$
(A.9)

se desea determinar la rotación que aplicada al referencial $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ lo transforma en el $\{\vec{k}, \vec{i}, \vec{j}\}$, es decir alinea \vec{v}_1 con \vec{k} , \vec{v}_2 con \vec{i} y \vec{v}_3 con \vec{j} .

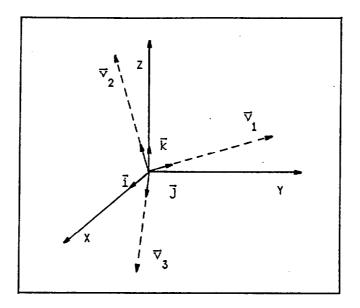


Fig. A.3. Rotación R que alinea el referencial $\{\overrightarrow{v}_1, \overrightarrow{v}_2, \overrightarrow{v}_3\}$ con el referencial absoluto $\{\overrightarrow{k}, \overrightarrow{1}, \overrightarrow{j}\}$.

A.2.2 Solución

Una posible solución al problema planteado consta de 3 pasos sucesivos agrupados en dos etapas. De manera resumida, esta solución quedaría como sigue:

- ETAPA 1 Rotar el triedro $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ hasta alinear \vec{v}_1 con \vec{k} . Dicha rotación la expresaremos como combinación de dos rotaciones sucesivas, una $R_x(\psi)$ en torno del eje OX y otra, $R_y(\vartheta)$, en torno del eje OY de acuerdo con los dos pasos siguientes:
 - 1. Determinar la matriz $R_x(\psi)$ tal que la proyección de \vec{v}_1 sobre el plano YZ, después de efectuada la rotación, esté alineada con el vector unitario \vec{k} .

 $R_{x}(\psi)$ $\vec{v}_{1}=\vec{v}_{1}$ ' proyección de \vec{v}_{1} ' sobre el plano YZ alineada con \vec{k} .

2. Determinar la matriz $R_y(\vartheta)$ tal que el vector resultante después de aplicarla a \vec{v}_1 coincida con \vec{k} .

Después de efectuar los dos pasos anteriores, el triedro $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ se habrá transformado en $\{\vec{v}_{1T}, \vec{v}_{2T}, \vec{v}_{3T}\}$ de manera que $\vec{v}_{1T} = \vec{k}$ y los vectores \vec{v}_{2T} y \vec{v}_{3T} estén contenidos en el plano XY (Fig. A.4).

$$\begin{bmatrix} R_{y}(\vartheta) & R_{x}(\psi) \end{bmatrix} \overrightarrow{v}_{1} = \overrightarrow{v}_{1T} = \overrightarrow{k}$$

$$\begin{bmatrix} R_{y}(\vartheta) & R_{x}(\psi) \end{bmatrix} \overrightarrow{v}_{2} = \overrightarrow{v}_{2T}$$

$$\begin{bmatrix} R_{y}(\vartheta) & R_{x}(\psi) \end{bmatrix} \overrightarrow{v}_{3} = \overrightarrow{v}_{3T}$$

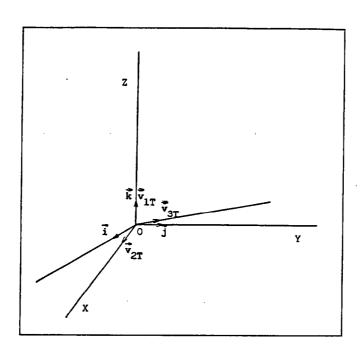


Fig. A.4. Posición del triedro inicial $\{\overrightarrow{v}1,\ \overrightarrow{v}2,\ \overrightarrow{v}3\}$ después de conseguir la alineación de $\overrightarrow{v}1$ con \overrightarrow{k} .

ETAPA 2 Al estar \vec{v}_{2T} y \vec{v}_{3T} contenidos en el plano XY sólo resta rotar el triedro $\{\vec{v}_{1T}, \vec{v}_{2T}, \vec{v}_{3T}\}$ en torno del eje OZ mediante $R_z(\varphi)$ de manera que, después de la transformación, \vec{v}_{2T} se alinee con \vec{i} y \vec{v}_{3T} con \vec{j} .

$$\frac{3.}{R_{z}}(\varphi) \overrightarrow{v}_{1T} = \overrightarrow{v}_{1T}$$

$$R_{z}(\varphi) \overrightarrow{v}_{2T} = \overrightarrow{1}$$

$$R_{z}(\varphi) \overrightarrow{v}_{3T} = \overrightarrow{j}$$

A.3 OBTENCION DE $R_{x}(\psi)$

Centremos nuestra atención en $\overrightarrow{v_1}$ exclusivamente. $\overrightarrow{v_1}$ viene expresado en la base $\{\overrightarrow{1}, \overrightarrow{j}, \overrightarrow{k}\}$ mediante 3 coordenadas (v_{1x}, v_{1y}, v_{1z}).

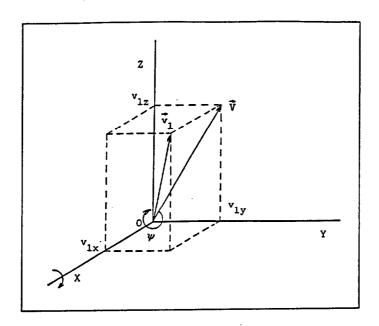


Fig. A.5. Rotación $\operatorname{Rx}(\psi)$ en torno del eje OX que aplicada a \overrightarrow{v} 1 alinea a su componente \overrightarrow{V} con \overrightarrow{k} .

Análogamente, \vec{v}_1 puede expresarse como suma de los vectores $\vec{v}_1 = \vec{v}_1 \cdot \vec{v}_1$ + \vec{v}_1 , siendo $\vec{v}_1 = \vec{v}_1 \cdot \vec{v}_2 \cdot \vec{v}_1$ + $\vec{v}_1 \cdot \vec{v}_2 \cdot \vec{v}_3$ la proyección de \vec{v}_1 sobre el plano YZ.

La rotación en torno al eje OX (especificado, en lo referente al sentido de giro, por el vector \vec{i}) $R_x(\psi)$, será tal que después de aplicarla alinee $R_x(\psi)\vec{V}$ con \vec{k} . De la Fig. A.5 se deduce que el ángulo de giro ψ vendrá caracterizado por las razones trigonométricas:

$$\cos \psi = \frac{v_{1z}}{V}$$

$$\sin \psi = -\frac{v_{1y}}{V}$$

$$\sin \psi = -\frac{v_{1y}}{V}$$
(A.10)

Teniendo en cuenta (A.8.a), la matriz $R_x(\psi)$ se expresa como:

$$R_{x}(\psi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & v_{1z}/V & -v_{1y}/V \\ 0 & v_{1y}/V & v_{1z}/V \end{bmatrix}$$
(A. 11)

Tras esta rotación, el vector \vec{v}_1 se habrá transformado en otro vector, $\vec{v}_1' = R_x(\psi)\vec{v}_1$, contenido en el plano XZ y de coordenadas $\vec{v}_1 = (v_{1x}, 0, V)$.

$$\vec{\mathbf{v}}_{1}' = \begin{pmatrix} \mathbf{v}_{1x}' \\ \mathbf{v}_{1y}' \\ \mathbf{v}_{1z}' \end{pmatrix} = \mathbf{R}_{\mathbf{x}}(\psi) \begin{pmatrix} \mathbf{v}_{1x} \\ \mathbf{0} \\ \mathbf{v} \end{pmatrix}$$
(A. 12)

Al ser \vec{v}_1 un vector unitario, también lo será su transformado según $Rx(\psi)$; es decir $|\vec{v}_1'|=1$.

A.4 OBTENCION DE Ry(v)

Téngase en cuenta que, por la forma de obtener $R_x(\psi)$, el vector \overrightarrow{v}_1 ' presenta un valor positivo en su coordenada z $(v_{1z}' = V)$. Determinemos a continuación el ángulo ϑ que caracteriza a la rotación en torno del eje OY y que hace coincidir $R_y(\vartheta)\overrightarrow{v}_1$ ' con \overrightarrow{k} . De la Fig. A.6 se observa que el ángulo ϑ ha de ser tal que sus razones trigonométricas vengan dadas por las relaciones:

$$\cos\theta = \frac{\mathbf{v}_{1z}'}{\left|\overrightarrow{\mathbf{v}}_{1}'\right|} = \mathbf{V}$$

$$\sin\theta = \frac{\mathbf{v}_{1x}'}{\left|\overrightarrow{\mathbf{v}}_{1}'\right|} = \mathbf{v}_{1x}$$
(A. 13)

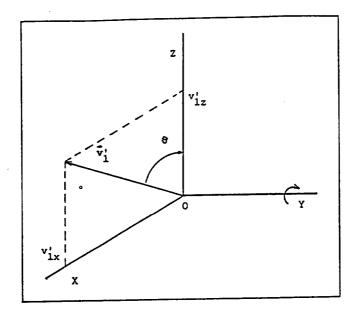


Fig. A.6. Rotación $Ry(\vartheta)$ en torno del eje OY que aplicada a V1' lo hace coincidir con K.

De acuerdo a (A.8.b) la matriz Ry(3) se expresa como:

$$R_{y}(\vartheta) = \begin{bmatrix} V & 0 & -v_{1x} \\ 0 & 1 & 0 \\ v_{1x} & 0 & V \end{bmatrix}$$
(A. 14)

Al finalizar la primera etapa y hacer coincidir \vec{v}_1 , después de las dos rotaciones anteriores, con el vector unitario \vec{k} , veamos cuales son las transformaciones experimentadas por el triedro inicial $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Combinando (A.11) y (A.14) y teniendo en cuenta las expresiones de \vec{v}_1 , \vec{v}_2 y \vec{v}_3 (A.9), se tiene:

$$\{\overrightarrow{v}_{1}, \overrightarrow{v}_{2}, \overrightarrow{v}_{3}\} \xrightarrow{R_{\mathbf{y}}(\mathfrak{G})R_{\mathbf{x}}(\psi)} \{\overrightarrow{v}_{1}, \overrightarrow{v}_{2T}, \overrightarrow{v}_{3T}\}$$

$$\overrightarrow{v}_{1T} = R_{\mathbf{y}}(\mathfrak{G})R_{\mathbf{x}}(\psi) \overrightarrow{v}_{1}; \qquad \overrightarrow{v}_{1T} = \begin{bmatrix} v_{1Tx} \\ v_{1Ty} \\ v_{1T} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$
(A. 15)

$$\vec{\mathbf{v}}_{2T} = \mathbf{R}_{\mathbf{y}}(\boldsymbol{\vartheta})\mathbf{R}_{\mathbf{x}}(\boldsymbol{\psi}) \quad \vec{\mathbf{v}}_{2}; \qquad \vec{\mathbf{v}}_{2T} = \begin{pmatrix} \mathbf{v}_{2Tx} \\ \mathbf{v}_{2Ty} \\ \mathbf{v}_{2Tz} \end{pmatrix} = \begin{pmatrix} \mathbf{v}_{2x} - \mathbf{v}_{1x}/\mathbf{v}(\mathbf{v}_{1y}\mathbf{v}_{2y} + \mathbf{v}_{1z}\mathbf{v}_{2z}) \\ 1/\mathbf{v}(\mathbf{v}_{1z}\mathbf{v}_{2y} - \mathbf{v}_{1y}\mathbf{v}_{2z}) \\ 0 \end{pmatrix}$$
(A. 16)

$$\vec{v}_{3T} = R_{y}(\vartheta)R_{x}(\psi) \vec{v}_{3}; \qquad \vec{v}_{3T} = \begin{pmatrix} v_{3Tx} \\ v_{3Ty} \\ v_{3Tz} \end{pmatrix} = \begin{pmatrix} Vv_{3x} - v_{1x}/V(v_{1y}v_{3y} + v_{1z}v_{3z}) \\ 1/V(v_{1z}v_{3y} - v_{1y}v_{3z}) \\ 0 \end{pmatrix}$$
(A. 17)

Conviene indicar que en aquellos casos en los que \overrightarrow{v}_1 esté contenido en el plano YZ no sería necesario efectuar ninguna rotación en torno del eje OX. Así, la matriz $R_x(\psi)$ podría igualarse a la matriz identidad. Igualmente, si \overrightarrow{v}_1 estuviese contenido en el plano YZ no sería preciso aplicar ninguna rotación en torno del eje OY y, por tanto, la matriz $Ry(\vartheta)$ coincidiría con la matriz identidad.

A.5 OBTENCION DE $Rz(\varphi)$

Una vez conocidas las posiciones del referencial $\{\vec{v}_{1T}, \vec{v}_{2T}, \vec{v}_{3T}\}$, y sabiendo que \vec{v}_{2T} y \vec{v}_{3T} están situados en el plano XY y que $\vec{v}_{1T} = \vec{k}$, estamos en condiciones de determinar la última rotación, $R_Z(\varphi)$, en torno del eje OZ que completa la solución al problema inicialmente planteado.

La caracterización del ángulo de giro φ se obtiene fácilmente con la ayuda de la Fig. A.7, en la que se muestra la posición de \vec{v}_{2T} y \vec{v}_{3T} , respecto de $\{\vec{i}, \vec{j}, \vec{k}\}$, vista desde el eje OZ y en sentido contrario a \vec{k} .

El ángulo φ será tal que :

$$\cos \varphi = \frac{\mathbf{v}_{2Tx}}{|\mathbf{v}_{2T}|} = \mathbf{v}_{2Tx}$$

$$\sin \varphi = \frac{\mathbf{v}_{2Ty}}{|\mathbf{v}_{2T}|} = \mathbf{v}_{2Ty}$$
(A. 18)

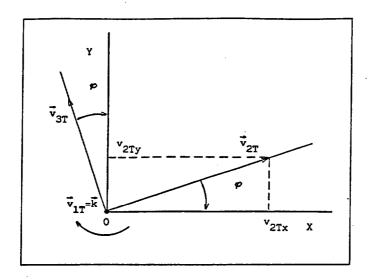


Fig. A.7. Rotación de valor φ en torno del eje OZ que alinea $\overrightarrow{V2T}$ con $\overrightarrow{1}$ y $\overrightarrow{V3T}$ con \overrightarrow{j} .

Así, y de acuerdo a (A.8.c), la expresión para $R_z(\phi)$ queda como:

$$R_{z}(\varphi) = \begin{bmatrix} v_{2Tx} & v_{2Ty} & 0 \\ -v_{2Ty} & v_{2Tx} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
(A. 19)

empleando aquí los valores vzxx y vzxy obtenidos en (A.16).

Por último, combinando en el orden correcto las matrices $R_x(\psi)$, $R_y(\vartheta)$ y $R_z(\varphi)$ se obtiene la matriz que transforma el referencial $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ en el $\{\vec{k}, \vec{1}, \vec{j}\}$ alineando entre sí los respectivos vectores.

Así, la matriz resultante $R(\psi,\vartheta,\varphi)$ será función de los ángulos ψ , ϑ , y φ asociados a las rotaciones individuales en torno a cada uno de los ejes OX, OY y OZ.

$$R(\psi, \vartheta, \varphi) = R_z(\varphi) R_y(\vartheta) R_x(\psi)$$
 (A. 20)

Evaluando el producto de esta matrices se obtiene la expresión para $R(\psi,\vartheta,\varphi)$ en función de las razones trigonométricas de los ángulos de rotación

$$R(\psi,\vartheta,\varphi) = \begin{bmatrix} \cos\varphi\cos\vartheta & \cos\varphi\sin\psi\sin\vartheta + \sin\varphi\cos\psi & -\cos\varphi\cos\psi\sin\vartheta + \sin\varphi\sin\psi \\ -\sin\varphi\cos\vartheta & -\sin\varphi\sin\psi\sin\vartheta + \cos\varphi\cos\psi & \sin\varphi\cos\psi\sin\vartheta + \cos\varphi\sin\psi \\ \sin\vartheta & -\sin\varphi\cos\vartheta & \cos\psi\cos\vartheta \end{bmatrix}$$

(A.21)

Más explícitamente, $R(\psi,\vartheta,\varphi)$ puede expresarse en términos de las coordenadas de \vec{v}_1 , \vec{v}_2 y \vec{v}_3 (A.9). Las razones trigonométricas de ψ , ϑ y φ pueden simplificarse considerablemente si se tiene en cuenta que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ forman un triedro ortonormal a derechas, es decir :

$$\vec{\mathbf{v}}_{\mathbf{n}} \cdot \vec{\mathbf{v}}_{\mathbf{n}} = \begin{cases} 0 & m \neq n \\ 1 & m = n \end{cases} \quad m, n : 1, 2, 3$$

$$\vec{\mathbf{v}}_{\mathbf{1}} \times \vec{\mathbf{v}}_{2} = \vec{\mathbf{v}}_{3}; \quad \vec{\mathbf{v}}_{2} \times \vec{\mathbf{v}}_{3} = \vec{\mathbf{v}}_{1}; \quad \vec{\mathbf{v}}_{3} \times \vec{\mathbf{v}}_{1} = \vec{\mathbf{v}}_{2};$$

Así, igualando a cero el producto escalar de \vec{v}_1 y \vec{v}_2 (\vec{v}_1 . \vec{v}_2 = 0), el valor de v2Tx en (A.16) puede reducirse a :

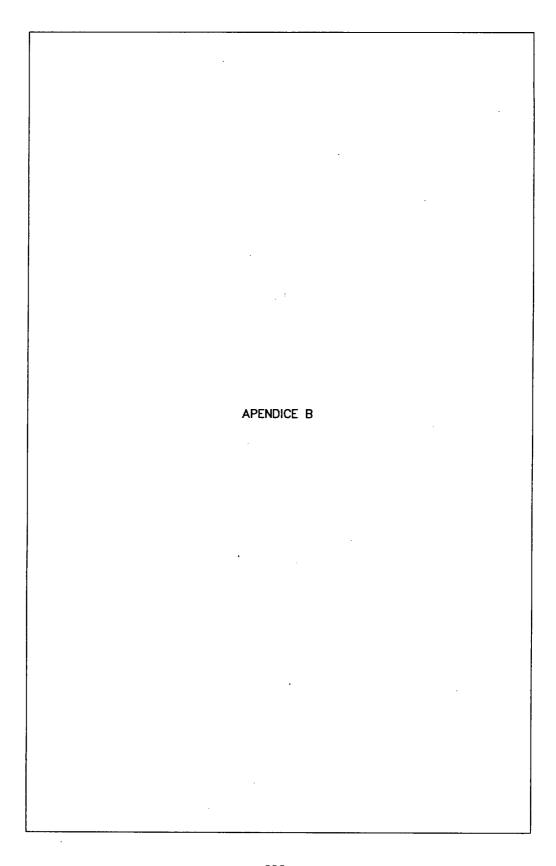
$$v_{2Tx} = \frac{v_{2x}}{V}$$
; $V = \sqrt{v_{1y}^2 + v_{1z}^2}$ (A. 22)

Igualmente, a partir de la expresión $\vec{v}_1 \times \vec{v}_2 = \vec{v}_3$, la coordenada vzty de (A.16) puede expresarse como :

$$V_{2Ty} = -\frac{V_{3x}}{V} \tag{A.23}$$

Con estas consideraciones se puede establecer el siguiente cuadro para las razones trigonométricas que aparecen en los elementos de la matriz de rotación $R(\psi,\vartheta,\varphi)$ de (A.21)

SI V = 0	SI V \neq 0 (v ₁ y = v ₁ z = 0)
cosψ = V1z/V	cos ψ = 1
$sen\psi = - v_1y/V$	$sen\psi = 0$
cosϑ = V	cosϑ = 0
senð = v1x	senϑ = V1x
cosφ = v2x/V	$\cos \varphi = -v_{1x}v_{2z}$
$sen \varphi = - v_{3x}/V$	$sen \varphi = v2y$



RELACION ENTRE LAS COORDENADAS DE UN PUNTO RESPECTO DE DOS BASES DE DIFERENTE DIMENSION.

Sea un espacio de dimensión 3 y la base ortonormal $\{\vec{i}, \vec{j}, \vec{k}\}$ con origen en el punto 0 (Fig. B.1). El lugar geométrico de todos los puntos (x,y,z) que satisfacen la ecuación:

$$\Pi \equiv ax + by + cz + d = 0; \quad a, b, c, d \in \mathbb{R}$$
(B.1)

definen un plano cuya normal viene dada por el vector de coordenadas (a,b,c).

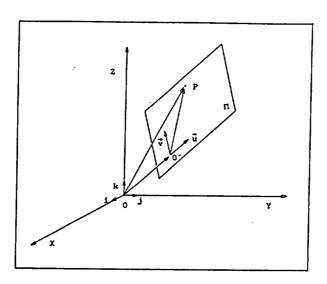


Fig. B.1. Descripción de un mismo punto P en dos bases distintas de dimensiones 2 y 3.

El plano Π es una variedad afín de dimensión 2; es decir, dado un punto O' y dos vectores $\{\overrightarrow{u}, \overrightarrow{v}\}$ linealmente independientes contenidos en el plano, cualquier otro punto P de dicho plano podrá obtenerse a partir de ellos en los siguientes términos:

$$\overline{O'P} = \lambda_{u} \dot{\overrightarrow{u}} + \lambda_{v} \dot{\overrightarrow{v}} \qquad \qquad \lambda_{u}, \ \lambda_{v} \in \mathbb{R}$$
(B.2)

siendo (λ_u, λ_v) las coordenadas del punto P del plano respecto de la base $\{\overrightarrow{u}, \overrightarrow{v}\}$ con origen en 0'.

Por otra parte, el mismo punto P es un punto del espacio afín de dimensión 3, y junto al punto O define un vector \overrightarrow{OP} que puede expresarse como combinación lineal de los vectores de la base $\{\vec{1}, \vec{j}, \vec{k}\}$.

$$\overrightarrow{OP} = \alpha_{x} \overrightarrow{i} + \alpha_{y} \overrightarrow{j} + \alpha_{z} \overrightarrow{k}$$
 (B.3)

Análogamente:

$$\overrightarrow{OO}' = o_{x}'\overrightarrow{1} + o_{y}'\overrightarrow{j} + o_{z}'\overrightarrow{k}$$
(B.4)

De la Fig. B.1 puede establecerse la siguiente relación vectorial:

$$\overrightarrow{OP} = \overrightarrow{OO'} + \overrightarrow{O'P}$$
 (B.5.1)

$$\vec{O} \cdot \vec{P} = \vec{OP} - \vec{OO}$$
 (B. 5. 2)

Consideremos, por sencillez, que la base $\{\vec{u}, \vec{v}\}$ del plano es también ortonormal, es decir:

$$\vec{\mathbf{u}} \cdot \vec{\mathbf{v}} = 0$$

$$|\vec{\mathbf{u}}| = |\vec{\mathbf{v}}| = 1$$
(B.6)

Veamos a continuación la relación que existe entre las coordenadas del punto P expresadas en la base del plano $\{\vec{u}, \vec{v}\}$, y las coordenadas de P expresadas en términos de la base $\{\vec{i}, \vec{j}, \vec{k}\}$. Los casos más interesantes y que mejor sirven a nuestros propósitos son:

Caso 1 Conocidas las coordenadas de O' y P respecto de la base $\{\vec{1}, \vec{j}, \vec{k}\}$ y conocida la expresión de los vectores \vec{u} y \vec{v} respecto de la anterior base, se desea determinar las coordenadas de P en la base $\{\vec{u}, \vec{v}\}$.

De (B.2), (B.5.2) y (B.6) se tiene:

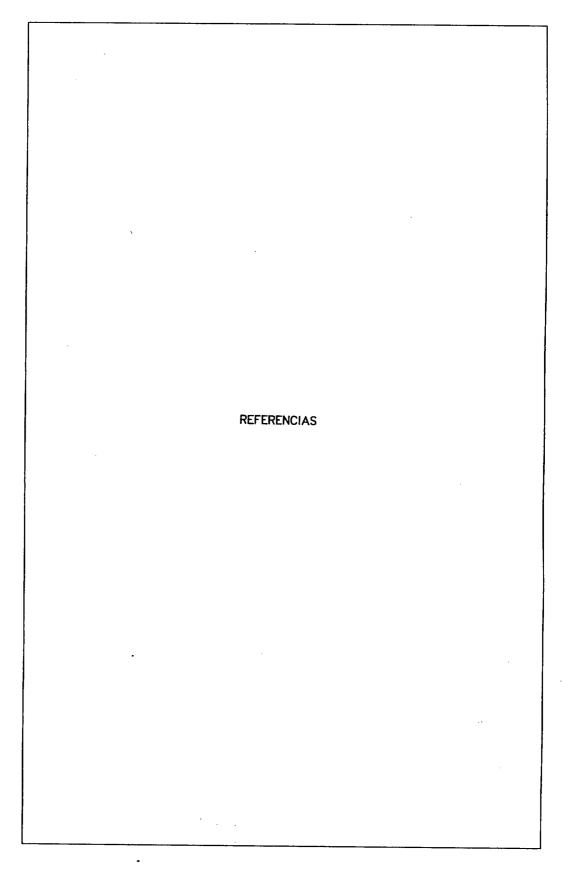
$$\overrightarrow{O'P} = \lambda_{u} \overrightarrow{u} + \lambda_{v} \overrightarrow{v} = (\alpha_{x} \overrightarrow{1} + \alpha_{y} \overrightarrow{J} + \alpha_{z} \overrightarrow{k}) - (o_{x} \overrightarrow{1} + o_{y} \overrightarrow{J} + o_{z} \overrightarrow{k})$$

$$\lambda_{u} = \overrightarrow{O'P} \cdot \overrightarrow{u}$$

$$\lambda_{v} = \overrightarrow{O'P} \cdot \overrightarrow{v}$$

Caso 2 Este es el caso contrario al anterior. Conocidas las coordenadas de P en la base del plano $\{\vec{u}, \vec{v}\}$ y conocida la expresión de \vec{u} , \vec{v} y $\vec{00}$ respecto de $\{\vec{i}, \vec{j}, \vec{k}\}$, se desea determinar las coordenadas de P en la base $\{\vec{i}, \vec{j}, \vec{k}\}$.

Esto se obtiene mediante la aplicación directa de la expresión (B.5.1).



- [ABU-84] Abu-Mostafa, Y. Psaltis, D., "Recognitive aspects of moment invariants", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-6, no. 6, pp. 698-706, Nov. 1984.
- [ABU-85] Abu-Mostafa, Y. Psaltis, D., "Image normalization by complex moments", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-7, no. 1, pp. 46-55, Jan. 1985.
- [AGGA-81a]Aggarwal, J.K., Davis, L.S., Martin, W.N., "Correspondence processes in dynamic scene analysis", <u>Proc.</u> <u>IEEE</u>, vol 69, no. 5, pp. 562-572, May. 1981.
- [AGGA-81b]Aggarwal, J.K., Martin, W.N., "Analyzing dynamic scenes containing multiple moving objects", en "Image Sequence Analysis", Huang, T.S. (Ed.), Springer-Verlag, 1981, pp. 355-380.
- [AGGA-81c]Aggarwal, J.K., Davis, L.S., Martin, W.N., "Survey: representation methods for three-dimensional objects", en "Progress in Pattern Recognition", L. Kanal and A. Rosenfeld (Eds.), Amsterdam, The Netherlands: North Holland, vol. 1, pp. 337-391, 1981.
- [ALTS-81] Altschuler, M.D., Altschuler, B.R., Taboada, J., "Laser electro-optic system for rapid three-dimensional (3-D) topographic mapping of surfaces", Optical Engineering, vol. 20, no. 6, pp. 953-961, Nov-Dec. 1981.
- [AMAT-83] Amat, J., Basáñez, L., Torras, C., "Sistemas de visión tridimensional en los robots industriales", <u>Mundo Electrónico</u>, n°. 130, pp. 71-78, Jun. 1983.
- [ANGU-85] Angulo, J.M., Iñigo, R., "Visión Artificial por Computador", Paraninfo, 1985.
- [ARFK-70] Arfken, G., "Mathematical Methods for Physicits", Academic Press., 1970.
- [ASAD-83] Asada, M., Tsuji, S., "Representation of three-dimensional motion in dynamic scenes", <u>Computer Vision</u>, <u>Graphics and</u>

- Image Processing, vol. 21, pp. 118-144, 1983.
- [ASAD-84] Asada , M., Yachida, M., Tsuji, S., "Analysis of three-dimensional motions in blocks world", <u>Pattern Recognition</u>, vol. 17, no. 1, pp. 57-71, 1984.
- [ASHD-86] Ashdowns, I.E., "Curve fitting with cubic splines", <u>Dr. Dobb's Journal</u>, pp. 24-83, Sep. 1986.
- [BADL-78] Badler, N., Bjacsy, R., "Three-dimensional representations for computer graphics and computer vision", <u>Computer Graphics</u>, vol. 12, pp. 153-160, 1978.
- [BALL-82] Ballard, D.H., Brown, C.M., "Computer Vision", Prentice Hall, Englewood Cliffs, NJ, 1982.
- [BARN-80] Barnard, S.T., Thompson, W.B., "Disparity analysis of images", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-2, no. 4, pp. 333-340, Jul. 1980.
- [BARN-82] Barnard, S.T., Fisher, M.A., "Computational stereo", Computing Surveys, vol. 14, no. 4, pp. 553-572. Dec. 1982.
- [BARN-83] Barnard, S.T., "Interpreting Perspective Images", <u>Artificial</u>
 <u>Intelligence</u>, vol. 21, pp. 435-462, 1982.
- [BAST-86] Bastuscheck, C.M., Schwartz, J.T., "Experimental implementation of a ratio image depth sensor", en "Techniques for 3-D Machine Perception", Rosenfeld A., (Ed.), North-Holland, 1986, pp. 1-12.
- [BEER-83] Beer, F., Russell, E., "Mecánica Vectorial para Ingenieros", McGraw-Hill, 1983.
- [BESL-85] Besl, P.J., Jain, R.C., "Three-dimensional object recognition", <u>Computing Surveys</u>, vol. 17, no. 1, pp. 75-145, Mar. 1985.
- [BHAN-82] Bhanu, B., "Surface representation and shape matching of 3-D objects", <u>Proc. PRIP-82</u>, <u>IEEE Computer Society on Pattern Recognition and Image Processing</u>, pp. 349-354, Las Vegas,

- Recognition and Image Processing, pp. 349-354, Las Vegas, 1982.
- [BHAN-84] Bhanu, B. "Representation and shape matching of 3-D objects", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-6, no. 3, pp. 340-351, May. 1984.
- [BINF-81] Binford, T.O., "Inferring surfaces from images", <u>Artificial</u>
 <u>Intelligence</u>, vol. 17, pp. 205-244, 1981.
- [BLOS-87] Blostein, S.D., Huang, T.S., "Error Analysis in stereo determination of 3-D point positions", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-9, no. 6, pp. 752-765, Nov. 1987.
- [BORN-75] Born, M., Wolf, E., "Principles of Optics", Pergamon Press, Oxford, 1975.
- [BOYC-83] Boyce, J.F., Hossack, W.J., "Moment invariants for pattern recognition", <u>Pattern Recogn.</u> <u>Letters</u>, vol. 1, pp. 451-456, Jul. 1983.
- [BOYE-88] Boyer, K.L., Kak, A.C., "Structural stereopsis for 3-D vision", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-10, no. 2, pp. 144-166, Mar. 1988.
- [BOYL-88] Boyle, R.D., Thomas, R.C., "Computer Vision. A First Course", Blackwell Scientific Publications, 1988.
- [BRAD-84] Brady, M., Yuille, A., "An extremun principle for shape from contour", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-6, no. 3, pp. 288-301, May. 1984.
- [BROW-87] Browse, R.A., "Feature-based tactile object recognition",

 IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-9, no.
 6, pp. 79-786, Nov. 1987.
- [CANT-82a]Cantón, M., "La Teoría de Momentos en el Análisis de Formas:

 Formulación Tensorial, Invarianzas y Aplicaciones", <u>Tesis</u>

 <u>Doctoral</u>, <u>Univer.</u> <u>La Laguna</u>, Oct. 1982.

- [CANT-82b]Cantón, M., Rubio, F., Moreno, R., "La teoría de momentos generalizados y su aplicación en problemas biocibernéticos", Proc. 1° Reunión Nac. Biocib., Real Academia de Ciencias, pp. 151-178, Madrid, 1982.
- [CANT-84] Cantón, M., Moreno Díaz, R., "Normalized moments for 3D characterization of visual patterns", <u>Int. Conf. on System Research Informatics</u>, Baden-Baden (RFA), Aug. 1-5, 1984.
- [CANT-86] Cantón, M., Gálvez, J.M., Rubio, E., "Normalización de momentos visuales 3-D frente a transformaciones afines: aplicaciones a formas sintetizadas", <u>Informática</u> Y <u>Automática</u>, vol. XIX, n°.2, pp. 26-31, 1986.
- [CARL-82] Carlson, W.E., "Techniques for the Generation of Three Dimensional Data for Use in Complex Image Synthesis", Ph. D.

 Dissertation, Dept. of Computer and Information Science,
 Ohio State University, Sep. 1982.
- [CHAR-85] Charniak, E., McDermott, D., "Introduction to Artificial Intelligence", Addison-Wesley Publishing Company, 1986.
- [CHAZ-80] Chazelle, B., Dobkin, D.P., "Detection is easer than computation", Proc. Twelfth Annual ACM Symposium On Theory Of Computing, Los Angeles, CA, pp. 146-153, Apr. 28-30, 1980.
- [CHEN-73] Chen, C.H., "Statistical Pattern Recognition", Hayden Book Comp., Inc., New Jersey. 1973.
- [CHEN-86a]Cheng, K., Idesawa, M., "A simplified method of data conversion from contour line surface model to mesh surface model", Proc.8th Int. Conf. Pattern Recognition, pp.582-585, Paris, Oct. 1986.
- [CHEN-86b]Cheng, K., Idesawa, M., "A simplified interpolation and conversion method of contour surface model to mesh surface model", <u>Journal of Robotics Systems</u>, vol. 3, no. 3, pp.

237-247, 1986.

- [CHIE-84] Chien, C.H., Aggarwal, J.K., "A volume/surface octree representation", <u>Proc. 7th Int. Conf. Pattern Recognition</u>, pp. 817-820, 1984.
- [CHRI-78] Christiansen, H.N., Sederberg, T., "Conversion of complex contour line definition into polygonal element mosaics", Computer Graphics, vol. 12, pp. 187-192, 1978.
- [CHRI-75] Christofides, N., "Graph Theory: An Algorithmic Approach",
 Academic Press, New York, 1975.
- [COOK-80] Cook, P.N. et al., "An algorithm for volume estimation based on polyhedral approximation", <u>IEEE Transactions on Biomedical Engineering</u>, vol. BME-27, no. 9, pp. 493-500, Sep. 1980.
- [COOK-81a]Cook, P.N., Batnitzky, S., Lee, K.R., Cook, L.T., Fritz, S.L., Dwyer, S.J.III, Charlson, E.J., "Three-dimensional reconstruction from serial sections for medical applications", Proc. 14th Hawaii Int. Conf. on System Sciences, vol. 2, pp. 358-389, 1981.
- [COOK-81b]Cook, P.N., "A study of three-dimensional reconstruction algorithms", <u>Automedica</u>, vol. 4, pp. 3-12, 1981.
- [COOK-81c]Cook, L.T., Cook, P.N., "Volume and surface area estimations", Automedica, vol. 4, pp. 13-23, 1981.
- [COUR-53] Courant, R., Hilbert, D., "Methods of Mathematical Physics. vol. I", Interscience, 1953.
- [DAVI-83] Davis, L.S., Janos, L., Dunn S.M., "Efficient recovery of shape from texture", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-5, no. 5, pp. 485-492, Sep. 1983.
- [DE FL-87]De Floriani, L., "Surface representations based on triangular grids", <u>The Visual Computer</u>, vol. 3, pp. 27-50, 1987.

- [DUDA-73] Duda, R.O., Hart, P.E., "Pattern Classification And Scene Analysis", John Wiley & Sons, New York. 1973.
- [DUDA-79a]Duda, R.O., Nilsson, N.J., Raphael, B., "State of technology in artificial intelligence", en "Research Directions in Software Technology", Wegner, P. (Ed.), MIT Press, 1979, pp. 729-749.
- [DUDA-79b]Duda, R.O., Nitzan, D., Barret, P., "Use of range and reflectance data to find planar surface regions", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-1, no. 3, pp. 259-272, Jul. 1979.
- [DUDA-77] Dudani, S.A., Breeding, K.J., McGhee, R.B., "Aircraft indentification by moment invariants", <u>IEEE Trans.</u>

 <u>Comput.</u>, vol. C-26, no. 1, pp. 39-45, Jan. 1977.
- [FABE-88] Faber, T.L., Stokely, E.M., "Orientation of 3-D structures in medical images", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-10, no. 5, pp. 626-633, Sep. 1988.
- [FAIR-88] Fairhurst, M.C., "Computer Vision for Robotic Systems. An Introduction", Prentice-Hall Inc., 1988.
- [FARI-88] Farin, G., "Curves and Surfaces for Computer Aided Geometric Design. A Practical Guide", Academic Press, Inc., 1988.
- [FAUG-86] Faugeras, O.D., Herbert, M., "The representation, recognition, and positioning of 3-D shapes from range data", en "Techniques for 3-D Machine Perception", Rosenfeld A., (Ed.), North-Holland, 1986, pp. 13-51.
- [FERN-87] Fernández, G., "Sistemas Expertos: Fundamentos y Aplicaciones", Notas del VII Curso de Automática en la Industria, (AEIA), Denia, 1-5 Junio, 1987.
- [FERR-86] Ferraté, G., y otros, "Robótica Industrial", Marcombo Boixareu Edt., 1986.

- [FERR-89] Ferrie, F.P., Levine, M.D., "Where and why local shading analysis works", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-11, no. 2, pp. 198-206, Feb. 1989.
- [FISH-36] Fisher, R.A., "The use of multiple measurements in taxonomic problems", Ann. Eugenics, vol. 7, pp. 179-188, 1936.
- [FOLE-84] Foley, J.D., Van Dan, A., "Fundamentals Of Interactive Computer Graphics", Addison-Wesley Publish. Comp. 1984.
- [FREE-88] Freeman, H., (Edt.) "Machine Vision: Algorithms, Architectures, and Systems", Academic Press, Inc., 1988.
- [FROB-83a]Frobin, W., Hierholzer, E., "Automatic measurement of body surfaces using rasterstereography. Part I: Image scan and control point measurement", Photogrammetric Engineering <a href="https://example.com/Ann.2016/1983. Remote Sensing, vol. 49, no. 3, pp. 377-384, Mar. 1983.
- [FROB-83b]Frobin, W., Hierholzer, E., "Automatic measurement of body surfaces using rasterstereography. Part II: Analysis of the rasterstereographic line pattern and three-dimensional surface reconstruction", Photogrammetric Engineering and Remote Sensing, vol. 49, no. 10, pp. 1443-1452, Oct. 1983.
- [FU-82] Fu, K.S., "Syntactic Pattern Recognition and Applications", Prentice-Hall, Englewood Cliffs, N.J., 1982.
- [FU-84] Fu, K.S., Rosenfeld, A., "Pattern recognition and computer vision", <u>IEEE Computer</u>, vol. 17, no. 10, pp. 274-282, 1984.
- [FUCH-77] Fuchs, H., Kedem, Z.M., Uselton, S.P., "Optimal surface reconstruction from planar contours", <u>Communications of the ACM</u>, vol. 20, no. 10, pp. 693-702, Oct. 1977.
- [GALV-89] Gálvez, J.M., Cantón, M., Rubio, E., Rubio, F., "The use of 3D moments for three-dimensional object orientation", Proc.

 2nd. Inter. Symposium On Systems Research Informatics And Cybernetics, Baden-Baden, West Germany, Aug. 11-16, 1989.
- [GAST-84] Gaston, P.C., Lozano-Pérez, T., "Tactile recognition and

- localization using object models: the case of polyhedral on a plane", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-6, no. 3, pp. 257-266, May. 1984.
- [GOLD-72] Goldstein, H., "Mecánica Clásica", Ed. Aguilar, 1972.
- [GRIM-81] Grimson, W.E.L. "From Images to Surfaces. A Computational Study of the Human Early Visual System", MIT Press, Cambridge, MA., 1981.
- [GRIM-86a]Grimson, W.E.L., Lozano-Pérez, T., "Model-based recognition and localization from sparse range data", en "Techniques for 3-D Machine Perception", Rosenfeld A., (Ed.), North-Holland, 1986, pp. 113-148.
- [GRIM-86b]Grimson, W.E.L., "Computing stereopsis using feature point contour matching", en "Techniques for 3-D Machine Perception", Rosenfeld A., (Ed.), North-Holland, 1986, pp. 75-111.
- [HABE-82] Haber, R., Abel, J.F., "Discrete transfinite mappings for the description and meshing of three-dimensional surfaces using interactive computer graphics", <u>Int. Journal for Numer. Meth. in Enginee.</u>, vol. 18, pp. 41-66, 1982.
- [HARA-80] Haralick, R.M., "Using perspective transformations in scene analysis", <u>Comput. Graphics Image Processing</u>, vol. 13, pp. 191-221, 1980.
- [HEND-83] Henderson, T.C., "Efficient 3-D object representations for industrial vision systems", <u>IEEE Trans. Pattern Anal.</u>

 <u>Machine Intell.</u>, vol. PAMI-5, no. 6, pp. 609-618, Nov. 1983.
- [HERN-87] Hernández, F.M., "Reconocimiento de Formas Tridimensionales mediante Caracterización Global de Vistas. Desarrollos Teóricos y Aplicaciones", Tesis Doctoral, Univ. Politécnica de Canarias, Jul. 1987.
- [HIER-82] Hierholzer, E., Frobin, W., "Automatic measurement of body surfaces using rasterstereography", <u>26th</u> <u>Annual</u>

- <u>International Symposium of the Inter. Soc. for Optical Engineering (SPIE)</u>, California, 24-27 Aug. 1982.
- [HORN-75] Horn, B.P.K., "Obtaining shape from shading information", en "The Psychology of Computer Vision", Winston P.H. (Ed.), McGraw-Hill Computer Science Series, 1975, pp. 115-155.
- [HORN-77] Horn, B.P.K., "Understanding image intensities", <u>Artificial</u>

 <u>Intelligence</u>, vol. 8, no. 2, pp. 201-231, 1977.
- [HU-61] Hu, M.K., "Pattern recognition by moment invariants", <u>Proc.</u>

 IRE., vol. 49, pp. 1428, Sep. 1961.
- [HU-62] Hu, M.K., "Visual pattern recognition by moment invariants",

 IRE Trans. Inform. Theory, vol. IT-8, pp. 179-187, Feb. 1962.
- [HUAN-81] Huang, T.S., "Image Sequence Analysis", Springer-Verlang, 1981.
- [IDES-86] Idesawa, M., Cheng. K., "3-D model reconstruction and processing for CAE", <u>Proc.</u> 8th <u>Int.</u> <u>Conf.</u> <u>Pattern</u> <u>Recognition</u>, pp. 220-225, París, Oct. 1986.
- [IKEU-81a] Ikeuchi, K., Horn, B.K.P., "Numerical shape from shading and occluding boundaries", <u>Artificial Intelligence</u>, vol. 17, pp. 141-184, 1981.
- [IKEU-81b] Ikeuchi, K., "Determining surface orientations of specular surfaces by using the photometric stereo method", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-3, no. 6, pp. 661-669, Nov. 1981.
- [IKEU-84] Ikeuchi, K., "Shape from regular patterns", Artificial Intelligence, vol. 22, pp. 49-75, 1984.
- [JARV-83] Jarvis, R.A., "A perspective on range finding techniques for computer vision", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-5, no. 2, pp. 122-139, Mar. 1983.

- [KAK-86] Kak, A.C., Boyer, K.L., Safranek, R.J., Yang, H.S., "Knowledge-based stereo and structural light for 3-D robot vision", en "Techniques for 3-D Machine Perception", Rosenfeld A. (Ed.), North-Holland, 1986, pp. 185-218.
- [KANA-83] Kanade, T., "Geometrical aspects of interpreting images as a three-dimensional scene", <u>Proc. IEEE</u>, vol. 71, no. 7, pp. 789-802, Jul. 1983.
- [KEPP-75] Keppel, E., "Approximating complex surfaces by triangulation of contour line", <u>IBM J. Res. Develop.</u>, vol. 19, pp. 2-11, Jan. 1975.
- [KIM-86] Kim, Y.C., Aggarwal, J.K. "Rectangular parallelepiped coding: a volumetric representation of three-dimensional objects", <u>IEEE Journal of Robotics and Automation</u>, vol. RA-2, no. 3, pp. 127-134, Sep. 1986.
- [LEVI-72] Levinthal, C., Ware, R., "Three dimensional reconstruction from serial sections", <u>Nature</u>, vol. 236, pp. 207-210, Mar. 1972.
- [LONG-81] Longuet, H.C., Higgins, "A computer algorithm for reconstructing a scene from two projections", Nature, vol. 293, pp. 133-135, Sep. 1981.
- [LOPE-84] López de Mántaras, R., "Representación y adquisición del conocimiento", en "Inteligencia Artificial. Introducción y Situación en España" FUNDESCO, 1984. pp. 59-69.
- [LUH-85] Luh, J.Y.S., Klaasen, J.A., 'A three-dimensional vision by off shell system with multi-cameras", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-7, no. 1, pp. 35-45, Jan. 1985.
- [LUO-88] Luo, R.C., Loh, H.H., "Tactile array sensor for object identification using complex moments", <u>Journal of Robotic Systems</u>, vol. 5, no. 1, pp.1-12, 1988.
- [MAIT-79] Maitra, S., "Moment invariants", Proc. IEEE, vol. 67, no.4

- pp. 697-699, Apr. 1979.
- [MARR-76a]Marr, D., "Artificial intelligence a personal view", Artificial Intelligence, vol. 9, pp. 37-48, 1976.
- [MARR-76b]Marr, D., "Early processing of visual information", <u>Proc. R.</u>
 <u>Soc. London</u>, vol. 275, pp. 483-534, 1976.
- [MARR-77] Marr, D., "Analysis of ocluding contour", <u>Proc. R. Soc.</u>
 <u>London</u>, vol. B.197, pp. 441-475, 1977.
- [MARR-78a]Marr, D., Nishihara, H.K., "Representation and recognition of the spatial organization of three-dimensional shapes", Proc. R. Soc. London, vol. B-200, pp. 269-294, 1978.
- [MARR-78b]Marr, D., Palm, G., Poggio, T., "Analysis of a cooperative stereo algorithm", <u>Biological Cybernetics</u>, vol. 28, pp. 223-239, 1978.
- [MARR-79] Marr, D., Poggio, T., "A theory of human stereo vision", <u>Proc. R. Soc. London</u>, vol. B-204, pp. 301-328, 1979.
- [MARR-82] Marr, D., "Vision", W.H. Freeman and Com., San Francisco, 1982.
- [MART-81a]Martin, W.N., Aggarwal, J.K., "Occluding contours in dynamic scenes", <u>Proc. IEEE Conference on Pattern Recognition and Image Processing</u>, Dallas, TX, pp. 189-192, Aug. 2-5, 1981.
- [MART-81b]Martin, W.N., Aggarwal, J.K., "Analyzing dynamic scenes",

 <u>Lab. for Image and Signal Anal.</u>, Univ. Texas, Austin, Rep.
 TR-81-5, Dec. 1981.
- [MART-81c]Martin, W.N., Aggarwal, J.K., "Occlusion in dynamic scenes analysis", en "Digital Image Processing", Simson, J.C. and Haralick R.M., (Eds.), D. Reidel Publishing Company, 1981, pp. 579-590.
- [MART-82a]Martin, W.N., Aggarwal, J.K., "Volumetric descriptions from dynamic scenes", <u>Pattern Recognition Letters</u>, vol. 1, no. 2,

- pp. 107-113, 1982.
- [MART-82b]Martin, W.N., Aggarwal, J.K., "Dynamic scenes and object descriptions", <u>Proc. IEEE International Conference on Acustics, Speech and Signal Processing</u>, Paris, pp. 859-862, May 3-5, 1982.
- [MART-83] Martin, W.N., Aggarwal, J.K., "Volumetric descriptions of objects from multiple views", <u>IEEE Trans. Pattern Anal.</u> <u>Machine Intell.</u>, vol. PAMI-5, no. 2, pp. 150-158, Mar. 1983.
- [MART-88] Martin, W.N., Aggarwal, J.K. (Eds.), "Motion Understanding.

 Robot and Human Vision", Kluwer Academic Publishers, 1988.
- [MATH-87] Mather, P.M., "Computer Processing of Remotely-Sensed Images", John-Wiley & Sons, 1987.
- [MIKH-78] Mikhailov, V.M., "Representation of surfaces in CADS", Cybernetics, pp. 683-687, 1978.
- [MIYA-84] Miyake, T., Doi, J., "Three-dimensional shape approximation system with a polyhedron", <u>ISP Japan</u>, vol. 25, pp. 745-754, Sep. 1984.
- [MUND-88] Mundy, J.L., "Industrial machine vision Is it practical?", en "Machine Vision. Algorithms, Architectures, and Systems", Freeman H. (Ed.), Academic Press, Inc., 1988, pp. 109-148.
- [NAG-82] "NAG FORTRAN LIBRARY Manual", Mark 10. Dec. 1982.
- [NAGE-81a]Nagel, H.H., Neuman, B., "On 3D reconstruction from two perspective views", <u>Proc. IJCAI-81, International Joint Conference on Artificial Intelligence</u>, pp. 661-294, 1981.
- [NAGE-81b]Nagel, H.H., "On the derivation of 3D rigid point configurations from image sequences", <u>Proc. IEEE Conference on Pattern Recognition and Image Processing</u>, Dallas, TX, pp. 103-108, Aug. 3-5, 1981.
- [NAGE-83] Nagel, H.H., "Overview on image sequence analysis", en

- "Image Sequence Processing and Dynamic Scene Analysis", NATO Advanced Study Institutes Series F, vol. 2, Springer-Verlag, Berlin, 1983.
- [NEUM-79] Neuman, W.E., Sproull, R.F., "Principles of Interactive Computer Graphics", 2° Edt., N.Y.: McGraw-Hill. 1979.
- [NEUM-80] Neuman, B. "Motion analysis of image sequences for object grouping and reconstruction", Proc. ICPR-80, International Recognition, pp. 1262-1265, 1980.
- [NEVA-76] Nevatia, R., "Depth measurement by motion stereo", <u>Computer</u>

 <u>Graphics and Image Processing</u>, vol. 5, pp. 203-214, 1976.
- [NEVA-82] Nevatia, R., "Machine Perception", Prentice-Hall Inc., 1982.
- [NISH-81] Nishihara, H.K., "Intensity, visible-surface and volumetric representations", <u>Artificial Intelligence</u>, vol. 17, pp. 265-284, 1981.
- [NITZ-88] Nitzan, D., "Three-dimensional vision structure for robot applications", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-10, no. 3, pp. 291-309, May 1988.
- [NOBO-88] Noborio, N., Fukuda, S., Arimoto, S., "Construction of the octree approximating three-dimensional objects by using multiple views", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-10, no. 6, pp. 769-782, Nov. 1988.
- [O'RO-79] O'Rourke, J., Badler, N., "Descomposition of three-dimensional objects into spheres", <u>IEEE Trans. Pattern Anal.</u>

 <u>Machine Intell.</u>, vol. PAMI-1, no.3, pp. 295-305, Jul. 1979.
- [OSHI-75] Oshima, M., Shirai, Y., "Representation of curved objects using three-dimensional information", <u>Proc. Second USA-JAPAN Computer Conference</u>, pp. 108-112, 1975.
- [OSHI-79] Oshima, M., Shirai, Y., "A scene description method using three-dimensional information", <u>Pattern Recognition</u>, vol. 11, pp. 9-17, 1979.

- [OSHI-83] Oshima, M., Shirai, Y., "Object recognition using three-dimensional information", <u>IEEE Trans. Pattern Anal.</u>

 <u>Machine Intell.</u>, vol. PAMI-5, no. 4, pp. 353-361, Jul. 1983.
- [OTA-85] Ota, Y., Arai, H., Tokumasu, S., Ochi, T., "An automated finite polygon division method for 3-D objects", <u>IEEE Computer Graphics and Applications</u>, vol. 61, pp. 60-70, Apr. 1985.
- [PATT-68] Patterson, T.N.L., <u>Maths.</u> <u>Comp.</u>, vol. 22, pp-847-856 and 877-881, 1968.
- [PAVL-80] Pavlidis, T., "Structural Pattern Recognition", Springer-Verlag, 2° Edt., 1980.
- [PAVL-82] Pavlidis, T., "Algorithms For Graphic And Image Processing", Computer Science Press. 1982.
- [PENT-84] Pentland, A. "Local shading analysis", <u>IEEE Trans. Pattern</u>

 <u>Anal. Machine Intell.</u>, vol. PAMI-6, no. 2, pp. 170-187, Mar.

 1984.
- [PERK-82] Perkins, W.J., Green, R.J., "Three-dimensional reconstruction of biological sections", <u>J. Bio. Med. Engineering</u>, vol. 4, pp. 37-43, Jan. 1982.
- [PERR-79] Perrin, J.C., Thomas, A., "Electronic processing of moiré fringes: application to moiré topography and comparison with photogrammetry", Applied Optics, vol. 18, no. 4, pp. 563-574, Feb. 1979.
- [PGPL-87] "PGPLOT Graphics Subroutines User's Manual". Centro de Cálculo de la Universidad de La Laguna. 1987.
- [POTM-78] Potmesil, M., Freeman, H., "Computer descriptions of curved-surface objects from multiple proyected-pattern images", <u>Tech. Report. CRL-58, Image Processing Lab. Rensselaer Polytech. Inst.</u>, Jun. 1978.

- [POTM-79] Potmesil, M., <u>Proc. IEEE Computer Society Conference on Pattern Recognition and Image Processing</u>, pp. 553-559, 1979.
- [QIUL-87] Qiulin, D., Davies, B.J., "Surface Engineering Geometric for Computer-Aided Design and Manufacture", Ellis Horwood Limited, 1987.
- [REDD-81] Reddi, S.S., "Radial and angular moment invariants for image identification", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-3, no. 2, pp. 240-242, Mar. 1981.
- [REQU-80] Requicha, A.A.G, "Representations of rigid solid objects",

 <u>Computer Surveys</u>, vol. 12, no. 4, Dec. 1980.
- [ROAC-80] Roach, J.W., Aggarwal, J.K., "Determining the movement of objects from a sequence of images", <u>IEEE Trans. Pattern</u>

 <u>Anal. Machine Intell.</u>, vol. PAMI-2, no. 6, pp. 554-562, Nov. 1980.
- [ROBE-65] Roberts, L.G., "Machine perception of three-dimensional solids", in: Optical And Electro-Optical Information Processing, Tippet, J.T. et al. (Eds.). MIT Press, Cambridge, MA, 1965, pp. 159-197.
- [ROME-73] Romeder, J., "Méthodes Et Programmes D'Analysis Discriminante", Edit. Dunod, París. 1973.
- [ROSE-76] Rosenfeld, A., Kak, A.C., "Digital Picture Processing", Academic Press. New York, 1976.
- [ROSE-86] Rosenfeld, A. (Ed.), "Techniques for 3-D Machine Perception", North-Holland, 1986.
- [ROSE-88] Rosenfeld, A., "Computer architectures for machine vision", en "Machine Vision: Algorithms, Architectures, and Systems", Freeman, H., (Edt.), Academic Press, Inc., 1988, pp. 97-101.
- [SADJ-80] Sadjadi, F.A., Hall, E.L., "Three-dimensional moment invariants", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-2, no. 2, pp. 127-136, Mar. 1980.

- [SAND-88] Sanderson, A.C., Weiss, L.E., Nayar, S., "Structured highlight inspection of specular surfaces", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-10, no. 1, pp-44-55, Jan. 1988.
- [SHAN-78] Shantz, M.J., McCann, G.D., "Computational morphology: three-dimensional computer graphics for electron microscopy", <u>IEEE Trans. Biomedical Engineering</u>, vol. BME-25, no. 1, pp. 99-103, Jan. 1978.
- [SHIR-72] Shirai, Y., "Recognition of polyhedra with a range finder", <u>Pattern Recognition</u>, vol. 4, no. 3, pp. 243-250, 1972.
- [SHIR-87] Shirai, Y., "Three-Dimensional Computer Vision", Springer-Verlag, 1987.
- [SLOA-81] Sloan, K.R., Hrechanyk, L.R., "Surface reconstruction from sparse data", <u>Proc. IEEE Computer Society on Pattern Recognition and Image Processing</u>, pp. 45-48, 1981.
- [SPIE-70] Spiegel, M.R., "Manual de Fórmulas y Tablas Matemáticas", McGraw-Hill Inc., 1970.
- [SRIH-80] Srihari, S.N., "Hierarchical representations for serial section images", Reprinted from 5th International Conference on Pattern Recognition, pp. 1075-1080, Dec. 1980.
- [SRIH-81] Srihari, S.N., "Representation of three-dimensional digital images", Computing Surveys, vol. 13, no. 4, Dec. 1981.
- [STEV-85] Stevens, L., "Artificial Intelligence", Hayden Book Com., 1985.
- [SUNG-78] Sunguroff, A., Greenberg, D., "Computer generated images for medical applications", <u>Computer Graphics</u>, vol. 12, pp. 196-202, 1978.
- [SUTH-74] Sutherland, I.E., Hodgman, G.W., "Reentrant polygon clipping", <u>CACM</u>, vol. 17, pp. 32-42, 1974.

- [TEAG-80] Teague, M.R., "Image analysis via the general theory of moments", <u>J. Opt. An.</u>, vol. 70, no. 8, pp. 920-930, Aug. 1980.
- [TECH-85] Tech. Tran. Consultants, Inc., "Special Report: Machine Vision Systems", 2° Edt., MacGraw-Hill, 1985.
- [TEH-88] Teh, C., Chin, R.T., "On image analysis by the methods of moments", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-10, no. 4, pp. 496-513, Jul. 1988.
- [TOUS-85] Toussain, G., "Computational Geometry", Machine Intelligence and Pattern Recognition 2. North-Holland. 1985.
- [TSAI-83] Tsai, R.Y., "Multiframe image point matching and 3-D surface reconstruction", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-5, no. 2, pp. 159-174, Mar. 1983.
- [TSAI-84] Tsai, R.Y., Huang, T.S., "Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces", <u>IEEE Trans. Pattern Anal. Machine.</u>

 <u>Intell.</u>, vol. PAMI-6, no. 3, Mar. 1983.
- [UDUP-83] Udupa, J.K., "Display of 3-D information in discrete 3D scenes produced by computerized tomography", <u>Proc. IEEE</u>, vol. 71, no. 3, Mar. 1983.
- [ULLM-79] Ullman, S., "The Interpretation of Visual Motion", MIT Press, Cambridge, MA, 1979.
- [UNDE-75] Underwood, S.A., Coates, C.L., "Visual learning from multiple views", <u>IEEE Transactions on Computers</u>, vol. 24, no. 6, pp. 651-661, 1975.
- [WANG-86] Wang, R-Y., "An approach to tree-classifier design based on hierarchical clustering", <u>International Journal of Remote Sensing</u>, vol. 7, pp. 75-88, 1986.
- [WANG-85] Wang, Y.F., Aggarwal, J.K., "Construction of surface representation from 3-D volumetric scene description", <u>Proc.</u>

- <u>IEEE Computer Society Conf. Comput. Vision and Pattern Recognition</u>, pp. 130-135, San Francisco, Jun. 1985.
- [WANG-87] Wang, Y.F., Mitiche, A., Aggarwal, J.K., "Computation of surface orientation and structure of objects using grid coding", <u>IEEE Trans. Pattern Anal. Machine Intell.</u>, vol. PAMI-9, no. 1, pp. 129-137, Jan. 1987.
- [WEBB-81] Webb, J.A., Aggarwal, J.K., "Visually interpreting the motion of objects in space", <u>IEEE Computer</u>, vol. 14, pp. 40-46, Aug. 1981.
- [WINS-75] Winston, P.H., "The Psychology of Computer Vision", McGraw-Hill Computer Science Series, 1975.
- [WINS-84] Winston, P.H., "Artificial Intelligence", Addison-Wesley Publish. Com., (2° Edt.), 1984.
- [WITK-81] Witkin, A.P., "Recovering surface shape and orientation from texture", <a href="https://example.com/artificial/linealinear.com/artificial/linealin
- [WOOD-78] Woodham, R.J., "Photometric stereo: A reflectance map technique for determining surface orientation from image intensity", Proc. SPIE, vol. 155, 1978.
- [WOOD-80] Woodham, R.J., "Photometric method for determining surface orientation from multiple images", Opt. Engineer., vol. 19, no. 1, pp. 139-144, 1980.