

**UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA**

**DEPARTAMENTO DE INFORMÁTICA**



**TESIS DOCTORAL**

**TRANSFORMACIONES DE CAMPO RECEPTIVO VARIABLE EN  
PROCESO DE IMÁGENES Y VISIÓN ARTIFICIAL: INTRODUCCIÓN  
DE CONCEPTOS, DESARROLLO TEÓRICO Y APLICACIONES**

**SANTIAGO CANDELA SOLÁ**

Las Palmas de Gran Canaria, 1987

UNIVERSIDAD POLITECNICA DE CANARIAS

FACULTAD DE INFORMATICA

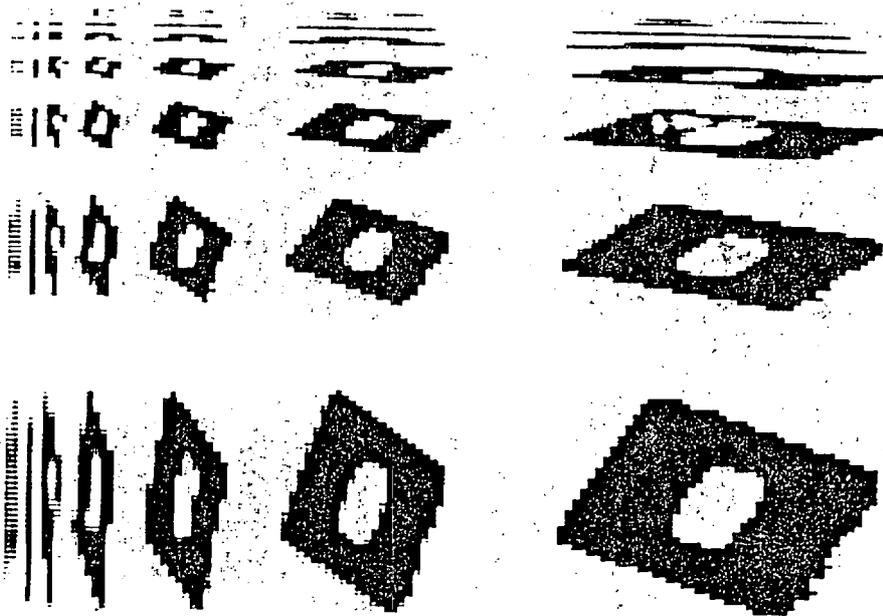
DEPARTAMENTO DE INFORMATICA Y SISTEMAS

TESIS DOCTORAL

TRANSFORMACIONES DE CAMPO RECEPTIVO VARIABLE

EN PROCESO DE IMAGENES Y VISION ARTIFICIAL

INTRODUCCION DE CONCEPTOS, DESARROLLO TEORICO Y APLICACIONES



*Autor: Santiago Candela Solá*

*Director: Roberto Moreno Díaz*

Noviembre 1987

DON ROBERTO MORENO DIAZ, CATEDRATICO DE  
UNIVERSIDAD EN EL AREA DE CIENCIA DE LA  
COMPUTACION E INTELIGENCIA ARTIFICIAL DEL  
DEPARTAMENTO DE INFORMATICA Y SISTEMAS DE LA  
UNIVERSIDAD POLITECNICA DE CANARIAS

CERTIFICA:

Que la memoria titulada TRASNFORMACIONES  
DE CAMPO RECEPTIVO VARIABLE EN PROCESO  
DE IMAGENES Y VISON ARTIFICIAL.  
INTRODUCCION DE CONCEPTOS, DESARROLLO  
TEORICO Y APLICACIONES, ha sido  
realizada por D. Santiago Candela Solá  
bajo mi dirección en el Departamento de  
Informática y Sistemas de la Universidad  
Politécnica de Canarias. Esta memoria  
constituye la Tesis que presenta para  
optar al grado de Doctor en Informática.

Las Palmas, a 23 de Noviembre de 1987

V° B°

EL DIRECTOR DEL DPTO. DE  
INFORMATICA Y SISTEMAS

P.O.



Victor Almeida Lozano

EL DIRECTOR DE TESIS

Roberto Moreno Díaz

DEDICADO A:

Mary

Bea y

Nestor

## AGRADECIMIENTOS

Mis primeras palabras de agradecimiento son para mi "maestro" Dr. Roberto Moreno Diaz, por haberme educado e introducido en las tareas de investigación. Bajo su intuición, trabajo y reflexiones se motivó, dirigió y creó esta tesis.

Agradezco a mis compañeros del Departamento de Informática y Sistemas de la Universidad Politecnica de Canarias, la colaboración, ayuda y disponibilidad que siempre me han prestado, y que de alguna forma se ha traducido en la realización de este trabajo. Deseo mencionar de forma especial a mi compañero y amigo Jose Antonio Muñoz Blanco, con el cual he convivido muchas horas de trabajo, y ha hecho que el trabajo sea productivo, ameno y alegre.

Deseo expresar mi agradecimiento a Vicente Umpierrez por su trabajo esmerado en las tareas de mecanografiado y edición del texto.

También quiero dar las gracias a todas aquellas personas que de forma directa o indirecta han hecho posible que se realice este trabajo.

Deseo terminar dando muchas gracias a mi familia por el apoyo moral que siempre he sentido, en especial a M<sup>a</sup>. Isabel por su ayuda, paciencia y comprensión en la realización de esta tesis.

## RESUMEN

Esta tesis titulada: Transformaciones de campo receptivo variable en proceso de imágenes y visión artificial, introducción de conceptos, desarrollo teórico y aplicaciones. Está estructurada en cinco capítulos, más un sexto que presenta un listado de la bibliografía consultada.

El primer capítulo titulado los problemas en proceso y reconocimiento de imágenes, trata sobre los conceptos generales en visión artificial, se define la visión como la operación de un sistema que proporciona, a partir de imágenes, una representación que es útil al que ve y que no está cargada de información irrelevante. Se hace énfasis de la medida en la que el proceso visual está determinado por los datos visuales o por la estructura de conocimientos. Se plantea la existencia de niveles en el aspecto funcional y estructural y en el segmento metodológico, y como el aspecto central en la visión computacional es llegar a métodos computacionales efectivos para pasar de imágenes o representaciones, mediante restricciones suficientes.

A continuación se hace una propuesta de un esquema para la visión artificial, donde el objetivo central es el de obtener información sobre formas a partir de las imágenes imponiendo restricciones que vienen determinadas por la propia naturaleza del mundo exterior, y no por la estructura del conocimiento impuesta a priori. Este esquema se completa con un procedimiento de reconocimiento visual, comandado principalmente por

conocimiento visual preestructurado y que echa mano de bases de datos conteniendo los modelos standard. Se presenta el concepto de visión multimodal, en el que el sistema realiza representaciones usando distintos modos visuales, donde se ha de llegar a un diagnóstico final, a partir de la conversación cruzada entre las modalidades.

En el siguiente apartado, sobre métodos convencionales, se presenta una revisión de las técnicas y métodos más interesantes utilizados en el reconocimiento de formas, a partir de la bibliografía publicada sobre el tema en los últimos años, se estudian distintas técnicas de reconocimiento, en cuanto a su implementación, velocidad, complejidad e independencia. Las distintas técnicas se agrupan en tres grandes categorías, características globales, características locales y características topológicas.

Por último se explica como se caracteriza una imagen, se explican los conceptos de muestreo uniforme y cuantización, se presenta con un ejemplo una imagen simple continua, y su imagen muestreada digital, en la que a cada punto se le asigna un valor que está relacionado con la intensidad de su área en el espacio objeto. Se explica como en algunos casos es posible mejorar la calidad de una imagen, por medio de un esquema adaptivo donde el proceso de muestreo dependa de las características de la imagen, ya que en general se requiere un muestreo fino en las cercanías de transiciones rápidas de niveles de grises, mientras que se puede emplear un muestreo ancho en regiones relativamente homogéneas. Es lo que se conoce con el nombre de muestreo no

uniforme. Cuando el número de niveles de grises es pequeño, es factible utilizar un proceso de cuantización de niveles desiguales espaciados, obteniéndose una cuantización no uniforme.

El segundo capítulo se denomina formalismos para el proceso de imágenes, donde se plantea un marco conceptual y teórico cuya potencialidad permite formalizar un modelo general de proceso de imágenes, se define el proceso de imágenes como un subsistema caracterizado por un espacio de entrada, un espacio de salida y una estructura relacional entre ellos. Se define en espacio de entradas como un espacio multidimensional compuesto por todas las características independientes necesarias para definir el sistema base estudio. La naturaleza del espacio de salida viene determinada por la caracterización del espacio de entrada y del nivel formal que tenga la estructura relacional. La estructura relacional es el enganche entre los espacios de entrada y salida, esta relación toma datos del espacio de entrada y en general del espacio de salida para producir la salida correspondiente en un instante. Se realiza una primera clasificación de los procesos en lineales y no lineales dependiendo de la estructura formal de la relación, y una segunda clasificación atendiendo al número de puntos del espacio de entrada que intervienen para generar un punto en el espacio de salida, teniendo los siguientes tipos de operaciones: Técnicas puntuales, técnicas locales y técnicas globales, presentándose varios ejemplos de cada una de ellas.

Se dedica un apartado al estudio de la descripción, y propiedades de las transformaciones globales, donde se comenta la

utilización de distintas transformaciones aplicadas a las imágenes para reducir la información necesaria para su reconocimiento. Se estudian la transformación de Karhunen Loeve, la transformada de Fourier, la transformación de Hadamard, y la transformada de Haar. En todas ellas, se define la transformación, se enumeran sus propiedades específicas, y se muestra su algoritmo rápido para su implementación. Se estudian con menos profundidad otras transformaciones como la de Fresnel, Hadamard Slant, Hadamard compleja, Maar Slant y la transformación coseno discreto.

Este capítulo finaliza con la realización de un estudio comparativo exhaustivo entre las transformaciones de Fourier, Hadamard y Haar. Para ello se diseña y realiza un experimento que consiste en la elección de un banco de imágenes, cálculo de los coeficientes transformados, selección de descriptores utilizando un criterio de bajo orden y otro de máximos relativos, reconstrucción de la imagen a partir de los coeficientes, cálculo de la distancia entre la imagen inicial y la reconstruida, distancia entre los descriptores y estudio de los resultados mostrándose estos mediante varias gráficas .

El tercer capítulo se denomina transformaciones de Campo receptivo variable y su papel en los Sistemas Perceptuales y se define el concepto de descripción completa de un entorno visual, como aquella que contiene todos los datos y sus propiedades, necesarios para cubrir unos objetivos. Se describen los dos métodos de plantear los problemas relativos a la descripción completa, que se corresponden con las formas clásicas de tratar

el problema de la visión artificial, la dualidad de los procesos dominados por los datos, frente a los dominados por el conocimiento, realizándose una descripción y comparación de estos.

Se realizan unas observaciones acerca del lenguaje formal apropiado a cada nivel metodológico en una descripción completa se define el concepto de lenguaje como un conjunto de átomos simbólicos y de estructuras relacionales con huecos a rellenar por átomos, con reglas de enganche de las estructuras relacionales, para formar frases. Se diferencian distintos niveles de lenguaje para la descripción del sistema, que van desde el nivel simbólico con varias tablas de significados hasta el nivel analítico básico, y se presenta el concepto de derivada de niveles, se realiza una comparación en analítica, entre visión artificial y comunicación visual, donde se observa que al presentar distintos objetivos las herramientas formales utilizadas no tienen porque coincidir.

A continuación se pasa al estudio de las transformaciones de campo receptivo variable, explicándose el concepto de campo receptivo, la definición de transformaciones de campo receptivo constante, para llegar a la definición de transformaciones de campo receptivo variable. Se realizan unas consideraciones acerca del carácter completo de una transformación con relación al campo receptivo y a la función, separándose lo que es campo receptivo y función, y se definen las transformaciones puras de campo receptivo.

Este apartado finaliza con la idea sugerida por la analítica sobre la descripción completa de un entorno visual y las propiedades de los campos receptivos, donde se define la independencia del conjunto de dominios y se definen las transformaciones de resolución progresiva (TRP), se hace una descripción detallada de los métodos de generación de una TRP y se estudia una clase particular de estas que denominamos transformaciones foveales, inspiradas en la estructura del "Moving average". Se ilustra las proyecciones de las TRP'S, entre todas las posibles se hace incapie en la denominada "ordenación natural". Una vez definida la transformación TRP, se analizan los algoritmos de cálculo secuencial rápido que permiten calcularlas, generándose los algoritmos bidimensionales como producto de algoritmos para una dimensión, por ser los núcleos separables.

Se termina el capítulo analizando las TRP detectoras de contraste y se ilustran distintas TRP'S con distintos criterios de ordenación, como son la representación fraccionada y la proyección radial.

El cuarto capítulo se titula reconocimiento visual directo en base a transformaciones de campo receptivo variable donde se espone una estructura general de sistemas de inferencia que toman decisiones acerca del mundo visual directamente apartir de los datos visuales; describiendose las partes básicas del sistema, la base de datos referencial y el dispositivo que contiene los algoritmos de inferencia. Se muestra con un esquema la estructura de un sistema visual no interactivo con el medio y la estructura de un sistema visual predeterminado.

Los sistemas reducidos de inferencia tienen su origen en la generalización de la propiedad de invarianza espacial de una transformación; por lo que se define al concepto de transformación invariante en función de la transformación de campo receptivo constante y de los funcionales generadores de descriptores. Se describe la idea de proyección columnar y los procedimientos generadores para una dimensión y para dos dimensiones, así como las representaciones directas estudiando los distintos casos en función de distintas proyecciones. Se dan las reglas para la determinación de dominios y transformaciones afines. A continuación, se ilustran transformaciones afines, partiendo de la expresión más general y dando valores para obtener transformaciones invariantes a traslaciones mediante el cálculo del centro de gravedad, se ilustran las invarianzas a giros, calculando el ángulo mediante la teoría apuntada por los momentos y se finaliza con las invarianzas frente a homotecias, presentándose para cada una de ellas ejemplos ilustrativos con imágenes.

Este capítulo finaliza con un apartado dedicado al estudio de distintas aplicaciones realizadas en la tesis, como a partir de la generación de una base referencial de datos. Se realiza un experimento consistente en una normalización del banco de imágenes, frente a traslaciones, rotaciones y amplificaciones, la aplicación de una transformación de resolución progresiva a las imágenes normalizadas para obtener las frases descriptoras, que conforman la base de datos referencial, realizándose una fase de

establecimiento de datos y una fase de reconocimiento. Se calculan las distancias entre frase descriptora incognita y las patrones, describiendose los distintos métodos utilizados, de mínima distancia, distancias ponderadas y producto distancias frecuencia. Terminando con un comentario de los resultados mediante un ploter de los resultados.

I N D I C E

## I N D I C E

1.- <u>LOS PROBLEMAS EN PROCESO Y RECONOCIMIENTO DE IMAGENES.</u>	
1.1 CONCEPTOS GENERALES EN VISION ARTIFICIAL.	1-2
1.1.1 PLANTEAMIENTOS EN VISION ARTIFICIAL.	1-2
1.1.2 PROPUESTA DE UN ESQUEMA PARA LA VISION ARTIFICIAL.	1-7
1.2 METODOS CONVENCIONALES.	1-11
1.2.1 INTRODUCCION.	1-11
1.2.2 TECNICAS DE RECONOCIMIENTO.	1-12
1.3 CARACTERIZACION DE IMAGENES.	1-19
1.3.1 MUESTREO UNIFORME Y CUANTIZACION.	1-19
1.3.2 MUESTREO Y CUANTIZACION NO UNIFORME.	1-24
2.- <u>FORMALISMOS PARA PROCESO DE IMAGENES.</u>	
2.1 GENERALIZACION DE METODOLOGIAS EN PROCESO DE IMAGENES.	2-2
2.1.1 MARCO GENERAL EN PROCESO DE IMAGENES.	2-2
2.1.2 TECNICAS PUNTUALES.	2-9
2.1.3 TECNICAS LOCALES.	2-19
2.1.4 TECNICAS GLOBALES.	2-28
2.2 DESCRIPCION Y PROPIEDADES DE TRANSFORMACIONES GLOBALES.	2-29
2.2.1 INTRODUCCION.	2-29
2.2.2 TRANSFORMACION DE KARHUNEN-LOEVE.	2-32
2.2.3 INCONVENIENTES DE LA TRANSFORMACION KARHUNEN-LOEVE.	2-34
2.2.4 TRANSFORMACIONES RAPIDAS.	2-35
2.2.4.1 DEFINICIONES.	2-36

2.2.4.2 TRANSFORMACIONES RAPIDAS.	2-39
2.2.5 TRANSFORMADA DE FOURIER.	2-43
2.2.5.1 DEFINICION.	2-43
2.2.5.2 PROPIEDADES DE LA TRANSFORMACION.	2-45
2.2.5.3 ALGORITMO DE CALCULO RAPIDO.	2-46
2.2.6 TRANSFORMACION DE HADAMARD.	2-48
2.2.6.1 DEFINICION.	2-48
2.2.6.2 PROPIEDADES DE LA TRANSFORMACION.	2-52
2.2.6.3 ALGORITMO DE CALCULO RAPIDO.	2-53
2.2.7 TRANSFORMADA DE HAAR.	2-55
2.2.7.1 DEFINICION.	2-55
2.2.7.2 PROPIEDADES DE LA TRANSFORMADA DE HAAR.	2-55
2.2.7.3 ALGORITMO DE CALCULO RAPIDO.	2-59
2.2.8 OTRAS TRANSFORMACIONES RAPIDAS.	2-60
2.2.8.1 TRANSFORMACION DE FRESNEL.	2-60
2.2.8.2 TRANSFORMADA DE HADAMARD-SLANT.	2-61
2.2.8.3 TRANSFORMACION DE HADAMARD COMPLEJA.	2-62
2.2.8.4 TRANSFORMADA DE HAAR-SLANT.	2-63
2.2.8.5 TRANSFORMADA DE COSENO DISCRETO.	2-64
2.3 ESTUDIO COMPARATIVO DE LAS MISMAS.	2-64

### 3.-TRANSFORMACIONES DE CAMPO RECEPTIVO VARIABLE Y SU PAPEL EN LOS SISTEMAS PERCEPTUALES.

3.1 TRANSFORMACIONES DE CAMPO RECEPTIVO VARIABLE (TCRV).	3-2
3.1.1 SOBRE EL CONCEPTO DE DESCRIPCIONES COMPLETAS DE UN ENTORNO VISUAL.	3-2
3.1.2 OBSERVACIONES ACERCA DEL LENGUAJE FORMAL APROPIADO A CADA NIVEL METODOLOGICO EN UNA DESCRIPCION COMPLETA.	3-5

3.1.3	VISION ARTIFICIAL VERSUS COMUNICACION VISUAL, AMBOS EN ANALITICA.	3-10
3.1.4	TRANSFORMACIONES DE CAMPO RECEPTIVO VARIABLE (TRV).	3-13
3.1.5	CONSIDERACIONES ACERCA DEL CARACTER COMPLETO DE UNA TRANSFORMACION CON RELACION AL CAMPO RECEPTIVO Y A LA FUNCION.	3-17
3.1.6	NUEVA VISITA A LAS DESCRIPCIONES COMPLETAS DE UN ENTORNO VISUAL SUGERIDAS POR LA ANALITICA.	3-20
3.1.7	PROPIEDADES DE LOS CAMPOS RECEPTIVOS.	3-22
3.2	TRANSFORMACIONES DE RESOLUCION PROGRESIVA (TRP).	3-25
3.2.1	METODOS DE GENERACION.	3-25
3.2.2	TRANSFORMACIONES FOVEALES.	3-29
3.2.3	OTRAS TRP'S.	3-37
3.2.4	PROYECCIONES DE LAS TRP'S.	3-38
3.3	ALGORITMOS DE CALCULO SECUENCIAL RAPIDO DE TRP.	3-41
3.3.1	REPRESENTACION MATRICIAL DE LAS TRP.	3-41
3.3.2	ALGORITMO DE CALCULO SECUENCIAL RAPIDO DE TRP.	3-45
3.3.3	TRP DETECTORA DE CONTRASTES.	3-52
3.4	ILUSTRACIONES DE TRP Y SU REPRESENTACION.	3-55
3.4.1	REPRESENTACION CRITERIO CR.	3-55
3.4.2	REPRESENTACION FRACCIONADA.	3-57
3.4.3	PROYECCION RADIAL.	3-58
4.-	<u>RECONOCIMIENTO VISUAL DIRECTO EN BASE A TRANSFORMACIONES DE CAMPO RECEPTIVO VARIABLE.</u>	
4.1	SISTEMAS REDUCIDOS DE INFERENCIA EN VISION-DECISION.	4-2
4.1.1	SISTEMAS REDUCIDOS DE INFERENCIA.	4-2

4.1.2 INTEGRACION DE UN SISTEMA REDUCIDO DE INFERENCIAS EN UN SISTEMA VISUAL.	4-3
4.2 INVARIANZA ESPACIAL. REPRESENTACION DIRECTA Y TRANSFORMACIONES AFINES.	4-7
4.2.1 SOBRE LA INVARIANZA ESPACIAL.	4-7
4.2.2 REPRESENTACIONES DIRECTAS.	4-10
4.2.3 REPRESENTACION DE DOMINIOS Y TRANSFORMACIONES AFINES.	4-12
4.3 ILUSTRACIONES DE TRANSFORMACIONES AFINES.	4-13
4.3.1 TRASLACIONES.	4-14
4.3.2 ROTACIONES.	4-16
4.3.3 HOMOTECIAS.	4-19
4.4 APLICACIONES.	4-22
4.4.1 GENERACION DE UNA BASE REFERENCIAL DE DATOS.	4-22
4.4.2 DESCRIPCION DEL EXPERIMENTO.	4-26
4.4.3 COMENTARIO DE LOS RESULTADOS.	4-35
5.- <u>CONCLUSIONES Y PRINCIPALES APORTACIONES.</u>	
5.1 APORTACIONES.	5-2
5.2 PROYECCION FUTURA Y CONCLUSION GENERAL.	5-5
6.- <u>BIBLIOGRAFIA.</u>	6-2

**CAPITULO:1. LOS PROBLEMAS EN PROCESO Y RECONOCIMIENTO  
DE IMAGENES.**

## 1.1 CONCEPTOS GENERALES EN VISION ARTIFICIAL.

### 1.1.1 PLANTEAMIENTOS EN VISION ARTIFICIAL.

El problema de la visión artificial, o de mecanización de los procesos de visión, trata, en general, de la definición de "máquinas" posibles (existentes o no) que, partiendo de imágenes proporcionen una descripción del mundo exterior. Así, un concepto que parece ser útil tanto en los sistemas naturales como en los artificiales, consiste en considerar la visión como la "operación de un sistema que proporciona, a partir de imágenes, una representación que es útil al que ve y que no está cargada de información irrelevante" (Marr 1976).

El proceso visual presenta dos características iniciales de importancia. Primero, el requisito de que la información extraída sea 'útil' para los fines de un sistema más amplio que incluye al visual como parte, y segundo, la reducción, en general drástica, de la información técnica contenida en las imágenes, con la esperanza de que se mantenga sólo aquella información que es relevante a los fines del sistema total.

Partiendo de que el sistema visual, artificial o natural, está integrado en un sistema total cuyos fines sirve, la investigación lógica en los mecanismos de visión debe partir de los requisitos y/o restricciones que vienen impuestas por esos fines. Ahora bien, aunque en el caso de la mayor parte de los

sistemas visuales artificiales el establecimiento de esas restricciones puede ser realizada a priori con una cierta precisión, en el caso de los sistemas naturales estos no son conocidos, y son, por consiguiente, tema también de investigación en visión. Desde el aspecto de la visión humana, la tarea podría consistir en el desarrollo de una teoría efectiva de como se construye la representación visual invariante a partir de las imágenes proyectadas en las retinas. La solución de los problemas asociados al reconocimiento visual exige, como ha sido reiteradamente demostrado, el uso de la información multisensorial para la elaboración de un modelo del medio, que en definitiva es algo así como una estructura de conocimiento de naturaleza relacional, donde las pistas iniciales proporcionadas por las modalidades sensoriales son usadas, a través de mecanismos cooperativos, para actualizar su estructura e indagar en el mundo exterior, requiriendo más información de las modalidades sensoriales.

Como consecuencia, una de las cuestiones que de forma más o menos explícita subyace en la investigación de la visión artificial es la de la medida en la que el 'proceso visual' está determinado por los 'datos visuales' o por la 'estructura de conocimiento'. Una creencia bastante antigua en visión artificial es que los fenómenos de percepción visual son análogos a los de 'resolución de problemas' en inteligencia artificial, y que, por consiguiente, están determinados por la 'estructura del conocimiento'. Ello conlleva a considerar que la función fundamental del 'proceso visual' es el establecimiento, la prueba

y la modificación de hipótesis acerca de lo que hay en la escena. Esta idea sigue siendo común en la visión por computador y tiene algunas contrapartidas en la psicología de la visión. El concepto básico es invocar el 'conocimiento especializado' que se dispone de la escena que se ve para ayudar en la segmentación de la imagen en regiones que se deben corresponder con los objetos que se esperan en la escena. Este tipo de enfoque tuvo buenos representantes en el laboratorio de Inteligencia Artificial de MIT y en el Stanford Research Institute y su expresión clara en los conceptos de "entender imágenes" (Amherst).

Marr presenta una alternativa al énfasis de "proceso determinado por el conocimiento", llamando la atención sobre que el "proceso determinado por los datos" es mucho más potente de lo que se creía, y puede racionalizar el problema de la segmentación de las imágenes. La idea central de Marr es que las imágenes han de ser usadas para obtener representaciones visuales y que la caracterización de formas en un espacio de representación es anterior a la interacción con la estructura de conocimiento. Históricamente, ha sido precisamente el problema de la segmentación de imágenes el que llevó a Marr y Poggio al concepto de los "bocetos" y representaciones de formas. Según sus argumentos, ha sido la creencia de los neurofisiólogos y psicólogos en la separación entre figura y fondo de una imagen, la que dió lugar a los intentos de implementar en computador los procesos de segmentación. Pero la segmentación ha probado ser una jungla que complica, en vez de ayudar, tanto en la teoría

como en la práctica: es imposible especificar o formular con precisión en términos de la imagen, e incluso en el mundo real, cuales son los propósitos exactos de la segmentación. Regiones que tienen importancia semántica no presentan siempre una distinción visual particular. La mayor parte de las imágenes son demasiado complejas para ser segmentadas, e incluso, imágenes muy sencillas no contienen suficiente información en el mosaico de intensidades, para ser separadas en objetos. El proceso visual no puede empezar por imponer hipótesis restrictivas acerca de la imagen, porque las restricciones impuestas no son generales, sino particulares y ciertas solo para un tipo de escenas. Si se basara el proceso visual esencialmente en un proceso determinado por la estructura de conocimiento, un sistema de visión generalizado debe tener acceso a un número extraordinariamente elevado de posibles hipótesis, y debe ser capaz de encontrar y aplicar justamente aquellas pocas que son exigidas por la situación particular. Así, para Marr, el principal problema del proceso visual es obtener la apropiada información y no el como manejar vastas cantidades de información de forma eficiente, aunque el pensar que es así da lugar a un gran esfuerzo dedicado al diseño de estructuras de control de programas.

La teoría computacional de la visión, es una de las más fuertes tendencias actuales en visión artificial. Los puntos básicos de las ideas de Marr se refieren a una doble existencia de niveles en el aspecto funcional y/o estructural y en el segmento metodológico.

En el aspecto funcional/estructural, la existencia de niveles corresponde al caso particular de la visión, a la distribución más general de niveles (que caracteriza a la Biocibernética) y en el reconocimiento de que cada nivel requiere sus propias herramientas formales que imponen limitaciones a lo que se puede obtener (Mira, Moreno Díaz). Los niveles funcionales van desde el bioquímico -o físico-, al neuronal o de componentes, al de operaciones lógicas, hasta el nivel simbólico, con el establecimiento de fines. Marr se refiere al nivel computacional, al nivel de los algoritmos y al nivel de la realización con hardware. Marr y sus seguidores, argumentan que al nivel computacional, las teorías de visión natural y artificial interactúan fuertemente, de tal forma que la visión por computador puede: a) proporcionar un lenguaje para expresar las teorías del proceso de datos visuales. b) dar la capacidad de establecer experimentos claramente, sin las confusiones que se suelen derivar de la única consideración a la visión humana. c) indicar que las teorías a distintos niveles son apropiadas para sugerir situaciones en el aspecto teórico experimentales, tanto en visión artificial como en la percepción humana.

En la visión computacional el aspecto central es el de llegar a métodos computacionales efectivos para pasar de 'imágenes' a 'representaciones', a través de la imposición de las 'restricciones' suficientes. Se ilustra con la búsqueda de métodos para pasar desde imágenes a superficies y desde superficies a formas, con el problema final de la representación

de formas para el reconocimiento.

### 1.1.2 PROPUESTA DE UN ESQUEMA PARA LA VISION ARTIFICIAL.

La teoría computacional de visión, más que una teoría, es una propuesta general de un procedimiento para investigar en visión artificial. El objetivo central es el obtener información sobre formas a partir de las imágenes, imponiendo restricciones que vienen determinadas por la propia naturaleza del mundo exterior, y no por la estructura del conocimiento impuesta a priori. Partiendo de las imágenes, la teoría computacional propone obtener un "boceto primario"; de aquí, un boceto en 2 1/2 dimensiones y finalmente, una representación de las formas en 3-D. La sospecha final de que éste es un enfoque apropiado para analizar el proceso visual les vino a Marr y Poggio del hecho de que es posible encontrar un procedimiento computacional efectivo para interpretar esteogramas formados por puntos al azar (Julezz, 1960) y que por, consiguiente, tal representación debe ser anterior al reconocimiento. El programa de "trabajo" de la teoría computacional de la visión se presenta a continuación, adaptado de Marr (1982). Según este esquema, el marco representacional presenta al menos cuatro niveles, en el que cada representación se ha de expresar en términos de unas primitivas propias. Las primitivas correspondientes a la imagen son el valor de intensidad en cada punto. El boceto primario tiene como finalidad explicitar información relevante acerca de la imagen bidimensional, esencialmente los cambios de intensidad, y su distribución geométrica y organización. Sus primitivas incluyen cruces de cero, terminaciones y discontinuidades, segmentos de

bordes, líneas virtuales, grupos, organización curvilínea y límites. El boceto 2 1/2 D tiene por objeto explicitar la orientación y profundidad aproximada de las superficies visuales, y los contornos y discontinuidades en estas magnitudes, en un sistema de coordenadas centrado en el observador. Sus primitivas incluyen la orientación local de superficies, distancia hasta el observador, discontinuidades en profundidad y en orientación de superficies. Finalmente, está la representación en forma de modelo 3-D, cuyo objetivo es describir formas y su organización especial referidas a un sistema de coordenadas centrado en el objeto, usando una representación modular jerarquizada que incluye descriptores volumétricos tanto como descriptores superficiales. Las primitivas son un conjunto de modelos 3-D dispuestos jerárquicamente, cada uno de los cuales posee un propio sistema de referencias intrínseco, al que se refieren los descriptores superficiales y de volumen.

El anterior esquema se completa con un procedimiento de reconocimiento visual, comandado fundamentalmente por conocimiento visual pre-estructurado y que echa mano de bases de datos conteniendo los modelos standard. La correspondencia entre las primitivas y los descriptores y las formas que representan es directa según lo anterior. En realidad, esto no es un requisito esencial aunque ayuda a un metaobservador a entender el proceso suyacente de una forma más próxima a la intuitiva. Las propiedades de las superficies y de las formas a la que se refiere el esquema anterior son propiedades deducibles y

expresables en un lenguaje formal muy próximo al de la geometría diferencial e integral. En esencia, las restricciones impuestas son debidas a implicar niveles simbólicos o semánticos hasta que se llega al reconocimiento.

La citada correspondencia no tiene en principio por que ser directa, y podría llevarse a cabo en un espacio transformado, donde ciertas primitivas de interés pueden ser determinadas y manejadas con mayor flexibilidad. De hecho, esto no pasa de ser una suposición y una esperanza. En muchos casos, el trabajar con descriptores procedentes de espacios transformados en sistemas de visión artificial descansa solamente en la adecuación posterior comprobada experimentalmente, tanto en tiempo de cómputo como en aciertos. Esto ha de tomarse con cierto cuidado, porque, normalmente, los investigadores no dan cuenta de los errores, sino de los aciertos.

En todo caso, es muy difícil, hoy por hoy, proporcionar un teorema negativo acerca de la utilidad de pasar a espacios transformados. Una transformada integral biunívoca que ha tenido gran aceptación en la representación visual ha sido la transformada de Fourier espacio-temporal. Incluso se la intentó justificar fisiológicamente por Campbell y otros, llegándose a afirmar que el objeto de la retina es realizar una descomposición en frecuencias espacio-temporales, que es transmitida al cortex. Posteriormente, se ha visto que esa afirmación no tiene más fundamento que la de Hubel y Wiesel acerca de la células corticales especializadas en detectar líneas y bordes orientados,

es decir, ninguno. El que un sistema tenga una cierta respuesta espectral selectiva, no implica nunca que el sistema esté codificando de acuerdo con una descomposición espectral.

Sin proseguir en la cuestión de que si algunos sistemas naturales de visión usan o no descriptores procedentes de transformaciones integrales, es cierto que estos descriptores tienen propiedades que los hacen atractivos, al menos para un sistema de visión artificial. Uno de ellos es la desaparición de los escotomas que aparecen en las representaciones punto por punto. Otra es la posibilidad de lograr invariantes en la representación de una forma rápida y elegante.

Como consecuencia de admitir la potencial utilidad de una representación múltiple, se llega a los conceptos correspondientes a la representación visual multimodal. En este caso, el sistema realiza representaciones usando distintos modos visuales. De la conversación cruzada entre las modalidades se ha de llegar a un diagnóstico final. Convergemos así en una propuesta de esquema de visión artificial, según el cual, las transformaciones básicas iniciales actúan directamente sobre las imágenes para proporcionar descriptores relevantes de los mismo, con las correspondientes etiquetas. De acuerdo con la idea básica del reconocimiento multimodal, las transformaciones deben utilizar descriptores invariantes, si bien los parámetros para determinar las invarianzas, deben, en general, proceder de una modalidad diferente. En todo caso, ha de tenerse en cuenta que

las transformaciones no corresponden a una única expresión, sino a una cascada de ellas.

## 1.2 METODOS CONVENCIONALES.

### 1.2.1 INTRODUCCION.

En este apartado se presenta una revisión de las técnicas y métodos más interesantes utilizados en el reconocimiento automático de formas. Para lo cual se ha realizado una extensa búsqueda de la bibliografía que en la última década se ha publicado sobre este tema.

Varios libros y artículos se han publicado sobre reconocimiento automático de formas, como podemos encontrar en los Proceedings of the International Joint Conference sobre Pattern Recognition y en International Systems, Man, y en Cybernetics Conference.

El estado del arte en reconocimiento de formas ha avanzado, desde el uso de primitivos esquemas para el reconocimiento de un limitado y estandar conjunto de formas, generadas mediante software, a la de sofisticadas aplicaciones técnicas para el reconocimiento de una gran variedad de formas reales, adquiridas mediante cámaras de video.

Las industrias de sistemas de visión artificial han optado por la definición de unas formas estandar que reducen la variación de las formas a fin de que los sistemas de visión

puedan obtener mejores resultados en el reconocimiento. Los sistemas de visión que se encuentran hoy día en la industria sólo pueden reconocer un tipo de piezas limitado. Los usuarios tienen que ajustarse al máximo a estos tipos predefinidos y proporcionar unas condiciones ambientales dentro de los límites permitidos con el fin de obtener una buena resolución.

### 1.2.2 TECNICAS DE RECONOCIMIENTO.

Las formas pueden ser reconocidas de diferentes maneras y a lo largo del tiempo los investigadores han desarrollado distintas técnicas que se aplica a la visión automática.

Las técnicas de reconocimiento varían mucho dependiendo de los descriptores elegidos para caracterizar las formas, el modo en que estos descriptores se obtienen y los métodos de clasificación empleados.

Presentamos a continuación varias de la distintas técnicas utilizadas en el reconocimiento de formas y se compara la bondad de cada una de estas técnicas basandose en los siguientes criterios.

- 1) Sensibilidad a la deformación de la imagen.
- 2) Facilidad de implementación de la técnica.

La sensibilidad a la deformación de la imagen de las formas se produce por los siguientes factores:

-Ruido. El cual causa que se pierdan trozos de líneas, que se unan piezas, aparezcan llenos los lazos, etc.

-Distorsión. Donde se incluyen variaciones locales, redondeo de las esquinas, alargamientos, ensanchamientos impropios, encogimientos, etc.

-Traslación. Desplazamiento de una forma y sus componentes.

-Rotación. Cambio en la orientación.

El ruido se introduce generalmente en el proceso de captura y digitalización de la imagen dependiendo de la luz, superficie de fondo y del contraste de la forma, las distorsiones son producidas principalmente por la cámara. Las traslaciones y rotaciones pueden resultar de la parte mecánica del sistema.

La implementación práctica de las distintas técnicas se analizan teniendo en cuenta los siguientes factores.

-Implementación de máscaras. Si es fácil o no la implementación de máscaras correspondientes a cada clase de patrón para hacer coincidir con las características obtenidas en una determinada técnica.

-Velocidad. La rapidez con que las formas son reconocidas.

-Complejidad. La dificultad y complejidad de implementar una técnica.

-Independencia. Si la técnica puede lograr características específicas sin utilizar técnicas suplementarias.

Todos estos factores dependen de la técnica utilizada, especialmente la velocidad depende fuertemente del hardware utilizado.

En la presentación de las distintas técnicas no se hace mención al hardware utilizado.

Las distintas técnicas pueden agruparse en tres grandes categorías:

- A -Características globales.
- B -Características locales.
- C -Características geométricas y topológicas.

#### A) CARACTERISTICAS GLOBALES

Esta categoría comprende técnicas las cuales extraen características que dependen de todos los puntos que están dentro de un rectángulo (llamado marco) que circunscribe a la pieza. Tales características no reflejan ninguna propiedad local, geométrica o topológica de la imagen.

### 1) Medidas de la coincidencia y correlación.-

Esta técnica toma como características el estado (blanco o negro) de todos los puntos que están dentro del "marco". Simplemente, mide la similitud entre la forma de entrada y las almacenadas como referencia mediante coincidencia y correlación de puntos o grupos de puntos dentro del marco.

### 2) Transformadas y desarrollo en series

Un problema que surge del hecho de tomar los puntos de la imagen como característica de la imagen, como en el caso anterior, es la gran dimensión del vector resultante de características. Se han realizado intentos para reducir la dimensionalidad y extraer características invariantes a algunas deformaciones globales (como traslaciones o rotaciones) mediante la utilización de transformadas y desarrollo de series. Entre las transformadas más utilizadas figuran las de Fourier, Walsh, Haar y Hadamard, y entre el desarrollo de series cabe destacar la de Karhunen-Loeve, Bessel y otros.

Las características globales pueden extraerse fácilmente y son esencialmente menos afectadas por los cambios locales.

Los métodos mencionados anteriormente se caracterizan por su baja sensibilidad al ruido. Además las transformadas y desarrollo en series permiten ciertas traslaciones o rotaciones. Sin embargo, un pequeño punto de suciedad o ruido puede desplazar el marco y afectar seriamente las posiciones de las características, y equivocar el reconocimiento. Así, el principal

inconveniente de las técnicas basadas en características globales son su dependencia con la posición y su alta sensibilidad a la distorsión, las cuales son fuentes de error usuales en el reconocimiento de formas.

Desde un punto de vista práctico, es relativamente fácil construir máscaras para estas técnicas; sin embargo el proceso de coincidencia incluye normalización en cuanto a tamaño y posicionamiento de la forma que puede resultar complejo. La velocidad de estas técnicas es lenta y a menudo no son autosuficientes y necesitan de otras técnicas complementarias para obtener los descriptores finales

### 3) Momentos

Los momentos de bajo orden de los puntos respecto a un centro, a menudo se elige el centro de gravedad de la forma, o un sistema predefinido de coordenadas, es otra de las técnicas utilizadas para caracterizar una forma.

### B) Características locales.

Otra forma de reducir la dimensión del conjunto de características es obtener características de la distribución de los puntos. Se han probado distintos tipos de distribuciones, correspondiendo a diferentes técnicas de reconocimiento.

1) Zonas.- El marco conteniendo la forma es dividido en varias zonas solapadas o no solapadas y la densidad de puntos en estas zonas forman y definen las características.

2) N-tuples.- La ocurrencia de elementos blancos o negros o la ocurrencia de estos elementos juntos son utilizados como características.

3) Característica Loci.- Para cada punto blanco en el fondo de la forma se generan dos vectores horizontal y vertical. El número de veces (un máximo de 2 se utilizaba para limitar la dimensión del conjunto de características). Las intersecciones de estas líneas con los segmentos de la forma son usadas como características.

4) Cruces y distancias.- Las características son obtenidas del número de veces que los segmentos son atravesados por vectores en determinadas direcciones, o la distancia de elementos o segmentos a un límite tal como el marco que contiene la forma.

Estas técnicas pretenden ser menos rígidas que los métodos de la sección anterior. Son tolerantes a distorsiones y a una cierta variación en el estilo, debido a que las características contienen una cierta información topológica de la forma. Esto es particularmente claro en el caso de características Loci, cruces y distancias.

En este grupo de técnicas, es relativamente difícil la realización de máscaras debido al gran número de posibles combinaciones de características y las variaciones de formas y

tamaños de los caracteres. Por otro lado, proporcionan una alta velocidad y baja complejidad de implementación.

### C) Características geométricas y topológicas

Estas técnicas están basadas en la extracción de características las cuales describen el interés geométrico o topológico del dibujo. Estas características pueden presentar propiedades globales y locales de la forma. Estas son sin duda alguna las técnicas más estudiadas por los investigadores; algunas de las características extraídas son:

- Línea y curvaturas en varias direcciones.
- Esquinas, intersecciones, bucles.

Estas propiedades en unión con análisis de contornos, relaciones entre líneas, propiedades de ángulos, son utilizadas para la caracterización de formas.

La principal ventaja de estas técnicas que utilizan propiedades geométricas y topológicas, frente a los otros métodos, es su alta tolerancia a distorsiones, así como permiten una cierta tolerancia a cierto grado de rotaciones y traslaciones. Esto explica que estas técnicas sean de las más estudiadas en la literatura.

Es más dificultoso la generación de máscaras para la técnica de clasificación. Pero una vez que éstas están implementadas, pueden procesar clusters a gran velocidad.

### 1.3 CARACTERIZACION DE IMAGENES.

La luz recibida de una escena por un sistema óptico da lugar a una imagen tri-dimensional. Esta imagen puede convertirse directamente en una señal eléctrica por un sensor, o puede ser fotográficamente registrada y posteriormente convertida. Matemáticamente, una imagen se define por una función  $f(x,y)$  de dos variables (las coordenadas en el plano de la imagen, correspondientes a las direcciones espaciales).

Los valores de la función son intensidades (brillos), o  $K$ -tuplas de valores de brillo en varias bandas espectrales. En el caso blanco y negro, los valores se llaman niveles de grises. Estos valores son reales, no-negativos y limitados. Esta función vale cero fuera de una región finita, ya que un sistema óptico tiene un campo visual limitado, de forma que la imagen es de tamaño finito; sin pérdida de generalidad, podemos suponer que esta región es rectangular. También podemos suponer que las funciones están analíticamente bien definidas, es decir, que son integrables y tienen transformada de Fourier, etc.

#### 1.3.1 MUESTREO UNIFORME Y CUANTIZACION.

De forma que el proceso computacional sea más fácil, una función de imagen  $f(x,y)$  debe digitalizarse tanto espacialmente como en amplitud. A la digitalización de las coordenadas espaciales  $(x, y)$  se le conoce como muestreo de imágenes,

mientras que a la digitalización en amplitud se le llama cuantización del nivel de gris.

Supongamos que aproximamos una imagen continua  $f(x,y)$  en muestras igualmente espaciadas colocadas en forma de un array  $N * N$  como se muestra en la siguiente ecuación, donde cada elemento del array es una cantidad discreta.

$$f(x,y) = \begin{pmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,1) & \dots & f(1,N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(N-1,0) & f(N-1,1) & \dots & f(N-1,N-1) \end{pmatrix} \quad (1)$$

La parte derecha de esta ecuación representa lo que comúnmente se llama imagen digital, mientras que a cada elemento del array se le llama elemento imagen (pixel). Es práctica común en el proceso digital de imágenes asignar tanto al valor de  $N$  como al número de niveles de grises, potencias enteras de dos; es decir:

$$N = 2^{*n} \quad \text{y} \quad G = 2^{*m}$$

donde  $m$  es el número de bits necesarios para codificar los  $G$  niveles de grises. Supondremos que los niveles discretos están igualmente espaciados entre 0 y  $L$  en la escala de grises. Según las ecuaciones anteriores el número de bits requeridos para almacenar una imagen digital es:

$$b = N * N * m \quad (2)$$

Por ejemplo, para una imagen de  $128 \times 128$  con 64 niveles de grises se requieren 98.304 bits.

Debido a que la ecuación 1 es una aproximación a una imagen continua, nos podemos preguntar cuantas muestras y niveles de grises son necesarios para una buena aproximación. La resolución, (es decir el grado de detalles discernibles) de una imagen depende fuertemente de  $N$  y de  $m$ ; cuanto mayores son éstos parámetros, más se aproxima el array digitalizado a la imagen original. Sin embargo la ecuación (2) muestra claramente el hecho de que el almacenaje y por consiguiente los requerimientos del proceso, aumentan rápidamente en función de  $N$  y  $m$ . Por todo esto, es de interés considerar los efectos que las variaciones en  $N$  y en  $m$  tienen sobre la calidad de la imagen. Es difícil de definir una "buena imagen" debido a que los requisitos de calidad varían según las aplicaciones.

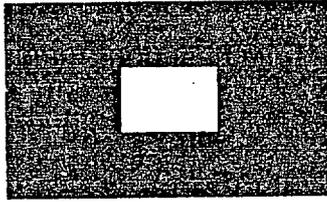
En la figura se muestra un ejemplo simple de imagen digital. La representación de la escena es una versión muestreada de la escena continua presente en el espacio objeto. Cada punto dentro de la representación digital corresponde a un área en el espacio objeto y a cada punto en la imagen digital se le asigna un valor digital que está relacionado con la intensidad de su área en el espacio objeto.

```

1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1
1 1 1 0 0 1 1 1
1 1 1 0 0 1 1 1
1 1 1 1 1 1 1 1
1 1 1 1 1 1 1 1

```

imagen



objeto

Se han tomado varias decisiones al generar la representación digital particular de la escena mostrada en la figura anterior. Una decisión concierne a la frecuencia de muestreo, puede observarse que ésta ha sido seleccionada de forma que el tamaño del objeto, aparezca en la imagen digital como un objeto de 2\*2 elementos. Una frecuencia más alta de muestreo produciría una matriz con más elementos que representaría al mismo cuadrado y una frecuencia menor significaría que el objeto no podría ser detectado en la representación digital de la escena.

La segunda decisión está relacionada con la representación de la intensidad dentro de la imagen muestreada. En la figura anterior se usa un único bit para representar la intensidad en cada posición muestreada. Las imágenes digitales en las cuales sólo se usa un bit para representar la intensidad son en principio poco interesantes, sin embargo hay un gran número de aplicaciones para este tipo de imágenes. Las imágenes de documentos a menudo se digitalizan usando un único bit para la intensidad, de forma que los caracteres tienen una intensidad de 0 y el fondo está digitalizado a una intensidad de valor 1. Este

tipo de imágenes se utilizan en muchas áreas incluyendo transmisión facsimile de textos y documentos, equipos de reconocimiento de caracteres ópticos y sistemas de almacenaje y recuperación de documentos digitalizados.

La siguiente figura contiene una representación digital diferente de la misma escena anterior. Se ha utilizado la misma frecuencia espacial de muestreo, pero el número de bits utilizado para representar la intensidad digital ha aumentado de 1 a 8. Pueden observarse varios efectos comunes que ocurren cuando se crea una representación digital muestreada de una escena continua. Los 4 pixels que representan el cuadrado negro uniforme tienen valores digitales de intensidad de 3,2,2,4. La mayoría de los sistemas que proporcionan una representación digital muestreada de una escena introducen un ruido aleatorio dentro de la imagen muestreada. Los sistemas que introducen ruido aleatorio provocan una fluctuación aleatoria alrededor de la intensidad digital media. Este efecto puede ser producido tanto por sistemas electrónicos como por procesos de conversión analógico digital efectuados al convertir la intensidad de la luz de entrada a un

255	255	255	255	255	255	255	255
255	255	255	126	132	255	255	255
255	255	119	3	2	120	255	255
255	255	123	2	4	119	255	255
255	255	255	121	118	255	255	255
255	255	255	255	255	255	255	255



imagen

objeto

El segundo efecto que puede observarse es la ausencia de una transición nítida en la imagen digital muestreada en los límites del cuadrado negro. Esta nueva figura representa de un modo más

real lo que ocurre cuando se muestrea un contorno nítido para crear una representación digital discreta de la escena. La resolución espacial dentro de una imagen digital muestreada será inferior que la resolución de la escena real continua.

En la figura la transición nítida de blanco a negro en la escena ocurre en la imagen digital sobre un rango de tres pixels. Una frecuencia mayor de muestreo proporcionaría una definición más concisa del límite de alto contraste. Aún habría pérdida de resolución, pero sería menor. Una frecuencia mayor, también generaría una mayor representación digital de la misma escena.

### 1.3.2 MUESTREO Y CUANTIZACION NO UNIFORME.

Para un valor fijo de  $N$  es posible en muchos casos mejorar la calidad de una imagen por medio de un esquema adaptativo, donde el proceso de muestreo dependa de las características de la imagen. En general, se requiere un muestreo fino en las cercanías de transiciones rápidas de niveles de grises, mientras que se puede emplear un muestreo ancho en regiones relativamente homogéneas. Sea por ejemplo, una imagen consistente en una cara sobre un fondo uniforme. Claramente se ve que el fondo aporta pocos detalles de información y es posible representarlo adecuadamente por un muestreo ancho. Por otra parte, la cara contiene más detalles considerables, y podemos usar en esta región las muestras adicionales no usadas en el fondo, mejorando de esta forma el resultado de la imagen total, particularmente si  $N$  es pequeño. En general, al distribuir las muestras debemos usar

una mayor concentración de estas en los límites de transición de niveles.

Cuando el número de niveles de grises es pequeño, a veces es factible usar un proceso de cuantización de niveles desigualmente espaciados. Puede utilizarse un método similar al anterior, sin embargo, debido a la baja capacidad del ojo al estimar tonalidades de grises cerca de cambios abruptos de niveles, la técnica en este caso es usar pocos niveles de grises en las cercanías de los límites. El resto de niveles pueden usarse en regiones donde las variaciones en los niveles de grises sean suaves, evitando ó reduciendo así los falsos contornos que a menudo aparecen en estas regiones si han sido cuantizados muy anchos.

**CAPITULO:2. FORMALISMOS PARA PROCESO DE IMAGENES.**

## 2.1 GENERALIZACION DE METODOLOGIAS EN PROCESO DE IMAGENES.

### 2.1.1 MARCO GENERAL EN PROCESO DE IMAGENES.

Se pretende plantear un marco conceptual y teórico cuya potencialidad permita formalizar un modelo general de proceso de imágenes.

Las herramientas conceptuales y formales que se utilicen para describir este marco teórico nos van a fijar los límites de validez de este modelo por lo que habría que poner especial énfasis y utilizar las herramientas formales más adecuadas y potentes.

Este marco teórico permite tener un estricto conocimiento del campo de validez, de una formulación concreta, determinando hasta donde alcanza la potencialidad de sus expresiones y los casos en los que no puede ser aplicado.

Definiremos el proceso de imágenes como un subsistema caracterizado por un espacio de entrada (I), un espacio de salida (O), y una estructura relacional entre ellos (R).

$$(I) \langle == \rangle (R) \langle == \rangle (O)$$

Estos espacios están definidos por la naturaleza del subsistema y por la instrumentación utilizada para registrar la actividad del subsistema considerado.

La estructura que une el espacio de entrada con el de salida define el nivel de proceso, y la función permitiendo describir los resultados y permitiendo realizar inferencias que estén acordes con los datos.

La semántica propia de un sistema puede repartirse entre los espacios de entrada, de salida y la estructura relacional, cuanto mayor sea la capacidad operativa de los símbolos codificados sobre las variables de entrada, menor será la complejidad del cálculo necesario para obtener el espacio de salida.

Esto nos obliga a realizar una buena elección de la naturaleza de los espacios de entrada y salida, así como del tipo de lenguaje usado para describir la estructura relacional entre ambos, si bien esta elección vendrá coaccionada por la tecnología utilizada en el subsistema.

En general, el espacio de entrada será un espacio multidimensional compuesto por todas las características independientes que son necesarias para definir el sistema bajo estudio. El espacio de las entradas viene determinado por la propia escena visual al excitar la instrumentación registradora de la escena. Esta multidimensionalidad en escenas estereoscópicas estará formada en el caso más amplio, por el espacio tridimensional, la dimensión tiempo y la dimensión color, o sea, un espacio de cinco dimensiones. En muchos casos prácticos esta multidimensionalidad se reduce, a si, eliminando el color

tendremos imágenes monocromáticas, si eliminamos la dimensión tiempo, tendremos imágenes estacionarias, si eliminamos una dimensión espacial tendremos imágenes planas (2-D). Sólo nos quedaremos con aquellas variables independientes que nos caracterizan nuestro caso particular de estudio.

La naturaleza del espacio de salida viene determinada por la caracterización del espacio de entrada y del nivel formal que tenga la estructura relacional, utilizada para describir el subsistema. La dimensionalidad del espacio de salida será igual o menor que el espacio de entrada, en muchos casos este espacio se reduce debido a que se realiza una codificación.

Por ejemplo una imagen bidimensional cromática (tres dimensiones) que al procesarla se obtiene una imagen monocromática (2 dimensiones), en la que el tono de gris tiene codificado un cierto color. Otro ejemplo de reducción de dimensionalidad es una imagen (3D) que al procesarla obtenemos una vista (2D).

La estructura relacional enlaza los espacios de entrada y salida. Esta relación tomando datos del espacio de entrada y en general, también del espacio de salida, producirá la salida correspondiente en un instante.

La elección de un lenguaje formal adecuado para esta estructura relacional es una tarea muy importante ya que además de limitar nuestro marco de actuación, nos va a facilitar o

complicar la interpretación de los resultados.

Existe una relación entre el aparato formal elegido, y el nivel de descripción del proceso. Así, si elegimos una formulación integral, podremos describir y realizar procesos de tipo lineal; para procesos no lineales deberemos elegir una estructura algorítmica que proporciona un nivel más alto de descripción.

En general, el espacio de entrada estará formado por un volumen (VI) de información, en el que estarán las dimensiones espaciales puras (imagen plana); si apilamos un conjunto de estas imágenes tendremos una película, donde la profundidad de apilamiento nos proporciona la dimensión tiempo. Podemos tener varias de estas películas correspondientes a distintas vistas, que nos darían otra dimensión espacial (imágenes 3D); podríamos añadir la dimensionalidad cromática si introducimos color; y así ir introduciendo más dimensiones que nos proporcionan los sensores; por lo que en general nuestro volumen de entrada será multidimensional o multiespectral, dependiendo de la aplicación concreta.

El trozo de película de entrada que interviene en la computación es lo que llamaremos "memoria de entradas".

Análogamente, el espacio de salida estará formado por un volumen (VO) formado con las distintas respuestas a lo largo de

un periodo de tiempo, que producirán una película de salida.

El volumen de salida utilizado en la computación es lo que denominamos la "memoria del sistema".

La estructura relacional está formada por un conjunto de elementos computacionales ( $E_1, E_2, \dots$ ), que alimentados con el volumen ( $V_I$ ) del espacio de entrada, y del volumen ( $V_O$ ), del espacio de salida, produce mediante su expresión relacional la imagen respuesta en un instante  $T$ .

En un caso particular puede suceder que el elemento computacional no utilice el volumen ( $V_O$ ) de salida; en este caso diremos que no existe realimentación del espacio de salida al espacio de entrada.

Este modelo general se irá simplificando en los casos particulares que estemos estudiando en cada momento. Si todos los elementos computacionales ( $E_i$ ) son iguales para todos los puntos de la imagen, existe invarianza espacial; también podemos encontrar invarianza temporal si la estructura que caracteriza a los elementos no varía con el tiempo.

Podemos realizar una clasificación de los procesos en lineales o no lineales, dependiendo de la estructura formal de la relación: si la herramienta formal elegida para la estructura relacional realiza cálculos de tipo lineal,  $y=ax+b$ , tendremos procesos lineales de imágenes; mientras que si se ha elegido una

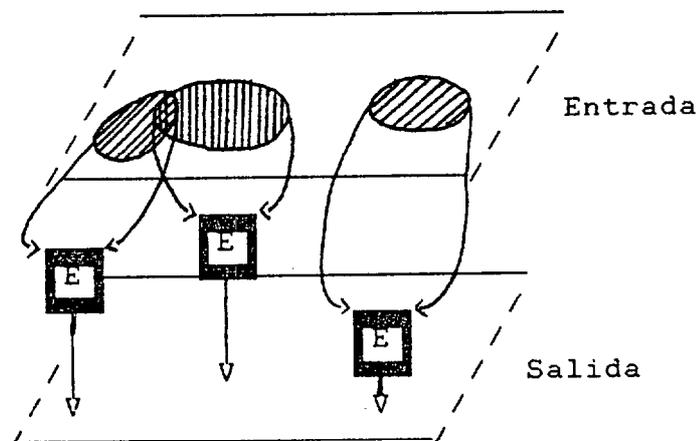
estructura no lineal para la relación, por ejemplo de tipo división o exponencial, tendremos procesos no lineales.

En muchos casos la complejidad del proceso será tal que se elegirá una estructura que llamamos "proceso por capas", para poder describir el sistema, donde cada capa está caracterizada por efectuar un subproceso específico, y de fácil formulación, de forma que la salida producida por una capa será la entrada de la siguiente capa, teniéndose una conexión en cascada.

Cada capa tendrá su propia y específica estructura relacional que caracterizará la función de la capa.

En este tipo de sistemas encuadraremos todos los procesos sobre imágenes con las dimensiones espaciales puras, en los que se ha prescindido de la dimensión tiempo.

Aquí el volumen de entrada (VI), se reduce a una sola imagen bidimensional discreta, el volumen de salida (VO) desaparece. La estructura relacional producirá un valor de salida mediante la imagen de entrada.



En lo que sigue, estudiaremos distintos procesos de imágenes bidimensionales espaciales no temporales, y siempre nos referiremos a este subgrupo de procesos, sin olvidarnos que existen los procesos temporales, cromáticos, 3D, y en general multiespectral.

Estudiaremos diferentes tipos de operaciones que pueden realizarse sobre imágenes digitales. Estas operaciones podemos clasificarlas atendiendo al número de puntos del espacio de entrada que intervienen en el cálculo de un punto del espacio de salida. Según este criterio tenemos los siguientes tipos de operaciones:

a) Operaciones puntuales:

El nivel de gris de salida de cada punto, depende sólo del nivel de gris de entrada en ese punto. Estas operaciones se usan extensamente en el tratamiento de escalas de grises y para la segmentación por clasificación de pixels. Puede haber más de una imagen de entrada; por ejemplo, podemos hacer la diferencia o

el producto de dos imágenes, punto a punto. En este caso, el nivel de salida en cada punto depende sólo del conjunto de niveles de entrada en el mismo punto.

b) Operaciones locales:

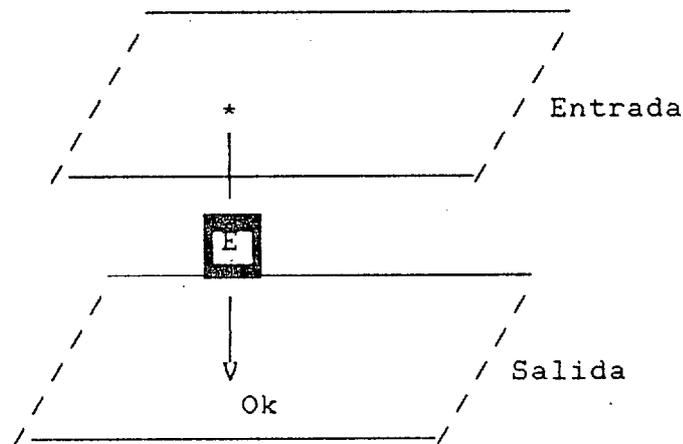
El nivel de salida en un punto depende sólo de los niveles de entrada en la vecindad del punto. Estas operaciones se usan para eliminación de borrosidad, limpieza de ruido, detección de flancos y detección de características locales entre otras aplicaciones.

c) Operaciones globales:

El nivel de salida en un punto, depende de todos los puntos del espacio de entrada, por lo que cada punto de salida contiene información relativa a la globalidad de los puntos. La principal aplicación de estas operaciones es en el campo del reconocimiento de formas, y transmisión de imágenes. Serán consideradas en la Sección 2.2 y 2.3.

### 2.1.2 TECNICAS PUNTUALES

Dentro de las técnicas puntuales, englobaremos todos aquellos procesos que se caracterizan porque un punto en el espacio de salida es función de un único punto del espacio de entrada



$$O_k = P(I_k)$$

$O_k$  = punto en el espacio de salida  
k

$I_k$  = punto en el espacio de entrada  
k

$P$  = proceso computacional de la estructura relacional

A continuación veremos distintos casos de procesos puntuales en el tratamiento de imágenes.

#### a) TRANSFORMACIONES EN LA ESCALA DE GRISES

Un procedimiento simple y efectivo para modificar el rango dinámico o el contraste en una imagen es alterar los valores de los niveles de grises, efectuando una transformación adecuada, lineal o no lineal, que dependerá del tipo de aplicación requerida, una transformación de una imagen  $f(x,y)$  puede expresarse formando una nueva función de imagen  $g(x,y) = P[f(x,y)]$  para cada  $(x,y)$  siendo  $P$  un proceso puntual.

Primeramente consideraremos procesos invariantes en el espacio, esto es, P no varía con la posición.

#### b) TRANSFORMACION LINEAL

Un método simple y útil podría ser realizar algún tipo de escalado. En el caso de que una imagen  $f(x,y)$ , no ocupe el rango total de niveles permitidos  $[z_1, z_k]$ , es decir si  $a \leq z = f(x,y) \leq b$  para todo  $(x,y)$ , siendo  $[a,b]$ , un subintervalo de  $[z_1, z_k]$ , la siguiente transformación lineal:

$$z' = g(x,y) = \frac{(z - a)}{(b - a)} (z_k - z_1) + z_1$$

expande los valores de la imagen sobre la totalidad del rango.

Este tipo de transformación suele ser necesario si la imagen original está subexpuesta, o bien, si el sistema de barrido no está ajustado de forma que ocupe la totalidad del rango permitido.

El realce lineal del contraste es adecuado cuando las imágenes tienen histogramas Gaussianos o casi Gaussianos, donde todos los valores de la intensidad caen dentro de un subconjunto estrecho del rango permitido de intensidad.

#### c) TRANSFORMACION LINEAL POR TRAMOS

La transformación anterior no es siempre la más adecuada, para imágenes caracterizadas por histogramas bimodales. De un modo mas general, podemos ampliar ciertas regiones en la escala

de grises a costa de comprimir otras, destacando detalles en las regiones ampliadas sin preocuparnos de la pérdida de información en las regiones comprimidas.

Podemos implementar cualquier transformación cuadrática, logarítmica, o completamente arbitraria con la única restricción que el resultado esté en el rango de valores permitidos.

También existen aplicaciones en procesos de imágenes donde conviene utilizar funciones de transferencia monótonas decrecientes o no monótonas. Por ejemplo si un sistema de presentación es altamente no lineal, en la parte oscura límite del rango de niveles, será útil una función de contraste inversa que transforme las regiones de baja iluminación de la imagen de entrada en valores mas bajos de intensidad.

#### d) UMBRAL BIDIMENSIONAL

Podemos pensar en otro tipo de procesos puntuales donde la estructura relacional no tenga invarianza espacial, esto es, la función  $P$  varía con la posición. Estudiemos un método, desarrollado por nuestro grupo, para uniformizar ruido de fondo, llamado umbral bidimensional.

En el caso que trabajemos con una distribución de ruido de fondo más o menos uniforme, podemos calcular un único umbral (de forma automática) para todos los puntos de la imagen.

No obstante hay casos en que dicha distribución de fondo no tiene un comportamiento uniforme. Luego en un caso más general, como el que acabamos de citar, habría que pensar en encontrar un umbral automático adaptativo a cada uno de los pixels de la imagen digitalizada.

En nuestro caso concreto, lo que hacemos es aproximar por el método de los mínimos cuadrados, la imagen a un plano inclinado, en el que la coordenada Z será igual al nivel de gris de cada punto y las coordenadas X e Y serán las correspondientes a cada pixel. Calculamos el valor mínimo de la coordenada Z del plano y ajustamos la distribución de ruido de fondo a un plano horizontal de valor  $Z = Z_{min}$ .

Este método de umbralizado automático tiene una especial utilidad en el caso de imágenes de iluminación no uniforme procedente de una fuente cercana, o en general creciente según un determinado sentido. Se utiliza un plano por ser la figura tridimensional más simple que pueda representar de forma aproximada la variación de los niveles del fondo de la imagen fundamentalmente a una iluminación no uniforme.

#### e) TECNICAS BASADAS EN LA MODIFICACION DEL HISTOGRAMA

Un caso especial e importante de proceso puntual, consiste en realizar una modificación del histograma. Este proporciona una descripción global de la apariencia de una imagen. Las técnicas de modificación del histograma consisten en transformar la imagen

original, de forma que el histograma de la imagen final tenga alguna forma predefinida. Discutiremos a continuación distintos procesos basados en el histograma.

#### f) ECUALIZACION DEL HISTOGRAMA

Supongamos que los valores de los pixels han sido normalizados de forma que  $0 \leq z \leq 1$ . Para cualquier valor de  $z$  en el intervalo  $[0,1]$ , centraremos nuestra atención en transformaciones de la forma  $z' = T(z)$ , las cuales producen un nivel  $z'$  para cada valor  $z$  de un pixel en la imagen original. Se supone que la función de transformación dada satisface las condiciones:

a)  $T(z)$  es una función univaluada y monótona creciente en el intervalo  $0 \leq z \leq 1$  y

b)  $0 \leq T(z) \leq 1$  para  $0 \leq z \leq 1$ .

La condición a) asegura el orden desde negro a blanco en la escala de grises, mientras que la condición b) garantiza un mapeado consistente con el rango de valores permitidos.

La transformación inversa de  $z'$  en  $z$ , se representa por:

$$z = T^{-1}(z'), \quad 0 \leq z' \leq 1$$

donde se supone que  $T^{-1}(z')$  también cumple las condiciones a y b con respecto a la variable  $z'$ .

Los niveles de grises en una imagen son cantidades aleatorias en el intervalo  $[0,1]$ . Suponiendo por el momento que son variables continuas, los niveles de grises originales y transformados están caracterizados por sus funciones densidades de probabilidad  $Pz(z)$  y  $Pz'(z')$  respectivamente.

A partir de la teoría elemental de probabilidad, se deduce que si  $Pz(z)$  y  $T(z)$  son conocidas y  $T^{*-1}(z')$  satisface la condición a), entonces la función densidad de probabilidad de los niveles de grises transformados, está dado por la relación:

$$Pz'(z') = \frac{Pz(z)}{\left| \frac{dz}{dz'} \right|_{z=T^{-1}(z')}}^{-1}$$

Si escogemos una función de transformación igual a la función de distribución cumulativa de  $z$ , es decir:

$$z' = T(z) = \int_0^z Pz(w) dw \quad 0 \leq z \leq 1$$

Tenemos que  $dz'/dz = Pz(z)$  y sustituyendo en la ecuación anterior:

$$Pz'(z') = \frac{Pz(z)}{Pz(z)} = 1 \quad 0 \leq z' \leq 1$$

La cual es una densidad uniforme en el intervalo de definición de la variable transformada  $z'$ . Hay que notar que este resultado es independiente de la función inversa de

transformación. Esto es importante porque no siempre es fácil obtener  $T^{-1}(z')$  analíticamente.

En términos de realce, este método de ecualización implica un incremento en el rango dinámico de los pixels que puede tener un efecto considerable en la apariencia visual de una imagen.

Aunque este método es bastante útil, no permite una aplicación interactiva, ya que su capacidad está limitada a la generación de un único resultado: una aproximación a un histograma uniforme.

#### g) ESPECIFICACION DIRECTA DEL HISTOGRAMA

A veces es necesario poder especificar histogramas particulares interactivamente, capaces de destacar ciertos rangos de niveles de grises en una imagen. Sean  $P_z(z)$  y  $P_u(u)$  las funciones densidad de probabilidad original y especificada respectivamente.

El algoritmo puede describirse como sigue:

- 1) Ecualizar los niveles de la imagen original:  $z'=T(z)$ .
- 2) Especificar la función de densidad deseada y obtener la transformación ecualizada de la misma:  $u'=G(u)$ .
- 3) Aplicar la transformación inversa,  $u=G^{-1}(z')$  a los niveles obtenidos en el paso 1.

Este procedimiento, puede expresarse en una única transformación a partir de la imagen original:

$$u = G^{-1} [T(z)]$$

La principal dificultad al aplicar el método anterior al realce de imágenes, consiste en ser capaz de construir un histograma significativo. Existen dos soluciones al respecto: la primera es especificar una función particular de densidad de probabilidad (Gaussiana, Rayleigh, logarítmica ...) y luego formar un histograma digitalizando la función dada, la segunda solución es especificar un histograma de forma arbitraria pero controlable, por ejemplo formando una cadena de líneas rectas conectadas. Después de obtener una forma determinada, se digitaliza y se normaliza al área unidad.

#### h) METODO DE APROXIMACION A UNA CUARTICA

En el caso de imágenes que presenten histogramas bimodales (objetos sobre un fondo), un método práctico para la eliminación de ruido de fondo consiste en aproximar la distribución del histograma, (dado su tipismo en este tipo de imágenes) a una curva de cuarto orden o cuártica cuya expresión general es:

$$Y(X) = AX^4 + BX^3 + CX^2 + DX + E$$

La aproximación la hacemos por el método de los mínimos cuadrados. En principio se toma como valor mínimo de dicha curva, el valor del umbral. Pero la experiencia nos demuestra en las imágenes procesadas que en dicho caso el mínimo queda excesivamente desplazado hacia el máximo del histograma

correspondiente a la información de la imagen, dando lugar a una pérdida de la misma después del procesado.

Para solventar este inconveniente se vió que si se elige el umbral como la media aritmética de las raíces de la curva (máximos y mínimos) no se produce dicha pérdida de información, debido a que el umbral se desplazaba hacia el máximo correspondiente al ruido de fondo. Por otra parte este procedimiento nos facilitaba el cálculo del umbral por programa ya que si el coeficiente del término de mayor grado de la curva es la unidad, entonces, el segundo coeficiente nos da la suma de las tres raíces (máximos y mínimos).

#### i) METODO DE APROXIMACION A DISTRIBUCIONES GAUSSIANAS

Como hemos apuntado anteriormente, en el caso que tengamos una cierta uniformidad en la distribución del fondo de la imagen, la curva del histograma generalmente presenta dos máximos separados por un valle.

Los dos máximos se asemejan con bastante fidelidad a distribuciones Gaussianas.

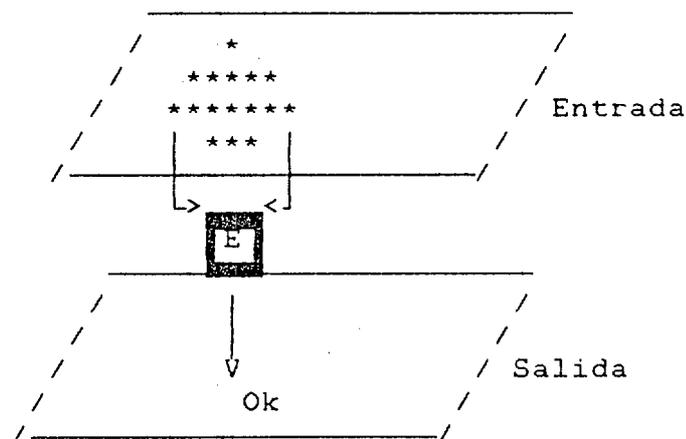
El método consiste en calcular un valor de umbral por debajo del cual, se encuentra el ruido de fondo y por encima, la información de la imagen.

En este caso el cálculo del valor del umbral se hace aproximando el primer pico (correspondiente al fondo) a una curva

de Gauss y calculando la desviación típica de dicha distribución. El valor del umbral entonces vendría dado por la suma de la desviación típica de dicha distribución con el valor de la mediana de la misma, que coincidirá de una forma bastante aproximada con el valle del histograma.

### 2.1.3 TECNICAS LOCALES

Con este nombre englobaremos todos aquellos procesos en los que el nivel de salida en un punto depende de las características locales de una región del espacio de entrada.



$$Ok = f(Ia, Ib, Ic, \dots In)$$

donde:

Ok = respuesta en un punto del espacio de salida

Ia, Ib, Ic, ... In = puntos de una región del espacio de entrada.

Todos los procesos en los que la estructura formal se puede implementar mediante una convolución, caen dentro de este grupo.

Típicos procesos que utilizan técnicas locales son los filtros espaciales.

Una imagen contiene información en un gran rango de frecuencias espaciales. Las transiciones graduales en los niveles de grises dentro de una imagen pueden interpretarse como componentes frecuenciales bajas. Variaciones rápidas del contraste, representan información correspondientes a altas frecuencias. Las texturas son un ejemplo de información de alta frecuencia.

Es posible realizar un análisis de las frecuencias espaciales presentes en una imagen digital utilizando distintas técnicas de filtrado. Veamos a continuación distintos tipos de ellos.

#### a) FILTROS PASO BAJO

Este tipo de filtros se caracteriza por dejar pasar sólo las bajas frecuencias espaciales de una imagen de entrada, esto es, sólo pasarán a la salida variaciones lentas de tonos de grises.

Un filtro paso bajo que sea lineal e invariante espacialmente, puede implementarse mediante una convolución.

Algunos ejemplos de núcleos de convolución para este tipo de filtro son:

1/9	1 1 1	1 1 1	1/10	1 2 1	1/(2+b)**2	1 b 1	b b*b b	1 b 1
	1 1 1	1 1 1		1 1 1		1 b 1	b b*b b	1 b 1

En el caso que sea necesario utilizar máscaras más grandes, es preferible realizar las operaciones de convolución indirectamente en el dominio de Fourier. El fundamento de las técnicas en el dominio frecuencial es el teorema de convolución.

#### b) PROMEDIADO

Una imagen puede estar sujeta a ruido e interferencias procedentes de varias fuentes, como pueden ser: ruido de los sensores eléctricos, ruido del granulado fotográfico y errores en los canales de transmisión. Una manera de eliminar o suavizar efectos anómalos es aplicar técnicas "ad hoc". Normalmente, el ruido que aparece en una imagen se manifiesta en forma de variaciones discretas aisladas de pixels que suelen no estar espacialmente correlacionadas. Esta observación es la base de muchas de los algoritmos de eliminación de ruido.

Un método elemental de suavizar el ruido consiste en realizar algún tipo de promediado sobre la imagen. Supongamos que los valores del ruido en cada punto sean muestras independientes escogidas de una distribución de valor medio cero y desviación standar. Supongamos también que estamos promediando sobre un conjunto de puntos cuyos niveles de grises son  $Z_1 \dots Z_n$  y sean los valores del ruido en esos puntos  $W_1 \dots W_n$ . Entonces, al promediar  $(Z_1 + \dots + Z_n)/n + (W_1 + \dots + W_n)/n$  el segundo término

puede considerarse como una muestra de una variable aleatoria de media  $\theta$  y desviación standar. De ésta forma el proceso de promediado ha reducido la amplitud de las fluctuaciones del ruido.

Podemos tratar de reducir el nivel de ruido mediante un promediado local, que consiste en ir reemplazando secuencialmente cada pixel en la imagen original por la intensidad promedio de sus vecinos en algún entorno del punto (incluyendo el propio punto). El tamaño del entorno a definir dependerá del grado deseado de reducción.

La dificultad básica de esta técnica de reducción de ruido es que tiende a degradar la imagen. Un método alternativo para evitar efectos de degradación consiste en realizar el promediado solamente en ciertos puntos seleccionados con algún procedimiento de umbralizado.

Podríamos formar  $g(x,y)$  según el siguiente algoritmo:

```
SI  $|f(x,y)-p(x,y)| > T$  ENTONCES  
     $g(x,y)=p(x,y)$   
SI NO  
     $g(x,y)=f(x,y)$   
FIN SI
```

donde  $p(x,y)$  es el promedio local en un entorno del punto  $(x,y)$  y  $T$  un umbral específico  $> 0$ .

Otra posibilidad es aplicar a la imagen un operador detector de bordes (o líneas) y promediar sólo en aquellos puntos en los que no están presentes estas características.

También podemos "pesar" los vecinos apropiadamente, p.e. podemos pesar cada vecino según la proximidad de su nivel de gris al punto dado.

### c) FILTRO MEDIANA

El filtro mediana es una técnica "ad hoc" que intenta resolver el conflicto entre eliminación de ruido y degradación de bordes. Se realiza desplazando una ventana predefinida sobre los puntos de una imagen y remplazando el punto central de la ventana por la mediana de los valores originales dentro de la ventana. Esto proporciona una imagen de salida, más suavizada que la imagen original.

El procedimiento clásico de suavizado de imágenes es usar un filtro lineal paso bajo y en muchos casos resulta ser el procedimiento más adecuado. Sin embargo, en ciertas situaciones el filtro mediana proporciona mejores resultados, siendo sus principales ventajas:

- 1) El filtro mediana preserva los bordes nítidos, mientras que un filtro lineal paso baja los degrada.

- 2) El filtro mediana es una técnica muy eficaz para eliminar el ruido impulsivo.

La mediana de una secuencia discreta  $a_1, a_2, \dots, a_n$ , para  $n$  impar es aquel valor de la secuencia, para el cual hay  $(n-1)/2$  elementos inferiores o iguales en valor y  $(n-1)/2$  elementos mayores ó iguales en valor.

En ciertos casos para preservar las líneas delgadas y no recortar esquinas es factible usar una ventana con forma de cruz.

Por último señalar que los filtros mediana son no lineales, por lo que su estructura formal será de tipo algorítmico.

#### d) FILTROS PASO ALTO

Los filtros paso alto se caracterizan por filtrar todas las frecuencias espaciales que no sean altas, esto es dejarán pasar los cambios bruscos en intensidad, por lo que se aplicarán especialmente en todos aquellos casos en que se quiera extraer o resaltar bordes en la imagen.

En el caso de que la operación de filtrado sea lineal e invariante respecto a traslaciones, este se puede implementar mediante una convolución.

Algunos núcleos típicos de convolución para filtros paso alto son los sugeridos por los mecanismos ON-OFF.

-1	-1	-1	1	-2	1
-1	9	-1	-2	5	-2
-1	-1	-1	1	-2	1

e) REALCE DE CONTRASTE POR DIFERENCIACION

Al promediar, es decir al realizar una operación de integración, se produce una degradación de bordes; esto sugiere que si realizamos operaciones de diferenciación podremos destacar ciertas características degradadas. El método diferencial más frecuentemente usado en aplicaciones de proceso de imágenes es el gradiente, que se define por la expresión:

$$\bar{V} f(x,y) = \frac{\partial f}{\partial x} \bar{i} + \frac{\partial f}{\partial y} \bar{j}$$

Un modo de implementar digitalmente la operación del gradiente, es mediante convoluciones con núcleos del tipo:

$$\begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix} \quad \begin{vmatrix} 1 & 1 \\ -1 & -1 \end{vmatrix}$$

Roberts

Lo importante en todas las posibles aproximaciones, es que el valor del gradiente sea proporcional a las diferencias entre los niveles de grises de pixels adyacentes. De esta forma, cabe esperar que el gradiente asuma valores relativamente grandes en bordes prominentes de una imagen y pequeños valores en aquellas regiones de cierta uniformidad, siendo cero, en regiones cuyo nivel de gris es constante. De la misma manera que en el caso del promediado, es posible combinar el gradiente con técnicas de umbralizado según la aplicación requerida.

Otro operador diferencial apropiado para extraer bordes de

una imagen es la Laplaciana, que puede implementarse mediante un núcleo del tipo:

$$\begin{array}{ccc} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{array}$$

Otra posibilidad es restarle a una imagen degradada un múltiplo de su Laplaciana, esto se puede realizar mediante una convolución con núcleo:

$$\begin{array}{ccc} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{array}$$

Se pueden definir Laplacianas digitales alternativas, usando diferentes entornos de vecinos o usando promedios pesados sobre los entornos.

#### f) PROCESOS ESPACIALES NO LINEALES DIVISORIOS

Este método consiste en definir dos zonas alrededor del punto a procesar, una zona más próxima al punto la denominamos centro C y otra zona que envuelve a esta, periferia P. La estructura que define este proceso consiste en aplicar el siguiente algoritmo: Se define un proceso que actúa en cada zona, sea éste la suma de todos los niveles de grises de la zona, al aplicar este proceso en el centro y en la periferia obtenemos los valores  $S_c$  y  $S_p$  respectivamente. Cada punto del espacio de salida se obtiene al aplicar la relación divisoria:

$$O(x,y) = \frac{a \cdot S_c}{1 + b \cdot S_p}$$

a y b son dos constantes de ajuste

Un umbralizado puede añadirse a la relación anterior para eliminar efectos de ruido.

Con la aplicación de este método se consigue una drástica detección de contrastes, comportándose como un filtro paso alto.

#### g) FILTRADO HOMOMORFICO

Un método útil en algunas ocasiones para reducir efectos de sombreado debido a las variaciones de iluminación, sin degradar detalles en una imagen, es realizar un filtrado paso alto conjuntamente con una transformación logarítmica de la escala de grises. Este método está basado en un modelo sobre imágenes, según el cual el nivel de gris  $f(x,y)$  puede expresarse en términos de sus componentes de iluminación y reflexión según la relación:

$$f(x,y) = i(x,y) * r(x,y)$$

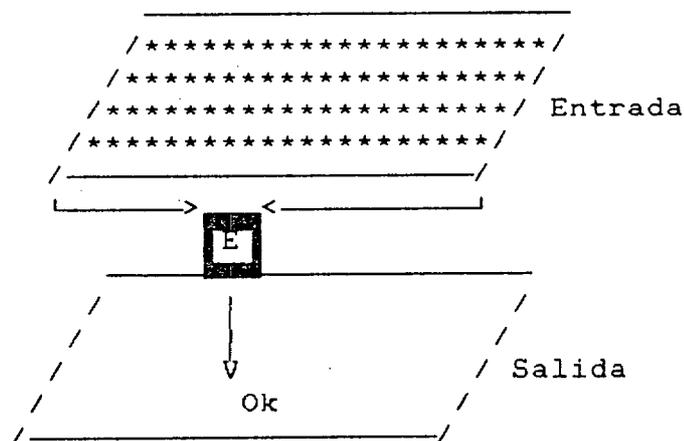
Donde la componente de iluminación  $i(x,y)$  está formada principalmente por frecuencias espaciales bajas y la componente de reflexión  $r(x,y)$  por las frecuencias altas. Debido a que las funciones  $i$  y  $r$  están combinadas multiplicativamente, no podemos reducir los efectos de la iluminación sobre la imagen simplemente destacando las altas frecuencias. Sin embargo, realizando una transformación logarítmica:

$$\log f = \log i + \log r$$

Tendríamos una combinación aditiva, y podríamos por tanto aplicar un filtro paso alto, para disminuir la contribución de logi y aumentar la correspondiente a logr, y luego tomar antilogaritmos para obtener la nueva versión filtrada de la imagen. De esta forma comprimimos el rango dinámico y simultaneamente realizamos el contraste.

#### 2.1.4 TECNICAS GLOBALES

En este grupo de técnicas englobaremos aquellas que vienen caracterizadas porque un punto  $O_k$  en el espacio de salida es función de todos los puntos del espacio de entrada, esto implica que cada punto resultado tiene información global de toda la imagen.



$$O_k = f(I)$$

$O_k$  = es un punto resultado en el espacio de salida

$I$  = todos los puntos del espacio de entrada

$f$  = es la estructura relacional que une el espacio de entrada con el espacio de salida.

Ejemplos de este tipo de técnicas son las transformaciones integrales, que por su importancia, pasamos a comentar en la siguiente sección.

## 2.2 DESCRIPCION Y PROPIEDADES DE TRANSFORMACIONES GLOBALES.

### 2.2.1 INTRODUCCION.

Una imagen cualquiera formada de puntos discretos encierra, en el sentido de la teoría de la información, una cantidad finita de información. Esta cantidad de información es muy grande, puesto que la imagen contiene un número elevado de puntos. El problema que se presenta entonces es el de guardar sólo lo esencial de la información y de encontrar las técnicas que permiten definir y de elegir este esencial.

Un método apropiado para esta selección es utilizar la transformación de Karhunen-Loeve. Pero presenta inconvenientes debidos a la cantidad de cálculos numéricos que necesita. Esta situación ha conducido al empleo de transformaciones menos óptimas, pero más adaptadas al cálculo en el ordenador, como son aquellas transformaciones ortogonales, que presentan un algoritmo rápido de cálculo. Entre estas transformaciones, tres son las más comunes: las transformadas de Fourier, de Hadamard y de Haar.

En los párrafos siguientes, vamos a comentar la utilización de distintas transformaciones aplicadas a las imágenes para reducir la información necesaria para su reconocimiento.

Los resultados que preceden pueden resumirse en la siguiente situación:

a) Es necesario tener una malla espacial fina y una cuantificación apretada de los niveles de gris para obtener una buena imagen.

b) Estas dos condiciones conducen a tratar una cantidad de información muy grande.

c) La utilización de las informaciones estadísticas ligadas a las características particulares de las imágenes tratadas permiten reducir la cantidad de información necesaria para reconocer una imagen.

La transformación de Karhunen-Loeve permite realizar esta reducción de la cantidad de información. Antes de describirla, es preciso definir la información transmitida por una variable aleatoria en el dominio estadístico.

#### Cantidad de información transmitida por una variable.

Sea una variable aleatoria  $X$  que puede representar, por ejemplo, la intensidad del gris en un punto determinado de una serie de imágenes. Esta variable posee una cierta función de distribución  $f(x)$ . Los primeros momentos de la función de distribución son el valor medio

$$x' = \int_{-\infty}^{+\infty} x f(x) dx$$

ya la varianza:

$$x'^2 = \int_{-\infty}^{+\infty} (x-x')^2 f(x) dx$$

La cantidad de información aportada por una variable "x" está relacionada directamente a la varianza de esta variable. Cuando la varianza es grande, implica que la cantidad de información es grande.

Consideremos, dos variables "x" e "y" de las que nosotros conocemos las medias y de las cuales una tiene una varianza pequeña ("x") y la otra una varianza elevada ("y"). El conocimiento de una realización "xi" de la variable "x" no aporta gran información, puesto que nosotros sabemos que este valor "xi" es muy vecino de la media "x" conocida, a consecuencia de la pequeña varianza. Por contra, el conocimiento de una realización "yi" de esta variable suministra una información mucho más importante.

La transformación de Karhunen-Loeve utiliza este resultado reemplazando los valores iniciales de las intensidades de gris de la imagen por un conjunto de variables independientes que son clasificadas después por varianzas decrecientes. La transformación conserva la cantidad total de información

contenida en la imagen. Eliminando las variables, cuyas varianzas son más pequeñas, se puede reducir el número total de variables a transmitir no reduciendo más que la mínima cantidad total de información.

### 2.2.2 TRANSFORMACION DE KARHUNEN-LOEVE.

Esta transformación está basada sobre los vectores y los valores propios de la matriz de correlación de la imagen original o de un conjunto de imágenes. Esta matriz de dimensiones  $N \times N$ , para una imagen cuadrada de  $N \times N$  tiene por elementos los valores medios:

$$[R] = E (z_i \cdot z_j) \quad i = 1, 2, \dots, N ; j = 1, 2, \dots, N$$

donde  $z_i$  y  $z_j$  son los valores de gris de los puntos "i" y "j" de la imagen.

Esta matriz puede determinarse experimentalmente a partir de las muestras de las clases de las imágenes a tratar y también se puede obtener de las expresiones analíticas deducidas de diferentes modelos. Estos modelos son, en general, invariantes para las traslaciones, es decir, tal que:

$$R(x_1, y_1, x_2, y_2) = R(dx, dy)$$

con

$$dx = x_1 - x_2$$

$$dy = y_1 - y_2$$

Las formas de R pueden ser, por ejemplo,

$$R(dx, dy) = e^{-[A dx + B dy]}$$

$$R(dx, dy) = e^{-AV dx^2 + dy^2}$$

$$R(dx, dy) = e^{-V a^2 dx^2 + 2bdxdy + c dy^2 + d}$$

Correspondiendo a los contornos de isocorrelación alrededor de un punto que son respectivamente unos paralelogramos, unos círculos y unas elipses.

A partir de la matriz de correlación [R], se define la matriz de transformación [A] tal que

$$[A]^T [R] [A] = \begin{bmatrix} l_1 & & & & \\ & l_2 & & & \\ & & \ddots & & \\ & & & l_{N-2} & \\ & & & & l_N \end{bmatrix}$$

donde  $l_1 \Rightarrow l_2 \dots \Rightarrow l_N$  son los valores propios de [R] puestos en un orden de valores decrecientes.

[A] está formada por los vectores propios de la matriz de correlación y es pues una matriz ortogonal, puesto que los vectores propios, para valores propios diferentes son

ortogonales. La transformación de Karhunen-Loeve de una imagen definida por el vector [f] de la serie de los grises de puntos es:

- en el sentido directo  $[F] = [f] [A]$
- en el sentido inverso  $[f] = [F] [A]**T$

La transformación conserva completamente la información de la imagen inicial. Los valores propios  $\lambda_1, \dots, \lambda_{N**2}$  corresponden a las varianzas de las nuevas variables obtenidas después de la transformación. Se puede reducir la cantidad de información a guardar para el reconocimiento de la imagen con una degradación lo más pequeña posible, utilizando sólo los M primeros valores propios de los  $N**2$  valores obtenidos. Se muestra que en este caso el error cuadrático medio que existe entre la imagen reconstruida con estos M valores y la imagen inicial es:

$$E = \sum_{K=M+1}^N \lambda_K^2$$

La transformación de Karhunen-Loeve es, pues, la que minimiza el error en el sentido de los mínimos cuadrados.

### 2.2.3 INCONVENIENTES DE LA TRANSFORMACIÓN DE KARHUNEN-LOEVE.

La transformación de Karhunen-Loeve permite utilizar lo mejor posible en el sentido de los mínimos cuadrados la información estadística que se puede obtener sobre la imagen estudiada. Presenta, sin embargo, tres inconvenientes de naturaleza diferente.

1. La determinación de la matriz de correlación y de la matriz de transformación conduce a cálculos con matrices  $N \times N$ . Para una imagen razonable con  $N=256$ , se tiene un tratamiento con matrices de  $65.536 \times 65.536$ , lo que posee serios problemas de análisis numérico y de cálculo. Se desconocen, en general, métodos de cálculo rápido.

2. El criterio utilizado para la reducción de la información es un criterio de mínimos cuadrados. Este criterio no es utilizable en muchos tipos de imágenes, porque las funciones realizadas por el cerebro para juzgar la degradación de una imagen por comparación a otra no se basan en los mínimos cuadrados.

3. El cálculo preciso de la matriz de transformación supone un conocimiento completo de la matriz de correlación, que es muy difícil de obtener.

Por estas diferentes razones, la transformación de Karhunen-Loeve es poco utilizada. Se la reemplaza, en general, por diferentes transformaciones ortogonales, que son, sin duda, menos óptimas, pero permiten cálculos mucho más rápidos sobre matrices de tamaño razonable.

#### 2.2.4 TRANSFORMACIONES RAPIDAS.

Para una imagen digitalizada de  $N \times N$  puntos, hemos visto

que es preciso tratar con matrices de  $N \times N$  utilizando la transformación de Karhunen-Loeve. Existe un conjunto de transformaciones menos óptima para la utilización de la información estadística contenida en una imagen, pero presentan propiedades que las hacen mas fácilmente aplicables a los cálculos numéricos.

#### 2.2.4.1 Definiciones.

La imagen digitalizada se considera como una matriz de  $N \times N$  puntos colocados en  $N$  filas y  $N$  columnas, el valor de cada punto representa el brillo de cada elemento correspondiente. Las transformaciones de esta imagen conducirán también a tablas de  $N \times N$  puntos.

Una transformación bidimensional se representa por una expresión:

Donde:

$$F(u, v) = \sum_x \sum_y f(x, y) a(x, y, u, v)$$

$f(x, y)$  corresponde a la imagen original;

$F(u, v)$  es la imagen transformada;

$a(x, y, u, v)$  es el núcleo de la transformación.

La transformación inversa, si existe, es:

$$f(x, y) = \sum_u \sum_v F(u, v) b(x, y, u, v)$$

El núcleo de la transformación se dice que es separable si

$$a(x, y, u, v) = a_x(x, u) a_y(y, v)$$

En este caso, la transformación puede efectuarse en dos tiempos; una vez sobre las líneas y otra sobre las columnas.

$$f_1(u, y) = \sum_x f(x, y) a_x(x, u)$$

y después

$$F(u, v) = \sum_y f_1(u, y) a_y(y, v)$$

La transformación es simétrica si  $a_x$  es idéntica a  $a_y$ .

Esta propiedad se cumple en general. Se pueden encontrar, sin embargo, ciertos casos en donde las transformaciones son diferentes. En el tratamiento de ciertas imágenes de radar, la transformada de "x" puede ser una transformada de Fourier y la de "y" una transformada de Fresnel. Para una transformación, supuesta separable y simétrica, se puede escribir, utilizando notación matricial:

$$[F] = [ [f] [A] ]^T [A]$$

Para la transformada directa,  $[A]^T$  es la matriz traspuesta de  $[A]$ . La transformada inversa se efectúa con la matriz  $[B]$  tal que:

$$[B] = [A]^{-1}$$

Se dice que la matriz  $[A]$  es unitaria si

$$[A]^{-1} = [A^*]^T$$

$[A^*]$  es la matriz conjugada de la matriz  $[A]$

Para una matriz unitaria real

$$[A]^{-1} = [A]^T$$

en una matriz ortogonal.

La transformación puede ser interpretada también como una descomposición espectral generalizada. En efecto, para unos valores dados  $u_0$  y  $v_0$  de las variables del espacio transformado, se tiene:

$$F(u_0, v_0) = \sum_x \sum_y f(x, y) a(x, y, u_0, v_0)$$

que es el producto escalar de la función a transformar  $f(x, y)$  por la función  $a(x, y, u_0, v_0)$  definido sobre el mismo espacio.  $F(u_0, v_0)$  representa la proporción de la función  $a(x, y, u_0, v_0)$  que está presente en la función  $f(x, y)$ .

En las transformaciones consideradas aquí se supondrá que las funciones  $a(x, y, u_0, v_0)$  son ortogonales, es decir:

$$\sum_x \sum_y a(x, y, u_i, v_i) a(x, y, u_j, v_j) = \delta_{ij} D_{ij}$$

Donde:

$\delta_{ij}$  es un número.

$D_{ij}$  representa la delta de Kronecker.

En este caso, hay conservación de la energía a partir de la transformación. Existe un gran número de transformaciones que presentan propiedades de ortogonalidad y de separabilidad.

Nos centramos en una clase más particular de estas transformaciones, la clase de las transformaciones llamadas rápidas.

#### 2.2.4.2 Transformaciones rápidas.

En el caso de una transformación separable, que nosotros supondremos siempre realizable, la transformación de una imagen se obtiene al efectuar sobre cada línea de la imagen, y después sobre cada columna, operaciones matriciales elementales del tipo:

$$[L] = [l] \times [A]$$

[L] = es la matriz línea transformada, N elementos.

[l] = es la matriz línea a transformar de N elementos.

[A] = es la matriz de transformación de N x N elementos.

En el caso general, tal multiplicación conduce a  $N \times 2$  operaciones aritméticas elementales (adiciones y multiplicaciones). Como en una imagen, es preciso hacer N veces estas operaciones (para cada línea) y recomenzar en seguida con las columnas, se tendrá un número de operaciones del orden de  $2N \times 3$ , lo que en general es muy elevado, pero menor que en la transformación de Karhunen-Loeve.

Para las transformaciones rápidas, el número de operaciones es mucho menor. En este caso, la matriz de transformación [A] puede ser factorizada en un producto de varias matrices.

$$[A] = [A1] [A2] \dots [An].$$

Las diferentes matrices [Ai] están caracterizadas por un

número importante de elementos nulos, lo que reduce considerablemente el número de operaciones aritméticas.

No es posible identificar de manera simple las matrices que son factorizables en matrices con muchos ceros. De igual forma, no hay algoritmos para determinar la mejor factorización posible de matrices factorizables. Los únicos métodos son la construcción de matrices por técnicas, de las cuales se sabe a priori que proporcionan matrices factorizables, o los ensayos de descomposición de una matriz dada a priori.

La figura da el ejemplo de la transformación de Hadamard utilizando como base las funciones de Walsh, para el caso de  $N=8$ . En este ejemplo, las matrices  $[A_i]$  son idénticas y  $n=3$ . Se ve claramente que cada multiplicación de una matriz línea por una columna de la matriz  $[A_i]$  se traduce por una única suma o resta en lugar de "N" operaciones en el caso general.

$$\begin{bmatrix}
 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\
 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\
 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\
 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\
 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\
 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\
 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1
 \end{bmatrix}
 =
 \begin{bmatrix}
 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1
 \end{bmatrix}^3$$

Factorización de la matriz de transformación de Hadamard con N=8 en tres matrices iguales con numerosos elementos nulos.

De la misma forma, en la siguiente figura se tiene la factorización de la transformación de Haar para N=8, en un producto de tres matrices diferentes. En este caso, las matrices tienen casi todos sus elementos nulos, lo que conduce a un número de operaciones aritméticas todavía menor que la transformación de Hadamard.

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix}$$

$$A1 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} \end{bmatrix}$$

$$A2 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sqrt{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sqrt{2} \end{bmatrix}$$

$$A3 = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

Factorización de la matriz de la transformación de Haar [H] en tres matrices [H] = [A1] [A2] [A3].

Good ha demostrado que para una clase de matrices factorizables, si N es de la forma

$$N = p^n$$

el número de operaciones elementales a efectuar para multiplicar una matriz línea por la matriz de transformación es

$$\frac{pN \log N}{p}$$

mientras que en el caso general es de  $N^2$ . Por ejemplo en el caso de la transformación de Fourier con  $N=8=2^3$ , el número de operaciones es de 48 en lugar de 64.

Para la transformación de Haar, el número de operaciones es

$$p(p^{*n} - 1) / (p-1).$$

Y para N=8 y p=2 se reduce a 14.

### 2.2.56 TRANSFORMADA DE FOURIER.

La transformada de Fourier es la más conocida de las transformaciones ortogonales, por el hecho del significado físico de sus componentes en el caso de tratamiento de senales que son función del tiempo. También pueden aplicarse a las imágenes, lo que conduce a la noción de frecuencias espaciales.

#### 2.2.5.1 Definición.

La definición de la transformada de Fourier directa e inversa viene dada por la expresión:

$$F(u, v) = \frac{1}{N} \sum_x \sum_y f(x, y) \exp \left[ - \frac{2(\pi)i}{N} (ux + vy) \right]$$

$$f(x, y) = \frac{1}{N} \sum_u \sum_v F(u, v) \exp \left[ \frac{2(\pi)i}{N} (ux + vy) \right]$$

La transformación es de núcleo separable y simétrico, por lo que se puede reemplazar por una secuencia de transformaciones unidimensionales, definida por la matriz de transformación.

$$[A] = \frac{1}{\sqrt{N}} \exp \left[ - \frac{2(\pi)iux}{N} \right]$$

Si hacemos

$$w = \exp \left[ - \frac{2(\pi)i}{N} \right]$$

$$[A] = \frac{1}{\sqrt{N}} \begin{matrix} & k \\ u & w \end{matrix}$$

con  $k=ux$ , siendo  $0 \leq u, x \leq N - 1$ .

La siguiente figura da la expresión de esta matriz para  $N=8$ .

$$\frac{1}{\sqrt{8}} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ w & w & w & w & w & w & w & w \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ w & w & w & w & w & w & w & w \\ 0 & 2 & 4 & 6 & 8 & 10 & 12 & 14 \\ w & w & w & w & w & w & w & w \\ 0 & 3 & 6 & 9 & 12 & 15 & 18 & 21 \\ w & w & w & w & w & w & w & w \\ 0 & 4 & 8 & 12 & 16 & 20 & 24 & 28 \\ w & w & w & w & w & w & w & w \\ 0 & 5 & 10 & 15 & 20 & 25 & 30 & 35 \\ w & w & w & w & w & w & w & w \\ 0 & 6 & 12 & 18 & 24 & 30 & 36 & 42 \\ w & w & w & w & w & w & w & w \\ 0 & 7 & 14 & 21 & 28 & 35 & 42 & 49 \\ w & w & w & w & w & w & w & w \end{bmatrix}$$

$$w = \exp \left[ - \frac{2(\pi)i}{8} \right]$$

### 2.2.5.2 Propiedades de la transformación.

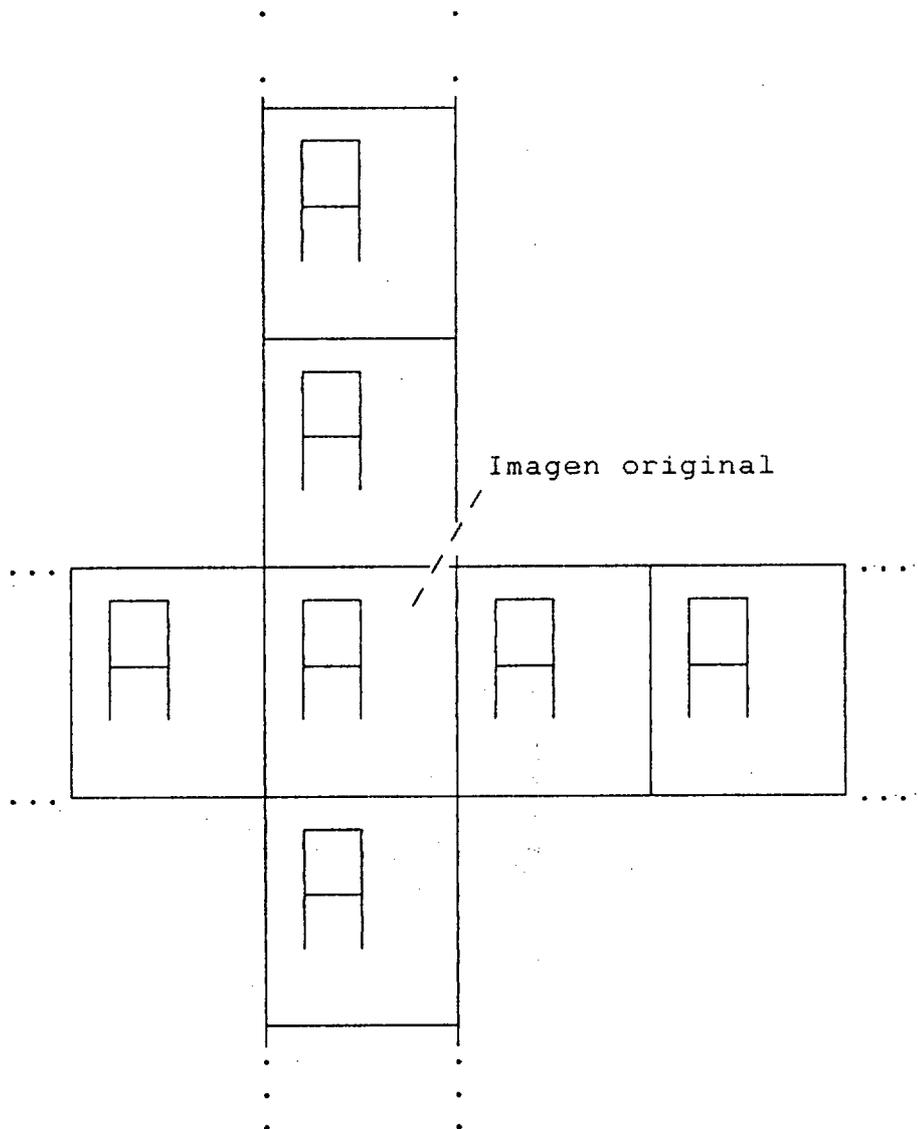
De todas las propiedades de la transformada de Fourier, hay que señalar dos características que es preciso tener en cuenta en su empleo:

a) La imagen inicial comporta  $N \times 2$  puntos y después de la transformación se encuentran  $N \times 2$  coeficientes complejos, por lo tanto  $2N \times 2$  coeficientes reales. Este efecto proviene de la hipótesis de la anulación de la parte imaginaria del brillo para cada punto de la imagen. Existe una relación entre los coeficientes de la transformada.

$$F(-u, -v) = F^*(u, v).$$

Basta con guardar  $N \times 2$  coeficientes reales (correspondientes a  $u \Rightarrow 0$ , por ejemplo), puesto que los  $N \times 2$  restantes se pueden deducir de estos últimos.

b) La transformada de Fourier discreta bidimensional es esencialmente una representación de Fourier de un campo bidimensional. La representación sólo es válida si el campo es periódico. La imagen original no es únicamente el conjunto de los  $N \times 2$  puntos dados, sino que comporta también las imágenes trasladadas vertical y horizontalmente de estos  $N \times 2$  puntos. Esta propiedad puede conducir a efectos parásitos de recubrimiento en ciertas operaciones sobre las transformadas de Fourier, con aparición de estructuras no reales.



### 2.2.5.3 Algoritmo de cálculo rápido.

El desarrollo del cálculo numérico de las transformadas de Fourier, que hacen intervenir un número elevado de puntos, como es el caso de las imágenes, está ligado a la descripción por Cooley y Tukey de un algoritmo de cálculo rápido. En lugar de tener  $N^2$  operaciones aritméticas para la transformación, el método de la transformada rápida sólo necesita  $2N \log_2 N$  operaciones, para  $N=2^n$ .

La tabla siguiente da los valores y el número de operaciones en los dos casos. Se ve claramente que el tratamiento de una imagen de de 256 x 256 puntos (que corresponde a una definición suficiente) es posible mediante la transformada rápida y prácticamente irrealizable por el método ordinario.

Dimensión imagen	Transformada ordinaria	Transf. rápida
8 x 8	4096	768
32 x 32	# 10**6	# 2.10**4
128 x 128	# 2,7.10**8	# 4,6 10**5
256 x 256	# 4,3 10**9	# 2,1 10**6

Número de operaciones aritméticas para la transformación de Fourier de una imagen.

La existencia de una transformada rápida de Fourier es debida a la simetría de la matriz de transformación [A], que no comporta, finalmente más que N términos diferentes para transformar N puntos. Los términos son las diferentes potencias.

$$w = \left\{ \exp \left[ - \frac{2(\pi)i}{N} k \right] \right\}$$

con  $k=0, 1, \dots, N-1$ .

$$\text{Pues: } \exp \left[ - \frac{2(\pi)i}{N} k \right] = \exp \left[ - \frac{2(\pi)i}{N} (k+N) \right].$$

La siguiente figura representa la matriz de transformación reducida para  $N=8$ . Esta matriz puede ser factorizada en un producto de matrices más simples, teniendo numerosos coeficientes nulos.

$$\frac{1}{\sqrt{8}} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ w & w & w & w & w & w & w & w \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ w & w & w & w & w & w & w & w \\ 0 & 2 & 4 & 6 & 0 & 2 & 4 & 6 \\ w & w & w & w & w & w & w & w \\ 0 & 3 & 6 & 1 & 4 & 7 & 2 & 5 \\ w & w & w & w & w & w & w & w \\ 0 & 4 & 0 & 4 & 0 & 4 & 0 & 4 \\ w & w & w & w & w & w & w & w \\ 0 & 5 & 2 & 7 & 4 & 1 & 6 & 3 \\ w & w & w & w & w & w & w & w \\ 0 & 6 & 4 & 2 & 0 & 6 & 4 & 2 \\ w & w & w & w & w & w & w & w \\ 0 & 7 & 6 & 5 & 4 & 3 & 2 & 1 \\ w & w & w & w & w & w & w & w \end{bmatrix}$$

Matriz de la transformada de Fourier con  $N=8$ , que tiene en cuenta la periodicidad de  $w^{**k}$ .

## 2.2.6 TRANSFORMACION DE HADAMARD.

### 2.2.6.1 Definición.

La transformación de Hadamard está definida por una matriz de transformación cuyos elementos están dados por



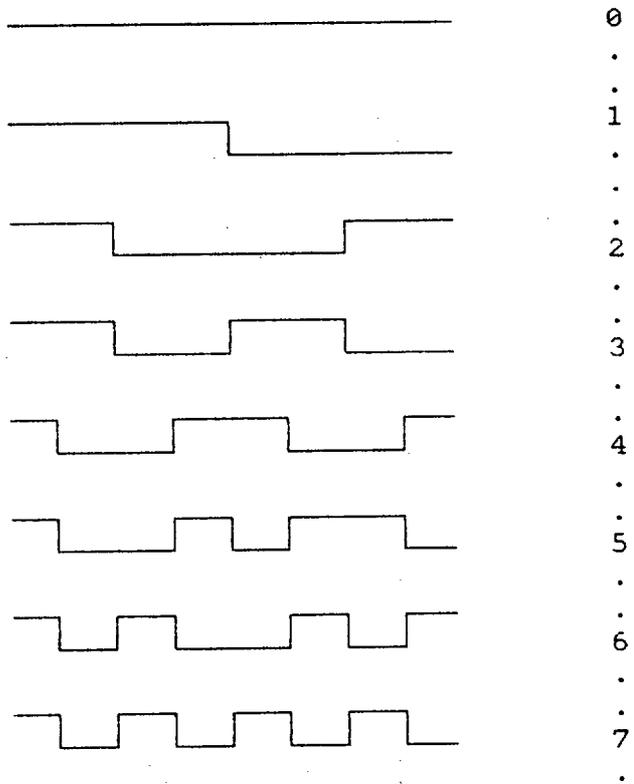
$$\frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \end{bmatrix}$$

Matriz de Walsh-Paley

$$\frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \end{bmatrix}$$

Matriz Walsh-Kaczmarz

Lo mismo que la transformada de Fourier corresponde a una descomposición de la función inicial en sumas de sinusoides, la transformación de Hadamard corresponde a una descomposición en funciones de Walsh que son funciones que sólo toman los valores  $\pm 1$ . Estas funciones son ortogonales, y están representadas para  $N=8$  en la siguiente figura.



Funciones de Walsh para N=8.

Por analogía con la frecuencia que aparece en la transformación de Fourier, se define para las funciones de Walsh la "secuencia". La secuencia es el número de cambios de signo de una función de Walsh en su intervalo de definición. Para las funciones de Walsh correspondientes a la transformación de Hadamard hasta  $N=2^n$  las secuencias van de 0 a (N-1). Las diferentes matrices de transformación de las figuras anteriores están caracterizadas por órdenes diferentes de secuencias para cada línea de la matriz.

### 2.2.6.2 Propiedades de la transformación.

La transformación de Hadamard está caracterizada por una matriz [A] real y simétrica, por lo que se tiene:

$$[A]**T = [A]$$

La transformada directa e inversa son, pues, idénticas.

La construcción más simple, corresponde a las matrices de Kronecker-Hadamard en que  $N=2**n$ . En este caso, si [HN] es la matriz de orden N, se tiene:

$$\begin{bmatrix} H2N \end{bmatrix} = \begin{bmatrix} HN & HN \\ HN & -HN \end{bmatrix}$$

Las funciones de Walsh que aparecen en la transformación de Hadamard presentan analogías con las funciones de seno y coseno de la transformada de Fourier. Se tienen dos sistemas de funciones ortonormales en el espacio de Hilbert  $L2 [0,1]$ . Se puede clasificar según su sistema las funciones de Walsh en funciones pares o impares, que son a veces llamadas "sal" y "cal". Existen fórmulas de adición o de sustracción de estas funciones, con formas análogas a las de las funciones trigonométricas.

Hay, sin embargo, diferencias fundamentales debidas al hecho de que las funciones circulares están ligadas al grupo topológico de los números reales, mientras que las funciones de Walsh están ligadas al grupo diádico que es el grupo topológico derivado del

conjunto de las representaciones binarias de los números reales. Esto se traduce para las funciones de Walsh por el empleo de operaciones aritméticas módulo 2 y la dificultad de interpretación de algunos resultados de transformación (filtrado, por ejemplo).

### 2.2.6.3 Algoritmo de cálculo rápido.

La transformación de Hadamard sólo necesita operaciones de suma y resta con números reales. El algoritmo de cálculo rápido corresponde a la descomposición de la matriz de transformación en un producto de matrices idénticas con gran cantidad de ceros. Es posible analizar de forma simple el proceso de cálculo rápido. En el caso de  $N=8$ , la transformación utiliza 8 valores "En" de la función a transformar y da el resultado siguiente:

$$T1 = E_1 + E_2 + E_3 + E_4 + E_5 + E_6 + E_7 + E_8$$

$$T2 = E_1 + E_2 + E_3 + E_4 - E_5 - E_6 - E_7 - E_8$$

$$T3 = E_1 + E_2 - E_3 - E_4 - E_5 - E_6 + E_7 + E_8$$

$$T4 = E_1 + E_2 - E_3 - E_4 + E_5 + E_6 - E_7 - E_8$$

$$T5 = E_1 - E_2 - E_3 + E_4 + E_5 - E_6 - E_7 + E_8$$

$$T6 = E_1 - E_2 - E_3 + E_4 - E_5 + E_6 + E_7 - E_8$$

$$T7 = E_1 - E_2 + E_3 - E_4 - E_5 + E_6 - E_7 + E_8$$

$$T8 = E_1 - E_2 + E_3 - E_4 + E_5 - E_6 + E_7 - E_8$$

Examinando las ocho expresiones, se observa que hay que calcular ocho variables intermedias.

$$\begin{array}{l}
 I_1 = E_1 + E_2 \\
 I_3 = E_5 + E_6 \\
 I_5 = E_1 - E_2 \\
 I_7 = E_5 - E_6 \\
 I_2 = E_3 + E_4 \\
 I_4 = E_7 + E_8 \\
 I_6 = E_3 - E_4 \\
 I_8 = E_7 - E_8
 \end{array}$$

luego

$$\begin{array}{l}
 I'_1 = I_1 + I_2 \\
 I'_3 = I_1 - I_2 \\
 I'_5 = I_5 + I_6 \\
 I'_7 = I_5 - I_6 \\
 I'_2 = I_3 + I_4 \\
 I'_4 = I_3 - I_4 \\
 I'_6 = I_7 + I_8 \\
 I'_8 = I_7 - I_8
 \end{array}$$

y, por fin

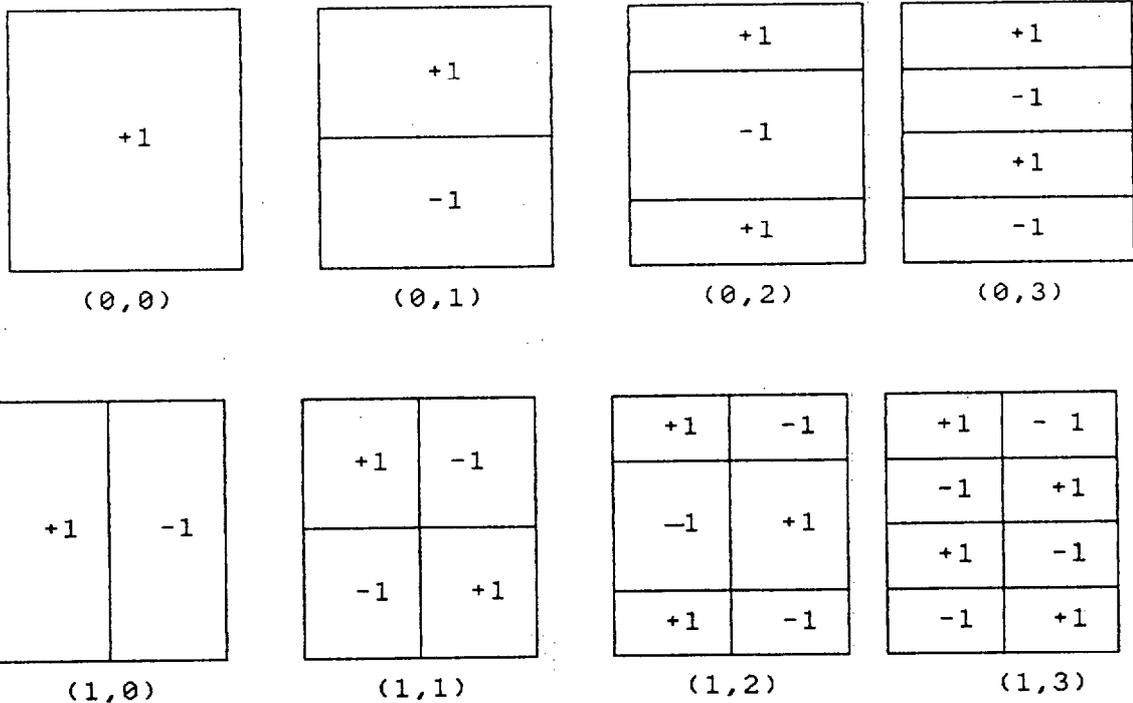
$$\begin{array}{l}
 I''_1 = I'_1 + I'_2 \\
 I''_3 = I'_3 + I'_4 \\
 I''_5 = I'_5 + I'_6 \\
 I''_7 = I'_7 + I'_8 \\
 I''_2 = I'_1 - I'_2 \\
 I''_4 = I'_3 - I'_4 \\
 I''_6 = I'_5 - I'_6 \\
 I''_8 = I'_7 - I'_8
 \end{array}$$

Es suficiente reagrupar los valores de  $I''_i$  para encontrar los  $T_i$ .



La matriz [HA] representa las líneas de la matriz de Haar numeradas de  $2^{k+1}$  a  $2^{k+2}$ . Las matrices [HD] son las matrices de las líneas de la matriz de Hadamard de 0 a  $2^k$  o de  $2^{k+1}$  a  $2^{k+1}$ .

Esta relación permite establecer la propiedad siguiente: los coeficientes de las transformadas de Haar y de Hadamard de índices comprendidos entre dos potencias sucesivas de dos tienen la misma energía total. Las dos transformadas se diferencian por las reparticiones diferentes de la energía sobre los coeficientes de la transformación. Estas diferencias pueden ser analizadas en las figuras siguientes que dan para cada elemento  $a(u,v)$  de las dos transformadas la ponderación de los elementos de la imagen inicial.



+1	-1	+1
----	----	----

(2,0)

+1	-1	+1
-1	+1	-1

(2,1)

+1	-1	+1
-1	+1	-1
+1	-1	+1

(2,2)

+1	-1	+1
-1	+1	-1
+1	-1	+1
-1	+1	-1

(2,3)

+1	-1	+1	-1
----	----	----	----

(3,0)

+1	-1	+1	-1
-1	+1	-1	+1

(3,1)

+1	-1	+1	-1
-1	+1	-1	+1
+1	-1	+1	-1

(3,2)

+1	-1	+1	-1
-1	+1	-1	+1
+1	-1	+1	-1
-1	+1	-1	+1

(3,3)

Significado de los elementos (u, v) de la transformada de Hadamard de una imagen. Los valores numéricos de cada zona representan la ponderación de los elementos de esta zona.

+1
----

(0,0)

+1
-1

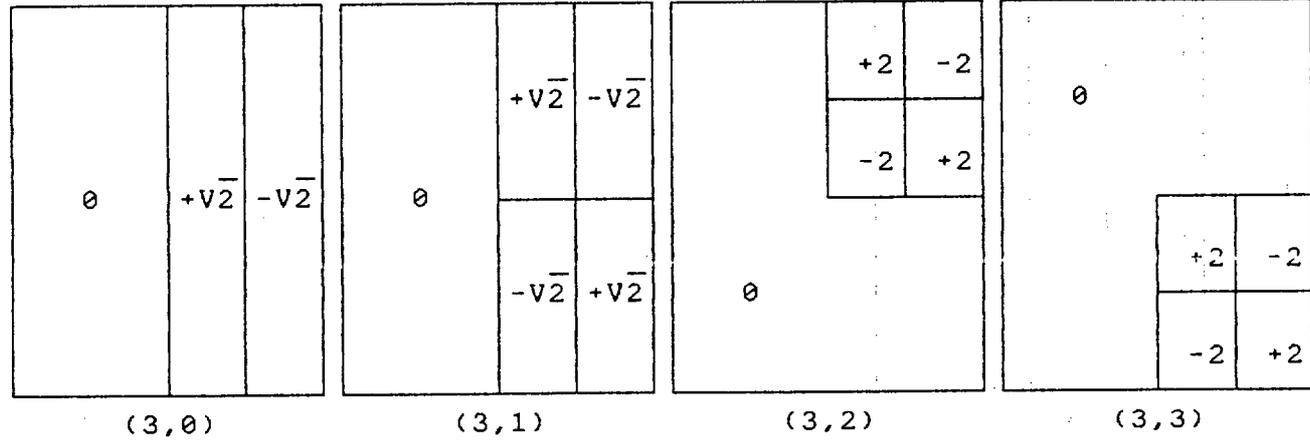
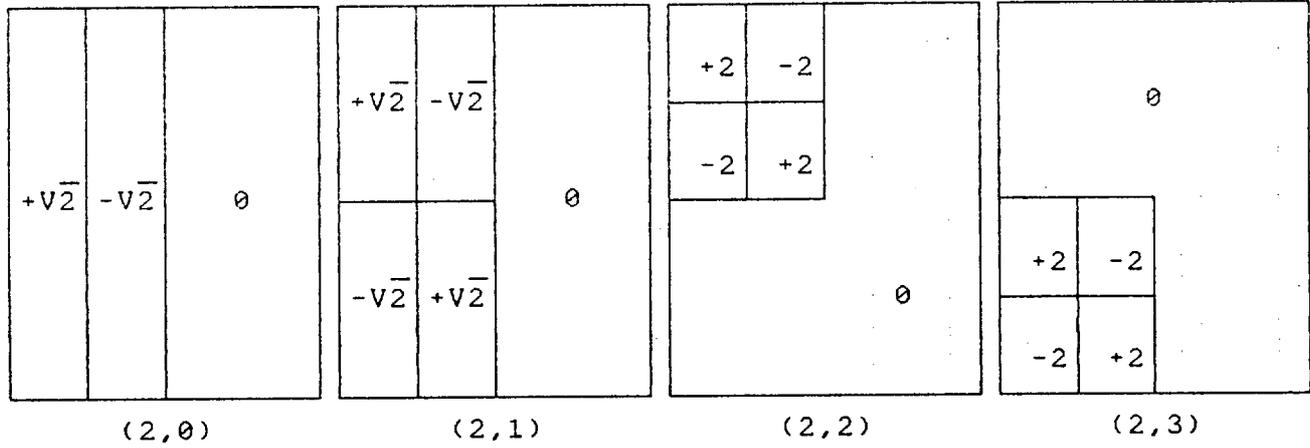
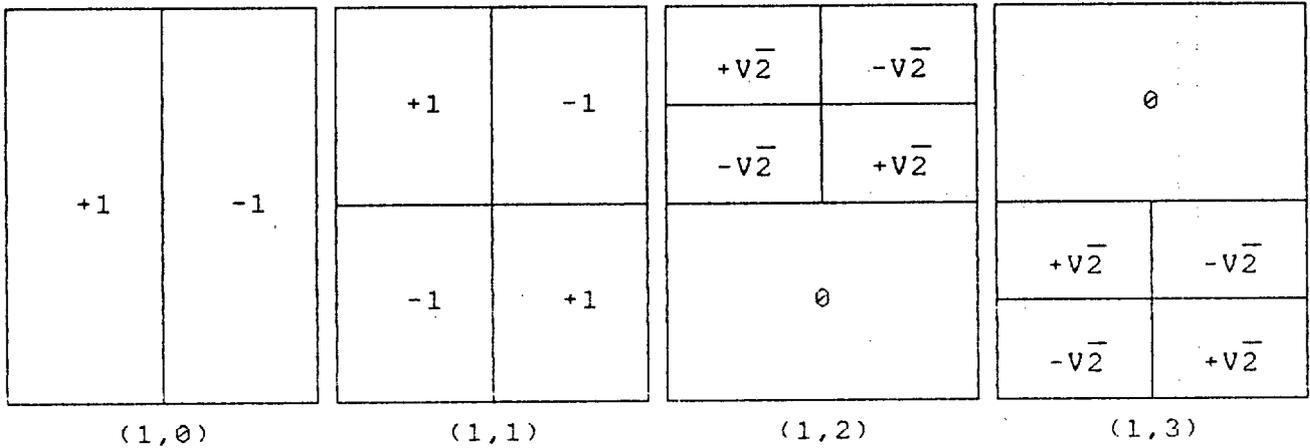
(0,1)

$+\sqrt{2}$
$+\sqrt{2}$
0

(0,2)

0
$+\sqrt{2}$
$-\sqrt{2}$

(0,3)



Significado de los elementos (u,v) de la transformada de Haar de una imagen.

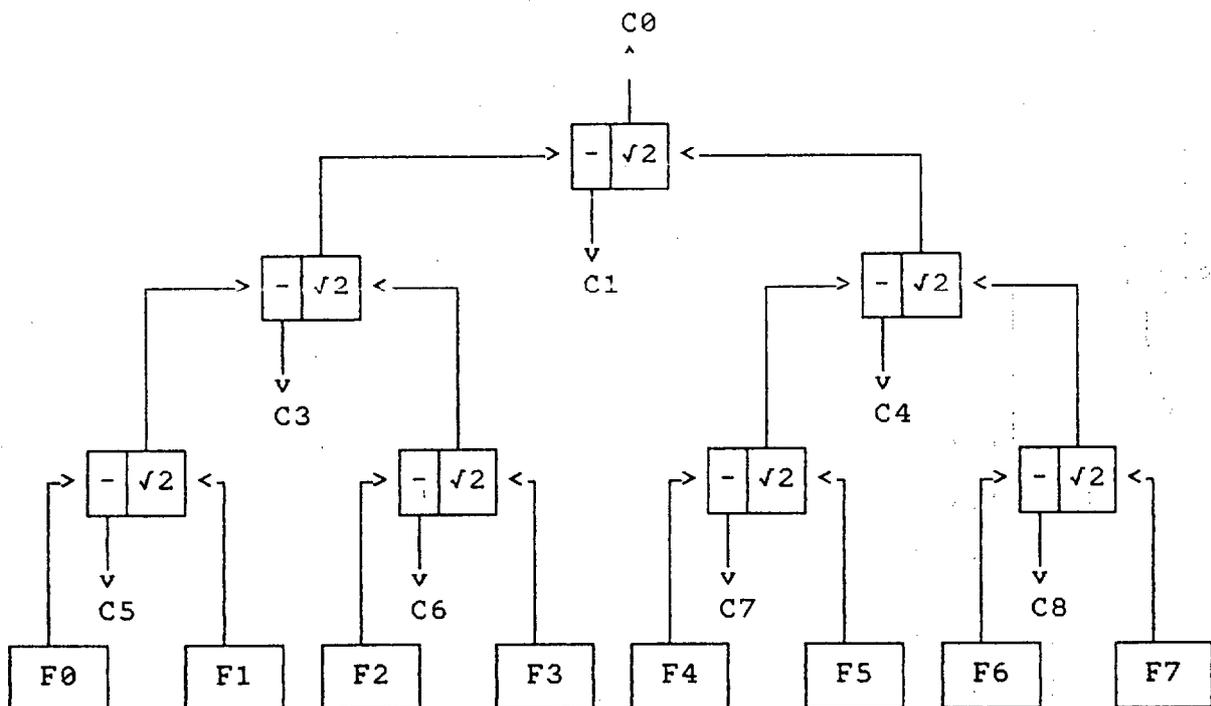
Salvo para los elementos de índice 0 y 1 que son calculados

de manera idéntica en las dos transformaciones, se ve que cada coeficiente de la transformada de Hadamard hace intervenir todos los puntos de la imagen; mientras que la transformada de Haar tiene coeficientes que proporcionan informaciones locales sobre las variaciones de brillos, siendo este caracter local tanto más acentuado a medida que los índices son más elevados.

### 2.2.7.3 Algoritmo de cálculo rápido.

El algoritmo de la transformada de Haar rápido se deriva de la factorización, de la matriz de transformación en matrices simples con gran número de ceros. Es parecido al de la transformada de Hadamard, pero con un número de operaciones menor.

El grafo de la transformación directa para N=8 es:



Para el caso en que el número de puntos es  $N=2^n$ , el número de operaciones a efectuar es  $2(N-1)$ . Por lo que es una de las transformaciones más rápidas.

### 2.2.8 OTRAS TRANSFORMACIONES RAPIDAS.

Estudiemos otras transformaciones más particulares utilizadas en aplicaciones físicas.

#### 2.2.8.1 Transformación de Fresnel.

La transformación de Fresnel se relaciona con la transformada de Fourier. Se expresa, en caso de imágenes, por la relación

$$F(u, v) = \sum_x \sum_y f(x, y) \exp[-2(\pi)i(u-x)^2 - 2(\pi)i(v-y)^2]$$

El cálculo nos lleva a la transformada de Fourier.

En efecto:

$$\exp[-2(\pi)i(u-x)^2] = \exp[-2(\pi)iu^2] \exp[-2(\pi)ix^2] \exp[+4(\pi)iux]$$

$$F(u, v) = \exp[-2(\pi)i(u^2 + v^2)] \sum_x \sum_y f(x, y) \exp[-2(\pi)i(x^2 + y^2) \exp[+4(\pi)i(ux + vy)]]$$

$$F(u, v) = \exp[-2(\pi)i(u+v)^2] \sum_x \sum_y g(x, y) \exp[+2(\pi)i(2ux+2vy)]$$

$$g(x, y) = f(x, y) \exp[-2(\pi)i(x^2 + y^2)]$$

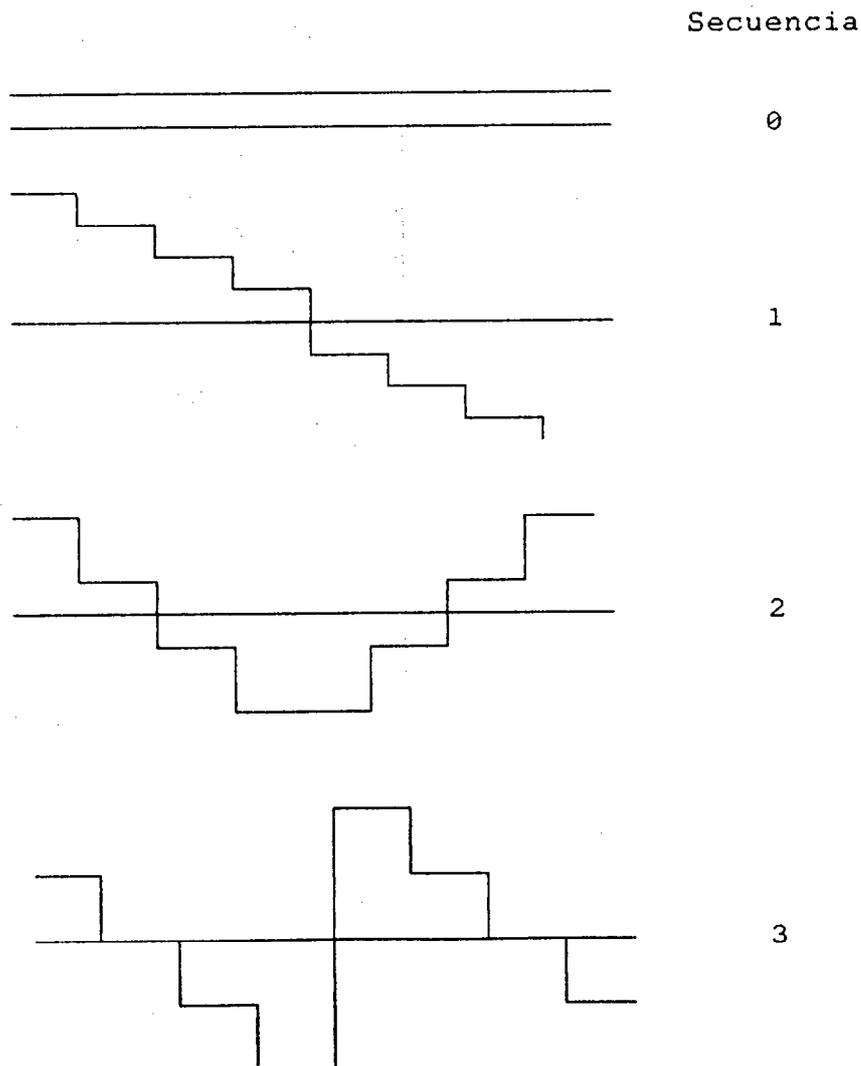
Lo que es también una transformación de Fourier de la función  $g(x, y)$ , estando el resultado multiplicado por

$$\exp [-2(\pi)i(u^{**2} + v^{**2})]$$

Esta transformación se aplica principalmente en óptica y en acústica.

### 2.2.8.2 Transformada de Hadamard Slant.

Esta transformada utiliza como funciones de base funciones parecidas a las de Walsh, que utiliza la transformada de Hadamard, pero presenta variaciones lineales sobre un intervalo de definición. La siguiente figura representa las cuatro primeras funciones de la transformación.



Primeras funciones de la transformación de Hadamard "slant".

Existe un algoritmo de "cálculo rápido" que ha sido estudiado para la transmisión de imágenes de televisión en color.

### 2.2.8.3 Transformada de Hadamard compleja.

Es una generalización de la transformada de Hadamard, introducida para tratar imágenes complejas. La matriz de transformación es

$$[A] = \frac{1}{\sqrt{N}} \begin{matrix} & \begin{matrix} \rightarrow & \rightarrow & \rightarrow [x] \end{matrix} \\ \begin{matrix} \rightarrow & \rightarrow \end{matrix} & \begin{matrix} 1 & u.x & u.[2] \end{matrix} \\ & \begin{matrix} (-1) & & (-i) \end{matrix} \end{matrix}$$

$u.x$  es el producto escalar de las representaciones binarias.

$\left[ \begin{smallmatrix} x \\ 2 \end{smallmatrix} \right]$  es la potencia entera mayor de dos en  $x/2$ .

La siguiente figura corresponde a la matriz de transformación bajo forma ordenada para  $N=8$ .

El orden es el de las secuencias, es decir, el número de vueltas del círculo de las raíces complejas de la unidad recorriendo los elementos de una columna y volviendo al punto (1.0). Las secuencias se escalonan de  $-N/2$  a  $N/2 - 1$ .

$$\frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -1 & -1 & i & 1 & 1 & 1 & -i & -1 \\ 1 & -i & -1 & i & 1 & -i & -1 & i \\ -1 & i & -i & i & 1 & -i & i & -i \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & i & -1 & 1 & -1 & -i & 1 \\ 1 & i & -1 & -i & 1 & i & -1 & -i \\ -1 & -i & -i & -i & 1 & i & i & i \end{bmatrix}$$

Existe un algoritmo de cálculo para esta transformada.

#### 2.2.8.4 Transformada de Haar Slant.

Esta transformación tiene las mismas características que la transformada de Hadamard slant, vista anteriormente, pero permite conservar el carácter local de la transformada de Haar, útil para el tratamiento de ciertas imágenes, y la rapidez de cálculo del algoritmo para la transformación de Haar.

### 2.2.8.5 Transformada de coseno discreto.

Esta transformación está definida para una secuencia de  $M$  puntos  $X(m)$  por:

$$\left[ \begin{array}{l} G(0) = \frac{\sqrt{2}}{2} \sum_{m=0}^{M-1} X(m) \\ G(k) = \frac{2}{M} \sum_{m=0}^{M-1} X(m) \cos \left[ \frac{(2m+1)k(\pi)}{2M} \right] \\ k = 1, 2, \dots (M-2). \end{array} \right.$$

El cálculo de esta transformada nos lleva a una transformada de Fourier de  $2M$  puntos. Presenta la ventaja de tener características muy parecidas a las de la transformación de Karhunen-Loeve para el reconocimiento de forma y para el filtrado de Wiener.

### 2.3 ESTUDIO COMPARATIVO DE LAS MISMAS.

Es conocida la aplicación de las transformaciones integrales en el campo de las comunicaciones, para la transmisión de señales vía satélite, transmisiones radar, etc. Esto nos hizo pensar en utilizar las transformaciones integrales en el área del reconocimiento de formas.

La primera pregunta que nos surgió es ¿que transformación integral utilizar? después de realizar una búsqueda

bibliográfica, en las Bases de Datos Internacionales con las palabras claves "Integral Transform", "Pattern Recognition", y "Vision", obtuvimos que las transformaciones más utilizadas eran las transformaciones ortonormales de Fourier, Hadamard, Hadamard-Walsh, Haar y Coseno, entre otras.

El problema del reconocimiento de formas, no exige, que las transformaciones sean ortonormales, tal vez, la característica más deseable es que los coeficientes transformados sean lo más discriminantes posible entre los clusters de las clases patrones, en los espacios transformados.

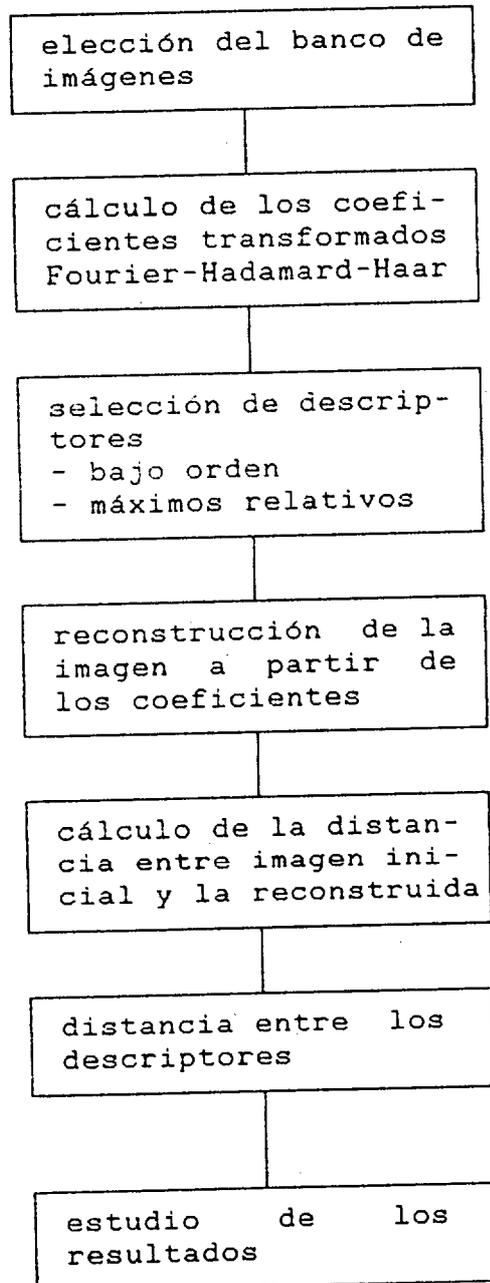
El encontrar en la bibliografía consultada, que las transformaciones ortonormales, eran las más descritas y que no existía un estudio a cerca de la discriminabilidad de los coeficientes, nos hizo elegir el subconjunto de las transformadas ortogonales dentro del conjunto de las transformadas integrales.

Una vez hecha esta elección, nos preguntamos que funciones ortonormales utilizar en nuestro experimento. Nos basamos en dos criterios para escoger nuestras funciones, un primer criterio fue que existiese información a cerca de la función y de sus propiedades y un segundo criterio que existiesen algoritmos rápidos de la función que nos permitiesen una implementación en el ordenador fácil y de cálculo rápido. Esta segunda condición la impusimos por trabajar con funciones bidimensionales y de gran número de elementos. Estos criterios nos llevaron a seleccionar

las transformadas integrales ortogonales de Fourier, Hadamard-Walsh y Haar.

El objetivo de nuestro experimento consiste en realizar un estudio comparativo de la potencialidad de estos tres tipos de descriptores integrales de bajo nivel en el campo de la visión artificial. Este estudio de la bondad de descripción de los coeficientes transformados se realiza mediante el cálculo de una función de error calculada localmente, comparando la imagen inicial preprocesada y la imagen reconstruida, mediante unos pocos coeficientes, el reconocimiento se efectúa mediante las distancias calculadas en el espacio de representación de los descriptores.

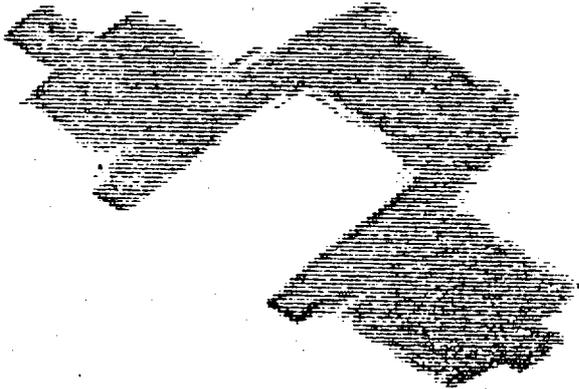
- Diagrama de flujo del experimento.



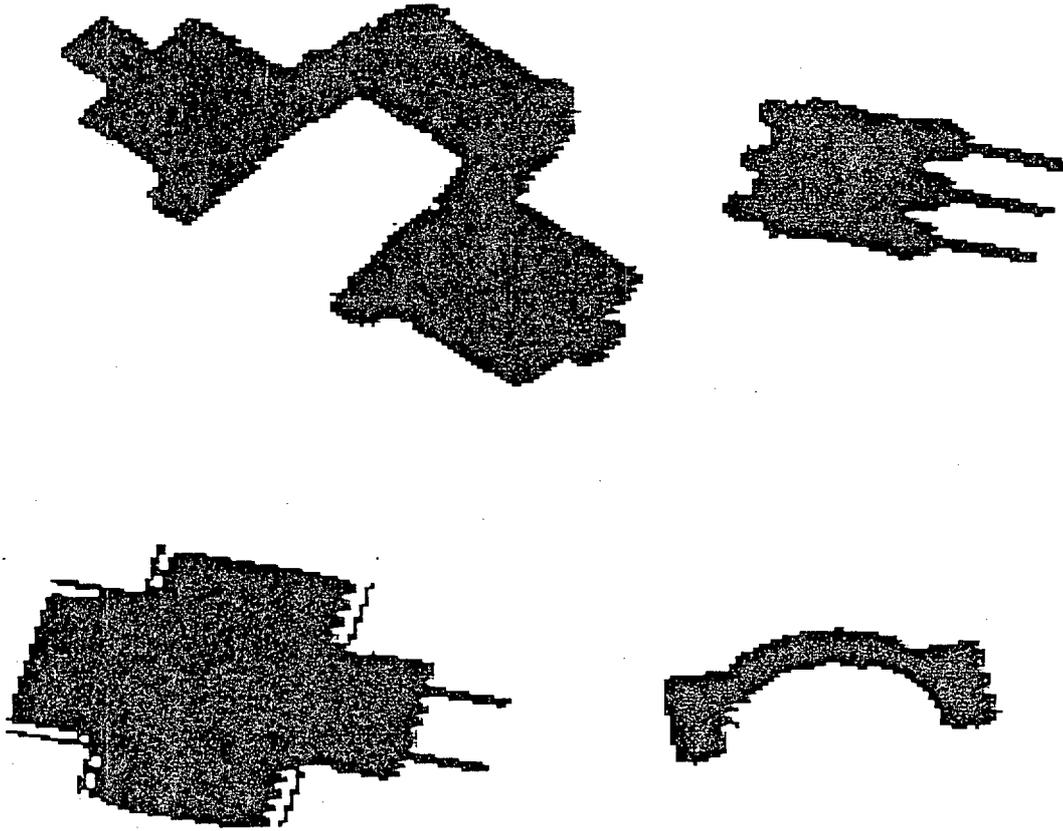
- elección del banco de imágenes.

El banco de imágenes que se escogió para realizar este experimento, está compuesto por diez tomas de imágenes diferentes de piezas reales industriales. Las imágenes de este conjunto son de 128x128 pixels y cada pixel tiene 8 bits. Estas piezas se

introducen en el ordenador mediante un sistema de adquisición compuesto por una cámara de estado sólido (CCD) y unas placas digitilizadoras



Inicialmente, las imágenes originales son preprocesadas en el ordenador para generar unas imágenes relativamente "limpias" donde van a tratarse los distintos algoritmos. Este preproceso, incluye filtrado espacial, umbralizado espacial para eliminar irregularidades en la iluminación y umbralizado automático de los niveles de gris basado en el cálculo del histograma.



- Cálculo de las transformadas.

Una imagen digitalizada es una matriz bidimensional de  $N \times N$  en la cual cada elemento corresponde con la intensidad de cada punto en la imagen original mediante un rango de niveles de gris entre 0 (blanco) y 255 (negro). La imagen transformada es otra matriz bidimensional de  $N \times N$ .

Para una imagen bidimensional la transformación integral directa e inversa en forma discreta viene dada por las ecuaciones

$$f(u,v) = \sum_{x=0}^{n-1} \sum_{y=0}^{n-1} f(x,y) * k(x,y,u,v)$$

y

$$f(x,y) = \sum_{u=0}^{n-1} \sum_{v=0}^{n-1} f(u,v) * h(x,y,u,v)$$

donde

$f(x,y)$  ---> es la imagen original.

$F(u,v)$  ---> es la imagen transformada.

$k(x,y,u,v)$  ---> núcleo directo de transformación.

$h(x,y,u,v,)$  ---> núcleo inverso de transformación.

Un núcleo es separable si se puede descomponer como producto de otros más sencillos.

$$g(x,y,u,v) = g_1(x,u) g_2(y,v)$$

Un núcleo es simétrico si  $g_1$  es funcionalmente igual a  $g_2$  en este caso la relación anterior se puede expresar como.

$$g(x,y,u,v) = g_1(x,u) g_1(y,v)$$

En este caso el cálculo de los coeficientes transformados puede realizarse en dos pasos.

$$\text{Paso 1) } f(u,y) = \sum_{x=0}^{u-1} f(x,y) g_1(x,u)$$

$$\text{Paso 2) } F(u,v) = \sum_{y=0}^{u-1} f(u,y) g_1(y,v)$$

Esto nos lleva a un algoritmo más sencillo de implementación.

Utilizando notación matricial, podemos expresar las ecuaciones de transformación en la forma.

$$[F]^T = [ [f]^T [A]^T ] [A]^T$$

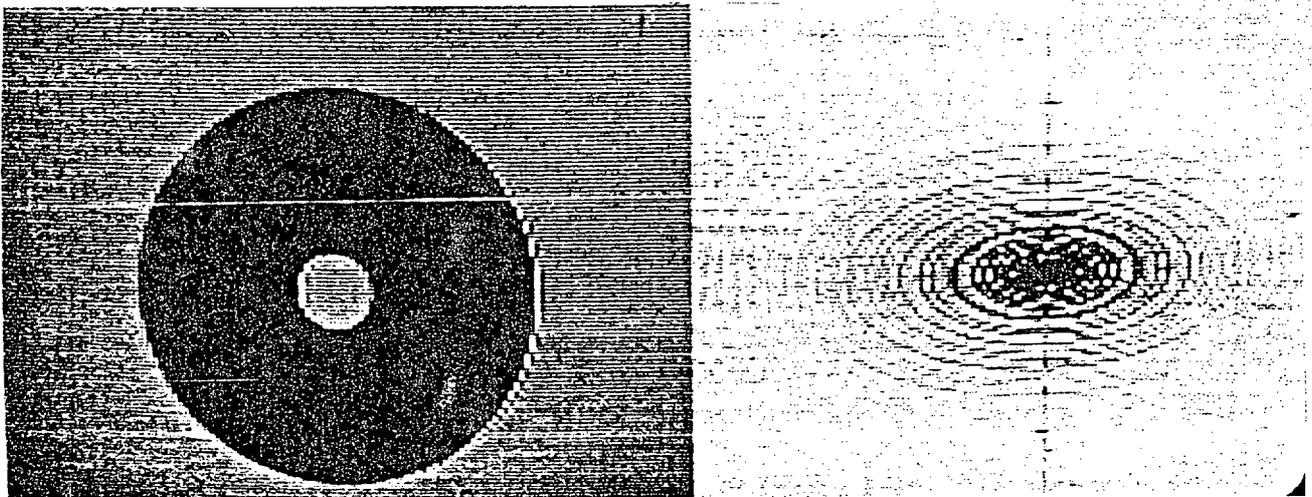
$$[f] = [ [F] [A] ] [A]^*$$

Donde T denota traspuesta y \* inversa.

Suponiendo  $N = 2^{**}n$  el núcleo de la transformada de FOURIER se define por.

$$[A] = 1/\text{SQRT}(n) \exp(-2\pi i u x / N) \quad u, x, [0..N-1]$$

El programa que implementa esta transformada de forma rápida para dos dimensiones es el (FFTMF.FOR) que produce dos ficheros uno conteniendo el módulo de los descriptores y otro conteniendo la fase de los descriptores.



El núcleo de la transformación HADAMARD-WALSH se define por:

$$[A] = 1/\text{SQRT}(N) \overline{u.x}^{k/2} (-1)^{x.k1} \quad u,x [0..N-1]$$

donde:

$u.x$  es el producto escalar de  $u$  y  $x$  en representación binaria.

El programa que implementa esta transformada es (FWT2D.FOR), dejando los coeficientes transformados reales en un fichero.

El núcleo de la transformada de HAAR se define por.

$$[A] = 1/\text{SQRT}(N) .2^{k/2} (-1)^{x.k1} d(k2,u1)$$

donde

$k$ - es la parte entera del logaritmo en base dos de  $u$

$$k1 = 2^{N-k-1}$$

$k2$ - es la parte entera de  $x/2^{N-k}$

$$u1 = u - 2^k$$

$d$ - es la delta de Kronecker

$\overline{x.k1}$  --> es el producto escalar de  $x$  y  $k1$  en representación binaria

El programa que implementa el algoritmo rápido de esta transformada es (TRKAAR.FOR) dejando la imagen transformada real en un fichero.

#### - Selección de descriptores.

Una vez calculados los distintos coeficientes transformados se pasa a seleccionar un subconjunto de estos, que se utilizan

posteriormente en la reconstrucción de la imagen original.

Dos metodos se utilizan para la selección de estos descriptores.

a) Descriptores de bajo orden.

Este método está basado en la selección de los coeficientes de bajo orden. Esta elección se realiza debido a que estos coeficientes contienen información global de las características de la imagen y contienen las frecuencias de bajo orden. El número de descriptores elegido varía entre 4 y 4096. El programa que efectúa esta selección es (SVENTAN.FOR).

b) Descriptores que son máximos relativos.

Este criterio se escoge debido a que el valor absoluto de estos descriptores contiene mayor energía, al igual que el caso anterior el número de descriptores varía entre 4 y 4096. El programa que realiza esta selección de descriptores es (RECMAXF.FOR).

- Reconstrucción de la imagen.

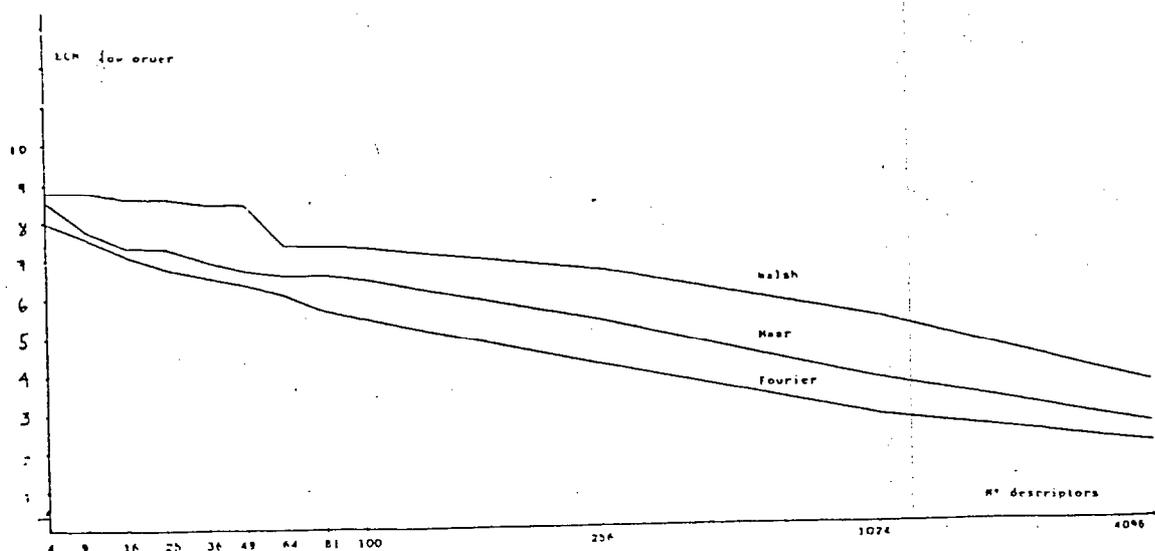
Una vez seleccionados un subconjunto de coeficientes como descriptores, se pasa, con estos, a reconstruir la imagen original, las transformaciones seleccionadas para este estudio tienen transformada inversa, y como la directa, su algoritmo

representación de descriptores más fácil puede efectuarse el reconocimiento.

Para cada pieza se calcula la distancia entre sus descriptores y el resto del conjunto de imágenes.

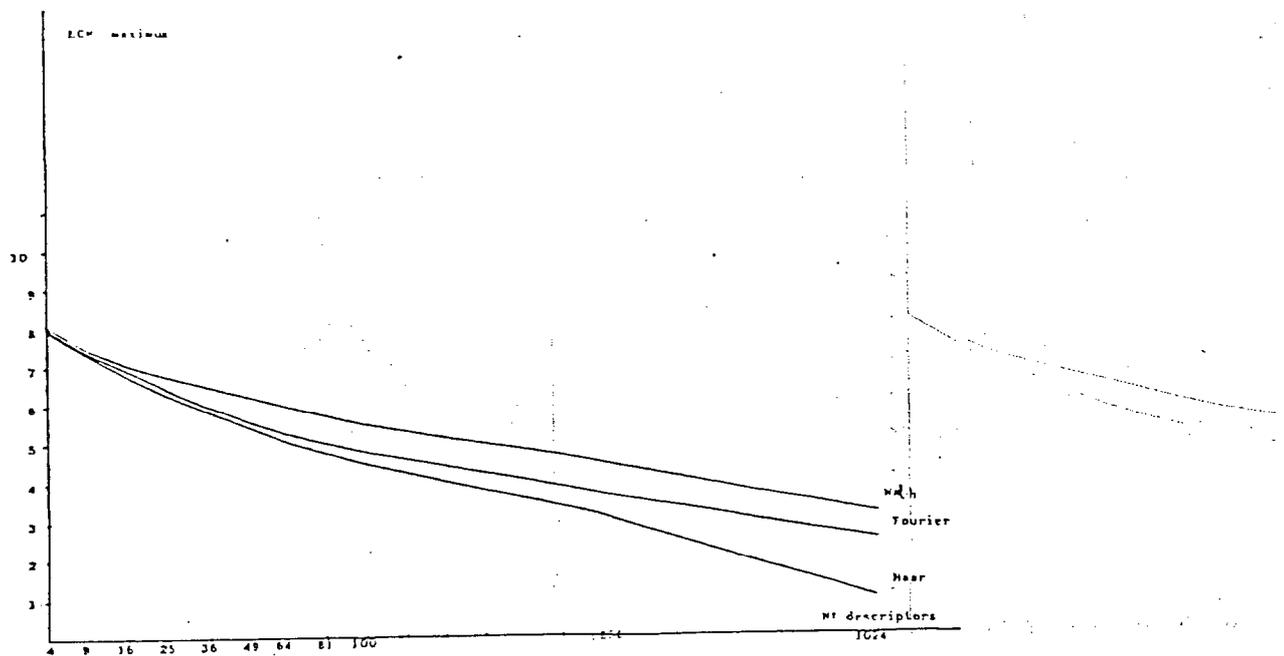
- Descripción de los resultados experimentales.

La figura muestra un gráfico, donde la coordenada horizontal, representa, el número de descriptores de bajo orden utilizados para la reconstrucción de la imagen original y la coordenada vertical el error, cuadrático medio, entre la imagen original y la reconstruida.

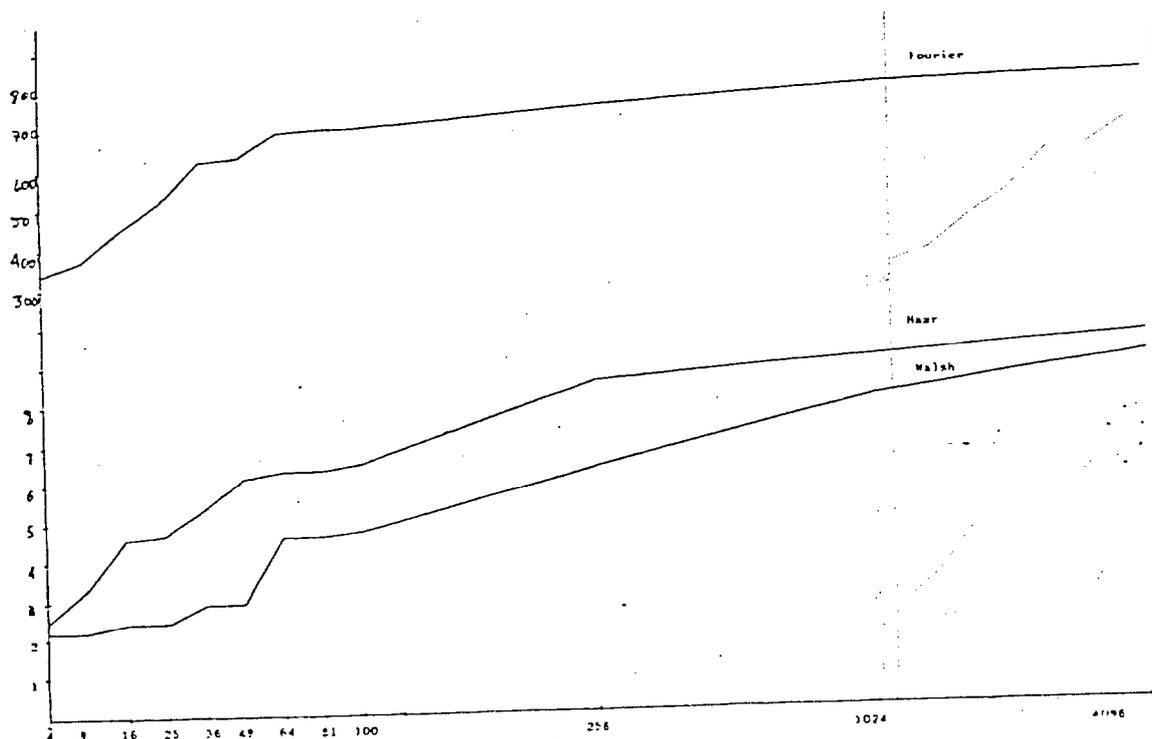


Este gráfico muestra que un incremento en el número de descriptores implica un decremento en el error cuadrático medio, y que el error cuadrático medio, utilizando los coeficientes de FOURIER es menor que utilizando los coeficientes de HAAR, y estos últimos producen un error menor que usando los coeficientes de HADAMARD-WALSH.

La figura muestra el error cuadrático medio versus el número de descriptores utilizados en la reconstrucción, pero en este caso, los descriptores se eligen utilizando el criterio del máximo. Este gráfico muestra que no existe una gran diferencia entre las transformadas para un número pequeño de descriptores sin embargo la transformada de Haar sobresale cuando se incrementa el número de descriptores.



La siguiente figura muestra las distancias medias, versus el número de descriptores utilizando los coeficientes de bajo orden para caracterizar la imagen. Como podemos ver, la distancia entre descriptores incrementa con el número de descriptores. Hay que tener en cuenta que la distancia entre descriptores calculada con la transformada de Fourier, se han utilizado los módulos de los descriptores.



El error cuadrático medio, calculado con los descriptores máximos, es menor, que el calculado con los descriptores de bajo orden, pero en el caso de coeficientes máximos, es necesario guardar la posición de estos, implicando que los algoritmos son un poco más costosos.

CAPITULO:3. TRANSFORMACIONES DE CAMPO RECEPTIVO VARIABLE  
Y SU PAPEL EN LOS SISTEMAS PERCEPTUALES.

### 3.1 TRANSFORMACIONES DE CAMPO RECEPTIVO VARIABLE.

#### 3.1.1. SOBRE EL CONCEPTO DE DESCRIPCIONES COMPLETAS DE UN ENTORNO VISUAL.

Un concepto básico en proceso de imágenes y en visión es el de "descripción completa".

Una descripción se admite que es "completa" en su entorno visual, si la descripción contiene todos los datos, y propiedades de los datos, necesarios para cubrir unos objetivos.

Por ello, en los planteamientos generales de visión artificial no es posible, a priori, establecer las condiciones necesarias y suficientes para decidir si una representación es completa o no en base a los datos visuales y a la naturaleza de la representación. Esto es, es preciso añadir el "propósito" de la representación y aún en el caso de que tal propósito se especifique, es imposible, por el momento, decidir, en general, acerca del carácter completo de la representación, sobre todo porque los objetivos y o propósitos se expresan de forma borrosa y son tales que incluso los límites de los grados de borrosidad son cambiantes.

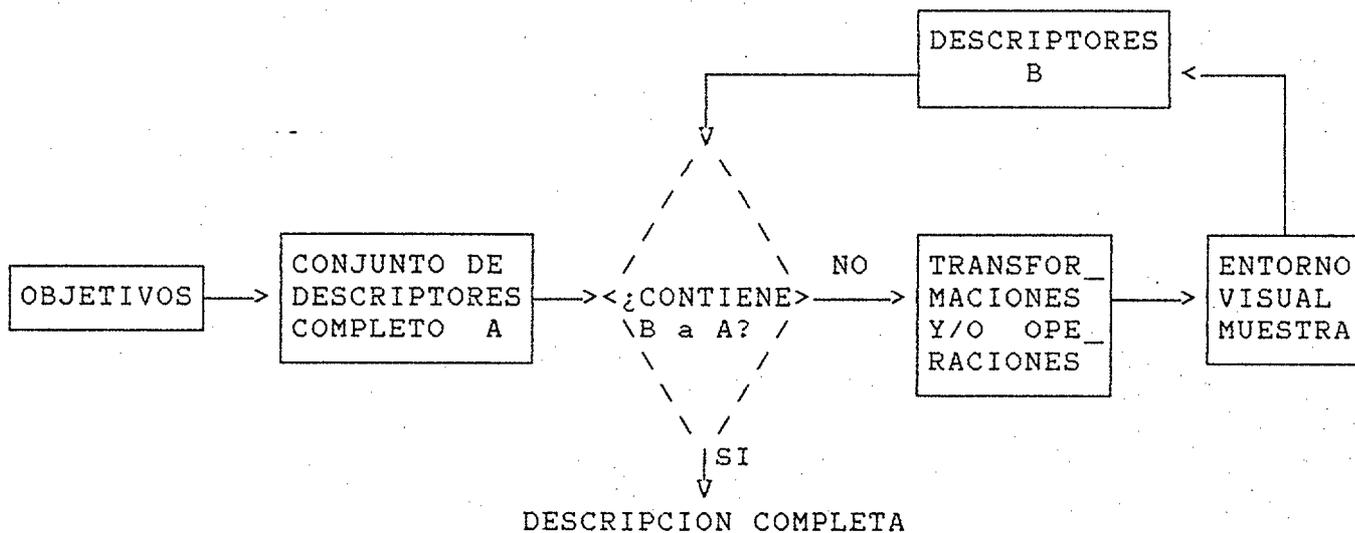
Hay dos métodos obvios de plantear los problemas relativos a la descripción completa, que, aunque en un contexto diferente, se corresponden con las dos formas clásicas de tratar el problema de la visión artificial, esto es, la dualidad de los procesos dominados por los datos frente a procesos dominados por el conocimiento (cap.1).

a) método 1. A partir de los objetivos (normalmente borrosos) se infiere un "conjunto descriptor" que se considera completo. De ahí, se pasa a la búsqueda de modos de descripción que obtengan el "conjunto descriptor" requerido. A continuación, se explora la implementación de operaciones y transformaciones que ejecuten los modos de descripción, verificandose su efectividad para obtener el conjunto descriptor desde un entorno visual muestra.

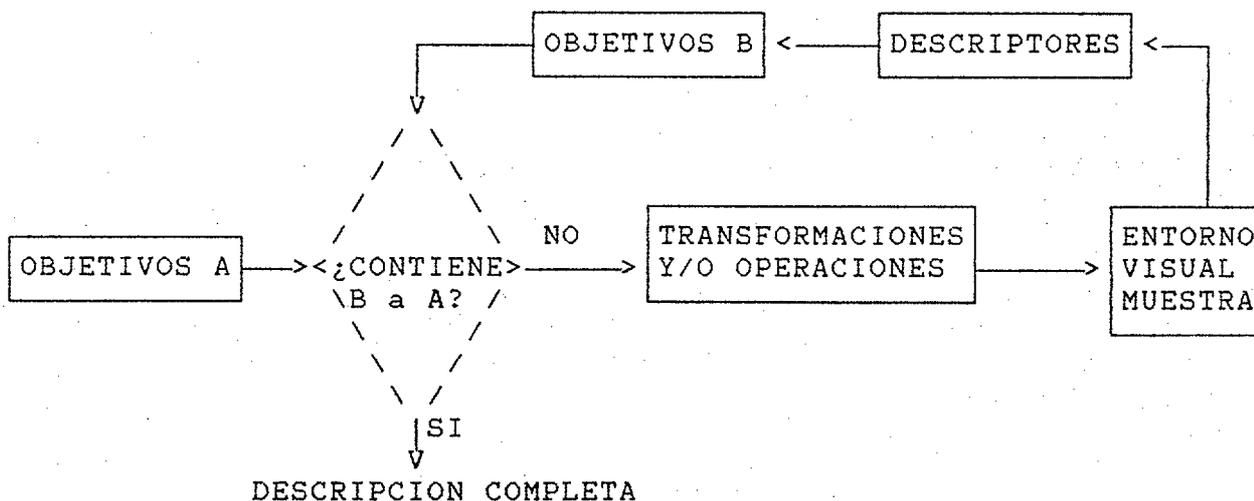
b) método 2. A partir de un entorno visual muestra, se exploran métodos descriptores diversos, proporcionados por transformaciones y/o operaciones sobre el entorno, inspiradas por la percepción natural, la física de los procesos, la analítica o la algorítmica. A continuación se verifica la adecuación del conjunto descriptor obtenido, para cubrir los objetivos, que han de haber sido dados previamente, o para definir clases de objetivos para los cuales la descripción lograda es completa, y ver si un objetivo dado está contenido en la clase. Tales métodos se ilustran en la figura 3.1.1.

Como se ve en la figura, existe una diferencia importante entre los dos métodos; en el método 1 se comparan descriptores mientras que en el método 2 se comparan objetivos para la decisión acerca de que la descripción sea completa. Dada la naturaleza, en general, más borrosa de los objetivos que la de los descriptores el método 2 es más vago y menos formalizable que el método 1, y puede dar resultados aceptables con más frecuencia

(debido precisamente a esa borrosidad) que el método A, sobre todo si la "aceptabilidad" se basa en la adecuación para resolver el problema, aunque con el riesgo de que el Sistema funcione no solo cubriendo los objetivos iniciales A, sino otros, B-A.



(a)



(b)

Fig.3.1.1. Ilustración grafica de los métodos de decisión acerca de que si una descripción es completa.

Como se ve, los dos metodos son operacionales, descansando el método 2, mucho más fuertemente que el 1, en la realización del número de experimentos de naturaleza apropiada.

Es decir, el método 1 puede ser completamente "preciso" es decir, determinista, si es el caso de que la borrosidad en el establecimiento del "conjunto completo de descriptores" desaparezca, y si  $A=B$ . En otras palabras, es conceptualmente más preciso comparar los descriptores necesarios para conseguir dos objetivos distintos, que comparar los objetivos mismos, al menos en visión.

### 3.1.2. OBSERVACIONES ACERCA DEL LENGUAJE FORMAL APROPIADO A CADA NIVEL METODOLOGICO EN UNA DESCRIPCIÓN COMPLETA.

En los esquemas metodologicos de la Sección anterior, resulta que, practicamente, el único lenguaje formal que está fuera de toda discusión es el lenguaje analitico que se usa para describir los datos fisicos que han de conformar el entorno visual. Es decir, a nivel de la definición del entorno visual, la correspondencia de las magnitudes fisicas y su representación debe realizarse por medio de la función de transducción de los sensores.

A partir de ahí, y yendo hacia atrás, es de esperar que el lenguaje apropiado vaya subiendo de nivel, sobre todo si se tiene en cuenta que los objetivos están normalmente expresados en un lenguaje de alto nivel, especialmente en los Sistemas de visión artificial con cierto grado de sofisticación.

Lo anterior permite establecer los dos límites, el superior y el inferior, entre los cuales han de moverse los lenguajes de descripción.

Evidentemente, el concepto de lenguaje (formal o no) es usado aquí en el sentido específico de la definición de un conjunto de "átomos simbólicos" y de estructuras relacionales, con huecos a rellenar por los átomos, con reglas de enganche de las estructuras relacionales, para formar "frases bien hechas" descriptivas del Sistema. En este sentido nos remontamos a los conceptos originales de Charles S. Pierre (ref.....), Russel (ref.....), Mc Culloch (ref....) y Mc Culloch y Moreno Díaz (ref...), sobre la Teoría de Relaciones como sustrato de una clase muy relevante de Lenguajes Formales. El enfoque corresponde a la generalización de la "descripción de Sistemas" frente a la de "descripción de operaciones".

Los niveles a los que nos referimos pueden, por lo tanto, ir desde el nivel simbólico, con diversas tablas de significados, hasta el nivel analítico básico, donde los "átomos" corresponden a variables y las "relaciones" a funciones en el sentido del análisis matemático (ref.....). El segundo nivel es el de los funcionales. El siguiente nivel supone un salto cualitativo, en cuanto a los "átomos" pueden ser obtenibles por funcionales, pero las estructuras relacionales ya no lo son, sino que éstas pertenecen a un nuevo campo donde hay asignados ya tablas de significación. En estas estructuras, las frases descriptoras son del tipo "Existe  $P_1$ , que guarda la relación  $R_1$  con  $P_2$  y ambos,

están conectados por R2 con P3, P4". Es decir, es un enganche de relaciones cuyos huecos están ocupados por "valores" obtenidos por funcionales. Por ejemplo, la notación simbólica descriptiva podría ser:

$$(P1 R1 P2) (R2 P3 P4)$$

donde R1 es una relación binaria y R3 es una relación triádica.

Como conjunto de tales frases, la descripción puede aceptarse como completa a efectos de los objetivos. Es decir, estaríamos en lo que se entiende como "Visión artificial estructurada".

En el siguiente nivel, los "átomos" ya no pertenecen a la analítica, si no a las estructuras relacionales descriptivas, tal que, en una descripción elemental tal como

$$P1 R P2$$

P1 y P2 son el resultado de operaciones relacionales del nivel anterior, que no es analítico.

Dos observaciones adicionales son pertinentes respecto de los lenguajes y los métodos de la figura 3.1.1.

Primero, Si los objetivos o los descriptores son expresables

en un lenguaje de cierto nivel, todos los procesos resultantes han de ser realizables a ese nivel, o a nivel inferior, tal que el bloque difuso denominado en esos esquemas TRANSFORMACIONES ha de copar con el enganche entre el nivel de los objetivos (o el de los descriptores) hasta llegar al nivel analítico básico del entorno visual.

Esta observación, que parece y es evidente, impone la precaución de que la "derivada de niveles" es siempre negativa cuando se procede, en lenguajes, desde los objetivos al entorno visual.

De hecho, lo anterior es justo una precaución que es prácticamente seguida por los esquemas de visión artificial de hoy día. Es decir, si los descriptores que definen unos objetivos son analíticos, no tiene sentido que se realice una transformación relacional sobre funcionales, por ejemplo. Y, también de hecho, queda por demostrar que la posible existencia de "derivadas de niveles" positivas en los caminos objetivos-entorno visual no sea pertinente.

Segundo, El concepto de descripción usado aquí no excluye a priori, esquemas adicionales de decisión que proporcionen una descripción "sintetizada", ya que de hecho, una decisión implica una descripción sintética. Aunque este asunto será tratado en las últimas Secciones del Capítulo, la pertinencia de la observación se refiere a que las estructuras de lenguajes descriptivos pueden contener relaciones condicionales, que incluyen una decisión. De

hecho, la existencia de relaciones condicionales "suaves" (es decir, no avocadas a la acción, sino a la descripción) es un mecanismo útil y apropiado para realizar un salto de nivel hacia arriba en el proceso de descripción de un Sistema. Como especificación adicional, podemos indicar que una relación condicional "suave" se corresponde con un Sistema de varias entradas y varias salidas con una regla de decisión y tal que, las siguientes restricciones se cumplan:

a) Las salidas son mutuamente exclusivas, es decir, en cada "instante de decisión" solo existe una.

b) El nivel del lenguaje de expresión de las salidas es el mismo o cuasi.

c) El nivel del lenguaje de expresión de las entradas es el mismo o cuasi.

Al ser, en definitiva, la estructura del Sistema la de un autómata que "tiene acceso" a lenguajes próximos, la suavidad de la relación condicional descansa en la existencia de un espacio de lenguajes que es "suave" en el sentido de la jerarquía de niveles de lenguajes descriptivos que mencionamos al principio de esta Sección.

### 3.1.3. VISION ARTIFICIAL VERSUS COMUNICACION VISUAL, AMBOS EN ANALITICA.

Admitamos que lo más que subimos en lenguajes es hasta los funcionales. Es decir, dentro de los esquemas metodológicos de la Sección 3.1.1, los objetivos y o los descriptores completos de dichos objetivos son expresables en términos de la analítica de funcionales, es decir, el Analisis.

Aún cuándo a tal herramienta se la escrutina, (con razon) con el ojo crítico de su utilidad, como lenguaje formal valido, a los niveles más abstractos o borrosos (como son en los Sistemas de Computación, o en su interacción, o en Sistemas neuronales o Sistemas sociales), su redondez formal, como lenguaje en los Sistemas artificiales, es paradigmática, puesto que goza de un prestigio histórico inigualable, que es paralelo al prestigio que se ha construido la Fisica dentro de los Sistemas naturales, usando dicha herramienta analítica.

Lo anteriormente expresado se debe entender en el sentido de que bajar en el nivel de lenguajes no supone en principio, el bajar la potencia descriptora, sino que puede ser todo lo contrario.

Dicha potencia descriptora, en lo que aquí respecta, se manifiesta en la facilidad en el establecimiento de objetivos y el de relacionales con el entorno visual, de forma "natural y elegante", para la obtención de "descriptores completos".

En el contexto de la visión artificial, según lo indicado en la Sección 3.1.1, deberemos establecer sólo los objetivos para una clase de problemas ya que la naturaleza del "entorno visual" nos viene dada.

A nivel analítico, los objetivos que se establecen con frecuencia es que una descripción se admite como completa cuando de ella se puede recuperar el entorno visual.

Pero, en el contexto indicado antes, es un contrasentido establecer como objetivo la "recuperación" del entorno visual. Evidentemente, (hablamos ahora de la Visión Artificial enfocada por científicos de lo Artificial) y en tal contexto, si el entorno visual es dato, ¿qué sentido tiene el construir un Sistema para recuperarlo?. La respuesta obvia es que el objetivo no tiene sentido, o lo que es equivalente, tiene una solución tautológica: mandar "cables", punto a punto, desde el entorno visual hasta el espacio de los objetivos.

Pueden haber problemas, como que los cables se entrecrucen, intercambien información, o que algunos se destruyan, o que capten datos de otros entornos no deseables. Pero, el punto que deseamos expresar con claridad es que esos problemas (graves, de gran importancia práctica y económica) no son conceptualmente relevantes para la visión, y en general, para cualquier fenómeno perceptivo.

Desde luego, un sistema exclusivamente dedicado a corregir

tales errores en comunicaciones es altamente razonable y justificable. Incluso en el Sistema nervioso, tales sistemas existen, pero en el sentido de guardianes o vigilantes (house keepers) que garantizan la fiabilidad de la comunicación, y no en el sentido indicado en 3.1.1. Ello ya no es visión (percepción), sino seguridad de Transmisión o de Comunicación.

Es notorio que los problemas derivados de la fiabilidad de las comunicaciones, por su naturaleza de copar con lo "natural", se han planteado de forma analítica (como ha sido la Teoría de la Comunicación de Shannon) y sólo recientemente (y, creemos que debido a un proceso de mimetismo en el uso de herramientas formales mirando las usadas en ciencias de la computación y no debido a una reflexión de cuál es la herramienta apropiada al campo), se empiezan a considerar lenguajes formales de nivel distinto.

Aún admitiendo que la finalidad de ciertos Sistemas visuales sean un puro "house keeping", que garanticen representaciones completas en el espacio de objetivos, que no son más que una reproducción del entorno visual, es apropiado, ilustrativo y sobre todo, inspirador, el considerar sus métodos.

Aquí, el requisito de que la "derivada de niveles" de lenguajes, cuando se marcha desde los objetivos hasta el entorno visual, sea negativa o nula, se rompe. Ello es debido a la acción de la estructura y funciones del Sistema natural (los canales de

comunicación).

#### 3.1.4 TRANSFORMACIONES DE CAMPO RECEPTIVO VARIABLE (TRV).

Dentro de la Teoría de Sistemas, una transformación corresponde a la acción de un sistema sobre el espacio de las entradas y quizá el de las salidas, para generar un nuevo "slide" del espacio de salidas.

En tal estructura, no es difícil definir y visualizar lo que se entiende por "campo receptivo". Por el momento, nos referiremos a sistemas que actúan solamente sobre el espacio de entradas, ya que generalizar para incluir el espacio de las salidas (anteriores) no es difícil. De hecho, muchos sistemas que actúan sobre ambos espacios, en principio, son "reducibles" a un sistema equivalente que actúa solo sobre el espacio de las entradas.

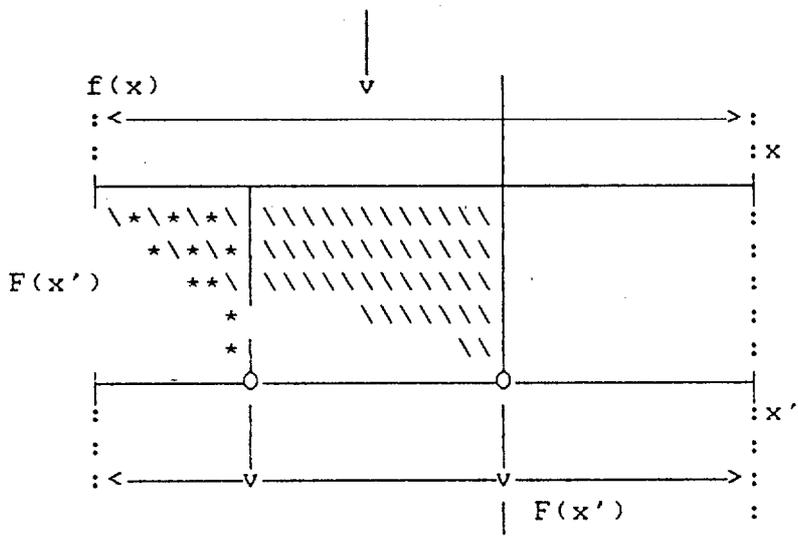
Para llegar a una definición de las TRV procederemos a definir las transformaciones de campo receptivo constante (TRC) y a su negación posterior. Trabajaremos en una dimensión, que es extensible a varias dimensiones sin esfuerzo conceptual adicional importante.

El sistema bajo consideración transforma  $f(x)$  en  $F(x')$ , donde  $x$  y  $x'$  son dominios que pueden coincidir, a través de una transformación arbitraria:

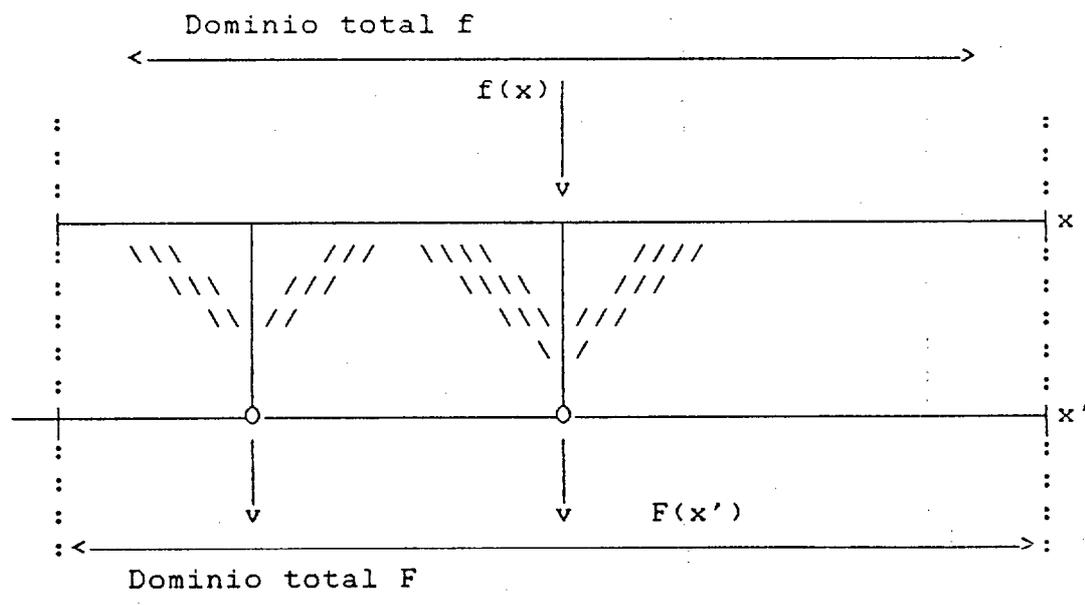
$$F(x') = T [f(x)]$$

Ahora bien, admitimos que existe un "dominio maximo" de definici3n de  $f(x)$ , en lo que respecta a la variable independiente  $x$ .

Por ejemplo, si  $x=t$  (el tiempo),  $x$  va de  $-\infty$  (3 0) a  $t$  (instante actual), igual que  $x'$ , tal que ?? "toma nota" de todos los valores posibles de  $f(t)$  antes de  $t$  para "decidir" su valor  $F(t)$ . Si la variable es espacial, el anterior "principio" de causalidad puede ser roto es decir  $x'$  puede ser tal que  $x' > x$ , tal como se ilustra en la figura 3.1.7.



(a)



(b)

Figura 3.1.7 Ilustración de sistema causal (a) y de sistema incausal (b).

Primeramente, hablaremos de "dominios totales" de las variables  $x$  y  $x'$  (que pueden ir de  $-\infty$  a  $+\infty$ , o de  $-\infty$  a  $x_0$ , o

viceversa o de  $x_1$  a  $x_2$ ). Este se corresponde al dominio de  $x$  donde  $f(x)$  (como familia) puede existir. Lo que esto quiere decir es claro en el caso de una retina con una imagen estatica, pero también queda claro si  $x$  es el tiempo.

Para una "operación"  $F(x')$ , en la posición  $x'$ , su dominio es la "zona" de  $x$  donde se adquieren valores de  $f(x)$  para tomar la decisión  $F(x')$ . Así, tenemos la siguiente condición (C1), obvia como condición suficiente:

C1. Establecido el dominio total de  $f(x)$ , una transformación  $F(x')$  es TRC si, para todo  $x'$ , el dominio de  $F(x')$  es el dominio total de  $x$ . (Condición suficiente).

Por ejemplo, todo sistema causal lineal en el tiempo realiza una transformación TRC, cosa que es muy fácil de demostrar.

Siguiendo con lo anterior, es fácil llegar a una condición más restrictiva para las TRC. Para ello, definimos el concepto de "tamaño" del campo receptivo de una transformación  $F(x')$ . Admitimos que  $x$  soporta una métrica de distancias tal que la distancia entre dos puntos  $x_1$  y  $x_2$  es el valor numérico  $d(x_1, x_2)$ . Así, pasamos a la condición más restrictiva de TRC siguiente:

C2. Dada una distancia  $d(x_i, x_j)$  medida en  $x$ , una transformación  $F(x')$  es TRC si  $F(x')$  decide sobre el dominio comprendido por  $(x_i < x < x_j)$  tal que  $d(x_i, x_j) = \text{cte}$  para todo  $x'$ .

Así, la generalización de lo anterior resulta en cierto modo trivial. En efecto:

Sea el dominio  $x$  y la definición de distancia  $d(x_1, x_2)$  entre dos "puntos" de  $x$ . Sea  $D$  el dominio definido por los intervalos determinados por las parejas  $(x_1x_2, x_3x_4, x_5x_6, \dots, x_lx_m)$ , es decir  $D=(d_{12}, d_{34}, \dots, d_{lm})$  donde  $D(, d_{lm}, )$  es una ristra que contiene las distancias  $d_{lm}$ , para los valores predefinidos de  $l$  y  $m$ .  $D$  es constante si cada  $d_{lm}$  lo es, uno a uno.

Entonces, en general, llegamos a la condición  $C_3$

$C_3$  Una TRC es una transformación  $F(x')$  tal que, para todo  $x'$ , su dominio cumple la condición  $D=cte$ .  $D$  es el campo receptivo.

No es difícil ver que las condiciones anteriores son tales que  $C_3 > C_2 > C_1$ . Hemos procedido en ese orden por razones de claridad.

Luego una TRV es una transformación donde  $D$  no es constante.

### 3.1.5 CONSIDERACIONES ACERCA DEL CARACTER COMPLETO DE UNA TRANSFORMACION CON RELACION AL CAMPO RECEPTIVO Y A LA FUNCION.

Al separar el campo receptivo de la función en sus raíces, se muestra que ambos tienen propiedades genuinas respecto de la

"complitud", que dependen de la propia naturaleza de ambos. Es decir, es posible demostrar teoremas del tipo:

Para toda  $f \in CR \in \{C1\}$  y toda  $F \in \{F2\}$ , entonces  $F \in \alpha$

o bien Para toda  $F \in \{F1\}$  y toda  $CR \in \{C2\}$ , entonces  $F \in \alpha$

Donde (ahí está el interés)  $\{C1\}$  y  $\{F1\}$  son mucho más amplios que  $\{C2\}$  y  $\{F2\}$ .

Dicho en palabras del lenguaje natural:

1) "Para una amplia clase de campos receptivos, si  $F$  pertenece a cierta clase restricta, entonces  $F$  es completa (analíticamente)".

2) "Para una amplia clase de funciones, si los campos receptivos pertenecen a una cierta clase restricta,  $F$  es completa."

Como regla, parece que los  $\{C1\}$  de los teoremas del tipo 1 corresponden a TRV, mientras que los de los teoremas de tipo 2,  $\{C2\}$ , son de TRC.

Es interesante estudiar los dos tipos de teoremas posibles anteriores desde el punto de vista de los grados de libertad admitidos para que se cumpla que  $F \in \alpha$ . La complitud requiere una cierta "constancia en el número de grados de libertad", tal como

se indicó en la Sección 3.1.1. En el sentido analítico, la constancia en el número de grados de libertad nos permite establecer una especie de "principio de conservación" que viene determinado por los campos receptivos y por los funcionales computados por  $F$ , que, además, reflejan una situación de "dualidad", en el sentido de que la completitud requiere aumentar el número de uno de ellos, si el otro disminuye.

Prosigamos de la siguiente forma: Admitamos la existencia de un conjunto de funcionales  $\{F\}$  tales que de ellos se selecciona un conjunto,  $A$ , "independiente", de coeficientes descriptores de las  $f(x)$ , tal que éstos actúan en cada dominio o campo receptivo. El concepto de independencia será precisado con posterioridad, aunque ahora nos referimos a la independencia funcional. Es decir, ninguno es obtenible como función de los otros. Sea un conjunto, de cardinal,  $B$ , de campos receptivos independientes en el sentido funcional anterior (respecto de los dominios o campos receptivos). Sea  $N$  el número de grados de libertad admitido a priori del espacio, o entorno, de entrada.

Entonces, una descripción es completa si

$$A + B = N$$

Para  $B=0$ ,  $A=N$ , tenemos las transformaciones globales puras clásicas. Sin embargo, aunque se han estudiado casos intermedios (eso sí, sin analizar la relación profunda entre las dos formas

duales de descripción) resulta que el caso limite  $B=N$ ,  $A=0$  no ha sido considerado excepto para la situación trivial de TRC, y cuando lo han sido para TRV, se ha hecho énfasis en las transformaciones empleadas, (Hadamard, Haar) lo cual según lo anterior es irrelevante, quiere decir, que da igual la transformación.

Ello nos lleva al concepto de las transformaciones puras de campo receptivo, TCR. Y por consiguiente, en este contexto se permite definir una TCR.

#### DEFINICION

Una TCR es una transformación  $F$  (TRV ó TRC) para la cual  $A=0$ ;  $B=N$ .

Unas TCR triviales son las  $TCR \in \{TRC\}$ , es decir, por ejemplo, la transformación base identidad, o, (en parte), el "moving average", como veremos en la Sección siguiente, aunque ahora parezca existir una inclusión lingüística (TCR= transformaciones puras de campo receptivo y TRC= transformación de campo receptivo constante).

#### 3.1.6 NUEVA VISITA A LAS DESCRIPCIONES COMPLETAS DE UN ENTORNO VISUAL SUGERIDAS POR LA ANALITICA.

Desde el punto de vista analítico, la descripción completa (que, según la heurística y el experimento, podrían ser "cortados" para un sistema de V.A. práctico), requiere "a priori", la conservación del número de grados de libertad, o del

número de "propiedades" independientes, N, del entorno visual.

Como hemos analizado en las Secciones anteriores, es importante en el proceso visual, no sólo la función sino el campo receptivo. De tal forma que si A son los grados de libertad del funcional, se requieren:

$$B = N / A \quad \text{en media}$$

campos receptivos independientes. O bien, si el número de campos receptivos dato independientes es B, requerimos "expandir"  $f(x)$  en cada campo para proporcionar un número

$$A = N / B \quad \text{en media}$$

de coeficientes independientes.

Supongamos que subimos de nivel en la estructura del Sistema. Ello quiere decir que cambiamos de lenguaje formal. Es decir, cambiamos el procedimiento formal para conseguir descriptores independientes de la imagen, o de trozos de la imagen. Sin embargo, en lo que respecta a los campos receptivos, el principio de conservación se mantiene, porque corresponde a una estructura básica que aparece en todos los niveles. Es decir, que si se parte de un entorno visual de dimensión N (N=número de variables independientes de interés) y se formulan A funcionales o extractores de propiedades con un lenguaje de cierto alto nivel

dado, con la condición de independencia a ese nivel, (1) entonces los correspondientes campos receptivos han de tener un grado de independencia  $B = N/A$ .

Esta conjetura tiene una particular incidencia en la representación visual a niveles altos, donde el número de descriptores semánticos se intenta reducir drásticamente. En puridad, esos procesos requerirían aumentar proporcionalmente el número de campos receptores donde se toman los datos para la decisión. Al menos que, a priori, se admita que una reducción drástica (del orden de  $N/A$ ) es, o puede ser apropiada para realizar un reconocimiento "aceptable" en ese nivel, según la aplicación buscada.

De hecho, eso es, asimismo una conjetura. Es decir, dado un problema definido por un entorno visual, se dice que es posible subir en lenguaje para resolverlo prácticamente, a expensas de los grados de libertad necesarios y proporcionables por campos receptivos, los cuales serían requeridos en puridad para una solución formal.

Nota (1): Dado un nivel, un conjunto de valores de los funcionales o extractores de propiedades es independiente si no existe una relación funcional o regla biunívoca que permita obtener unos de otros.

### 3.1.7 PROPIEDADES DE LOS CAMPOS RECEPTIVOS.

Tal como, indicamos antes, una transformación pura de campos

receptivos (TCR) es tal que el número de dominios independientes de  $F(x')$  es igual al número de grados de libertad de  $f(x)$ ,  $N$ .

En las Secciones siguientes consideraremos que  $x$  y  $x'$  son discretas, que en ciertos casos puede generalizarse con facilidad al continuo.

Para la situación de completitud analítica, el número de grados de libertad es el de pixels de la retina. Veamos con más detalle el concepto de "independencia" de los campos receptivos.

Un conjunto  $\{D\}$  de campos receptivos es independiente si ninguno de ellos puede obtenerse como una función (en el sentido analítico) de otros del conjunto. El concepto de "función" requiere que se especifiquen las operaciones entre dominios, que, en esencia desde el punto de vista que aquí nos ocupa, son las operaciones entre conjuntos.

Es decir, un dominio  $D$  (o campo receptivo  $D$ ) es un subconjunto de una Retina. Definimos la unión de dos dominios en el sentido usual de teoría de conjuntos.

Entonces:

#### DEFINICION

Un conjunto  $\{D\}$  es independiente si ningún  $D \in \{D\}$  puede obtenerse por unión de miembros de  $\{D\}$ .

Además si la dimensión  $B$  de  $\{D\}$  ( $B = \text{cardinal}\{D\}$ ) es igual a  $N$ , entonces  $\{D\}$  es completo y cualquier transformación con esos campos receptivos es una transformación pura de campos receptivos (TCR) y la función (o funciones) que se realicen son irrelevantes respecto a la complitud.

Como indicamos antes, un dominio puede ser caracterizado por distancias  $d_{mn}$ . Consideremos los dominios tales que cada uno es caracterizable por una única distancia. Es decir, sea  $\{D\}$  un conjunto de  $M$  dominios  $D_i$ , con distancias  $d_i$ . Estas distancias pueden ordenarse de menor a mayor y si algunas son iguales, por la proximidad al origen del menor elemento de  $D$ . Llamaremos  $C_r$  a este criterio de ordenación. Ello nos permite la siguiente definición, base de lo que será tratado en la sección 3.2.

#### DEFINICION.

Una Transformación de Resolución Progresiva (TRP) es una en la que:

- a)  $\{D\}$  es independiente.
- b)  $\{D\}$  es ordenado, con el criterio  $C_r$
- c)  $\{D\}$  es completo ( $\text{cardinal}\{D\} = N$ )

Es fácil ver que la transformación Hadamard es una TRP, y que las propiedades relevantes de dicha transformación se deben, en esencia, a que es TRP y no a la forma de las funciones (o núcleos) empleados.

3.2 GENERACION DE TRANSFORMACIONES DE RESOLUCION PROGRESIVA (TRP).

3.2.1 METODOS DE GENERACION.

Consideraremos en esta Sección métodos sistematicos para generar TRP's, que en definitiva, consisten en hacer N particiones de la retina, independientes y ordenables.

Estas particiones pueden representarse en una matriz de NxN, donde cada fila representa uno de los N dominios de {D}. La partición más elemental es la Identidad, representable con unos en la diagonal principal:

$$\begin{array}{c}
 \text{Retina} \\
 \leftarrow \text{-----} \rightarrow \\
 \begin{array}{l}
 (1 \ 0 \ 0 \ 0 \ \dots \ 0) \ D1 \\
 (0 \ 1 \ 0 \ 0 \ \dots \ 0) \ D2 \\
 (0 \ 0 \ 1 \ 0 \ \dots \ 0) \ . \\
 (\dots \ \dots \ \dots \ \dots) \ . \\
 (0 \ 0 \ 0 \ \dots \ 1 \ 0) \ . \\
 (0 \ 0 \ 0 \ \dots \ 0 \ 1) \ DN
 \end{array}
 \end{array}$$

En donde, de inmediato, se ve que cumple las condiciones de TRP.

Una forma sencilla de generar una TRP's de gran interés es tomar la Identidad como generadora y proceder de la forma siguiente:

- a) Tomar los dominios  $I_i$  de  $I$  alternativamente (el  $I_1, I_3, I_5, \dots$ ), es decir  $N/2$  elementos, para generar  $D_1, D_2, \dots, DN/2$ .



Que el {D} es un conjunto independiente resulta del procedimiento de partición seguido. Para cada dimensión d se ve que, por construcción, los Di son independientes. Y en general, todo Di contiene al menos un elemento Ij que no está contenido en los demás.

Que {D} es completo resulta de la condición de independencia y de que se generan N dominios, ya que el número total es:

$$N/2 + N/4 + N/8 + \dots + N/N = N$$

(se admite que N es potencia de 2, para generar enteros).

En cuanto a la ordenación, la que se propone coincide con la del criterio Cr.

Este procedimiento generador es de interés ya que proporciona, asimismo, una forma de generar la transformada rápida correspondiente, como se verá en la Sección siguiente.

Consideremos otra ilustración, correspondiente a los campos receptivos del "moving average" para una retina discreta. Si el promedio se hace sobre un campo receptivo de dimensión d=cte (entera), entonces la regla de partición es la siguiente:

a) El primer elemento es el (1 0 0 0 0 0 0 0.....0 0)

b) El segundo elemento es el (1 1 0 0 0 0 0 0.....0 0)

c) Proseguir hasta el elemento d  $(1\ 1\ 1\ 1\dots 1\ 0\ 0\dots 0\ 0)$   
 $\leftarrow d \rightarrow$

d) Los siguientes elementos consisten en desplazar el "campo" de en la retina hasta que el pixel de orden N es 1, es decir hasta la situación  $(0\ 0\dots 0\ 0\ 1\ 1\ 1\ 1\dots 1)$   
 $\leftarrow d \rightarrow$

e) A partir de la situación d), generar elementos desplazando a la derecha el dominio y eliminando unos:

$(0\ 0\ 0\ 0\ \dots\dots 0\ 0\ 1\ 1\ 1\ 1\dots 1)$   
 $\leftarrow d-1 \rightarrow$   
 d-1  
 (. . . . . )  
 $(0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ \dots 0\ 0\ 0\ 1)$

Que los dominios son independientes es trivial. Que el conjunto es completo es obvio ya que existen:

d elementos en las fases de a)..c)

N-2d elementos en la fase d) y

d elementos en la fase e)

Es decir  $d + N - 2d + d = N$

El orden ha de ser alterado, sin embargo, para cubrir el criterio Cr. Una vez alterado, el "moving average" es una TRP.

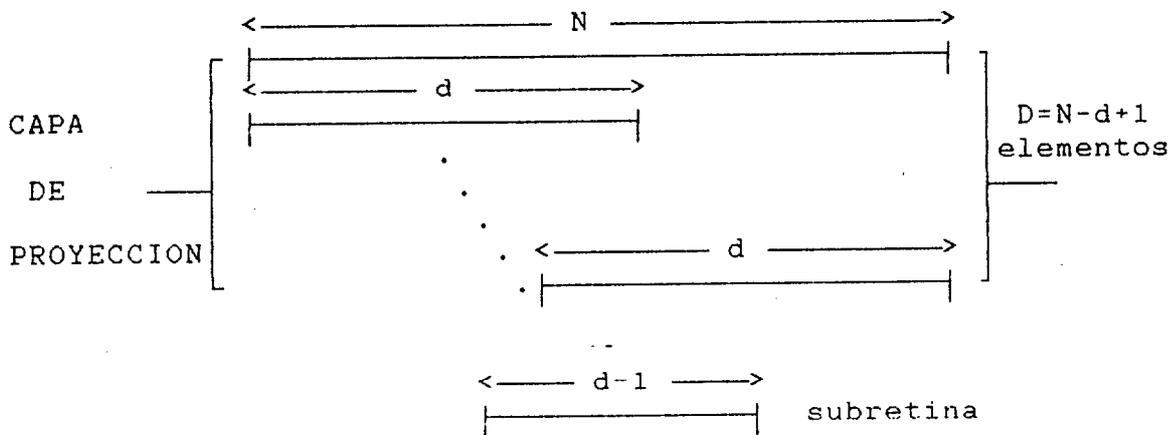


Figura 3.2.1

### 3.2.2 TRANSFORMACIONES FOVEALES.

Una clase particular de TRP's, que están inspiradas en la estructura del "moving average", son las transformaciones que podemos denominar "foveales", ya que para una retina, proporcionan la máxima resolución "directa" en la parte central, o "fovea". Estructuralmente, el concepto puede parecer desorientador, porque, de hecho todos los pixels son recuperables, al ser la TRP completa, independientemente de su distancia al centro de la retina. Sin embargo, representando directamente la transformada, resulta que, precisamente, la máxima resolución corresponde a la parte central. Hemos de indicar que, en general, el criterio de representación de la transformada es un asunto de gran importancia como se deduce de las distintas discusiones de representación de una TRP ya aceptada, la de Hadamard, y será considerado con detalle más adelante.

La idea de TRP foveales parte del hecho de que en una partición completa global de una retina podemos jugar con los grados de libertad iniciales, para deducir, o imponer, campos receptivos o dominios independientes, a nuestra elección. La idea se ilustra imponiendo la existencia de todos los campos receptivos de una dimensión dada,  $d$ , sobre la retina, de forma que sean independientes (como en el caso del "moving average"). Para una dimensión dada,  $d$ , existen

$$D = N - d + 1 \quad (\text{figura 3.2.1})$$

dominios linealmente independientes. Estos  $D$ , aumentados con un número  $d-1$  de dominios linealmente independientes de dimensión 1 arbitrarios (es decir, dominios correspondientes a la transformación identidad), forman una TRP. Es decir:

#### TEOREMA

Dada una retina de  $N$  grados de libertad y una dimensión  $d < N$  arbitraria, existe una TRP formada por todos los dominios linealmente independientes de dimensión  $d$ , más  $d-1$  dominios unitarios distintos y arbitrarios.

En efecto: los  $N-d+1$  dominios de dimensión  $d$  son independientes, por construcción. Asimismo lo son los  $d-1$  dominios unitarios distintos y arbitrarios.

Formemos la matriz  $A$ , de  $N \times N$ , por yuxtaposición de las

matrices

a, de dimensión  $(N-d+1)*N$  (dominios de dimensión d)

b, de dimensión  $(d-1)*N$  (dominios unitarios)

Es decir

$$A = \begin{bmatrix} a \\ b \end{bmatrix}$$

Sea Y una matriz vertical de  $N+1$  componentes de valores arbitrarios y distintos de cero.

Sea

$$Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

Veamos que la ecuación  $Y = A X$

tiene siempre solución inversa  $X = A^{-1} Y$

En efecto, si  $d_{ik}$  son las componentes de  $b$  ( $i=1,2,\dots,d-1$ ;  $k=1,2,\dots,N$ ), se tiene la solución  $Y_i = X_k$ , de  $y_2 = b x$ , para  $d-1$  valores.

Sustituyendo estos valores en la ecuación

$$y_1 = a X$$

y pasando al primer miembro los terminos ya calculados de  $X_k$ , resulta la ecuación reducida

$$y = a'x$$

donde  $a'$  es una matriz de dimensiones  $(N-d+1) \times (N-d+1)$ , obtenida de  $a$  eliminando las columnas donde  $d_{ik}$  es 1. Pero  $|a'| \neq 0$  y el conjunto de dominios es independiente. q.e.d.

Puesto que los dominios unitarios pueden ser cualesquiera los podemos elegir de forma que estén "centrados", es decir que su unión proporcione un dominio de dimensión única situado en el centro de la retina. Este "dominio unión" puede ser considerado como una subretina, a la que es aplicable de nuevo el teorema anterior, o bien cualquier otra TRP.

Una TRP con las propiedades anteriores es lo que entendemos como transformación foveal.

Existen procedimientos sistemáticos para conseguir las TRP foveales como el siguiente:

Definimos el número de subconjuntos de dominios de dimensiones diferentes que se van a admitir. A cada subconjunto se le denomina "capa de proyección". Sea  $K$  el número de tales capas, tales que el cardinal de cada capa es constante.

Se tiene que dicho cardinal es

Parte entera de 
$$\left[ \begin{array}{c} N \\ - \\ K \end{array} \right] = \left[ \begin{array}{c} N \\ - \\ K \end{array} \right] E$$

Para la primera capa se tiene una dimensión

$$d_1 = N - \left[ \begin{array}{c} N \\ - \\ K \end{array} \right] E + 1$$

e iterativamente, para la capa  $i$

$$d_i = d_{i-1} - \left[ \begin{array}{c} N \\ - \\ K \end{array} \right] E$$

En general si  $N/K$  no es entero, la última capa proporciona una subretina de dimensión:

$$N - (K-1) \left[ \begin{array}{c} N \\ - \\ K \end{array} \right] E$$

Como ilustración, consideremos una retina de  $N=128$ , con 4 capas de proyección. Se tiene que el cardinal de cada capa es:

$$\left[ \begin{array}{c} 128 \\ - \\ 4 \end{array} \right] E = 32$$

Es decir, cada subconjunto de dominios está formado por 32 elementos.

La dimension de los elementos de las capas son:

$$d1 = 128 - 32 + 1 = 97$$

$$d2 = 97 - 32 = 65$$

$$d3 = 65 - 32 = 33$$

$$d4 = 33 - 32 = 1$$

y las de las subretinas:

$$\text{SUB1} \quad 97 - 1 = 96$$

$$\text{SUB2} \quad 65 - 1 = 64$$

Otro ejemplo es el siguiente: Una retina de 128, con 6 capas

Se tiene 
$$\left[ \begin{array}{c} N \\ - \\ K \end{array} \right] E = 21$$

y la subretina final será de  $128 - 5 \cdot 21 = 23$  dominios unitarios

Se tiene en este caso

$$d1 = 128 - 21 + 1 = 106$$

$$d2 = 106 - 21 = 85$$

$$d3 = 85 - 21 = 64$$

SUBRETINAS  
105

84

63

$$\begin{array}{rclcl} d4 & = & 64 - 21 & = & 43 & 42 \\ d5 & = & 43 - 21 & = & 22 & 23 \\ d6 & = & 22 - 21 & = & 1 & \end{array}$$

es decir, 5 capas de 21 dominios por capa, más una capa final de 23 dominios unitarios.

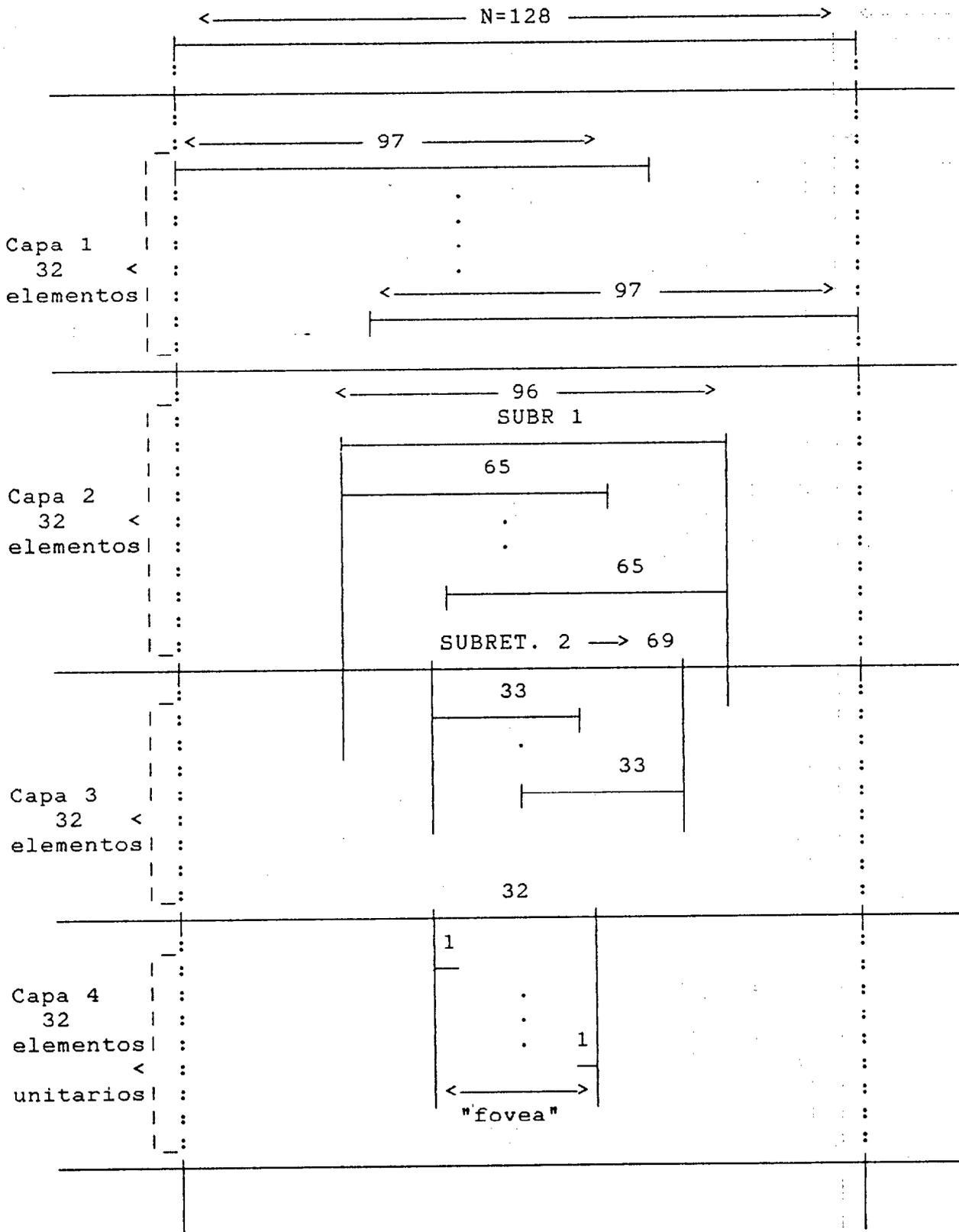


figura 3.2.2.

### 3.2.3 OTRAS TPR's.

El problema de la partición de la retina en campos receptivos o dominios independientes tiene otros planteamientos que son de interés en lo que respecta a la generación de proyecciones por capas en los Sistemas naturales, concretamente en corteza, dado que asimismo estas particiones sugieren procedimientos en visión artificial.

Por ejemplo, en la TPR's foveales se puede fijar el tamaño de la subretina final y el número de capas para generarlas. Una ilustración es la siguiente: una retina de  $N$ ,  $k$  capas y una subretina final de dimensión  $d$ .

Podemos proceder admitiendo que el proceso es equivalente a aplicar una transformación foveal a una retina de  $N - d$  grados de libertad.

Fijamos arbitrariamente el número de elementos por capa con la condición de que sumen  $N-d$ . Sean  $n_i$  esos números. Entonces, las dimensiones son:

$$d_1 = N - n_1 + 1$$

$$d_i = d_{i-1} - n_i$$

Por ejemplo, sea  $N = 128$ ,  $d=9$  y tres capas. Una partición "equilibrada" es:

n1 = 40  
 n2 = 40  
 n3 = 39

con lo que

	SUBRETINAS
d1 = 87	86
d2 = 47	46
d3 = 8	9

Otra situación de interés es el de las particiones al azar, que dejaremos abierto para estudiarlo en el futuro. Esto se refiere a la generación al azar de conjuntos de campos receptivos (situación que podría parecer que se da en las retinas naturales). Por ejemplo, consideremos de partida un conjunto de dominios o campos receptivos posibles de dimension  $M > N$ . Puede tratarse por ejemplo, de elegir al azar  $N$  de los  $M$  de tal manera que el conjunto resultante sea independiente. O bien, imponer la restricción adicional de complitud. Asimismo, si se expresan las reglas de decisión con un "costo" temporal, se puede preguntar, evolutivamente, ¿cual es el tiempo consumido por cada rama de decisión para llegar a una "retina" de descripción completa?

#### 3.2.4 PROYECCIONES DE LAS TRP'S.

Una cuestión importante es el de la forma de proyectar una transformación, o bien de representar los resultados de la misma. Para la mayor parte de las transformaciones en las que los grados de libertad vienen determinados por la función y no por el campo receptivo, suele existir una "ordenación" natural para realizar tal proyección. Es decir, para una transformación

$$F(u,v) = T (f(x,y))$$

se define el espacio  $u, v$  y se van representando los valores de  $F$  correspondientes. Tal es, por ejemplo la transformación que nos da el módulo de la transformada bidimensional de Fourier.

$$A(w_1, w_2) = | TF(f(x,y)) |$$

donde las nuevas variables de representación, continuas o discretas, son  $w_1$  y  $w_2$ .

Para las TRP's y en general las TCRV's existe un método sistemático, que está basado en que los conjuntos de dominios independientes que definen los campos receptivos sean ordenables, (según, p.e., el criterio  $C_r$  de la sección 3.1.8), ya que la ordenación resultante puede usarse para la proyección.

Sin embargo, la existencia de un criterio de ordenación  $\{D\}$  implica a priori, que dicho criterio sea el apropiado para la proyección TRCV o TRP.

Según se verá prácticamente en las Secciones siguientes, un método de proyección apropiado para muchas TRP's es el siguiente:

- a) Generar subconjuntos de dominios (o campos receptivos) de la misma dimensión.
- b) Ordenarlos según el criterio de la Sección 3.1.8 respecto

de su distancia al origen.

c) Proyectar los resultados de la transformación en un espacio donde las coordenadas siguen ese orden.

Las variantes de este procedimiento se desarrollarán en la Sección 3.4.

Para las TRP foveales puede interesar, sin embargo, un procedimiento de proyección diferente. Refiriendonos al esquema de la figura 3.2.1, resulta natural admitir que la subretina actual se proyecta centralmente, mientras que las correspondientes a los dominios {D}, se reparten alrededor de la fovea, tal como se ilustra en la figura 3.2.4.

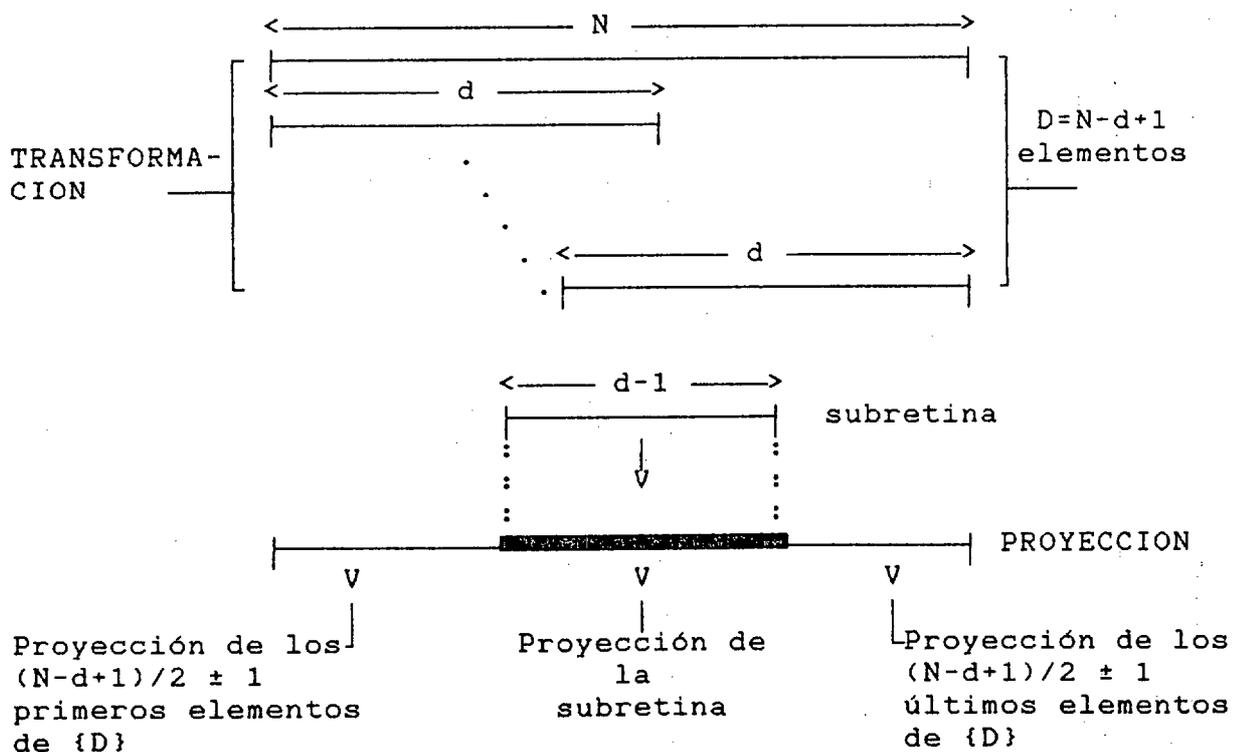


Figura 3.2.4

Este procedimiento de proyección puede extenderse al caso genérico de varias capas de proyección, es decir, a estructuras como la de la figura 3.2.2.

### 3.3 ALGORITMOS DE CALCULO SECUENCIAL RAPIDO DE LAS TRP.

#### 3.3.1 REPRESENTACION MATRICIAL DE LAS TRP.

Tal como indicamos en la Sección 3.2.1, las particiones de campos receptivos para generar una TRP pueden representarse en una matriz de  $N \times N$ , para una retina lineal. Nos referimos en principio, a la TRP de dicha Sección (figura 3.3.1). Si  $I$  es una matriz de  $N \times 1$ , correspondiente a la imagen de entrada, e  $I'$  es la de la salida, es decir:

$$I = \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{bmatrix} \quad I' = \begin{bmatrix} I'_1 \\ I'_2 \\ \vdots \\ I'_N \end{bmatrix}$$

Eligiendo como funcional el valor medio, resulta:

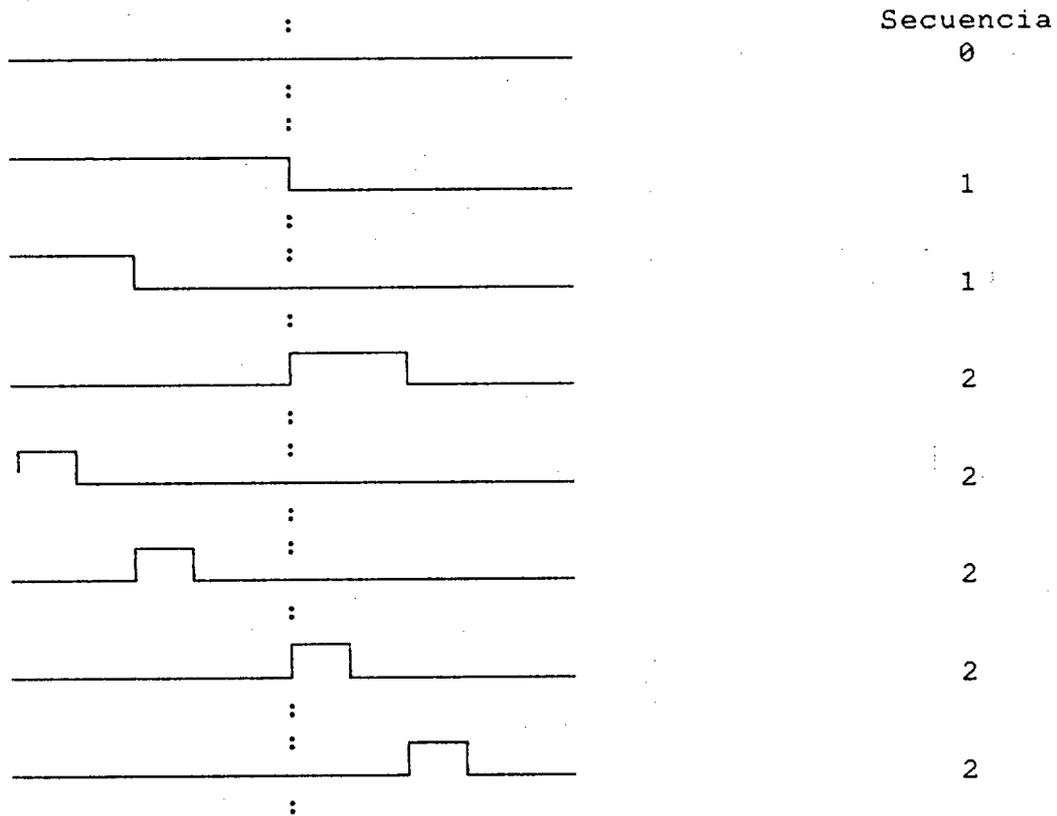
$$I' = M \cdot I, \quad \text{con } |M| \neq 0$$

$$\text{y por lo tanto} \quad I = M^{-1} \cdot I'$$

$$M = 1/N \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & \dots & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \dots & \dots & 0 & 0 & 0 \\ \vdots & \vdots \\ 1 & 1 & 0 & 0 & 0 & \dots & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ \vdots & \vdots \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \dots & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Figura 3.3.1

En la siguiente figura 3.3.2 representamos la forma de estas funciones RP, para el caso de N=8.



Funciones RP para N=8

Figura 3.3.2

La extensión a dos dimensiones se realiza facilmente.

En efecto, sean  $I(i,j)$  e  $I'(k,l)$  las matrices de  $N \times N$  que

representan las imagenes de entrada y salida. Entonces la transformación TRP en dos dimensiones con el funcional "valor medio" es:

$$I'(i,j) = \sum_{kl} M(i,k,j,l) I(k,l)$$

Existen varias formas de generar el "cubo cuatridimensional" matricial que corresponde a una TRP en dos dimensiones. Una forma es "cruzar" dos valores de la TRP en una dimensión, con la de la otra, para generar un conjunto de N "cubos" matriciales, de N\*N\*N, donde cada cubo permite computar una fila de la I'. El proceso se ilustra, en general en la figura 3.3.3.

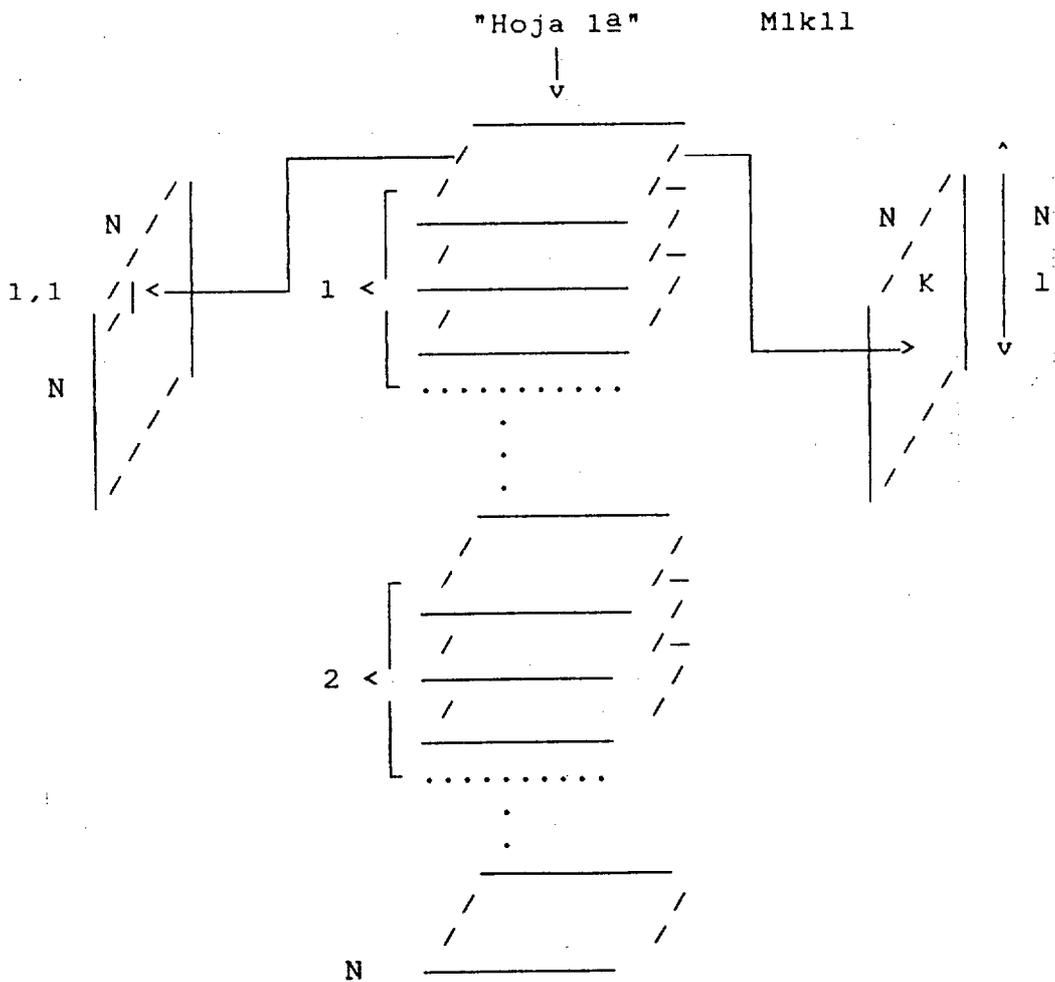


figura 3.3.3

La primera "hoja"  $M(1,k,1,1)$  se obtiene "cruzando" la primera fila  $M(1,k)$  consigo misma por productos internos, es decir:

$$M(1,k,1,1) = M(1,k) * M(1,1)$$

y en general,

$$M(i,k,j,1) = M(i,k) * M(j,1)$$

luego:

$$I'(i,j) = \sum_{kl} M(i,k) M(j,1) I(k,l) = \sum_k M(i,k) \sum_l I(k,l) M(1,j)^T$$

Es decir, en lenguaje matricial, la transformación es:

$$I' = M I M^T = T(I)$$

Este tipo de transformación tiene la propiedad de que la transformada de una imagen traspuesta es la traspuesta de la transformada, es decir,

$$I' = T(I)^T = M I^T M^T$$

En efecto, trasponiendo la ecuación anterior, resulta

$$I' = M(I M) = (I M)M = M I M$$

Es decir, el considerar filas o columnas es irrelevante, siempre que se tomen en su orden (primero filas y después columnas en cada producto matricial, o viceversa).

La ecuación de generación:

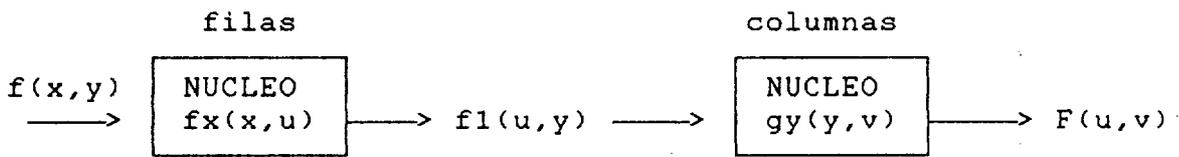
$$M(i,k,j,l) = M(i,k)*M(j,l)$$

proporciona una condición suficiente de generación de un conjunto de dominios {D} en dos dimensiones que es independiente y completo si, en una dimensión, M(i,j) lo es. Notese que en principio, conjuntos {D} completos en dos dimensiones pueden obtenerse "cruzando" A y B en una dimensión, con la condición de complitud, es decir:

$$M(i,k,j,l) = A(i,k)*B(j,l)$$

### 3.3.2 ALGORITMO DE CALCULO SECUENCIAL RAPIDO DE TRP.

Segun lo anterior, las transformadas TRP's en dos dimensiones, tienen núcleos que son factorizables en producto de otros más sencillos iguales, lo que hace que la transformada de una matriz pueda realizarse de dos pasadas, una por filas y posteriormente otra por columnas.



Estos algoritmos rápidos reducen el número de operaciones aritméticas de  $N+2$  a solamente  $2(N-1)$ , para  $N=2+n$

La figura representa una matriz reducida para  $N=8$ .

		F0	F1	F2	F3	F4	F5	F6	F7	
H =	1/8	1	1	1	1	1	1	1	1	C0
	1/4	1	1	1	1	0	0	0	0	C1
	1/2	1	1	0	0	0	0	0	0	C2
	1/2	0	0	0	0	1	1	0	0	C3
		1	0	0	0	0	0	0	0	C4
		0	0	1	0	0	0	0	0	C5
		0	0	0	0	1	0	0	0	C6
		0	0	0	0	0	0	1	0	C7

matriz para  $N=8$

Figura 3.3.4

Esta matriz puede expresarse como producto de matrices más sencillas con gran número de ceros.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}
 \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}
 \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

H1

H2

H3

$$H = H1 * H2 * H3$$

Figura 3.3.5

Una característica atractiva de las transformadas TRP's es que efectúan operaciones muy sencillas de tipo suma; podemos analizar los pasos del proceso de cálculo para un caso sencillo de 8 elementos (N=8).

Sea  $F_0, F_1, F_2, F_3, F_4, F_5, F_6, F_7$ , los valores de la función original.

Se procederá a calcular el valor medio por parejas de valores.

$$A_0 = (F_0 + F_1) / 2$$

$$A_1 = (F_2 + F_3) / 2$$

$$A_2 = (F_4 + F_5) / 2$$

$$A_3 = (F_6 + F_7) / 2$$

Con estos nuevos valores  $A_i$  se procederá a realizar el mismo cálculo del valor medio entre par de puntos.

$$B_0 = (A_0 + A_1) / 2$$

$$B_1 = (A_2 + A_3) / 2$$

El proceso se realiza recursivamente hasta llegar a un sólo valor, en que se finaliza.

$$C_0 = (B_0 + B_1) / 2$$

Con lo que se obtienen todos los valores transformados

$$C_0 = (B_0 + B_1) / 2 = (A_0 + A_1 + A_2 + A_3) / 4 = (F_0 + F_1 + F_2 + F_3 + F_4 + F_5 + F_6 + F_7) / 8$$

$$C_1 = B_0 = (A_0 + A_1) / 2 = (F_0 + F_1 + F_2 + F_3) / 4$$

$$C_2 = A_0 = (F_0 + F_1) / 2$$

$$C_3 = A_2 = (F_4 + F_5) / 2$$

$$C_4 = F_0$$

$$C_5 = F_2$$

$$C_6 = F_4$$

$$C_7 = F_6$$

Este cálculo, se visualiza en forma de árbol en la siguiente figura 3.3.6:

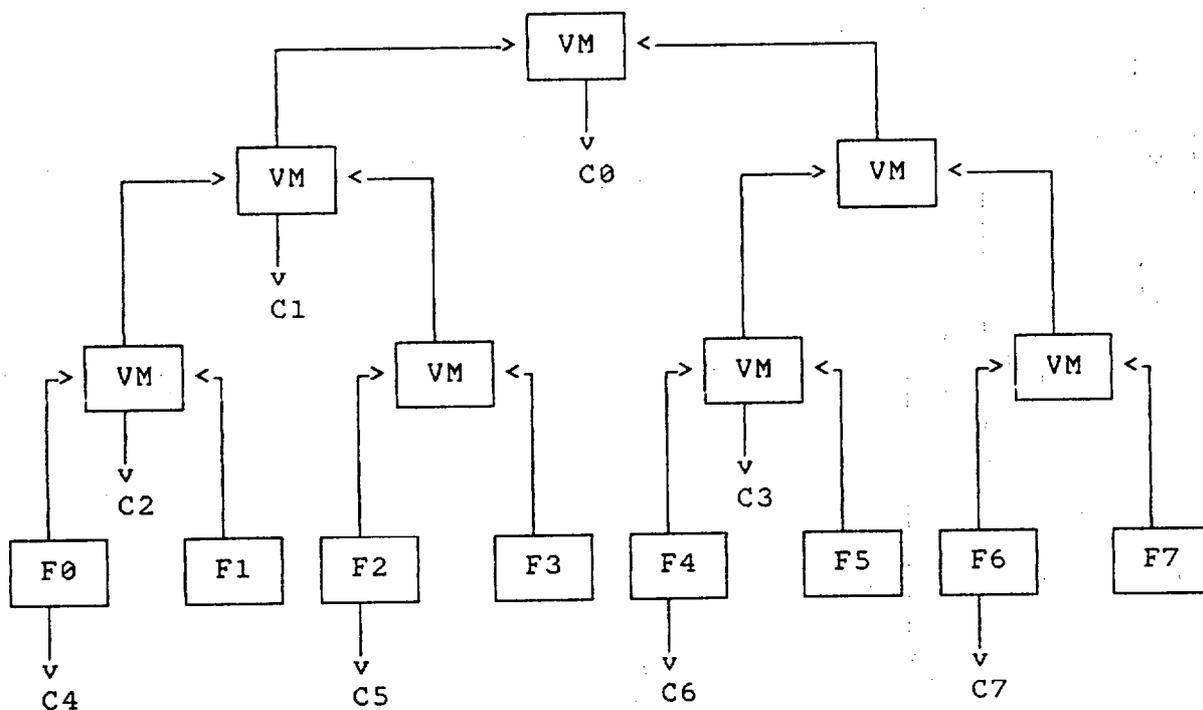


Figura 3.3.6

Donde cada nodo padre realiza la operación "VM" valor medio entre sus nodos hijos.

En forma algoritmica el cálculo de las TRP'S se podría escribir de la siguiente forma

Mientras n° elementos del vector > 1 hacer

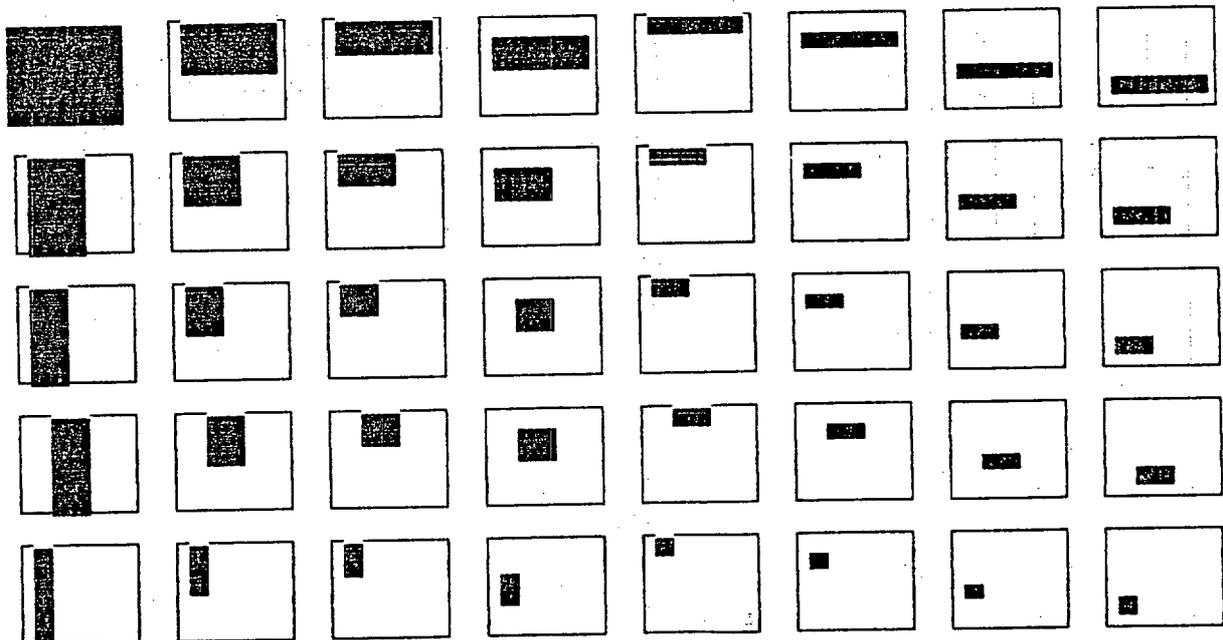
tomar los elementos impares del vector como coeficientes.

generar un nuevo vector promediando parejas de elementos.

Fin mientras

tomar unico elemento del vector como coeficiente.

A continuación presentamos en la figura 3.3.7 la forma que tienen los núcleos de las transformadas TRP'S para imagenes bidimensionales de 8\*8 elementos donde el color negro denota el valor uno y el color blanco el valor cero.



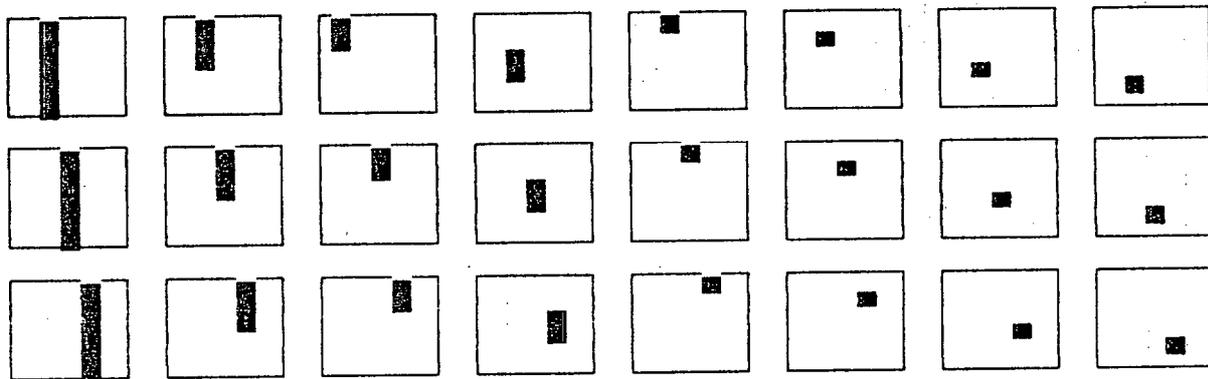


Figura 3.3.7

El núcleo de la transformada inversa puede expresarse mediante la siguiente matriz de la figura:

	C1	C2	C3	C4	C5	C6	C7	C8	
0	0	0	0	1	0	0	0	0	F1
0	0	1	0	-1	0	0	0	0	F2
0	0	0	0	0	1	0	0	0	F3
0	1	-1	0	0	-1	0	0	0	F4
0	0	0	0	0	0	1	0	0	F5
0	0	0	1	0	0	-1	0	0	F6
0	0	0	0	0	0	0	1	0	F7
1	-1	0	-1	0	0	0	-1	0	F8

matriz para N=8

Figura 3.3.8

Esta matriz puede expresarse como producto de matrices más sencillas con gran número de ceros.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}
 \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1 \end{bmatrix}
 \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 1 \end{bmatrix}$$

H1

H2

H3

$$H = H1 * H2 * H3$$

Figura 3.3.9

Lo que hace que esta transformación sea susceptible, igualmente, de algoritmo rapido, este, se ilustra en forma de arbol en la siguiente figura:

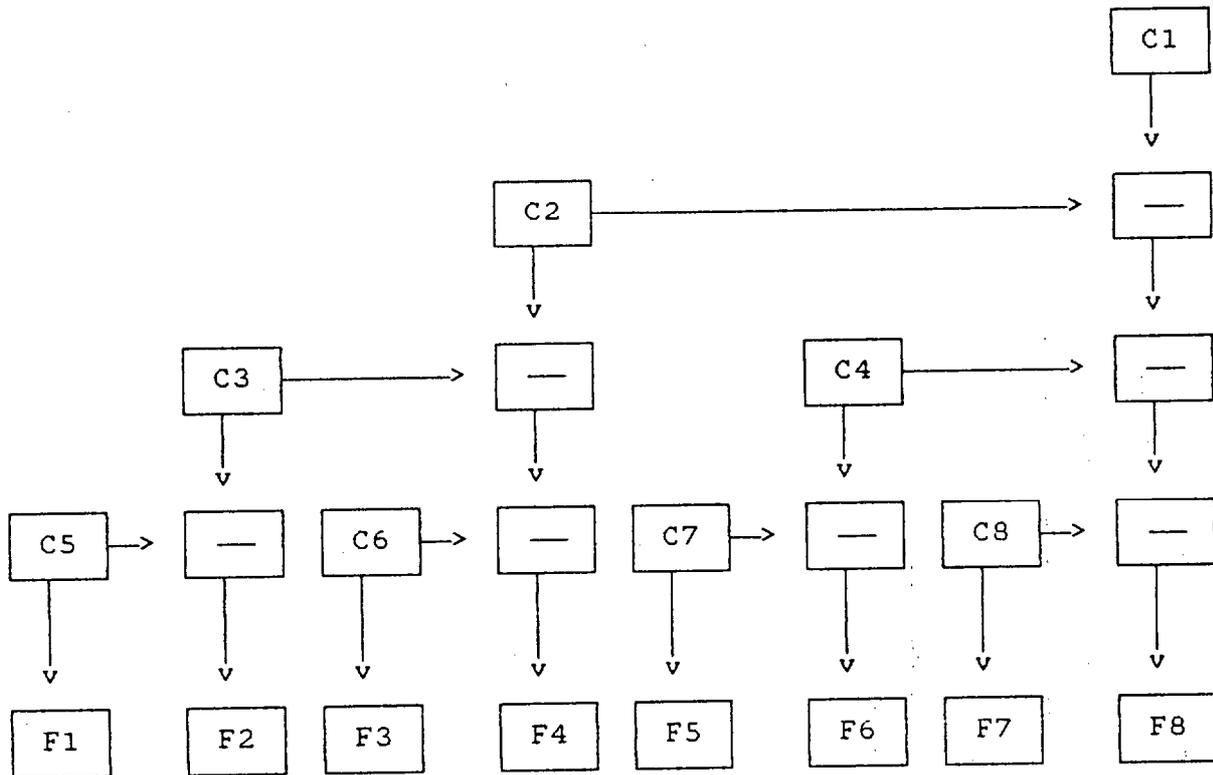


Figura 3.3.10

Los valores originales  $F_i$ , pueden ser recuperados facilmente mediante los coeficientes transformados  $C_i$ , mediante las expresiones.

$$\begin{aligned}
 F_8 &= C_1 - C_2 - C_4 - C_8 \\
 F_7 &= C_8 \\
 F_6 &= C_4 - C_7 \\
 F_5 &= C_7 \\
 F_4 &= C_2 - C_3 - C_6 \\
 F_3 &= C_6 \\
 F_2 &= C_3 - C_5 \\
 F_1 &= C_5
 \end{aligned}$$

### 3.3.3 TRP DETECTORA DE CONTRASTES.

Una TRP de interés, donde el funcional único es un funcional detector de contrastes, se obtiene de la TRP de la figura 3.3.1, tomando, para cada vector fila, un grupo de elementos igual al del campo receptivo y dándole el valor -1. Esta transformación se ilustra en la figura 3.3.11, para una dimensión.

	F1	F2	F3	F4	F5	F6	F7	F8	
	1	1	1	1	1	1	1	1	C1
	1	1	1	1	-1	-1	-1	-1	C2
	1	1	-1	-1	0	0	0	0	C3
	0	0	0	0	1	1	-1	-1	C4
	1	-1	0	0	0	0	0	0	C5
	0	0	1	-1	0	0	0	0	C6
	0	0	0	0	1	-1	0	0	C7
	0	0	0	0	0	0	1	-1	C8

matriz para N=8

Figura 3.3.11

De donde se obtienen las expresiones de los coeficientes transformados en función de los valores de entrada:

$$C1 = F1+F2+F3+F4+F5+F6+F7+F8$$

$$C2 = F1+F2+F3+F4-(F5+F6+F7+F8)$$

$$C3 = F1+F2-(F3+F4)$$

$$C4 = F5+F6-(F7+F8)$$

$$C5 = F1-F2$$

$$C6 = F3-F4$$

$$C7 = F5-F6$$

$$C8 = F7-F8$$

Esta matriz admite una descomposicion en producto de otras matrices mas sencillas, con mayor numero de ceros, como las de la figura.

$$\begin{matrix}
 \begin{bmatrix}
 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1
 \end{bmatrix} &
 \begin{bmatrix}
 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
 \end{bmatrix} &
 \begin{bmatrix}
 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
 \end{bmatrix} \\
 H1 & H2 & H3
 \end{matrix}$$

$$H = H1 \cdot H2 \cdot H3$$

Figura 3.3.12

Lo que hace que esta transformada pueda implementarse mediante un algoritmo rapido, de facil formulacion.

Este cálculo, de coeficientes transformados se visualiza en forma de arbol en la siguiente figura:

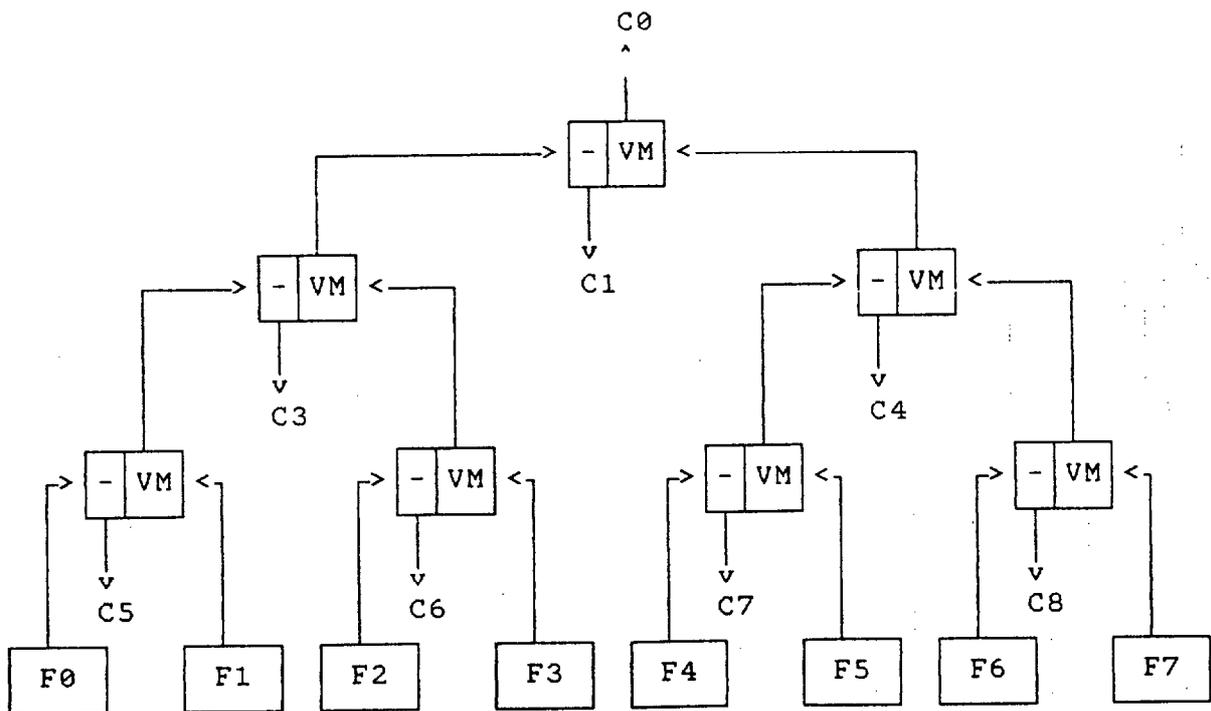


Figura 3.3.13

Donde cada nodo padre realiza dos operaciones con sus nodos hijos, la operación valor medio, y la operación diferencia.

En forma algorítmica el cálculo de las TRP'S se podría escribir de la siguiente forma

Mientras n° elementos del vector > 1 hacer

    -- calcular coeficientes como diferencia entre pares de elementos.

    generar un nuevo vector promediando parejas de elementos.

Fin mientras

tomar unico elemento del vector como coeficiente.

Una ilustración interesante, es representar graficamente los nucleos de esta transformación detectora de contrastes, veamos la forma de los primeros nucleos.

1
---

1
-1

1
-1
0

0
1
-1

1	-1
---	----

1	-1
-1	1

1	-1
-1	1
0	

0	
1	-1
-1	1

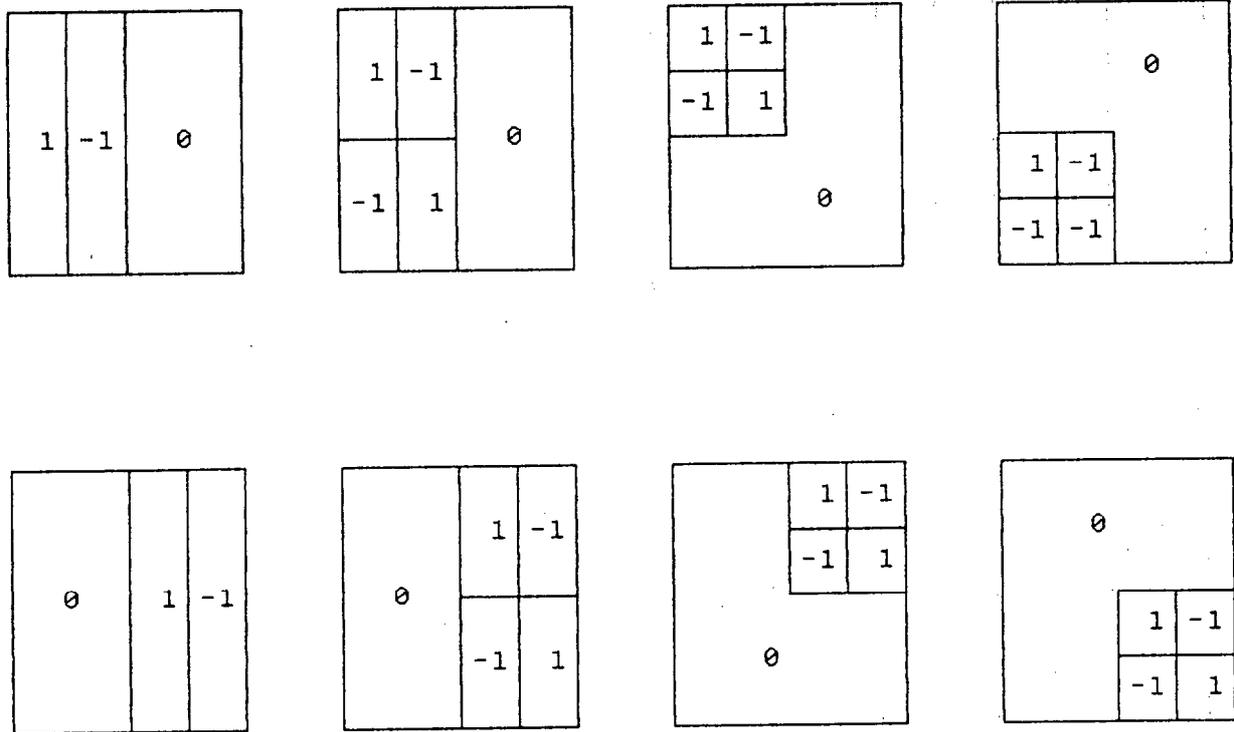


Figura 3.3.14

### 3.4 ILUSTRACIONES DE TRP Y SU REPRESENTACION.

De las distintas posibilidades de proyección de los resultados de una transformación TRP, que tienen a su vez distintos significados para su posterior tratamiento en la generación de frases descriptoras, elegimos tres modalidades.

#### 3.4.1 REPRESENTACION CRITERIO CR.

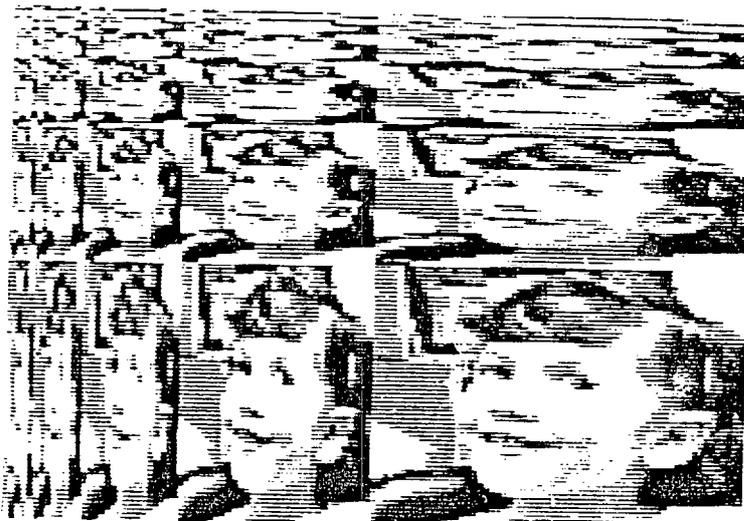
La representación según el criterio Cr de la Sección 3.1.7 se muestra en las ilustraciones siguientes:

Veamos primeramente una representación con una transformada de resolución progresiva, donde los dominios se han ordenado por distancias, y a igualdad de distancias, por cercanía al origen.

La figura representa en la parte (a) la imagen original, y la parte (b) la representacion de los coeficientes.



(a)



(b)

Figura 3.3.15

La siguiente ilustración representa una transformación de representación progresiva diferencial, donde se ha seguido el mismo criterio de ordenación Cr.

La imagen de la izquierda representa la imagen original, mientras que la de la izquierda representa los coeficientes, donde se puede apreciar la propiedad detectora de contrastes que tiene este tipo de nucleos.

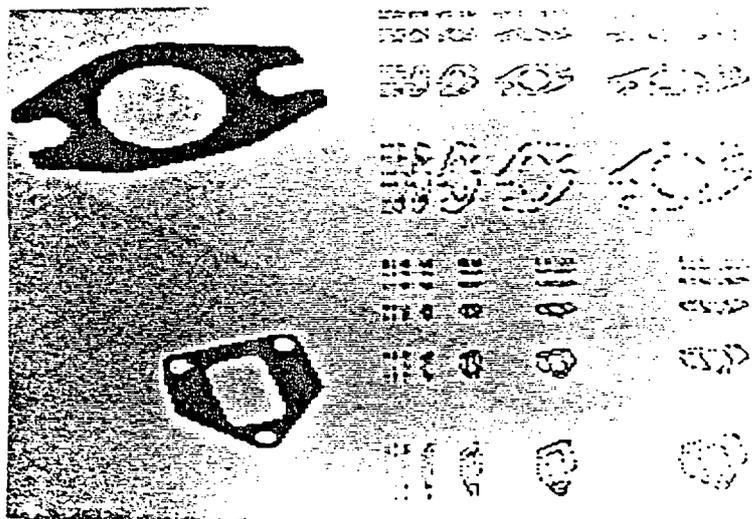
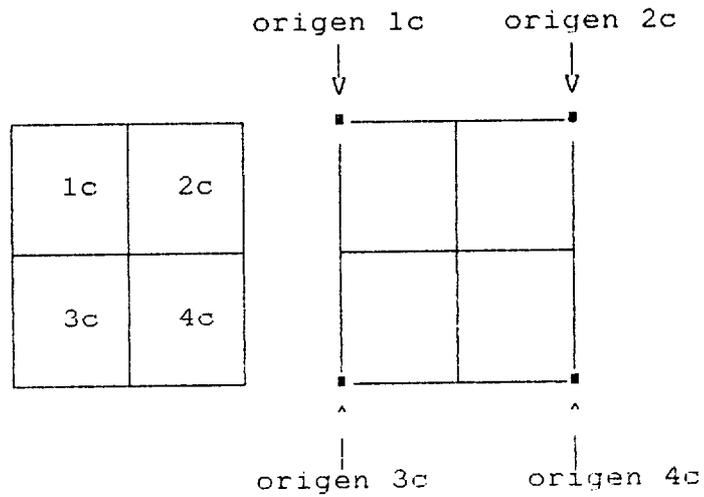


Figura 3.3.16

### 3.4.2 REPRESENTACION FRACCIONADA.

Esta representación fraccionada, consiste en dividir la imagen de entrada en cuatro cuadrantes, y aplicar a cada cuadrante una transformación TRP, ordenando los coeficientes con el criterio Cr, y cambiando la situación de los orígenes de cada fracción a los vértices de la imagen inicial, tal como se ilustra a continuación:



Los resultados corresponden a una representación centrada.

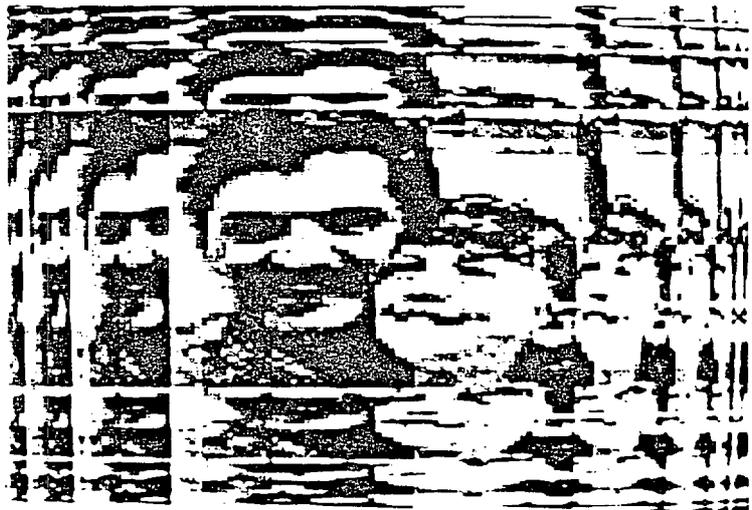


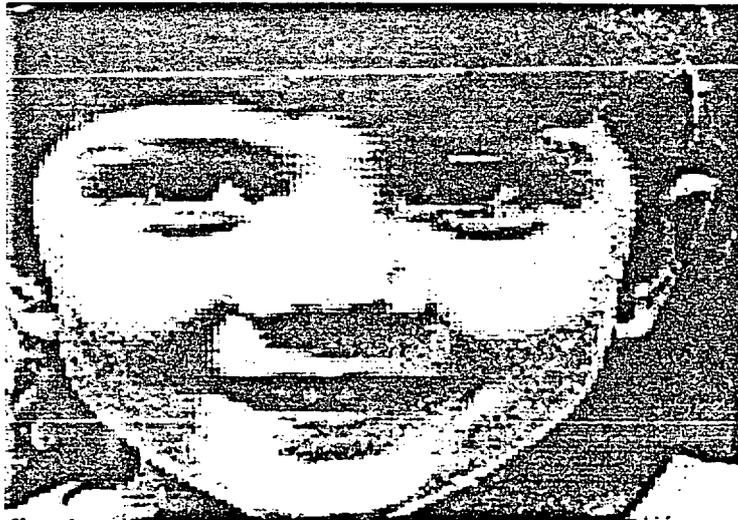
Figura 3.3.17

### 3.4.3 PROYECCION RADIAL.

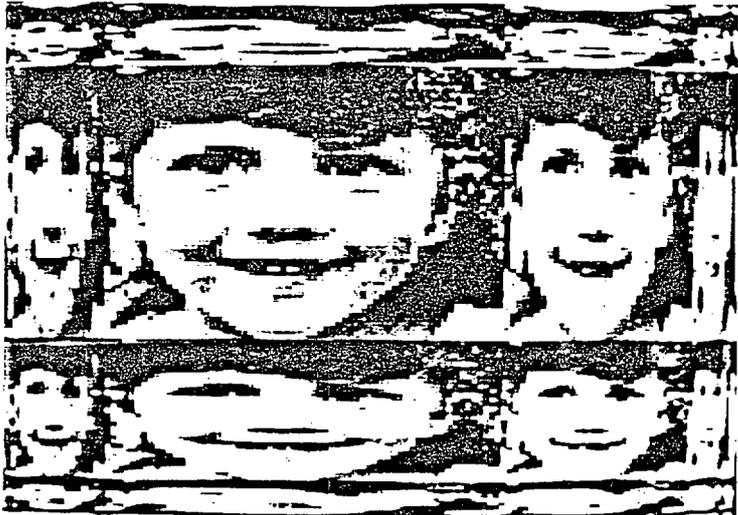
En este caso, se parte del subtrozo de más resolución que se sitúa en el centro, y alternativamente, se van proyectando los elementos, en orden creciente de dimensión del dominio, a uno y otro lado.

Este tipo de proyección proporciona una representación centrada.

Los resultados de la representación se muestran a continuación, donde la figura (a) es la original y la figura (b) la procesada:



(a)



(b)

Figura 3.3.18

**CAPITULO:4. RECONOCIMIENTO VISUAL DIRECTO EN BASE A TRANSFORMACIONES  
DE CAMPO RECEPTIVO VARIABLE.**

#### 4.1 SISTEMAS REDUCIDOS DE INFERENCIA EN VISION-DECISION.

En esta sección expondremos una estructura general de sistemas de inferencia que toman decisiones acerca del mundo visual directamente a partir de los datos visuales. Este trabajo ha sido realizado en colaboración con J.A. Muñoz y bajo la orientación del director de tesis.

##### 4.1.1 SISTEMAS REDUCIDOS DE INFERENCIA.

Un sistema reducido de inferencia (SRI) tiene la naturaleza de un Sistema Experto con los siguientes componentes:

a) Una base de datos referencial, que contiene un conjunto de frases descriptoras de referencia, a cada una de las cuales hay asignada una etiqueta. Esta base de datos está construida a priori, en base de conocimiento inyectado, y/o se construye a partir de un proceso de aprendizaje, donde se inyecta parcialmente el conocimiento que permite etiquetar las frases descriptoras.

b) Un dispositivo o máquina de inferencia, que a partir de ciertas reglas (normalmente, de naturaleza if..then..else) acepta una frase descriptora incógnita y, tras la interacción con la base de datos referencial, produce una decisión o diagnóstico.

La estructura relacional de estos dos componentes se muestra en la figura 4.1.

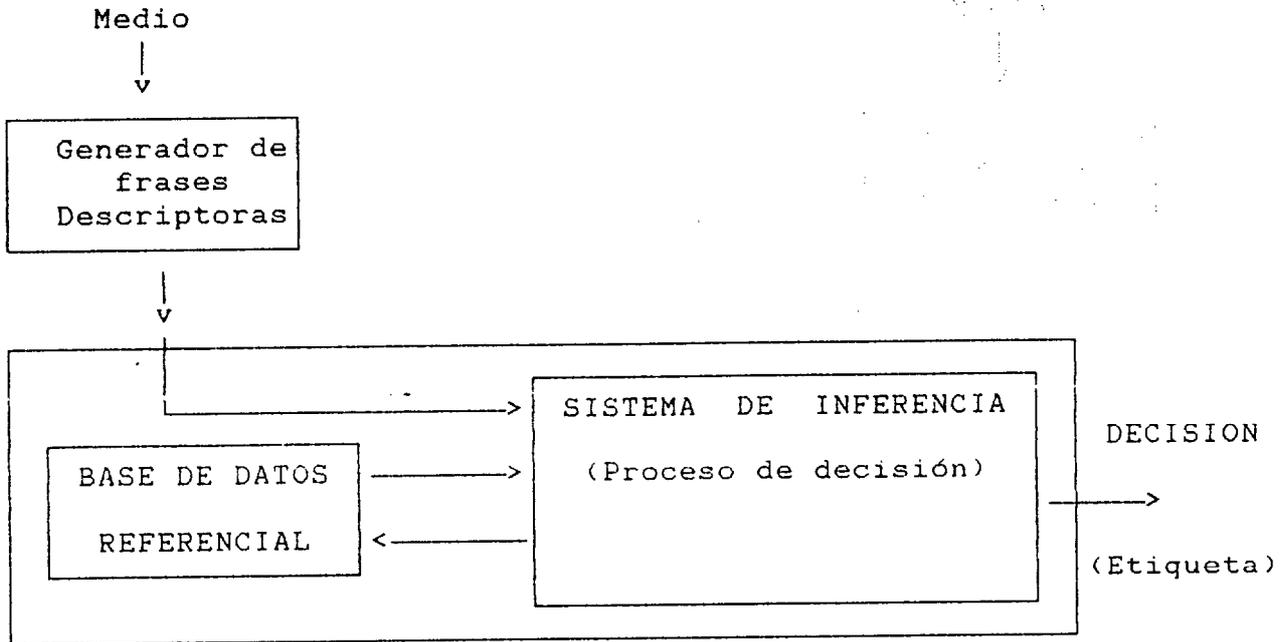


Figura 4.1

#### 4.1.2 INTEGRACION DE UN SISTEMA REDUCIDO DE INFERENCIAS EN UN SISTEMA VISUAL.

Consideraremos el caso de que el Sistema Visual no sea interactivo con el medio. Tal situación corresponde a la estructura básica de la figura 4.1, ampliada con las especificaciones sobre la naturaleza del generador de frases descriptoras y tres mecanismos de inyección de conocimiento (por aprendizaje o por conocimiento de procedencia puramente exterior). Hemos de notar que, en cualquier caso, la inyección de conocimiento requiere la acción del sistema exterior: en el aprendizaje, éste conocimiento se utiliza como asesor acerca de la corrección de la decisión para modificar los distintos bloques funcionales, mientras que en el caso del conocimiento exterior impreso éste se implanta directamente en dichos bloques. La situación típica y general es híbrida.

En el caso de que el Sistema Visual no sea interactivo con el medio, la estructura computacional correspondiente se muestra en la figura 4.2

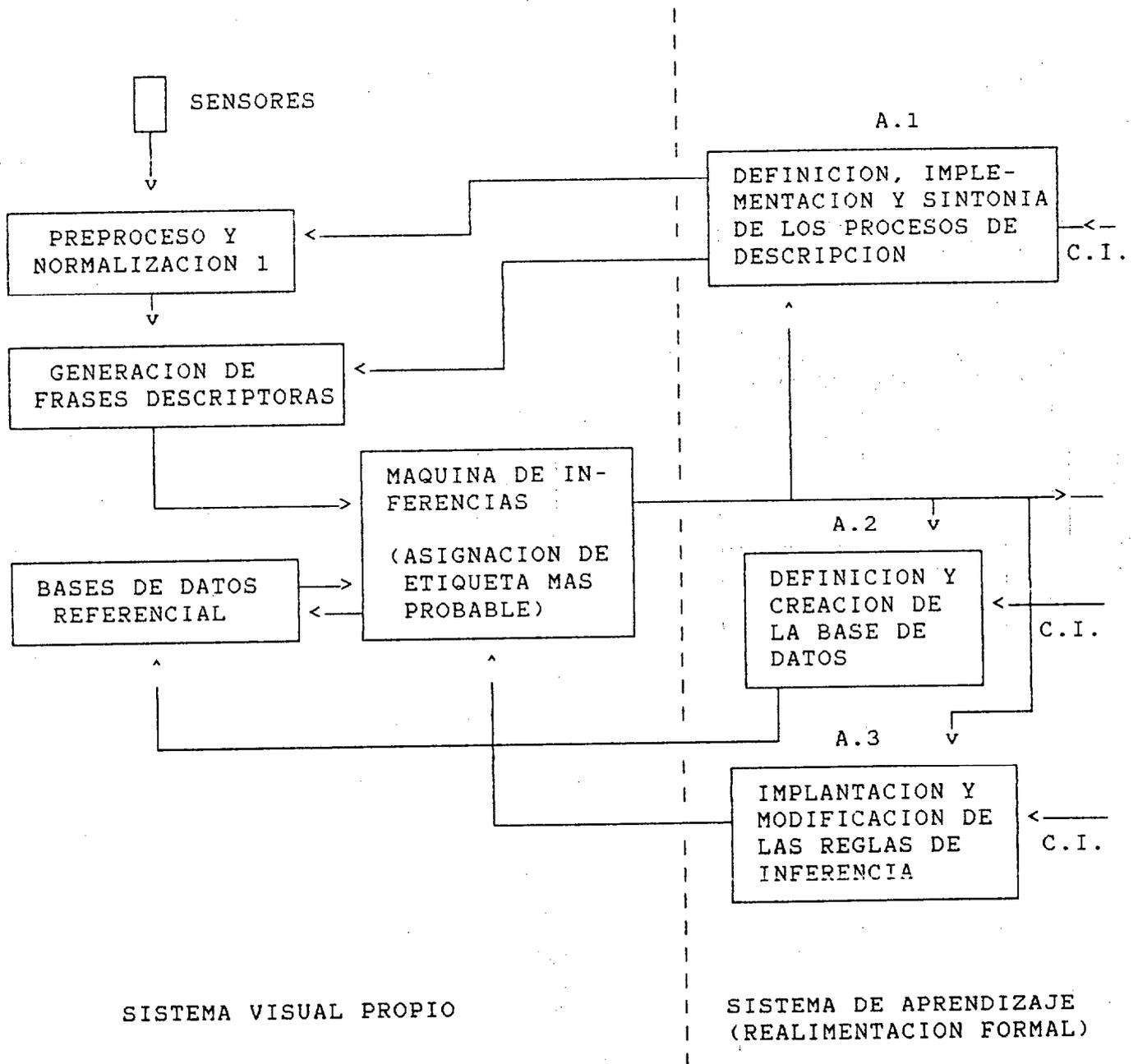


Figura 4.2. Estructura del Sistema Visual no Interactivo con el Medio

Los módulos operacionales del Sistema Visual propio son los indicados anteriormente y ahora desglosamos el proceso de generación de frases descriptoras en:

- a) Adquisición de la imagen.
- b) Preproceso y Normalización.
- c) Generación de las frases descriptoras al nivel elegido (analítico, algorítmico, simbólico), o a varios niveles.

De esta forma, es posible esquematizar la función de un sistema visual propio sin aprendizaje por experiencia, es decir, que funcione con conocimiento inyectado puro. Este esquema se muestra en la figura 4.3, y corresponde a la función realizada por la práctica totalidad de los sistemas actuales de reconocimiento visual (Sistema Visual Predeterminado).

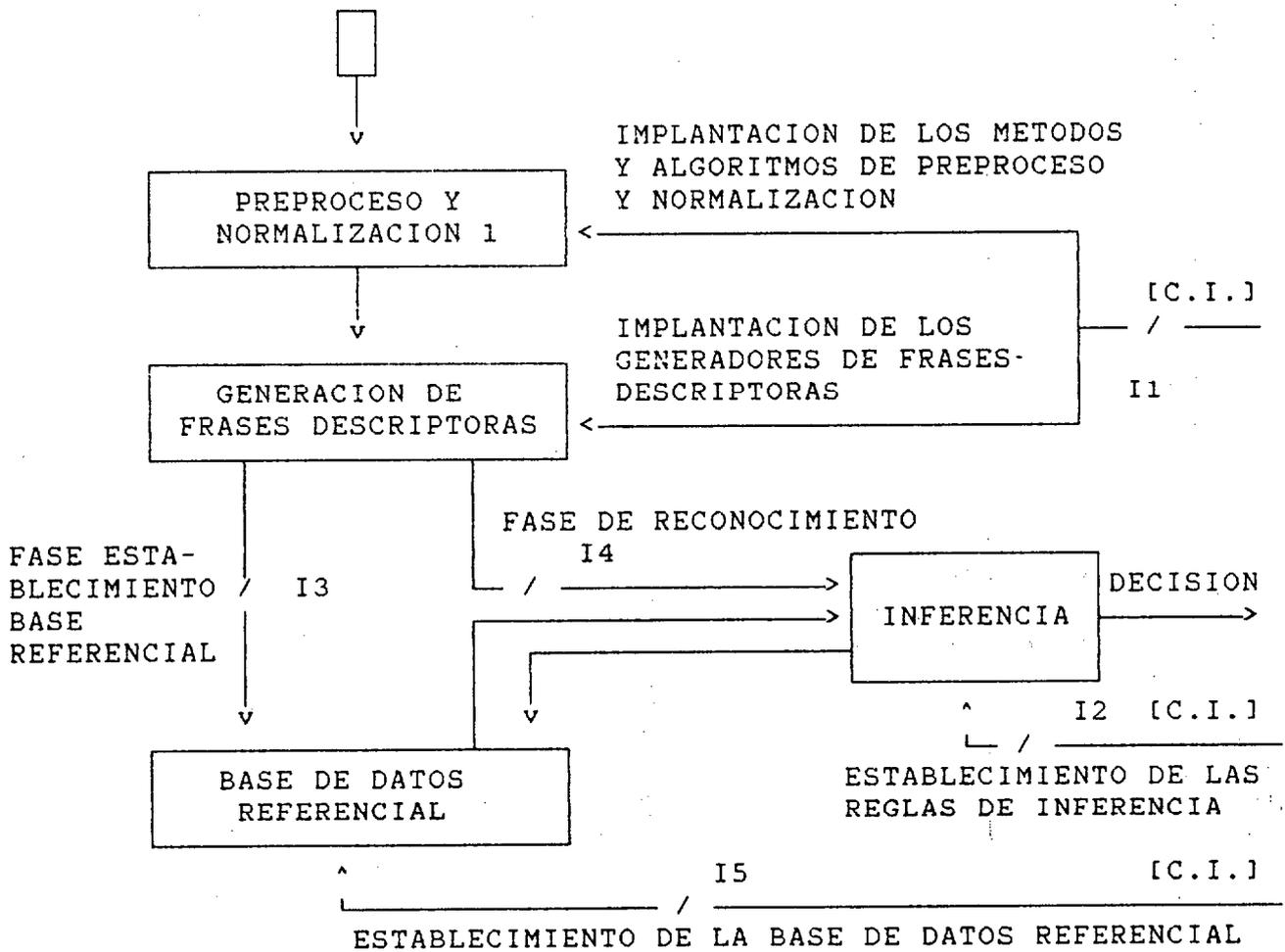


Figura 4.3. Estructura de un Sistema Visual Predeterminado

En estos sistemas se distingue la fase de establecimiento e implantación y la fase decisión. La primera fase corresponde a mantener cerrados los interruptores simbólicos I1, I2, I3 e I5 y abierto el interruptor simbólico I4. En la segunda fase de decisión (o de operación) el sistema queda desenganchado del medio, por la apertura de los interruptores simbólicos I1, I2, I3 e I5, quedando enganchado al mismo solamente a través del cierre del interruptor simbólico I4.

## 4.2 INVARIANZA ESPACIAL, REPRESENTACION DIRECTA Y TRANSFORMACIONES AFINES.

### 4.2.1 SOBRE LA INVARIANZA ESPACIAL.

Los conceptos relativos a la generación de bases referenciales de datos directos usables en el proceso de reconocimiento visual, es decir, en los sistemas reducidos de inferencia, tienen su origen en la generalización de la propiedad de invarianza espacial de una transformación.

#### DEFINICION 4.1

Una transformación es espacialmente invariante si es una TRC y los funcionales generadores de descriptores sobre cada campo receptivo son los mismos.

Al ser la transformación una TRC, el conjunto  $\{D\}$  es tal que las dimensiones  $d_i$  de  $D_i$  son las mismas, para todo  $i$ . El requisito que los  $D_i \in \{D\}$  sean de  $d_i = cte.$  obliga a que  $\{D\}$  no puede ser completo. Si el  $\text{cardinal}\{D\} = B$ , el número adicional de descriptores a generar por los funcionales generadores es  $N - B$ , siendo  $(N - B) / B$  el número de descriptores independientes a obtener por campo receptivo (en media).

Por ejemplo, si la dimensión de todos los  $D_i$  es  $d$ , según la Sección 3.2.2, tenemos que  $\text{cardinal}\{D\} = N - d + 1$ . Por consiguiente, en media, se han de calcular  $N / (N - d + 1)$  descriptores funcionales en cada campo receptivo, para una descripción completa.

Ordenemos los  $D_i \in \{D\}$  de acuerdo con el criterio  $C_r$  de la Sección 3.1.8 ( $i=1,2,\dots,N-d+1$ ).

Sean  $F_j$  (ordenados,  $j=1,2,\dots,N/(N-d+1)$ ) los funcionales correspondientes a cada campo receptivo.

Procedamos a una proyección columnar (ver figura 3.2.4).

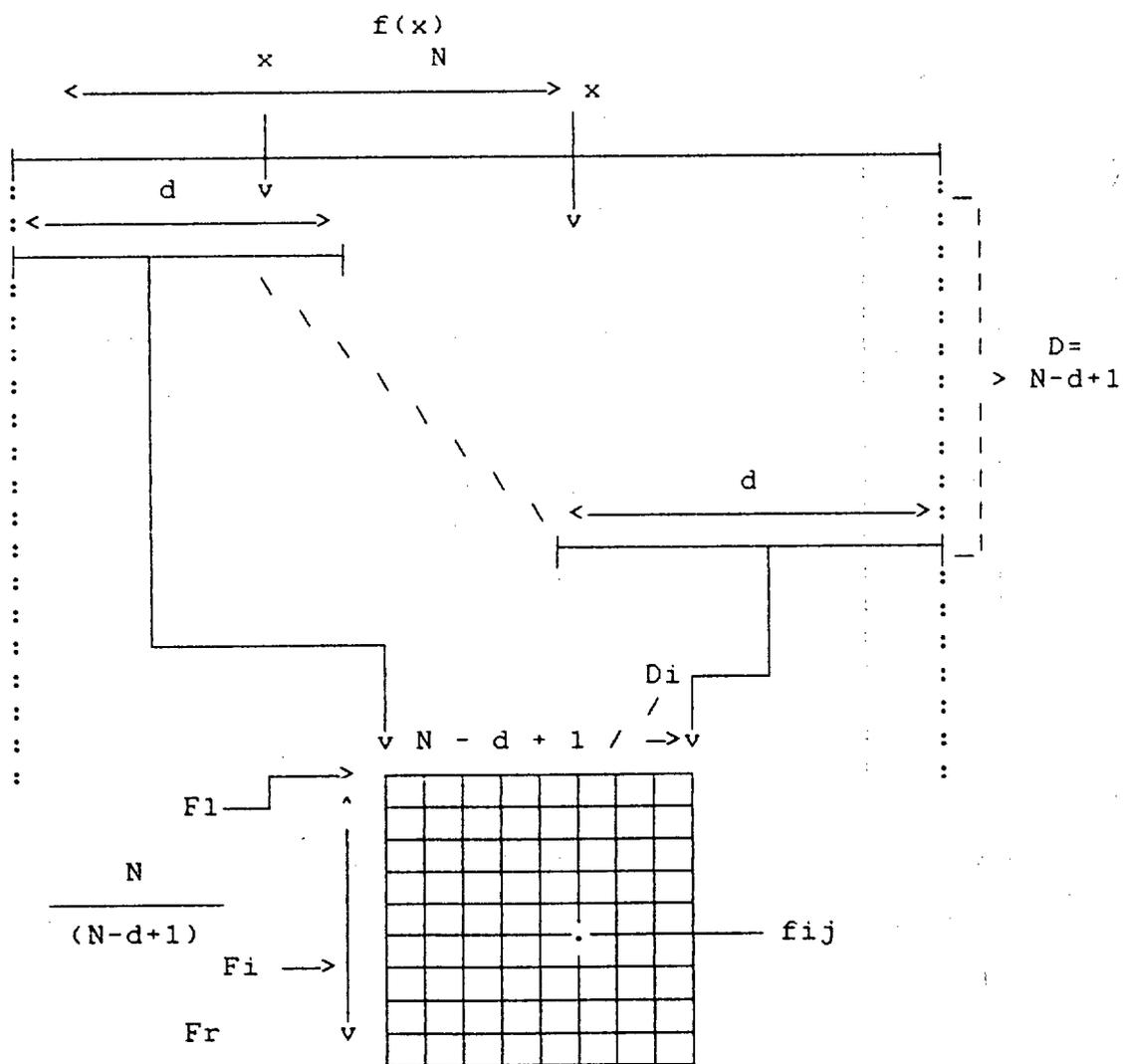


Figura 4.4

Tenemos dos cuestiones. Primero, el aplicar el funcional  $F_j$  al campo receptivo  $i$  equivale a "desplazar" la imagen  $f(x)$  hacia la izquierda, en cantidad  $i$  y aplicar el funcional (nótese que cada  $D_i$ , es de hecho, una subretina). Es decir que el valor  $f_{ij}$  correspondiente a aplicar el funcional  $j$  a la función  $f(x)$  en el campo receptivo  $D_i$  es:

$$f_{ij} = F_j[f(x-i)]$$

Nótese que no decimos nada acerca de la naturaleza del funcional y por consiguiente, lo anterior es consecuencia de la estructura del campo receptivo y de la condición de invarianza espacial.

Nótese que, para cada nivel del funcional  $F_j$  (es decir, para cada  $j$ ) existe una "compresión" de  $f(x)$ , por una especie de homotecia de razón  $N/N-d+1$ .

Desde el punto de vista funcional, ello equivale a actuar sobre la variable  $x$ , para obtener una nueva variable  $x'$  (con un rango igual al de  $i$ ), al tiempo que se actúa sobre  $F$  para obtener, con el funcional  $F_j$ , el valor de  $f_{jx'}$ .

Refiriéndonos a la figura 3.2.4, en la representación columnar tenemos una línea horizontal, por ejemplo, la  $j$ . Nos movemos en el subíndice  $i=x'$ . El proceso para construir  $f_{jx'}$  es el siguiente:

### PROCEDIMIENTO 1.

- a) Tomar un valor de  $x'$
- b) Determinar el dominio  $Dx' = Dx'(x)$
- c) Calcular el valor del funcional  $F_j(f(x))x'$  es decir, el valor de  $F_j$  en el dominio  $Dx'$ , sobre  $f(x) \rightarrow f_jx'$
- d) Adjudicar al "punto"  $jx'$  el valor  $\phi_j(x') = f_jx'$

El procedimiento anterior es una forma generalizada de obtener transformaciones con invarianza espacial.

Por lo ilustrativo y puesto que lo vamos a necesitar más adelante, veamos la forma que toma el PROCEDIMIENTO 1, para dos dimensiones:

### PROCEDIMIENTO 2. (Procedimiento 1 en dos dimensiones)

- a) Tomar una pareja  $(x', y')$ .
- b) Determinar el dominio  $Dx'y' = Dx'y'(x, y)$ .
- c) Calcular el valor del funcional  $F_j[f(x, y)]$  en el dominio  $Dx'y'$ , que es  $f_jx'y'$ .
- d) Adjudicar al punto  $jx'y'$  el valor  $\phi_j(x', y') = f_jx'y'$ .

#### 4.2.2 REPRESENTACIONES DIRECTAS

Consideramos primeramente la naturaleza del funcional  $F_j$ , y ello según los grados de libertad de  $f(x)$  y  $\phi_j(x')$ . (Es decir, los puntos c) y d) de los procedimientos 1 y 2).

Obviamente, tal como se ha planteado la cuestión en 4.2.1, los grados de libertad de  $f(x)$  son mayores que los de  $\varnothing_j(x')$ . Si nos liberamos de las condiciones de complitud y admitimos redundancia, podría ser al revés. Es decir, que, a priori, la aplicación de los procedimientos 1 y 2 de la Sección anterior no están restringidos por la condición de complitud.

Consideremos los tres casos (casos C)

- a) grados de libertad de  $f(x) >$  grados de libertad de  $\varnothing_j(x')$
- b) grados de libertad de  $f(x) <$  grados de libertad de  $\varnothing_j(x')$
- c) grados de libertad de  $f(x) =$  grados de libertad de  $\varnothing_j(x')$

Eligiendo los funcionales más sencillos.

#### CASOS C.

Caso a) En este caso hay una proyección de varios valores de  $f(x)$  en un único valor de  $\varnothing_j(x')$ .  $\varnothing_j(x')$  puede ser, el valor medio de  $f(x)$  en el dominio  $Dx'(x)$ , o el valor máximo, o cualquier otro de  $f(x)$  dentro de  $Dx'(x)$ .

Caso b) Existe una proyección de un único valor de  $f(x)$  en varios de  $\varnothing_j(x')$ . La solución funcional elegida es adjudicar ese valor de  $f(x)$  a todos los  $\varnothing_j(x')$  que lo "generan".

Caso c) Un único valor de  $f(x)$  se proyecta en un único de  $\varnothing_j(x')$  y a  $\varnothing_j(x')$  se le adjudica ese valor.

#### DEFINICION 4.2

Una Representación Directa es una representación obtenida por el procedimiento 1 o 2 y los CASOS C anteriores.

#### 4.2.3 DETERMINACION DE DOMINIOS Y TRANSFORMACIONES AFINES.

Nos referimos ahora a los puntos a) y b) de los procedimientos 1 y 2 de la Sección 4.2.1.

Consideramos el caso síntesis de los CASOS C de la Sección anterior, es decir cuando:

Grados de libertad de  $f(x)$   $\neq$  grados de libertad de  $\emptyset(x')$ .

Logicamente, en la situación real, tanto  $x$  como  $x'$  son discretos e identificables por enteros, con lo que los grados de libertad se corresponden con el número de "pixels" independientes de  $f(x)$  y  $\emptyset_j(x')$ .

Los puntos a), b) de los procedimientos 1 y 2 son:

- a) Tomar un  $x'$  o  $(x', y')$
- b) Determinar el Dominio  $Dx'(x)$ , [ $Dx'y' = Dx'y'(x, y)$ ]

La determinación del dominio  $Dx'(x)$  se puede realizar, en principio, siguiendo cualquier regla. Si ponemos la condición de biunicidad, la regla será una función. Si la regla es lineal, la determinación del dominio corresponde a una transformación afin de coordenadas. Ello nos lleva a la siguiente

### DEFINICION 4.3

Una transformación afín de coordenadas es una regla de determinación del dominio  $Dx'(x)$ , dada por una expresión lineal.

Para una dimensión, la regla es:

$$\begin{aligned} & x' = \text{Entero}(ax + b) \\ \text{o bien} & x' = \text{Entero}(a'x' + b') \end{aligned}$$

Para dos dimensiones, la regla es

$$\begin{aligned} & x' = \text{Entero}[Ax + B] \\ \text{o bien} & x' = \text{Entero}[A'x' + B'] \end{aligned}$$

donde  $A$  y  $A'$  son matrices de  $2 \times 2$  y  $B$  y  $B'$  son matrices de  $2 \times 1$ .

De interés particular son las representaciones directas con transformaciones afines, que serán usadas en casos prácticos, según se indica en las próximas Secciones.

### 4.3 ILUSTRACIONES DE TRANSFORMACIONES AFINES

Según lo anterior, una transformación afín es una transformación lineal cuya expresión general tiene la forma.

$$TA \equiv \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} c & d \\ e & f \end{bmatrix} \begin{bmatrix} x+a \\ y+b \end{bmatrix}$$

donde  $a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$  y  $f$  son constantes y  $\Delta$  el determinante.

$$\begin{vmatrix} c & d \\ e & f \end{vmatrix} \neq 0$$

El estudio de este tipo de transformaciones y sus invariantes algebraicos fueron iniciados por Boole, Cayley y Sylvester hace más de cien años. Aquí consideraremos las transformaciones afines con representación directa. Tal como apuntamos en la sección 4.1.2, en la representación directa que conserve el número de grados de libertad, el procedimiento es sencillo, mientras que cuando dichos números no coinciden, es preciso introducir reglas de transformación de los funcionales. Como veremos, las traslaciones y giros son del primer tipo, mientras que las homotecias son del segundo.

#### 4.3.1 TRASLACIONES

En el caso particular  $c=f=1$ ,  $e=d=0$  la expresión general de la transformada afín toma la forma:

a)            Transformación  
                  de  
                  dominios             $\begin{vmatrix} x' = x + a \\ y' = y + b \end{vmatrix}$

b) Transformación funcional  $\begin{vmatrix} f(x'y') = f(x y) \end{vmatrix}$

que expresa traslaciones.

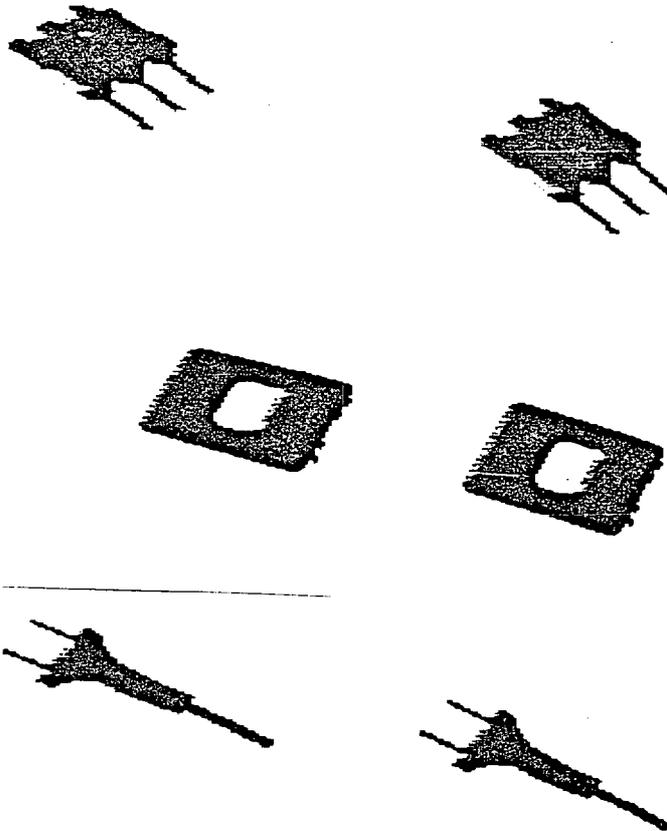
Cuando se trabaja con imágenes, que representan piezas aisladas, una transformación que tiene especial interés es la traslación de la pieza al origen de coordenadas, esto es el caso

que a y b coinciden con el centro de gravedad de la pieza  $X_g, Y_g$ .

Las coordenadas del centro de gravedad de una pieza, es de fácil obtención como lo apunta la teoría general de momentos, ya que precisamente una de las propiedades de los momentos de bajo orden es que nos dan las coordenadas del centro de gravedad, así que estas pueden ser calculadas mediante las expresiones.

$$X_G = M_{10} / M_{00} \quad Y_G = M_{01} / M_{00}$$

Aplicando este tipo de transformación, conseguimos que las imágenes sean invariantes frente a cualquier tipo de traslación de coordenadas.



#### 4.3.2 ROTACIONES

La expresión de una transformación que gire la imagen un ángulo  $\theta$  se consigue dando a las constantes los siguientes valores.

$$a=b=0$$

$$c=d=\text{Cos } \theta$$

$$e= \text{Sen } \theta$$

$$f= -\text{Sen } \theta$$

que nos dan la expresión:

a)

$$\text{Transformación de dominio} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \text{Cos } \theta & \text{Sen } \theta \\ -\text{Sen } \theta & \text{Cos } \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

b)

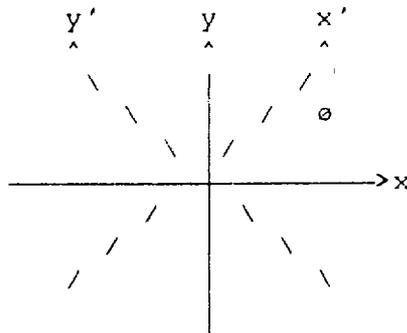
$$\text{Transformación funcional} \quad f(x' \ y') = f(x \ y)$$

donde

$$J = \begin{vmatrix} \text{Cos } \theta & \text{Sen } \theta \\ -\text{Sen } \theta & \text{Cos } \theta \end{vmatrix} = +1$$

Para conseguir invarianza algebraica respecto a rotaciones debemos calcular el parametro  $\theta$ . Una forma fácil de cálculo es la que aporta la teoría de momentos. Cuando una imagen se caracteriza sólo con momento de hasta segundo orden, esta es equivalente a una elipse homogénea, con un tamaño, orientación y excentricidad definidos. Los ejes  $(x' \ y')$  solidarios a esta elipse (figura) y por tanto, a la imagen, se denominan ejes principales. Es posible calcular el ángulo  $\theta$  que forman los ejes

(x' y') y (x,y) en función de los momentos determinados respecto a (x,y). Si a continuación se rota la imagen un ángulo  $\theta$  en el sentido adecuado, habremos llevado la imagen a una posición estandar en la que coinciden los ejes (x', y') y (x,y); y de este modo habremos normalizado frente a rotaciones.



El ángulo  $\theta$  se determina mediante los momentos de bajo orden normalizados respecto al centro de gravedad, utilizando la expresión:

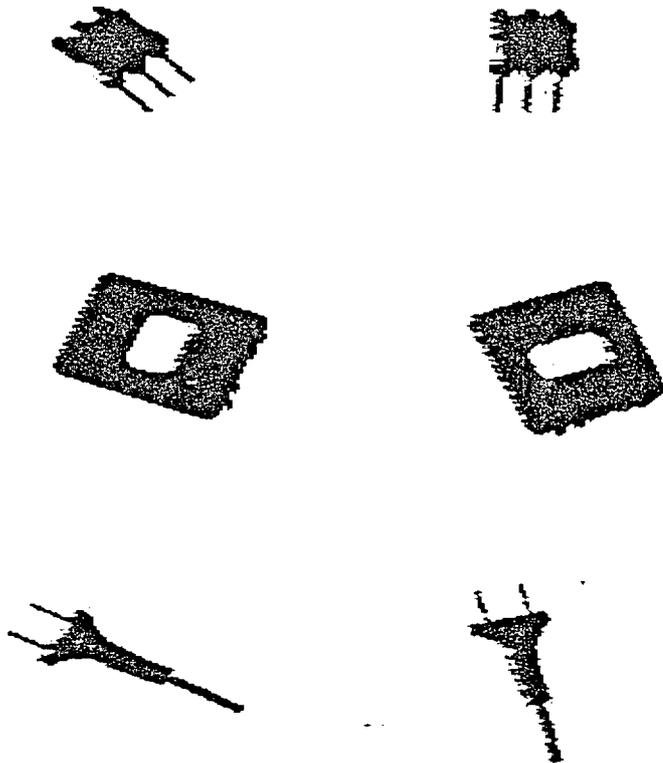
$$\theta = \frac{1}{2} \arctg \frac{2 M_{11}}{M_{20} - M_{02}}$$

Existe una ambigüedad en el cálculo del ángulo  $\theta$ , a partir de la expresión anterior, la cual puede evitarse de diversas formas, entre otras, la siguiente consiste en elegir  $\theta$  de tal forma que este sea el ángulo entre el eje x y el semieje mayor de la elipse a ( $a \geq b$ ) y eligiendo el valor principal del arcotangente, es decir:

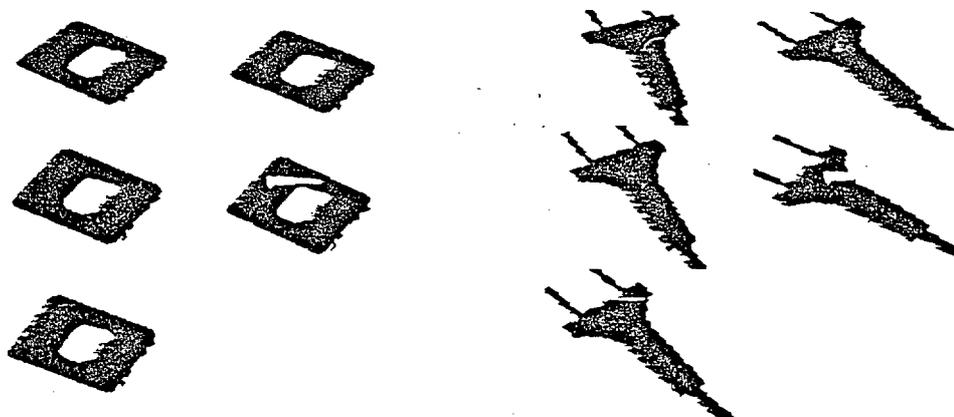
$$-\pi/2 \leq \arctg \theta \leq \pi/2$$

con estas consideraciones, obtenemos para  $\theta$  los valores de la tabla:

M20-M02	M11	$\theta$	
$\emptyset$	$\emptyset$	$0^\circ$	$a = 2M11 / (M20 - M02)$
$\emptyset$	+	$45^\circ$	
$\emptyset$	-	$-45^\circ$	
+	$\emptyset$	$0^\circ$	
-	$\emptyset$	$-90^\circ$	
+	+	$(\frac{1}{2} \arctg a)$	$(0^\circ < \theta < 45^\circ)$
+	-	$(\frac{1}{2} \arctg a)$	$(-45^\circ < \theta < 0^\circ)$
-	+	$(\frac{1}{2} \arctg a) + 90^\circ$	$(45^\circ < \theta < 90^\circ)$
-	-	$(\frac{1}{2} \arctg a) - 90^\circ$	$(-90^\circ < \theta < -45^\circ)$



Realizadas varias pruebas con distintas piezas, se observa que debido a la resolución empleada en estas imágenes, el ángulo calculado mediante momentos tiene un cierto error, que se acentúa en aquellas piezas que tienen formas alargadas.



#### 4.3.3 HOMOTECIAS

Haciendo que las constantes  $a$ ,  $b$ ,  $d$ ,  $e$ , sean cero la expresión general de la transformada afín se reduce a la ecuación:

a)

$$\begin{array}{l} \text{Transformación} \\ \text{de} \\ \text{dominio} \end{array} \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} c & 0 \\ 0 & f \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

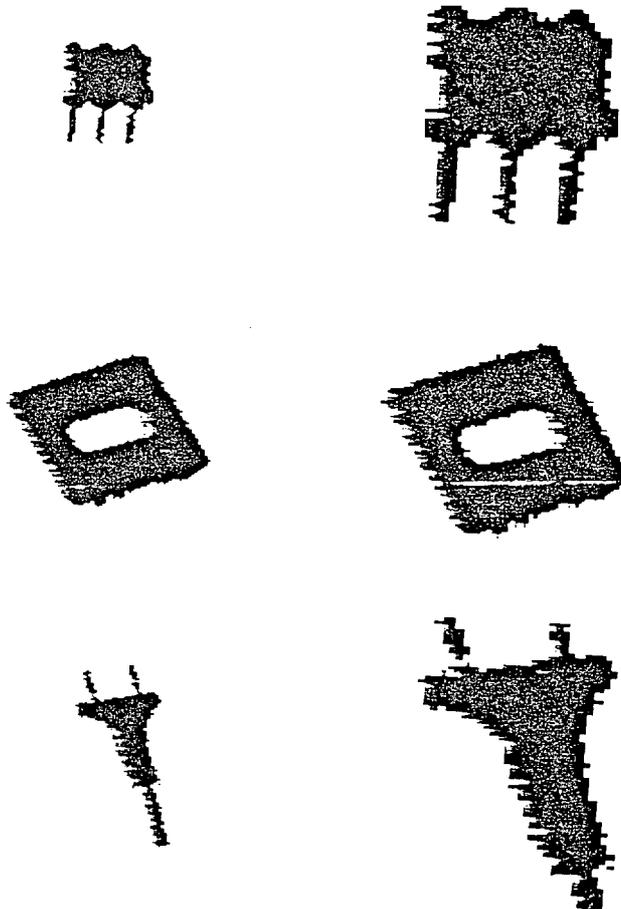
que representa la transformación de dominios para una homotecia. En el caso particular que las constantes  $c$  y  $f$  sean iguales, la relación de homotecia es la misma tanto para el eje  $x$ , como para el eje  $y$ .

b) Transformación funcional.

Los criterios de transformación funcional, se pueden conseguir imponiendo invarianza respecto a ampliaciones, calculando el factor de homotecia, (c) imponiendo la condición de que la imagen transformada tenga un área predefinida constante  $A_k$ , esto es

$$c = \text{SQRT}(A_k/A)$$

donde A es el área de la pieza calculada mediante el momento M00.



Mencionemos los casos.

1) Cuando  $c > A$ , existe una expansión apareciendo nuevos puntos, tendremos que definir un criterio para "rellenar" funcionalmente estos puntos que antes no existían. Un criterio utilizado, consiste en barrer la imagen transformada y asociar a cada punto el valor del pixel de la imagen original calculado por las expresiones  $x=x'/c$ ;  $y=y'/c$ , esto es, se está efectuando una aplicación inyectiva en la que a varios pixels de la imagen transformada le corresponde un único punto en la imagen original.

b) Caso en que  $C < A$ , existe una contracción, por lo que hay puntos que desaparecen, tendremos que definir un criterio para ver como estos puntos que desaparecen, los hacemos intervenir, en la imagen transformada. Un criterio consiste en hacer un enrejado en la imagen original, y calcular para un punto transformado la media de los valores de una cuadrícula del enrejado.

#### 4.4 APLICACIONES

##### 4.4.1 GENERACION DE UNA BASE REFERENCIAL DE DATOS.

Para realizar los experimentos que a continuación se explican, se generó una base de datos a partir de piezas industriales, reales (fig. 4.4.1) tales como enchufe, transformador, perno, rueda dentada, etc.

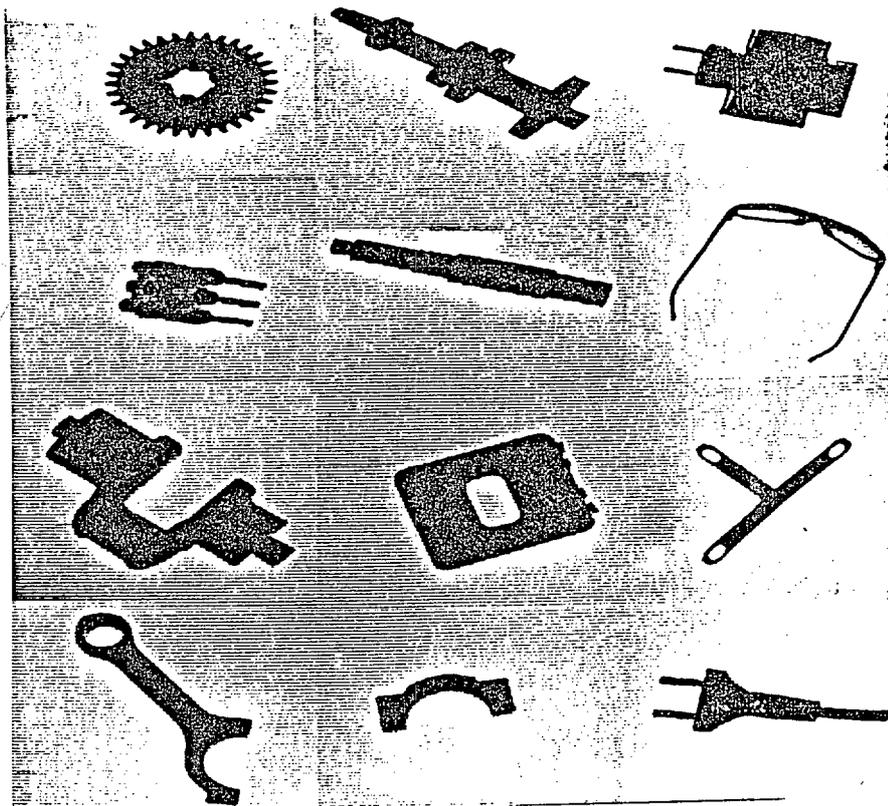


Figura 4.4.1

Mediante una cámara de estado sólido (CCD), y unas placas digitalizadoras, se toman imágenes digitalizadas de 512x512 pixels y cada pixel con 2\*\*12 niveles de gris.

De cada pieza se realizaron diez vistas, entre toma y toma se variaba el zoom de la cámara para conseguir variaciones en amplitud, se movía la pieza para conseguir traslaciones y se giraba la pieza, dando como resultado que las vistas sean todas distintas.

Sólo se mantiene constante la luz ambiental, ya que en este experimento no se tiene en cuenta los cambios en intensidad, puesto que, como veremos a continuación las imágenes se someten a un preproceso de umbralizado.

Tras comprobar que la frecuencia de muestreo de la cámara varía en horizontal y vertical con una relación aproximada de dos o tres, se corrige esta deformación de la imagen tomando 128x128 puntos, con la misma proporción anterior, de los originales 512x512.

Una vez corregidos estos errores de distinta frecuencia de muestreo, mediante el programa que adquiere (TOMA\_FOTO.FOR), se realiza a cada imagen un preproceso, que consiste en un umbralizado de la imagen; de los distintos programas de umbralizado desarrollados por el grupo, se eligió el método de la cuartica, que consiste en el cálculo del histograma de la imagen, (programa HISTO.FOR) frecuencia de todas las tonalidades de

grises, versus tonos de gris (fig. 4.4.2) y aproximar la envolvente del histograma a una cuartica, por presentar el histograma dos lomas, una mayor con forma gaussiana correspondiente al fondo, otra más pequeña correspondiente a la pieza, y un valle entre estas.

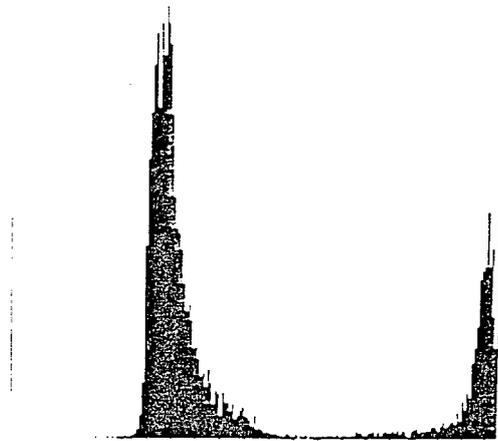
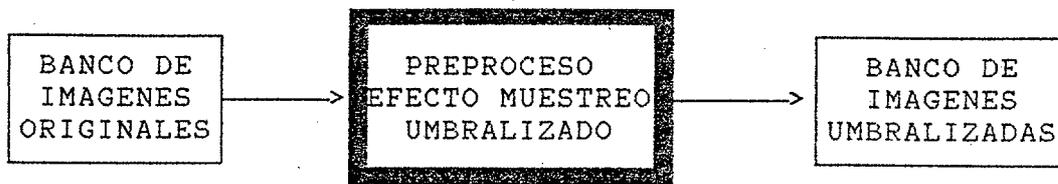


Figura 4.4.2

Calculando el mínimo de la cuartica (el valle) y cortando por este tono de gris, podemos separar el fondo de la imagen de la pieza en cuestión y umbralizar esta última a un valor. Este método tiene la desventaja del costo computacional, como el cálculo del histograma, la aproximación a una cuartica, el cálculo del mínimo, corte y umbralizado, pero tiene la ventaja de ser automático, esto es, no interviene el operador, y los resultados son aceptables como lo ilustra la fig. 4.4.3.



Al finalizar este preproceso obtuvimos un banco de imágenes de doce piezas con diez vistas de cada una, (cada imagen consiste de 128x128 enteros) y están umbralizadas en intensidad a dos niveles, cero el fondo, y uno la pieza. Estas imágenes mediante los correspondientes programas de almacenamiento se guardan en ficheros sobre discos. Para poder utilizarlas posteriormente en experimentos; estos ficheros se denominan MOTXXYY.UMB donde XX es un número de 1 a 12 que indica la pieza e YY es un número de 1 a 10, que delata la vista.

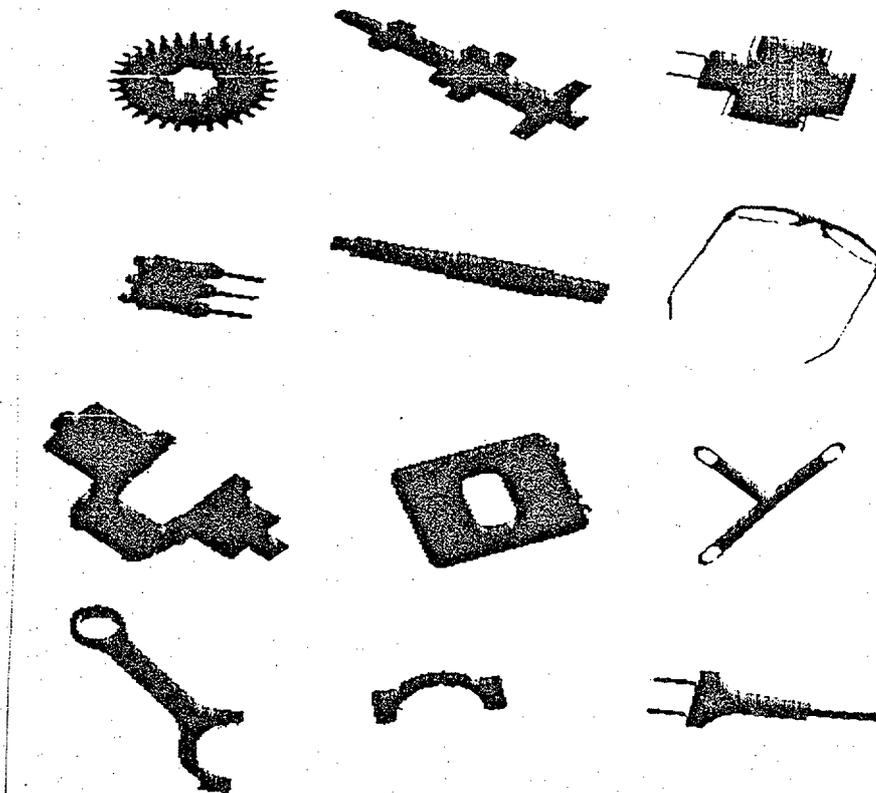
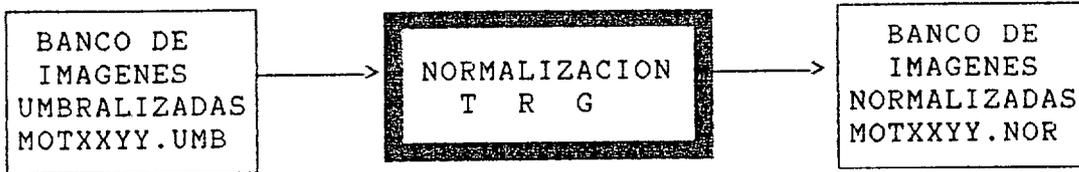


Figura 4.4.3

#### 4.4.2 DESCRIPCION DEL EXPERIMENTO.

##### - Normalización.



Las distintas imágenes (MOTXXYY.UMB) que forman el banco de imágenes umbralizado sufren un proceso de normalización. (Programa NORMALIZATRA.FOR), frente a traslaciones, calculando el centro de gravedad mediante los momentos de bajo orden  $XG=M10/M00$   $YG=M01/M00$ , frente a giros calculando el ángulo mediante los momentos normalizados de segundo orden  $\frac{1}{2} \cdot \text{ARCTG}(M11/(M20-M02))$ , y por último una normalización frente a homotecias obligando que el área final sea una constante AK, creandose un banco de imágenes normalizadas con el nombre MOTXXYY.NOR.

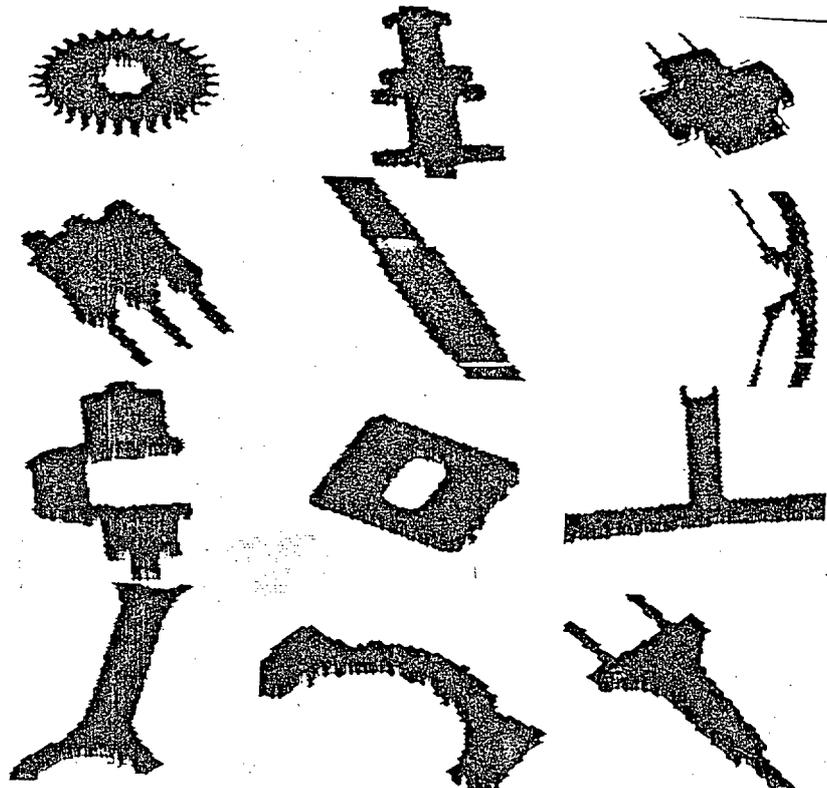
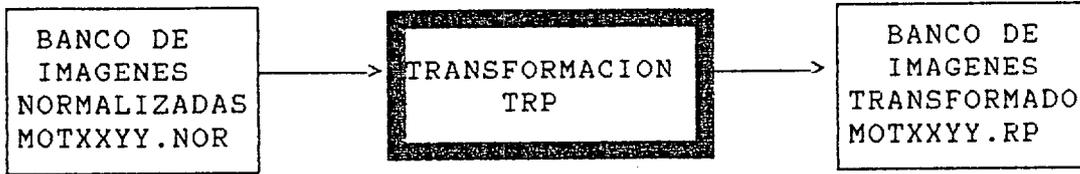


Figura 4.4.4

- Cálculo de la transformada de resolución progresiva (R.P.).



A este banco de imágenes normalizadas, se le calcula la transformada R.P. (Programa TRP.FOR) produciendo un nuevo banco de imágenes transformadas MOTXXYY.RP. Moviendonos sobre la diagonal principal de una transformada R.P. obtenemos imágenes con distinta resolución. La mayor es de 64x64 coeficientes y la llamamos resolución 5, la siguiente submatriz de la diagonal es de 32x32 coeficientes y la denominamos resolución 4, la siguiente 16x16 resolución 3, la siguiente 8x8 resolución 2 y por último 4x4 resolución 1, las siguientes submatrices 2x2 y 1x1 no se tienen en cuenta en el experimento por contener muy pocos coeficientes.

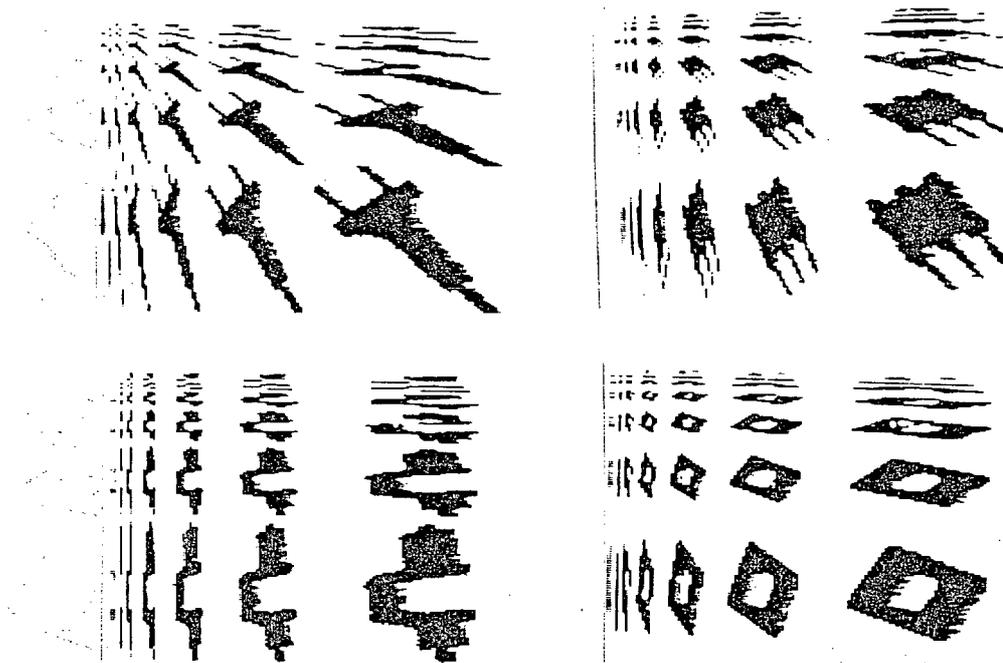
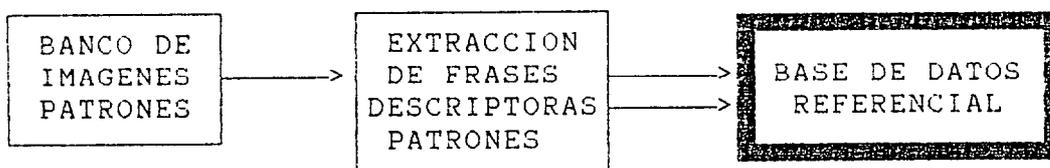


Figura 4.4.5

- Fase de establecimiento de la base referencial de datos.

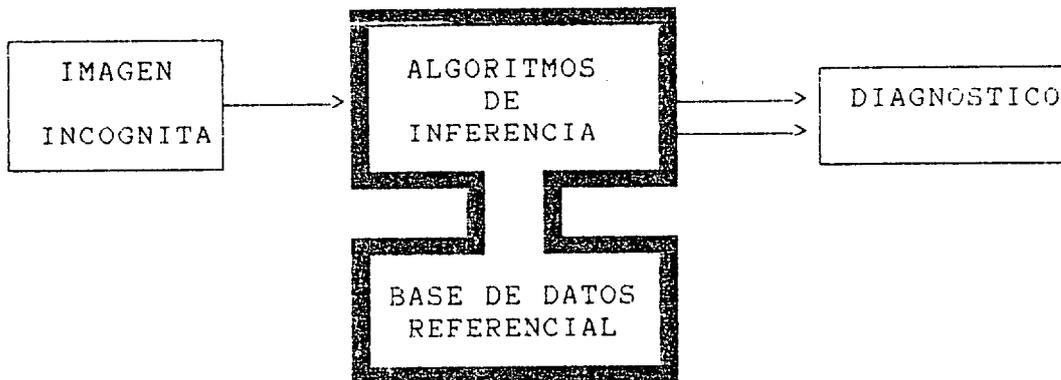
En nuestro caso de estudio, esta acción consiste en elegir una imagen normalizada cualquiera, de las diez vistas que componen el set de una pieza, y tomarla como "clase patron" o referencia y asignarle una etiqueta que denota la pieza que representa. Repitiendo esta elección para las doce piezas que componen nuestro banco de imagenes, tendremos un conjunto de frases descriptoras patrones con sus respectivas etiquetas, que seran los datos a almacenar en nuestra base de datos referencial. Cerrados I3, I5, y el resto abiertos.



En nuestra implementación, los datos almacenados en la base referencial tienen una estructura sencilla de array, se elige esta estructura por tener pocas clases patrones. Esto obliga a utilizar algoritmos de búsqueda secuenciales de fácil implementación pero de tiempo de respuesta no optimizado. Un tema que sale fuera de los fines de esta tesis, es el estudio de estructuras de datos mas complejas para las frases descriptoras, de forma que los algoritmos de búsqueda no sean secuenciales, si no, que admitan una estructura arborea, dando lugar a que las consultas en la base de datos tengan un tiempo de respuesta menor.

- Fase de reconocimiento.

En la fase de reconocimiento, en la cual, el sistema admite frases descriptoras incognitas y los algoritmos de inferencia tras una consulta en la base de datos referencial producen un "diagnostico" o resultado. Esto se muestra a continuación, y corresponde a mantener abiertos los interruptores simbólicos I1, I2, I3, I5, y cerrar I3 en la figura 4.3.



Este tipo de consultas en la base referencial es "borrosa" ya que la coincidencia de la frase incognita, con alguna de las frases patrones no es exacta, en el sentido matematico, la decision debera tomarse con la condicion de "frase mas parecida", o "mas probable".

Esta interrogación a la base de datos, con busqueda "mas parecida", se implementa utilizando un criterio de distancia entre frases descriptoras. En nuestro caso una frase descriptora

esta formada por una matriz, que representa una imagen de una cierta resolución, (matrices de la diagonal principal de una transformada TRP). Por lo que esta distancia puede ser el error cuadrático medio entre matrices, esto es:

$$\text{distanci} = \text{err.cuad} = \sqrt{\sum_x \sum_y (P(x,y) - I(x,y))^2}$$

donde P representa una frase patron e I es la frase incognita.

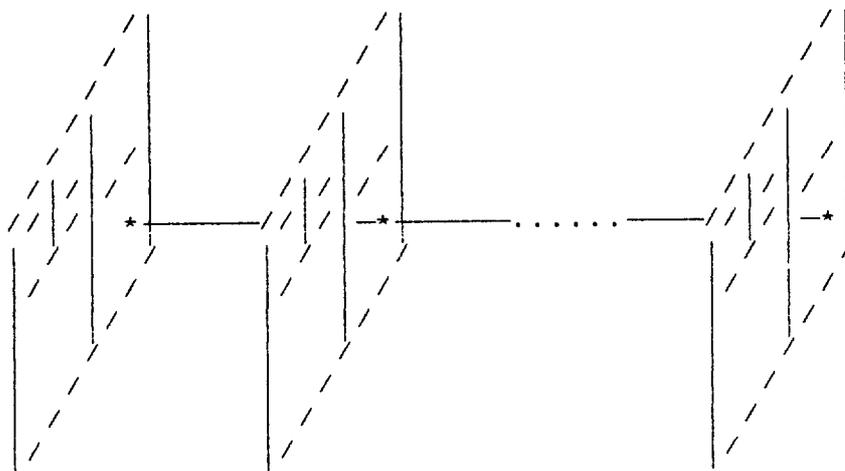
- Cálculo de las distancias todos contra todos.

En nuestro estudio hemos calculado todas las distancias de de todas las imagenes con todas, para las distintas resoluciones 6, 5, 4, 3, 2, 1 se leen, las siguientes submatrices de la diagonal y se calculan las distancias de una pieza con todas las demás.

Esto lo realiza el programa ERRMACH.FOR que alimentandose de todas las transformadas R.P. (MOTXXYY.RP) genera unas matrices triangulares, una para cada precisión, donde un elemento de la matriz  $D(i,j)$  expresa la distancia entre la imagen i y la imagen j. Estas matrices se almacenan en los ficheros ERRMACH.PRX, donde "X" indica la resolución.

resolución	1	matriz de	4*4	de la diagonal principal
resolución	2	matriz de	8*8	de la diagonal principal
resolución	3	matriz de	16*16	de la diagonal principal
resolución	4	matriz de	32*32	de la diagonal principal
resolución	5	matriz de	64*64	de la diagonal principal

resolución 6 matriz de 128\*128 imagen total normalizada



Visualización del calculo de distancias entre submatrices.

#### -Clasificación.

Una vez calculadas todas las distancias de todas las piezas respecto a todas las demás se pasa a la clasificación, para clasificar se han probado tres criterios. 1) La menor distancia, 2) la mayor distancia ponderada tomando K vecinos, y 3) la mayor distancia ponderada con frecuencia. El programa que realiza esta clasificación es (KVECENR.FOR) que alimentandose de las matrices de distancia. (ERMDCH.PRX) realiza la clasificación dejando los resultados en el fichero ERRMACH.LRX.

##### 1) Minima distancia.

Este criterio consiste:

- a) calcular todas las distancias de la frase descriptora con las frases patrones  $d_1, d_2, \dots, d_N$  (N numero de patrones).

b) calcular la minima distancia que apunta a la pieza patron resultado.

En algunos casos, puede interesar "suavizar" el criterio de minima distancia, para ello se pueden ensayar otros criterios como el de distancias ponderadas o el del producto distancia frecuencia.

## 2) Distancias ponderadas.

El esqueleto del algoritmo para implementar este criterio tiene la forma:

- a) - elegir un numero de vecinos K.
- b) - calcular todas las distancias.
- c) - legir las K menores distancias.
- d) - ponderar estas minimas distancias.
- e) - calcular el maximo de las distancias ponderadas.

a) El primer paso consiste en elegir un numero K de vecinos, la eleccion de este numero K se hara de forma experimental, en nuestro estudio se tomo para K el valor 10. Un caso limite consiste en tomar  $K=1$ , con lo cual el metodo de las distancias ponderadas se convierte en el criterio de minima distancia descrito anteriormente.

b) Una vez elegido el numero de vecinos K, se calculan todas las distancias, entre la frase incognita y las frases patrones,  $d_1, d_2, \dots, d_N$ .

c) De todas las distancias se calculan la K menores distancias,  $d_1 \leq d_2 \leq \dots \leq d_K$ , que apuntan a las piezas patrones mas parecidas a nuestra pieza incognita.

d) El siguiente paso consiste en elegir un criterio de ponderación de las distancias minimas, entre los posibles, se utiliza el siguiente:

$$P_i = \begin{cases} (d_K - d_i) / (d_K - d_1) & \text{para } i \neq 1 \\ 1 & \text{para } i = 1 \end{cases}$$

donde:

$d_K$  - es la distancia maxima dentro del conjunto de minimas distancias.

$d_1$  - es la distancia minima dentro del conjunto de minimas distancias.

Como se puede comprobar estos pesos van decreciendo desde el valor uno para  $i=1$ , hasta cero para  $i=K$ .

e) El ultimo paso consiste en multiplicar las distancias por sus pesos correspondientes, creandose un nuevo conjunto producto  $w_1, w_2, \dots, w_K$ . donde:

$$w_i = P_i \cdot d_i$$

Una vez realizado este paso, se calcula el máximo de todos los  $w_i$ , que nos apuntará a la pieza "mas parecida".

### 3) Producto distancia frecuencia.

Otro criterio utilizado, a la hora de producir el diagnostico de "pieza mas parecida", es el llamado producto distancia frecuencia. Este metodo se utiliza cuando se tienen varios patrones con la misma etiqueta. Esto es si M es el numero de frases patrones de la base referencial, y N el numero de etiquetas que representan, se cumple que  $M > N$ .

a) El primer paso consiste en calcular las distancias de la frase incognita a todas las frases de referencia, sean estas  $d_1, d_2, \dots, d_M$ , (donde aqui M denota el numero de frases referencia).

b) Se eligen K distancias minimas dentro del conjunto anterior de distancias,  $d_1, d_2, \dots, d_K$ .

c) Se calcula la frecuencia de cada etiqueta patron dentro de las K menores distancias, sean estas  $f_1, f_2, \dots, f_N$  (N es el numero de etiquetas referencia), donde  $f_i$ , es la frecuencia de aparicion de la etiqueta i, dentro de las K minimas distancias.

d) El siguiente paso consiste en efectuar el producto de las distancias por su correspondiente frecuencia, es decir, si la distancia  $d_i$  apunta a la etiqueta j, se efectua el producto:

$$P_i = d_i * f_j$$

e) Una vez calculados estos productos, se calcula el máximo de ellos, y este apunta al patrón "más parecido".

En nuestro estudio se experimentó con estos criterios, obteniéndose como resultado, que el producto distancia frecuencia, produce una tasa de error mayor en el reconocimiento, por lo que se desechó este método. Los resultados obtenidos con los criterios de mínima distancia, y distancias ponderadas, no difieren apreciablemente para nuestro banco de datos, por lo que se apunta como mejor criterio el de mínima distancia, por presentar menor complejidad de cálculo.

#### 4.4.3 COMENTARIO DE LOS RESULTADOS.

Se visualizó un plotter de los resultados comparando número de errores con las distintas resoluciones. figura 4.4.6 donde se observa que para este banco de imágenes se consigue un alto porcentaje de aciertos con la resolución 4, obtenida con submatrices de 32x32. A medida que disminuimos la resolución encontramos que algunas piezas incógnita, nos dan una distancia cero con varias piezas patrón, lo que impide hacer una buena clasificación para esta resolución.

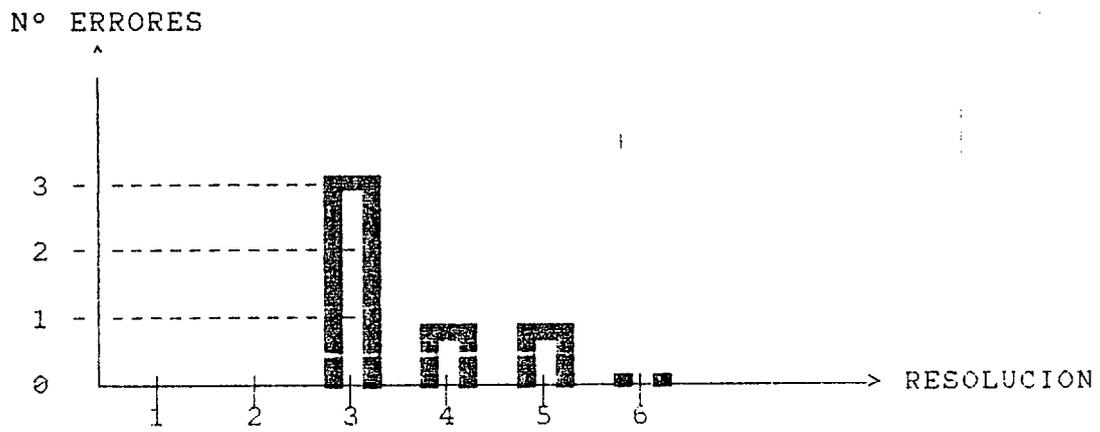


Figura 4.4.6

**CAPITULO:5. CONCLUSIONES Y PRINCIPALES APORTACIONES.**

Las principales aportaciones de esta Memoria, las conclusiones y la proyección futura de estos trabajos son los siguientes:

### 5.1 APORTACIONES.

La raíz de las aportaciones es de naturaleza metodológica. De hecho, al principio, el trabajo se planteó como el uso de las transformaciones típicas del campo de la transmisión de señales, y una cascada de ellas para generar espacios de representación, actualmente, bases referenciales de datos, para el reconocimiento visual. Pero, al considerar una familia de estas transformaciones, nos percatamos que las propiedades "deseables" que aparecían en algunas de ellas, no eran resultado de las transformaciones (funcionales) propias, sino consecuencia de una "partición" del campo receptivo. Ello nos llevó a replantear la cuestión desde sus principios, analizar el significado de las representaciones visuales completas e introducir, como consecuencia, toda una serie de nuevos conceptos, de una gran riqueza formal, relacionados con lo que denominaremos transformaciones de campo receptivo variable. La primera de la lista de transformaciones, que sigue, corresponde a la primera etapa. El resto, a la segunda.

1ª Modelo generalizado de proceso de imágenes, basado en la reformulación, para proceso visual, de un esquema generalizado de computación neuronal previo, que permite la sistematización de todos los tipos de procesos visuales existentes, y que aparecen

así dentro de su propio contexto, desde las transformaciones "puntuales" a las "globales".

2ª Significado de las representaciones visuales completas, con las siguientes contribuciones:

a) Relación entre la definición de complitud - de una representación visual y los objetivos del Sistema y niveles de especificación de la necesaria relación.

b) Condiciones de los lenguajes descriptivos a distinto nivel en la cadena visión-objetivos, donde las transformaciones son el medio de enganche entre niveles.

c) Especificación del sistema visión-objetivos como la de un autómata que tiene acceso a lenguajes próximos y produce una salida en el lenguaje de nivel superior.

3ª Conceptos relativos a las Transformaciones de Campo Receptivo Variable. Lo anterior, como consecuencia de la complitud respecto del número de campos donde se toma la información y las operaciones que se realizan sobre los campos, lleva a introducir el concepto de transformaciones de campo receptivo variable y la subfamilia de transformaciones completas, analíticamente, con las siguientes contribuciones:

a) Definición de las Transformaciones de Campo Receptivo

Variable.

b) Dualidad entre [función  $\leftrightarrow$  campo receptivo] en una transformación completa, desde el punto de vista analítico.

c) Definición de las propiedades de los campos receptivos y de las transformaciones de resolución progresiva (TRP).

d) Métodos de generación de TRP y TRP foveales, de un interés potencial considerable en la teoría de la proyección cortical.

e) Desde el punto de vista computacional, se desarrollan los algoritmos rápidos correspondientes.

#### 4a Conceptos relativos al Reconocimiento Visual.

Aparte de la estructuración de un sistema percepción - decisión como un sistema reducido de inferencias (realizado en colaboración de J.A. Muñoz), son de resaltar las siguientes contribuciones:

a) Definición de los conceptos de invarianza espacial de una transformación, introduciendo el concepto de proyección columnar y especificando la dualidad funcional-campo receptivo par las TRP, el concepto de proyección o representación directa.

b) Se llega de forma natural y elegante a las transformaciones afines, se formulan y se implementan.

## 5.2 PROYECCION FUTURA Y CONCLUSION GENERAL.

La aplicación futura de este trabajo se deduce del numero de definiciones y teoremas, que abren un considerable campo experimental, así como unas posibilidades adicionales importantes en la teoría de la proyección y representación en corteza de los datos perceptivos. Existen dos conclusiones prácticas de interés especial, relativas al estudio comparativo que se realizó sobre las transformaciones globales ( Capitulo 2° ) y a la experimentación del Capitulo 4°. Esto se presenta en forma de conjetura, puesto que fué lo que guió dichos experimentos y lo que los resultados sostienen, pero no prueban.

### Conjetura.

a) La naturaleza de las transformadas no es tan relevante en reconocimiento visual, y por consiguiente, en un amplio rango de aplicaciones, la generación de la base referencial de datos directa (después de las transformaciones afines) y la estructuración de una máquina de inferencias que es un sencillo "matching" supera, en resultados, cualquier otro procedimiento.

b) La optimización, en tiempo y recursos, se obtiene con representaciones completas progresivas que enfatizan el papel del campo receptivo, frente a la función, y de las cuales, las TRP son un caso particular.

La Conclusión General se relaciona, como en el caso de las aportaciones (5.1), con la metodología. En visión artificial debe existir un compromiso básico entre los funcionales a realizar y los campos receptivos, ( En otras palabras, lo que se hace sobre los datos y de donde se toman.), para mantener una "constancia" de los grados de libertad impuestos por los objetivos del Sistema. Este trabajo enfatiza el papel de los campos receptivos, frente a los funcionales clásicos, de manera que, en analítica, las particiones de los campos receptivos, "per se", pueden generar una descripción completa, con las propiedades que se han usado para justificar los funcionales, es decir, completitud, posibilidad de "recorte" en el número de grados de libertad y seguridad frente a los escotomas, o a la desaparición local de datos.

**CAPITULO:6.BIBLIOGRAFIA.**

## B I B L I O G R A F I A

- AHMED, N.; NATARAJAN, T.; RAO, K. R.  
DISCRETE COSINE TRANSFORM.  
KANSAS STATE UNIV., MAHATTAN, USA.  
IEEE TRANS. COMPUT (USA) V.C-23. N.1, P 90-3, JAN 74.
- ALEXANDRIDIS, N. A.; KLINGER, A.  
WALSH ORTHOGONAL FUNCTIONS IN GEOMETRICAL FEATURE EXTRACTION.  
UNIV. CALIFORNIA, LOS ANGELES, USA.  
IEEE TRANS. ELECTROMAGN COMPAT V.EMC-13, N 3, AUG 71, P 18-25.
- ANDREWS, H. C.  
SOME UNITARY TRANSFORMATIONS IN PATTERN RECOGNITION AND IMAGE  
PROCESSING.  
UNIV. SOUTHERN CALIFORNIA, LOS ANGELES, USA.  
INFORMATION PROCESSING 71 PROCEEDINGS OF THE CONGRESS 1971  
V 1, P 155-60.
- ANDREWS, H. C.  
MULTIDIMENSIONAL ROTATIONS IN FEATURE SELECTION.  
UNIV. SOUTHERN CALIFORNIA, LOS ANGELES, USA.  
IEEE CONFERENCE RECORD OF THE SYMPOSIUM ON FEATURE  
EXTRACTION AND SELECTION IN PATTERN RECOGNITION P 10-18.
- ANDREWS, H. C.  
ORTHOGONAL TRANSFORMS AND FEATURE SELECTION IN PATTERN  
RECOGNITION.  
UNIV. SOUTHERN CALIFORNIA, LOS ANGELES, USA.  
PROCEEDINGS 4TH HAWAII INTERNATIONAL CONFERENCE ON  
SYSTEM SCIENCES , P 675-7, 1971.
- ARES, L., VERDUZCO, R., RUBIO ROYO F, MIRA, J (1979)  
DETECCION DE CONTRASTES EN LUMINANCIA Y COLOR.  
AUTOMATICA; 7, 23-28:
- BACKER, E.  
AUTOMATIC RECOGNITION OF HANDWRITTEN NUMERALS.  
TECHNOL. UNIV. DELFT, NETHERLANDS.  
7TH INTERNATIONAL CONGRESS ON CYBERNETICS. NAMUR. BELGIUM.
- BAILA, I; OSOSKOV, G. A.  
UTILIZATION OF THE WALSH-HADAMARD TRANSFORMATION FOR  
CALCULATING FAST SLANT TRANSFORMATION.  
JOINT INSTITUT FOR NUCLEAR RESEARCH. DUBNA (USSR). LAB. OF  
COMPUTING TECHNIQUES AND AUTOMATION.
- BAILA, I.; OSOSKOV, G. A.

DATA COMPRESSION AT FILMLESS READOUT OF THE STREAMER CHAMBER INFORMATION. 2. ANALYSIS OF TWO BASIC TYPES OF DATA COMPRESSION ALGORITHMS.

JOINT INSTITUT FOR NUCLEAR RESEARCH. DUBNA (USSR). LAB. OF COMPUTING TECHNIQUES AND AUTOMATION.

BENNET, R. S.

A COMPARISON OF LINEAR BASES FOR PULSED SIGNALS.  
JOHNS HOPKINS UNIV BALTIMORE MD CARLYLE BARTON LAB.  
GRAI7618.

BLYUMIN, S. L.

SELECTING A POTENTIAL FUNCTION.  
AUTOM. AND REMOTE CONTROL (USA) VOL. 35, N 9, PT2 1522-4 SEPT.  
1.974

BOHNER, M.; KAZMIERCZAK, H.; SCHIFKOWITZ, U.

STANDARD METHODS OF IMAGE PROCESSING. PART 3: NUMERICAL ESTIMATE OF SERIES DEVELOPMENTS FOR PURPOSE CLASSIFICATION OD STANDARMETHODEN DER BILDVERARBEITUNG. TEIL 3: NUMERISCHE BEWERTUNG VON RIEHENENTWICKLUNGEN FUER DIE ZIELKLASSIFIZIERUNG. NATIONAL AERONAUTICS AND SPACE ADMINISTRATION, WASHINGTON, D. C.  
GRAI8316

BOLIVAR, O. (1983)

CLUSTERIZACION DE TEXTOS ADQUIRIDOS POR UNA CADENA CAMARA ORDENADOR.  
UNIVERSIDAD DE LA LAGUNA. TENERIFE.  
TESINA DE LICENCIATURA.

BROLLEY, J. E.

EDGE AND LINE ENHANCEMENT BY ADAPTIVE LATTICE FILTERING.  
SPIE ELECTRO-OPTICAL TECHNICAL SYMPOSIUM AND WORKSHOPM,  
HUNTSVILLE, AL, USA, 22 MAY 1.979.  
GRAI7922

BROOKS, M. J.

RATIONALIZING EDGE DETECTORS.  
DEPT OF COMPUTER SCI., UNIV. OF ESSEX, COLCHESTER, ENGLAND.  
COMPUT. GRAPHICS AND IMAGE PROCESS (USA) V8. N 2,  
P 277-85, OCT 78.

CALVERT, T. W.

NOOORTHOGONAL PROJECTIONS FOR FEATURE EXTRACTIO IN PATTERN RECOGNITION.  
CARNEGIE-MELLON UNIV. PITTSBURGH, PA., USA.  
PROCEEDINS OF THE IEEE 1969 SYSMPOSIUM ON ADAPTIVE PROCESSES DECISION AND CONTROL 5P 1969.

CANTON, M., RUBIO ROYO, F., MORENO-DIAZ, R. (1982).

LA TEORIA DE MOMENTOS GENERALIZADOS Y SU APLICACION EN PROBLEMAS BIOCIBERNETICOS.  
PROC. 1 REUNION NAC. BIOCIB. MADRID. REAL ACADEMIA DE CIENCIAS. P 151-178.

CANTON, M. (1982)  
LA TEORIA DE MOMENTOS EN EL ANALISIS DE FORMAS: FORMULACION  
TENSORIAL, INVARIANZAS Y APLICACIONES.  
TESIS DOCTORAL  
UNIVERSIDAD DE LA LAGUNA. TENERIFE.

CANTON, M. (1983).  
RECONSTRUCCION DE FORMAS A PARTIR DE UN CONJUNTO FINITO DE  
SUS MOMENTOS.  
2 SIMPOSIUM. ING. BIOMEDICA. MADRID OCT-83, E.T.S.I.  
TELECOMUNICACION. P 475-481.

CHEN, C. H.  
ON STATISTICAL AND STRUCTURAL FEATURE EXTRACTION.  
SOUTHEASTERN MASSACHUSETTS UNIV NORTH DARTMOUTH DEPT OF  
ELECTRICAL ENGINEERING.  
GRAI7621.

CHEN, CHI-AU; FAN, MAISIE.  
A COMPUTER SOFTWARE PACKAGE FOR STATISTICAL PATTERN RECOGNITION.  
SOUTHEASTERN MASSACHUSETTS UNIV NORTH DARTMOUTH DEPT OF  
ELECTRICAL ENGINEERING.  
GRAI7423.

CHEN, PI-FUAY, SEEMULLER, WILLIAM.  
DETECTION OF SIGNAL SIGNATURES OF CARTOGRAPHIC FEATURES.  
ARMY ENGINEER TOPOGRAPHIC LABS, FORT BELVOIR, VA.  
GRAI8014.

CHEN, C. H.  
SEISMIC PATTERN RECOGNITION.  
DEPT. OF ELECTRICAL ENGG., SOUTHEASTERN MASSACHUSETTS UNIV.,  
NORTH DARTMOUTH, MA, USA.  
GEOEXPLORATION (NETHERLANAS) V 16, N 1-2, P 133-46, APR-78

CONNELL, DAVID B.; KLINGBAIL, KERMIT N.; JACKSON, RICHARD A.  
MULTICS OLPARS OPERATING SYSTEM. VOLUME I.  
PATTERN ANALYSIS AND RECOGNITION CORP ROME NY.  
GRAI7707.

COOPER, J. A.  
ORTHOGONAL TRANSFORMATIONS OF DIGITAL DATA.  
SANDIA LABS., ALBURQUERQUE, N. MEX.  
GRAI7812.

COPPINGER, PAUL A.  
FUNCTIONAL ANALYSIS USING WALSH FUNCTIONS.  
HOUSTON UNIV TEX DEPT OF MECHANICAL ENGINEERING.  
GRAI7309.

CRETTEZ, J.-P.  
A PSEUDO-COSINE TRANSFORM FOR HEXAGONAL TESSELLATION WITH AN  
HEPTARCHICAL ORGANIZATION.  
INST. DE PROGRAMMATION, UNIV. PIERRE AND MARIE CURIE, PARIS,  
FRANCE.

PROCEEDINGS INT. ASSOC. PATTERN RECOGNITION. IEEE  
PROCEEDINGS OF THE 5TH INTERNATIONAL CONFERENCE ON  
PATTERN RECOGNITION P 192-4, 1980.

DAVIES, A. (1972).  
ON THE DEFINITION AND GENERATION OF WALSH FUNCTIONS.  
IEEE TRANSACTIONS ON COMPUTERS, FEB 1972, 187-189.

DA FONSECA, J., MORENO-DIAZ, R., MIRA, J. (1983).  
CONDITIONS FOR A THEORY OF DIAGNOSTICS.  
PROC. 1983 CONF. ON GENERAL SYSTEMS RESEARCH, DETROIT,  
MICH. INTERSYSTEMS PUB. P 675-678.  
COLEGIO UNIVERSITARIO. APT 550. LAS PALMAS. SPAIN.

DE FIGUEIREDO, RUI J. P.  
FEATURE EXTRACTION TECHNIQUES FOR CLASSIFICATION AND  
IDENTIFICATION OF SPECTRAL SIGNATURES.  
RICE UNIV HOUSTON TEX DEPT OF ELECTRICAL ENGINEERIGN.  
GRAI7621.

DE FIGUEIREDO, R. J. P.  
OPTIMAL LINEAR AND NONLINEAR FEATURE EXTRACTION BASED ON THE  
MINIMIZATION OF THE INCREASED RISK OF MISCLASSIFICATION.  
RICE UNIV., HOUSTON, TEX. INST. FOR COMPUTER SERVICES AND  
APPLICATIONS.  
GRAI7524.

DE LA NUEZ, I. (1982)  
SISTEMA DE INTERCONEXION DE DOS ORDENADORES DE DISTINTA  
ESTRUCTURA PARA LA TRANSMISION, ALMACENAMIENTO Y PROCESO  
DE MATRICES DE DATOS.  
LAS PALMAS ETSII. UNIV POLITECNICA.  
PROYECTO FIN DE CARRERA.

D'ENTREMONT, ROBERT P.  
PERFORMANCE OF THE DISCRETE FOURIER TRANSFORM SATELLITE IMAGERY  
CLASSIFICATION TECHNIQUE.  
AIR FORCE GEOPHYSICS LAB., HANSCOM AFB, MA.  
GRAI8113.

DIRILTEN, H.  
APPLICATIONS OF GROUP THEORY TO PATTERN RECOGNITION.  
TEXAS TECH UNIV., LUBBOCK, USA.  
UNIV. MICROFILMS AND ARBOR. MICH, USA.

DOI, Y.; KAWABE, S.  
RECOGNITION OF 3D OBJECTS BY MEANS OF LASER BEAM SECTION.  
UNIV. TOKYO, BUNKYO-KU, JAPAN.  
TRANS.SOC. INSTRUM. AND CONTROL ENG. (JAPAN) V 9,  
P 16-21, 1973.

DROMEY, R. G.  
A FAST WALSH TRANSFORM SUBROUTINE.  
LA TROBE UNIV. BUNDOORA, AUSTRALIA.  
DECUSCOPE (USA) V. 12, N 3, P 5-6, 1973.

DUNCAN, J. S.; FREI, W.  
VERY LARGE SCALE INTEGRATION (VLSI) APPROACH TO FEATURE  
EXTRACTION.  
INST. OF PHYS. AND IMAGING SCI., UNIV. OF SOUTHERN CALIFORNIA,  
MARINA DEL REY, CA., USA.  
PROC. SPIEINT. SOC. OPT. ENG. (USA) V 359, P 378-85

ELLIOTT. DOUGLAS F.  
CLASS OF GENERALIZED CONTINUOUS ORTHOGONAL TRANSFORMS.  
ROCWELL INT CO, ANAHEIM, CALIF.  
IEEE TRANS. ACOUST SPEECH SIGNAL PROCESS .  
V ASSP-22, N4, AUG-74, P 245-254.

ESKENAZI, R.; WILF, J. M.  
LOW-LEVEL PROCESSING FOR REAL-TIME IMAGE ANALYSIS.  
JET PROPULSION LAB., PASADENA, CA.  
GRAI7926.

FALCON, A., MENDEZ, J. (1983).  
PROCESO Y SEGMENTACION CROMATICA EN UNA CADENA CAMARA  
COLOR-COMPUTADOR.  
ETSII. UNIV. POLITECNICA LAS PALMAS. SPAIN.  
II SIMP. ING. BIOMEDICA. MADRID, OCT-1983  
E.T.S.I.T. P 505-509.

FALCON, A (1983)  
PROCESO Y EXTRACCION DE PROPIEDADES CROMATICAS EN FORMAS  
VISUALES.  
E.T.S.I.I. LAS PALMAS. UNIV. POLITECNICA. SPAIN.  
TESIS DOCTORAL.

FINO, B., ALGAZI, V. (1973).  
SLANT HAAR TRANSFORM.  
PROCEEDINGS LETTERS IEEE MAY-1974, 653-654.  
ELECTRONICS RESEARCH LABORATORY, COLLEGE OF ENGINEERING  
UNIVERSITY OF CALIFORNIA, BERKELEY, CALIF. 94720

FINO, W.  
ETUDE EXPERIMENTALE DU CODAGE D'IMAGES PAR LES TRANSFORMATIONS  
DE HAAR ET DE HADAMARD COMPLEXE.  
ANNALES DES TELECOMMUNICATIONS. P185-208.  
UNIVERSITY OF STANFORD, CALIFORNIA, USA.

FORTES, J. (1982)  
REALIZACION Y PUESTA EN FUNCIONAMIENTO DE UN SISTEMA  
CAMARA ORDENADOR, CON DISEÑO HARDWARE Y SOFTWARE DE  
ADQUISICION, PREPROCESO Y PRESENTACION DE IMAGENES  
DIGITALES.  
E.T.S.I.I. LAS PALMAS. UNIV. POLITECNICA. SPAIN.  
TESIS DOCTORAL.

FREI, W.; CHUNG-CHING CHEN.  
FAST BOUNDARY DETECTION: A GENERALIZATION AND A NEW ALGORITHM.  
DEPT. OF RADIOLOGY, UNIV. OF SOUTHERN CALIFORNIA, MARINA DEL R

REY, CA, USA.

IEEE TRANS. COMPUT (USA) V C26, N 10, P 988-98, OCT-77.

GERSON, DONALD.

COMPUTER ESTIMATION OF THE PRESENCE OF SEA ICE IN SATELLITE PICTURES.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH, ARLINGTON, VA.  
GRAI7519.

GONZALEZ, M., MORENO-DIAZ, R. (1983).

MODELOS DE PROCESOS RENTINALES DEL COLOR AL NIVEL ANALITICO.

E.T.S.I.I. LAS PALMAS. UNV. POLITECNICA. SPAIN.

II SIMP. ING. BIOMEDICA. MADRID. OCT-83

E.T.S.I.T. P 239-243.

GONZALEZ, R., WINTZ, P. (1977).

DIGITAL IMAGE PROCESSING  
ADDISON WESLEY.

GONZALEZ, R. C.; HARRIS, J. M.

FEATURE EXTRACTION AND RECOGNITION OF TWO-DIMENSIONAL DATA BY THE METHOD OF MOMENTS.

TENNESSEE UNIV KNOXVILLE DEPT OF ELECTRICAL ENGINEERING.  
GRAI7712.

GRANOVSKAYA, R. M.; VOLKOV, YA. I.

CURVE DETECTION NETWORKS.

PROBL. BIONIKI (USSR) NO. 7 53-64 1.971.

GREBLICKI, W.; RUTKOWSKI, L.

DENSITY-FREE BAYES RISK CONSISTENCY OF NONPARAMETRIC PATTERN RECOGNITION PROCEDURES.

INST OF ENGG. CYBERNETICS, TECHN. UNIV. OF WROCLAW, WROCLAW, POLAND.

PROC. IEEE (USA) V 69, N 4, P 482-3, APR 81.

GREBLICKI, W, RUTKOWSKA, D.

AN ORTHOGONAL SERIES ESTIMATE OF TIME-VARYING REGRESSION.

ANNALS OF THE INSTITUTE OF STATISTICAL MATHEMATICS

V. 35, N 2,A, 1983P 215-228.

GUREVICH, I. B.

SELECTION OF ENSEMBLE OF FEATURES RECOGNITION SYSTEMS BY THE VOTING PRINCIPLE.

SOV AUTOM. CONTROL (USSR) VOL. 7, NO. 5 43-42 SEPT.-OCT. 1.974.

HANUISE, G.

IMAGE PROCESSING BY COMPUTER.

OFFICE NATIONAL D'ETUDES ET DE RECHERCHES AEROSPATIALES, PARIS (FRANCE).

GRAI7701. FR ISSN 0078-3781.

HARALICK, ROBERT M.

TEXTURE-TONE STUDY WITH APPLICATION TO DIGITIZED IMAGERY.

ARMY ENGINEER TOPOGRAPHIC LABS., FORT BELVOIR, VA.  
GRAI7512

HARALICK, ROBERT M.  
USE OF FEATURE EXTRACTION TECHNIQUES FOR THE TEXTURE AND CONTEXT  
INFORMATION IN ERTS IMAGERY: SPECTRAL AND TEXTURAL PROCESSING OF  
ERTS IMAGERY.  
NASA EARTH RESOURCES SURVEY PROGRAM, WASHINGTON, D. C.  
GRAI7505.

HARALICK, R. M.  
THE DIGITAL EDGE.  
DEPT OF ELECTRICAL ENNG. AND COMPUTER SCI., VIRGINIA POLYTECH.  
INST. AND STATE UNIV., BLACKSBURG, VA, USA.  
IEEE COMPUTER SOCIETY CONFERENCE ON PATTERN RECOGNITION  
AND IMAGE PROCESSING P 285-91, 1981.

HUECKEL, M. (1973).  
A LOCAL VISUAL OPERATOR WHICH RECOGNIZES EDGES AND LINES.  
JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY,  
V 29, N 4, OCT-73, 634-647.  
STANFORD UNIVERSITY, STANFORD, CALIFORNIA.

HUECKEL, M (1971).  
AN OPERATOR WHICH LOCATES EDGES IN DIGITIZED PICTURES.  
JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY,  
V 18, N 1, JAN-71, P 113-125.  
STANFORD UNIVERSITY, STANFORD, CALIFORNIA.

IJJIMA, TAIZO  
ANALYTICAL EXPRESION OF AN ISOLATED FIGURE AND  
ITS APPLICATION.  
TOKYO INST OF TECHONOL, JPN.  
SYST COMPUT CONTROLS V12, N 3, JUN 81, P 1-8.

KAILATH, THOMAS; COVER, THOMAS.  
SYSTEM MODELING AND STATISTICAL DATA PROCESSING.  
STANFORD UNIV CALIF DEPT OF ELECTRICAL ENGINEERING.94305.  
GRAI7525.

KAILATH, T.  
SIGNAL PROCESSING IN THE VLSI ERA.  
MODERN SIGNAL PROCESSING AND VLSI. (ED) S.Y.KUNG, H.  
WHITEHOUSE AND T. KAILATH, PRENTICE-HALL 1984.  
INFORMATION SYSTEMS LABORATORY, DEP OF EE, STANFORD  
UNIVERSITY, STANFORD, CA 94305.

KELLIHER, EDWARD G.  
DIRECT ELECTRONIC TRANSFORMS FOR FEATURE EXTRACTION.  
UNDERSEA RESEARCH CORP FALLS CHURCH VA.  
GRAI7815.

KITTLER, J.  
FEATURE SELECTION METHODS BASED ON THE KARHUNEN-LOEVE EXPANSION.  
DEPT. OF ELECTRONICS, UNIV. SOUTHAMPTON, SOUTHAMPTON, ENGLAND.

PATTERN RECOGNITION. THEORY AND APPLICATION. P 61-74, 1977.

KLINGER, A.  
IMAGES DECOMPOSITION: ORTHOGONAL TRANSFORMS.  
UNIV. OF CALIFORNIA, LOS ANGELES, CA, USA.  
PROCEEDINGS OF THE WORKSHOP ON PICTURE DATA DESCRIPTION  
AND MANAGEMENT P 130, 1977.

KOHONEN, TEUVO; NEMETH, GABOR; BRY, KALLE-J.; JALANKO, MATTI;  
CLASSIFICATION OF PHONEMES BY LEARNING SUBSPACES.  
RIITTINEN, HEIKKI.  
HELSINKI UNIV. OF TECHNOLOGY, ESPOO (FINLAND). DEPT. OF  
TECHNICAL PHYSICS.  
GRAI8302.

KOHENEN, T.; REUHKALA, E.; MAKISARA, K.; VAINIO, L.  
ASSOCIATIVE RECALL OF IMAGES.  
DEPT OF TECH. PHYS., HELSINKI UNIV. OF TECHNOL., OTANIEMI,  
FINLAND.  
BIOL. CYBERN. (GERMANY) V 22, N 3, P 159-68, 1976.

KOWALSKI, B. R.; BENDER, C. F.  
AN ORTHOGONAL FEATURE SELECTION METHOD.  
DEPT. OF CHEM., UNIV. OF WASHINGTON, SEATTLE, WA, USA.  
PATTERN RECOG V 8, N1, P 1-4, 1976.

KOWALSKI, B. R.; BENDER, C. F.  
THE HADAMARD TRANSFORM AND SPECTRAL ANALYSIS BY PATTERN  
RECOGNITION.  
COLORADO STATE UNIV., FORT COLLINS, USA.  
ANAL. CHEM. (USA) V 45, N 13, P 2234-9, NOV 73.

KSIENSKI, A. A.; LIN, Y. T.; WHITE, L. J.  
LOW FREQUENCY APPROACH TO TARGET IDENTIFICATION.  
OHIO STATE UNIV COLUMBUS DEPT OF ELECTRICAL ENGINEERING.  
GRAI7513.PROC. IEEE. VOL. 63. N12, DEC-75, 1651-1659.

KSIENSKI, AHARON. A., LIN, H. (1981).  
OPTIMUM FREQUENCIES FOR AIRCRAFT CLASSIFICATION.  
IEEE TRANS. ON AEROSPACE AND ELECTRO. SYSTEMS,  
V AES-17, N 5, SEP-81, 655, 665.  
THE OHIO STATE UNIVERSITY ELECTROSCIENCE LABORATORY,  
COLUMBUS, OH 43212.

KURZYNSKI, M.  
THE OPTIMAL STRATEGY OF TREE CLASSIFIER.  
PATTERN RECOGNITION V 16, N 1, P 81-87, 1983.  
TECHNICAL UNIVERSITY OF WROCLAW, CONTROL SYSTEMS GROUP,  
WYBRZEZE WYSPIANSKIEGO 27, 50-370, WROCLAW, POLAND.

LACKEY, R., MELTZER, D.  
A SIMPLIFIED DEFINITION OF WALSH FUNCTIONS.  
IEEE TRANS. ON COMPUTERS, FEB-71, P 211-213.  
DEPARTMENT OF ELECTRICAL ENGINEERING, OHIO STATE UNIVERSITY  
COLUMBUS, OHIO.

LAINIOTIS, D. G.; HENDERSON, T. L.  
APPLICATION OF STATE VARIABLE TECHNIQUES TO OPTIMAL FEATURE  
EXTRACTION.  
TEXAS UNIV AUSTIN ELECTRONICS RESEARCH CENTER.  
PROCEEDINGS OF THE IEEE V 56, N 12, P 2175-2, DEC 68.

LOWITZ, G. (1983)  
CAN A LOCAL HISTOGRAM REALLY MAP TEXTURE INFORMATION?  
PATTERN RECOGNITION V 16, N 2, P 141-147, 1983.  
MATRA ESPACE, RUE DES COSMONAUTES, 31400, TOULOUSE, FRANCE.

LOWITZ, G. (1983)  
STABILITY AND DIMENSIONALITY OF KARHUNEN-LOEVE MULTISPECTRAL  
IMAGE EXPANSION.  
PATTERN RECOGNITION V 10, P 359-363, 1978.  
MATRA ESPACE, RUE DES COSMONAUTES, 31400, TOULOUSE, FRANCE.

MACANALLY, R.; LIVANOS, A.  
THE SPACE BANDWIDTH OF MULTI-CHANNEL COHERENT PROCESSES.  
CALIFORNIA INST. TECHNOL., PASADENA, USA; FEB. 1.975.  
GRAI7511.

MARR, R. B.  
OVERVIEW OF IMAGE RECONSTRUCTION.  
BROOKHAVEN NATIONAL LAB., UPTON, NY.  
GRAI8021.

MAURER, MICHAEL LEE.  
A DIMENSIONALITY REDUCTION TECHNIQUE FOR ENHANCING INFORMATION  
CONTEXT.  
NAVAL POSTGRADUATE SCHOOL, MONTEREY, CA.  
GRAI8103.

MENDEZ, J., FALCON, A. (1983).  
CRITERIOS DE SEMEJANZA Y DISTANCIA EN EL RECONOCIMIENTO  
DE FORMAS.  
ETSII LAS PALMAS. UNIV. POLITENCNICA. SPAIN.  
PROC. II SIMP. ING. BIOMEDICA, MADRID OCT 83, ETSIT P 489-493

MENDEZ, J (1983)  
CONTRIBUCIONES A LA TEORIA DE LA CARACTERIZACION Y  
RECONOCIMIENTO DE FORMAS. APLICACIONES.  
ETSII LAS PALMAS. UNIV. POLITECNICA. SPAIN  
TESIS DOCTORAL.

MIRA, J., MERINO, R., MARTIN, P. (1972)  
UNA GEOMETRIA DUAL COMO MODELO DE PERCEPCION.  
PROC. CONGR. NAC. AUTOMATICA, VOL II, MADRID. AEIA, 1213-1244

MIRA, J., MORENO-DIAZ, R., DELGADO, A. (1982).  
COOPERATIVE PROCESSES EN CEREBRAL DYNAMICS.  
APPLICATIONS OF INFORMATION AND CONTROL SYSTEMS. LAINIOTIS  
TZAUNES (EDS.) HOLLAND: REIDEL PUB. CO. P 273-280.

MIRA, J., MORENO-DIAZ, R. (1982)  
UN MARCO TEORICO PARA INTERPRETAR LA FUNCION NEURONAL A  
ALTOS NIVELES.  
PROC. I REUNION NAC. DE BIOCIBERNETICA. MADRID. REAL  
ACADEMIA DE CIENCIAS. P 151-178.

MIRA, J., MORENO-DIAZ, R., DELGADO, A. (1983).  
A THEORETICAL PROPOSAL TO EMBODY COOPERATIVE DECISION  
IN THE NERVOUS SYSTEM.  
PROC. 1983 CONF. ON GENERAL SYSTEMS RESEARCH, DETROIT,  
MICH. INTERSYSTEMS, P 687-690.

MIRA, J., DELGADO, A., MORENO-DIAZ, R. (1983).  
ON COOPERATIVE LANGUAGES FOR SURVIVAL.  
10TH INT. CONGRESS ON CYBERNETICS. NAMUR, BELGIUM  
P 22-27, AUG-83.

MORENO-DIAZ, R. (1965).  
AN ANALITICA MODEL OF THE GROUP TWO GANGLIONAL CELLS IN THE  
FROG'S RETINA.  
INSTRUMENTATION LABORATORY, CAMBRIDGE: MIT.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R. (1966).  
CONCEPTUAL MODEL OF THE FROG'S TECTUM.  
SENSORY, DECISION AND CONTROL SYSTEMS, R-548, SUTRO, L (ED)  
INSTRUMENTATION LABORATORY, CAMBRIDGE: MIT. 30-40.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R. (1983).  
LAS TEORIAS RETINAL Y DE REDES NEURONALES.  
PROC. II SIMP. ING. BIOMEDIC. ETSIT, MADRID, P 209-219.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R., BEVER, J. (1967).  
VISUAL DATA PROCESING.  
SENSORY, DECISION AND CONTROL SYSTEMS, R-548, SUTRO, L (ED)  
INSTRUMENTATION LABORATORY CAMBRIDGE: MIT, 32-36.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R. (1967).  
VISUAL PROCESSING IN ANIMALS.  
SENSORY, DECISION AND CONTROL SYSTEMS, R-565, SUTRO, L (ED)  
INSTRUMENTATION LABORATORY CAMBRIDGE: MIT, 62-65.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R. (1968).  
CONTRAST DETECTORS.  
NEURAL NETWORKS. CAIANIELLO, E. (ED). BERLIN HEIDELBERG  
NEW YORK: SPRINGER 86-88.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R., RUBIO ROYO, F., MIRA, J. (1969).  
APLICACION DE LAS TRANSFORMACIONES INTEGRALES AL PROCESO  
DE DATOS EN LA RETINA.

AUTOMATICA, 2-5, 7-17.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R., RUBIO, E. (1979).  
A THEORETICAL MODEL FOR LAYERED VISUAL PROCESSING.  
INT. J. BIO-MED. COMP 10, 231-243.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R., RUBIO, E., NUNEZ, A. (1980).  
A LAYERED MODEL FOR VISUAL PROCESSING IN AVIAN RETINA.  
BIOLOG CYBERNETICS, 38, 85-89.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R., SANTANA, O. (1979).  
COMPUTER PROGRAMAS TO IMPLEMENT RETINAL MODELS.  
INT. J. BIOMED. COMP. 10, 217-229.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R., RUBIO, E. (1980).  
A MODEL FOR NONLINEAR PROCESSING IN CAT'S RETINA.  
BIOL. CYBERNETICS, 37, 25-31.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R., RUBIO ROYO, F., RUBIO, E. (1980).  
A THEORETICAL PROPOSAL TO ACCOUNT FOR VISUAL COMPUTATION  
IN A FROG'S RETINA.  
INT. J. BIO-MED. COMP. 11, 415-426.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R., RUBIO ROYO, F. (1981).  
TOWARDS A THEORY OF VISUAL PROCESSING IN VERTEBRATE RETINA.  
APPLIED SYSTEMS AND CYBERNETICS. E. LASKER (ED). NEW YORK  
PERGAMON 1859-1864.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R. MIRA, J., DELGADO, A. (1983).  
SOBRE LA GENERACION Y TRANSFORMACION DE SENALES CON REGLAS  
DE DECISION ALGORITMICAS DE NATURALEZA ARBITRARIA.  
REV. REAL ACADEMIA CIENCIAS MADRID, LXXVII, 2, 470-480.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R., MIRA, J., DELGADO, A. (1983).  
HACIA UNA TEORIA SOCIAL DE REDES NEURONALES.  
PROC. II SIMP. NAC. BIOINGENIERIA. MADRID, ETSIT, 215-220.  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MORENO-DIAZ, R., MIRA, J., FORTES, J., DELGADO, A. (1983).  
PROCESO DE DATOS VISUALES E INTEGRACION MULTISENSORIAL  
EN ROBOTS.  
PROC. JORNADAS DE ROBOTICA Y FABRICACION FLEXIBLE, CSIC, MADRID  
COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

MOSS, R. H.; ROBINSON, C. M.; POPPELBAUM, W. J.  
ONLINE RECOGNITION (OLREC): A NOVEL APPROACH TO VISUAL PATTERN  
RECOGNITION.

DEPT. OF ELECTRICAL ENGG., UNIV. OF MISSOURI-ROLLA, ROLLA, MO,  
USA.

PATTERN RECOGNITION (GB) V 16, N 6, P 535-50, 1983.

MURAKAMI, J.; YOSHIMOTO, C.

ESTIMATIONS OF FOURIER, WALSH AND HAAR SYSTEMS BY TRUNCATION  
ERROR AND ALIASING.

FACULTY OF ENGG., HOKKAIDO UNIV., SAPPORO, JAPAN.

SYST. COMPT. CONTROL (USA) V 5, N 2, P 45-52.

MUZIO, J. C.; HURST, S. L.

THE COMPUTATIONS OF COMPLETE AND REDUCED SETS OF ORTHOGONAL  
SPECTRAL COEFFICIENTS FOR LOGIC DESIGN AND PATTERN RECOGNITION  
PURPOSES.

DEPT OF COMPUTER SCI., UNIV. OF MANITOBA, WINNIPEG, CANADA.

COMPUT. AND ELECTR. ENG. (USA) V 5, N 3, P 231-49, SEP-78.

NAGIN, PAUL; HANSON, ALLEN R.; RISEMAN, EDWARD M.

REGION EXTRACTION AND DESCRIPTION THROUGH PLANNING.

MASSACHUSETTS UNIV AMHERST DEPT OF COMPUTER AND INFORMATION  
SCIENCE.

GRAI7723.

NAKAMURA, Y.; SUNOHARA, S.; FURUYA, M.

AN OPTIMAL ORTHOGONAL EXPANSION FOR CLASSIFICATION OF PATTERNS.

FACULTY OF ENGG., SHINSHU UNIV., NAGANO, JAPAN.

IEEE TRANS COMPUT. V C-26, N 12, DIC 77, P 1288-1292.

NEAGOE, V.; COSMA, D.; PODRUMARU, D.; PRODESCU, V.

OF THE BOUNDARY.

FACULTY OF ELECTRONICS AND TELECOMMUNICATIONS, POLYTECH. INST.

OF BUCHAREST, BUCHAREST, ROMANIA.

CANADIAN COMMUNICATIONS AND ENERGY CONFERENCE P 332-6, 1982.

NEAGOE, V. E.

USING LEGENDRE POLYNOMIALS TO INTRODUCE A NEW ORTHOGONAL  
TRANSFORM FOR SIGNIFIANT FEATURE SELECTION.

FACULTY OF ELECTRONICS AND TELECOMMUNICATIONS, POLYTECH. INST OF  
BUCHAREST, BUCHAREST, ROMANIA.

IEEE COMPUTER SOCIETY CONFERENCE ON PATTERN RECOGNITION  
AND IMAGE PROCESSING P 177-82, 1982.

NUNEZ, A. (1981).

TEORIA UNIFICADA DEL PROCESO Y TRANSMISION DE DATOS VISUALES  
EN LA RETINA.

TESIS DOCTORAL. MADRID. ETSIT. UNIV. POLITECNICA.

LAS PALMAS SPAIN. ETSII. UNIV. POLITECNICA.

OJA, E.; KUUSELA, M.

THE ALSM ALGORITHM-AN IMPROVED SUBSPACE METHOD OF  
CLASSIFICATION.

DEPT. OF APPL. MATH., UNIV. OF KUOPIO, KUOPIO, FINLAND.

PATTERN RECOGNITION (GB) V 16, N 4, P 421-7, 1983.

OTSU, N.

PICTURE REPRESENTATION BY MARGINAL EIGENVECTORS.  
ELECTROTECH. LAB., CHIYODA-KU, TOKYO, JAPAN.  
BULL ELECTROTECH. LAB (JAPAN) V 39, N 12, P 58-73.

OZAWA, K. (1983).  
CLASSIC: A HIERACHICAL CLUSTERING ALGORITHM BASED ON  
ASYMMETRIC SIMILARITIES.  
PATTERN RECOGNITION V 16, N 2, P 201-211 1983.  
OSAKA ELECTRO-COMMUNICATION UNIVERSITY, NEYAGAWA-SHI,  
OSAKA 572, JAPAN.

PETERSON, LARRY J.  
CASCADES OF TRANSFORMATIONS OF TWO DIMENSIONAL DISTRIBUTION  
FUNCTIONS.  
ILLINOIS UNIV URBANA BIOLOGICAL COMPUTER LAB.  
USGRDR7001.

PONCIN, J., SCHWARTZ, P. (1972)  
PROPIETES STATISTIQUES DE TRNASFOMEES ORTHOGONALES D'IMAGES.  
ANNALES DES TELECOMUNICATIONS. T.27, N 5-6, P 173-184.

RAO, K. R.; AHMED, N.  
ORTHOGONAL TRANSFORMS FOR DIGITAL SIGNAL PROCESSING.  
DEPT. OF ELECTRICAL ENGG., UNIV. OF TEXAS, ARLINGTON, TX., USA.  
1976 IEEE INTERNATIONAL CONFERENCE ON ACOUSTICS, SPEECH AND  
SIGNAL PROCESSING P 136-40.

REITER, B.  
ORTHOGONAL SUBSPACES FOR MULTI-CLASS PATTERN RECOGNITION.  
UNIV. HAWAII, HONOLULU.  
PROCEEDINGS OF THE 5TH HAWAII INTERNATIONAL CONFERENCE ON  
SYSTEM SCIENCE P 398-400, 1972.

ROHDE, FREDERICK W.  
INVESTIGATION OF DISCRETE FUCNTION TECHNOLOHGY FOR TOPOGRAPHIC  
SCIENCES.  
ARMY ENGINEER TOPOGRAPHIC LABS FOR BELVOIR VA.  
GRAI7907.

ROSENFELD, A., KAK, A. (1976).  
DIGITAL PICTURE PROECESSING  
ACADEMIC PRESS, NEW YORK.

ROSENFELD, A., WESZKA, J. (1980).  
PICTURE RECOGNITION.  
DIGITAL PATTERN RECOGNITION, K,S, FU (ED), SPRINGER-VERLAG  
P 135-166.

ROTHSTEIN, J.  
GENERALIZED ORTHOGONAL REGRESSION IN PATTERN RECOGNITION.  
COMPUTER AND INFORMATION SCI., OHIO STATE UNIV., COLUMBUS, OH,  
USA.  
PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON CYBERNETICS  
AND SOCIETY P 572-6, 1972.

RUBIO, E. (1979).  
 MODELOS RETINALES GENERALIZADOS.  
 TESIS DOCTORAL, TENERIFE, UNIV. DE LA LAGUNA.  
 ETSII. LAS PALMAS. UNIV PLITECNICA.

RUBIO, E., SANTANA, O., MORENO-DIAZ R. (1981).  
 CAMERA-COMPUTER SIMULATION OF RETINAL PROCESSES.  
 APPLIED SYSTEMS AND CYBERNETICS. E. LASKER (ED). NEW YORK  
 PERGAMON 1865-1869.

RUBIO ROYO, F. (1968).  
 PROCESO DE DATOS VISUALES EN LA RETINA  
 TESIS DOCTORAL, MADRID, UNIV. COMPLUTENSE, FAC. CIENCIAS.

SADJADI, F. A.; HALL, E. L.  
 A MULTISOURCE FEATURE SELECTION.  
 IMAGE AND PATTERN ANALYSIS LAB., DEPT. OF ELECTRICAL ENGG.,  
 UNIV. OF TENNESSEE, KNOXVILLE, TN, USA.  
 PROCEEDINGS OF SOUTHEASTCON 1979 P 373-6, 1979.

SANTANA, O. (1979).  
 LENGUAJES PARA LA SIMULACION DE PROCESOS VISUALES.  
 TESIS DOCTORAL. TENERIFE. UNIV. DE LA LAGUNA.  
 COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

SANTANA, O., FORTES, J., DE LA NUEZ, I. (1981).  
 OBJET ISOLATION AND RECOGNITION IN A CAMERA-COMPUTER CHAIN.  
 APPLIED SYSTEMS AND CYBERNETICS. E. LASKER (ED) NEW YORK  
 PERGAMON, 2405-2410.  
 COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

SANTANA, O., MENDEZ, J., MORENO-DIAZ, R. (1981).  
 MOMENTOS NORMALIZADOS PARA EL PROCESO DE DATOS VISUALES.  
 REV. REAL ACAD. CIENCIAS MADRID, LXXV, 1, 287-289.  
 COLEGIO UNIVERSITARIO, APT 550, LAS PALMAS, SPAIN.

SCHILLER, H.; WILLE, M. J.  
 SOME REMARKS CONCERNING THE USAGE OF WALSH TRANSFORMATION IN  
 PICTURE PROCESSING.  
 DEUTSCHE AKADEMIE DER WISSENSCHAFFEN ZU BERLIN, ZEUTHEN  
 (GERMAN D. R.) INST. FUER HOCHENERGIEPHYSIK.  
 GRAI7912. PHE 76-14.

SHANKS, J. (1969).  
 COMPUTATION OF THE WALSH-FOURIER TRANSFORM.  
 IEEE TRANS. ON COMPUTERS, P 457-459, MAY-69.  
 PAN AMERICAN PETROLEUM CORPORATION TULSA, OKALA 74102, USA.

SHI-KUO CHANG; CHOW, C. K.  
 THE RECONSTRUCTION OF THREE-DIMENSIONAL OBJECTS FROM TWO  
 ORTHOGONAL PROJECTIONS AND ITS APPLICATION TO CARDIAC  
 CINEANGIOGRAPHY.  
 ACAD. SINICA, TAIPEI, TAIWAN, CHINA.  
 IEEE TRANS. COMPUT. (USA) V C-22, N 1, P 18-28, JAN 73.

SHORE, JOHN E.  
COMPLETE ORTHONORMAL SET OF TWO-DIMENSIONAL HAAR-LIKE FUNCTIONS.  
NAVAL RESEARCH LAB WASHINGTON D. C.  
GRAI7307. NRL REPORT 7471.

SIVAK, GARY.  
THE HARR TRANSFORM: ITS THEORY AND COMPUTER IMPLEMENTATION.  
ARMY ARMAMENT RESEARCH AND DEVELOPMENT COMMAND DOVER NJ FIRE  
CONTROLL AND SMALL CALIBER WEAPON SYSTEMS LAB.  
GRAI7923.

STARKS, SCOTT A.; DE FIGUEIREDO, RUI J. P.  
A NEW APPROACH TO STRUCTURE PRESERVING FEATURE EXTRACTION.  
RICE UNIV HOUSTON TEX DEPT OF ELECTRICAL ENGINEERING.  
GRAI7723.

STOCKMAN, GEORGE C.; LAMBIRD, BARBARA A.; LAVINE, DAVID  
KNOWLEDGE-BASED IMAGE ANALYSIS.  
ARMY ENGINEER TOPOGRAPHIC LABS., FORT BELVOIR, VA.  
GRAI8123.

SULLIVAN, JERRY.  
ACCURATE LEAST-SQUARES TECHNIQUES USING THE ORTHOGONAL  
FUNCTION-APPROACH.  
NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION, WASHINGTON.  
GRAI8024.

TAI-SEN YU. KING-SUN FU. (1983).  
RECURSIVE CONTEXTUAL CLASSIFICATION USING A SPATIAL  
STOCHASTIC MODEL.  
PATTERN RECOGNITION V 16, N 1, P 89-108. 1983.  
SCHOOL OF ELECTRICAL ENGINEERING, PURDUE UNIVERSITY,  
WEST LAFAYETTE, IN 47907, USA.

THERRIEU, C. W.  
A GENERALISED APPROACH TO LINEAR METHODS OF FEATURE EXTRACTION.  
MIT. LEXINGTON, MASS., USA; 30 DEC. 1.974.  
PROCEEDINGS OF THE CONFERENCE ON COMPUTER GRAPHICS, PATTERN  
RECOGNITION AND DATA STRUCTURE. P 312-15.

TOMITA, S.; NOGUCHI, S.  
SELF-CLASSIFICATION FOR PATTERNS BY THE KARHUNEN-LOEVE  
ORTHOGONAL SYSTEM.  
TOHOKU UNIV., SENDAI, JAPAN.  
ELECTRON COMMUN. JAPAN. V 55, N 9, SEP 72, P 22-29.

TOMITA, S.; NOGUCHI, S.; OIZUMI, J.  
THEORY OF FEATURE EXTRACTION FOR PATTERNS BY THE KARHUNEN-LOEVE  
ORTHOGONAL SYSTEM.  
TOHOKU UNIV. SENDAI, JAPAN.  
SYS. COMPUT. CONTROL (USA) V 1, N 6, P 55-6, DEC 70.

TOMITA, S.; NOGUCHI, S.; OIZUMI, J.  
DISCRIMINATION FOR HANDWRITTEN CHARACTERS BY THE KARHUNEN-LOEVE  
ORTHOGONAL SYSTEM.

TOHOKU UNIV., SENDAI, JAPAN.  
PROCEEDINGS 4TH HAWAII INTERNATIONAL CONFERENCE ON SYSTEM  
SCIENCES. P 684-5, 1971.

TOMITA, S.  
ESTIMATION OF UNKNOWN PARAMETERS ON A MULTI-CLASSIFIED PROBLEM BY  
THE KARHUNEN-LOEVE ORTHOGONAL SYSTEM.  
FACULTY OF ENGNG., YAMAGUCHI UNIV., UBE, JAPAN.  
SUS. COMPUT. CONTROL (USA) V 10, N 6, P 13-21, DEC-79.

WADBROOK, D. G.; WOOLONS, D. S.  
IMPLEMENTATION OF 2-DIMENSIONAL WALSH TRANSFORMS FOR PATTERN  
RECOGNITION.  
UNIV. SUSSEX, BRIGHTON, ENGLAND.  
ELECTRON LETT (GB). V 8, N 5, P 134-6, MAR-72.

WATANABE, S.  
PATTERN RECOGNITION AS INFORMATION COMPRESSION.  
UNIV. HAWAII, HONOLULU.  
1ST INTERNATIONAL JOINT CONFERENCE ON PATTERN  
RECOGNITION P 25-32, 1973.

WATANABE, S.; PAKVASA, N.  
SUBSPACE METHOD IN PATTERN RECOGNITION.  
UNIV. HAWAII, HONOLULU, USA.

WESZKA, JOAN S.; ROSENFELD, AZRIEL.  
A COMPARATIVE STUDY OF TEXTURE MEASURES FOR TERRAIN  
CLASSIFICATION.  
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH, ARLINGTON, VA.  
GRAI7519.

WINTZ, PAUL A.; HUANG T. S.  
IMAGE ANALYSIS AND MODELING.  
ROME AIR DEVELOPMENT CENTER, GRIFFISS AFB, NY; DEFENSE ADVANCED  
RESEARCH PROJECTS AGENCY, ARLINGTON, VA.  
GRAI7602.

WALSH FUNCTIONS. 1.972-MARC..  
NATIONAL TECHNICAL INFORMATION SERVICE. SPRINGFIELD, VA.  
GRAI8212. (CITATIONS FROM THE INTERNATIONAL AEROSPACE  
ABSTRACTS DATA BASE).

MACHINE HOLOGRAPHY, PRELIMINARY REPORT.  
FORD AEROSPACE AND COMMUNICATIONS CORP., NEWPORT BEACH, CA.  
AERONUTRIC DIV.  
GRAI7913.