

**UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA**

**DEPARTAMENTO DE INFORMÁTICA Y SISTEMAS**



**TESIS DOCTORAL**

**ESTUDIO DE LA VARIABILIDAD DE LOS CAMPOS RECEPTIVOS  
EN LAS TRANSFORMACIONES DE CAMPOS DE DATOS.  
IMPLICACIONES EN TRANSFORMACIONES DE IMÁGENES Y  
TEORÍA RETINAL**

**JUAN CARLOS QUEVEDO LOSADA**

Las Palmas de Gran Canaria, 1997

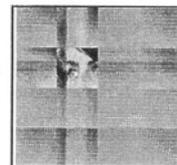
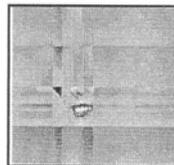
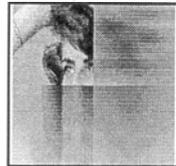
UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Facultad de Informática

Departamento de Informática y Sistemas

## TESIS DOCTORAL

**ESTUDIO DE LA VARIABILIDAD DE LOS CAMPOS  
RECEPTIVOS EN LAS TRANSFORMACIONES DE  
CAMPOS DE DATOS. Implicaciones en  
Transformaciones de Imágenes y Teoría Retinal**



AUTOR: Juan Carlos Quevedo Losada

DIRECTORES: Dra. Olga Bolívar Toledo

Dr. Roberto Moreno Díaz, jr.

Junio, 1997

# UNIVERSIDAD DE LAS PALMAS DE GRAN CANARIA

Doctorado en Informática

Departamento de Informática y Sistemas  
Programa de Percepción Artificial y Aplicaciones

## ESTUDIO DE LA VARIABILIDAD DE LOS CAMPOS RECEPTIVOS EN LAS TRANSFORMACIONES DE CAMPOS DE DATOS. Implicaciones en Transformaciones de Imágenes y Teoría Retinal

Tesis Doctoral presentada por D.  
JUAN CARLOS QUEVEDO LOSADA

Dirigida por los Drs.  
Dña. OLGA BOLÍVAR TOLEDO  
D. ROBERTO MORENO DÍAZ, jr.

Los Directores

El Doctorando

Las Palmas de Gran Canaria, 9 de Junio de 1997

*A mi Familia*

## AGRADECIMIENTOS

Quisiera expresar mi más sincero agradecimiento a los codirectores de esta tesis, los Drs. Olga Bolívar y Roberto Moreno, por su inestimable ayuda y colaboración en el seguimiento y culminación de la misma.

También me gustaría agradecer a los miembros del Departamento de Informática y Sistemas la colaboración prestada, y, es especial, a José A. Muñoz por el apoyo y gran interés que ha demostrado en todo momento.

Finalmente, y no por ello menos importante, quiero dar las gracias a Sonia por su paciencia y ayuda incondicional durante el desarrollo de este trabajo.

# **RESUMEN**

---

En este trabajo titulado "*Estudio de la Variabilidad de los Campos Receptivos en las Transformaciones de Campos de Datos. Implicaciones en Transformaciones de Imágenes y Teoría Retinal*" presentamos distintos desarrollos relacionados con el concepto neurofisiológico de Campo Receptivo, estudiando la influencia de su variabilidad en las transformaciones algebraico-analíticas que combinan campos receptivos y funcionales. Lo hemos dividido en 4 capítulos y dos apartados referentes a las Conclusiones y Referencias bibliográficas.

En el Capítulo 1 hacemos referencia a un conjunto de antecedentes en Transformaciones de Imágenes donde exponemos distintos conceptos creados y desarrollados en trabajos anteriores y que constituyen los cimientos de los siguientes capítulos. Además de los conceptos teóricos expuestos, presentamos algunos resultados prácticos de la aplicación de éstos a una imagen en particular como ejemplo. Se expone los Filtros de Newton y los aplicamos para obtener núcleos utilizables en la normalización de las intensidades. Destacamos también, dentro de este apartado, la utilización de estos núcleos en la operación de filtrado de una imagen para la extracción de bordes.

En el Capítulo 2, hemos desarrollado y aplicado la Transformación Foveal Dirigida (TFD), como primer paso en la modelización de los mecanismos de atención que rigen gran parte del proceso visual. Como punto de partida, exponemos su antecedente en la Transformación Foveal para continuar con su desarrollo. La base fundamental se centra en la posibilidad de dirigir la fovea a cualquier punto del campo de datos, dándole mayor resolución y funcionando como un atractor de interés. Concluimos este tema proponiendo áreas de aplicabilidad de esta transformada como son utilizarlas como método criptográfico de imágenes y como mecanismo de atención. Finalmente, hacemos mención de su incidencia en la Teoría Retinal.

En el Capítulo 3, presentamos un estudio relacionado con la representación de

imágenes y campos receptivos, dándole una enfoque *natural* y *artificial*. En primer lugar, y resumiendo lo expuesto en el Capítulo 1, definimos el concepto de Campo Receptivo desde el punto de vista *natural* para a continuación aterrizar en lo *artificial* y su relación con los funcionales. Esto nos conduce a plantear y estudiar la influencia de distintos parámetros que *rodean* a las transformaciones de campos receptivos. Para ello, utilizamos el Reconocimiento de Formas, que como es sabido, constituye la base de la visión artificial, donde obtenemos distintas conclusiones que reflejamos al final de este trabajo. Además, nos sirve de base para desarrollar un método de descripción automática de formas mediante la localización de los campos receptivos.

En el Capítulo 4, estudiamos la variabilidad de los campos receptivos y su conectividad con la complitud. Desarrollamos un método para la generación de funcionales y lo aplicamos basándonos en distintas transformadas clásicas. A parte de éstas, desarrollamos otros tipos de transformadas como son las Diagonales y las basadas en los Filtros de Newton. Por último presentamos un marco de aplicabilidad exponiendo su incidencia en la compresión de imágenes.

# ÍNDICE

---

**CAPITULO 1: ANTECEDENTES EN TRANSFORMACIONES DE IMÁGENES**

1.1 Generación de Espacios de Representación	1
1.2 El Sistema Visual	6
1.3 Codificación Retinal	7
1.4 Convoluciones Generalizadas	11
1.4.1 Convoluciones Algebraicas. Filtros Digitales	12
1.4.2 Convoluciones Analíticas	15
1.4.3 Convoluciones Algorítmicas	22
1.5 Filtros de Newton	25
1.6 Transformaciones en Campos de Datos	31

**CAPITULO 2: TRANSFORMACIONES FOVEALES GENERALIZADAS**

2.1 Antecedentes	36
2.2 Transformación Foveal Dirigida (TFD)	40
2.3 Aplicaciones de la TFD	45
2.3.1 TFD Como Método Criptográfico	45
2.3.2 TFD Como "Mecanismo de Atención"	52
2.4 Incidencia en Teoría Retinal	55

**CAPITULO 3: REPRESENTACIONES DE IMÁGENES A PARTIR DE LAS TRANSFORMACIONES DE CAMPOS RECEPTIVOS**

3.1 Campos Receptivos y Funcionales	57
3.2 Influencia de los Parámetros Característicos de las Transformaciones de Campos Receptivos en el Reconocimiento de Formas	62
3.2.1 Introducción	62
3.2.2 Banco de Imágenes	63
3.2.3 Frases Descriptores: Particiones y Funcionales	68
3.2.4 Estudio Experimental	69
3.3 Descripción Automática de Formas Utilizando los Campos Receptivos	92

---

**CAPITULO 4: VARIABILIDAD DE LOS CAMPOS RECEPTIVOS. COMPLITUD. GENERACIÓN DE FUNCIONALES**

4.1 Introducción	96
4.2 Métodos de Generación	101
4.3 Basados en la Transformada de Haar	104
4.4 Basados en la Transformada de Hadamard	108
4.5 Basados en las Transformaciones Diagonales	110
4.6 Basados en los Filtros de Newton	114
4.7 Incidencia en la Compresión de Imágenes	116
Conclusiones y Principales Aportaciones	122
Referencias	125

# **CAPÍTULO 1**

## **ANTECEDENTES EN TRANSFORMACIONES DE IMÁGENES**

## 1.1 GENERACIÓN DE ESPACIOS DE REPRESENTACIÓN

La generación de espacios de representación, basada en la expansión de un conjunto de coeficientes de imagen en un conjunto alternativo, se basa en una conjetura que, aunque no haya sido reducida a teoremas formales, proporciona pistas para construir sistemas que resuelvan situaciones prácticas concretas. Un caso particular que ha dado un juego considerable en los procesos visuales y que fue importada, en cierto modo, por el papel que jugó y juega en los procesos auditivos y en los problemas de transmisión y comunicaciones es la Transformada de Fourier y sus variantes tipo Fourier-Bessel. Otro caso es el de los momentos donde a partir del espacio canónico se llega a otro por potencias de las coordenadas que han de ser mantenidas dentro de cierto rango. Esta técnica de representación alternativa tiene un origen no perceptual, más relacionado con la física.

Un conjunto clásico de espacios de representación, con una gran variedad de aplicaciones en el Proceso Digital de Imágenes, lo constituye el campo de las Transformadas de Imágenes, las cuales, al ser representaciones discretas, pueden ser consideradas modelos deterministas de las propias imágenes, lo cual las hace útiles en aplicaciones de extracción de características. Sin embargo, las aplicaciones más extendidas en el campo de las Transformadas de Imágenes, pertenecen al ámbito de la codificación y restauración.

### **Codificación y Transformaciones**

El objetivo final de las técnicas de codificación es reducir al máximo el número de bits necesario para llevar a cabo la representación digital de una imagen. Quizás, la forma más simple y drástica de compresión de datos es el muestreo de imágenes limitadas en banda, por el cual un número infinito de pixels por unidad de área se reduce a una muestra, sin pérdida ninguna de información, por lo que el número de muestras por unidad de área

se reduce infinitamente.

Tanto la compresión como el cifrado son dos técnicas basadas en la manipulación de la representación de los datos, pero preservando el significado de los mismos. Son, pues, técnicas semánticas y representan cambios en la forma de codificar los mismos.

De acuerdo con la teoría de la información, la codificación de una imagen puede ser de dos tipos: codificación de la fuente y codificación del canal. En la codificación de la fuente, se trata de eliminar la información redundante, mientras que en la codificación del canal se añade a la señal una cierta redundancia para protegerla contra el ruido.

En general, la información superflua puede ser estadísticamente redundante o subjetivamente redundante. La redundancia estadística es aquella que puede ser eliminada de los datos sin destruir la información; esto implica que los datos originales pueden ser recuperados. La redundancia subjetiva puede ser eliminada sin que cause molestia al observador, pero es irreversible en el sentido de que la información original no puede ser recuperada. La idea básica de la compresión es la de transformar la señal de imagen correlada, en otra incorrelada, susceptible de ser representada por un número de bits menor que la señal original. La transformación debe ser lineal e invertible.

Las principales técnicas de codificación de imágenes pueden ser clasificadas en diversas categorías [Net80]: PCM, codificación predictiva, codificación por transformada, codificación interpolativa y extrapolativa, codificación estadística, codificación híbrida y otras. Estas clases pueden subdividirse dependiendo de si los parámetros del codificador son fijos (codificación fija) ó si cambian en función del tipo de datos a codificar (codificación adaptativa).

Para nuestros propósitos podemos subdividir las categorías en dos:

- Técnicas de codificación espacial
- Técnicas de codificación por transformada

En la codificación por transformada, se aplica una transformada 2D discreta, lineal e invertible, a la señal de imagen previamente dividida en bloques, obteniéndose una matriz transformada. Se seleccionan un conjunto de coeficientes, se cuantifican y se almacenan o transmiten. La recuperación de la señal original se logra mediante la transformada inversa. En este tipo de codificación se plantean tres problemas básicos: la elección de la transformada discreta decorrelante, la elección de los coeficientes a seleccionar y la fidelidad de la reconstrucción.

Desde el punto de vista de la selección de bloques, las transformadas pueden clasificarse en dos grandes grupos:

- Transformadas por bloques.
- Transformadas solapadas.

En la codificación mediante transformada por bloques, la imagen de dimensión  $N \times N$  se divide en  $B$  bloques de  $M$  muestras, efectuándose una transformación mediante una matriz ortogonal de orden  $M$  sobre cada bloque, obteniéndose  $M$  coeficientes transformados. Debido a la independencia de cada bloque, se producen discontinuidades en la imagen reconstruida, dando lugar al "efecto de bloques". De todas las transformadas por bloques lineales y ortonormales, sólo la KLT (Transformada de Hotelling) produce exactamente coeficientes incorrelados. Además la KLT minimiza el error cuadrático medio (ecm) entre la imagen original y la reconstruida. El inconveniente que presenta es su complejidad computacional, determinada por los autovectores de la matriz de  $N^2 \times N^2$  [Gir77].

La Transformada Coseno Discreta (DCT), que es un sustituto aceptable para la KLT discreta, presenta unos coeficientes que no están exactamente incorrelados [Rao90]. Con objeto de reducir el efecto de bloques, surgieron las transformadas solapadas. Con una transformada solapada, se transforman  $L$  muestras del bloque de entrada en  $M$  coeficientes, siendo  $L$  la longitud de las funciones base LT, con  $L > M$ . Como se ha de mantener la frecuencia de muestreo se computan  $M$  nuevos coeficientes para cada  $M$  nuevas muestras de entrada. Una vez calculada la LT de un bloque se llevan las  $M$  muestras del bloque al buffer de  $L$  muestras. De esta forma habrá un solape de  $L-M$  muestras en el cómputo de transformadas de bloques consecutivos.

Resulta de interés la consideración de la ortogonalidad, lo que ha dado lugar al desarrollo de las transformadas solapadas ortogonales [Cas85]. De entre éstas, cabe destacar las Transformadas Solapadas Moduladas (MLT) y, a su vez, las Expandidas (ELT). Otra clase de LT son las Transformadas Solapadas Jerárquicas (HLT) basadas en el concepto de descomposición multirresolución de una señal; de hecho, la HLT puede considerarse como una Transformada Discreta en Onditas.

Las transformadas en onditas (WT) surgen del procesamiento de señales por descomposición en canales multifrecuencia y se derivan del Análisis de Fourier. Presentan un gran interés en el proceso de imágenes y en visión artificial, ya que los modelos multicanales se han empleado para explicar procesos de bajo nivel en la corteza visual.

Moreno Díaz y Candela [Can87] han puesto de manifiesto el hecho de que las transformaciones típicas presentan propiedades adecuadas para generar espacios de representación no tanto debido a los funcionales propios, sino más bien como consecuencia de una partición del campo de datos en los denominados campos receptivos. Dichos autores enfatizan el concepto de complitud, y consideran que en Visión Artificial debe existir un compromiso básico entre los funcionales a realizar y los campos receptivos, es decir, entre

lo que se hace sobre los datos y donde se toman éstos, para mantener una constancia de los grados de libertad impuestos por los objetivos del sistema .

Lo anterior les ha llevado a la introducción de las Transformaciones de Resolución Progresiva (TRP), cuya propiedad inmediata consiste en la obtención de un espacio de salida, que representa al espacio de entrada con varias resoluciones. Esto equivale a disponer en el dominio transformado, de subimágenes con distinta resolución, equivalentes a un muestreo variable [Rov92].

Teniendo en cuenta que el *desideratum* de todo sistema de proceso de imágenes es el Sistema Visual humano, parece coherente antes de aplicar cualquier esquema clásico de codificación, preprocesar la forma canónica de una imagen, teniendo en cuenta las propiedades y características de los Sistemas Naturales.

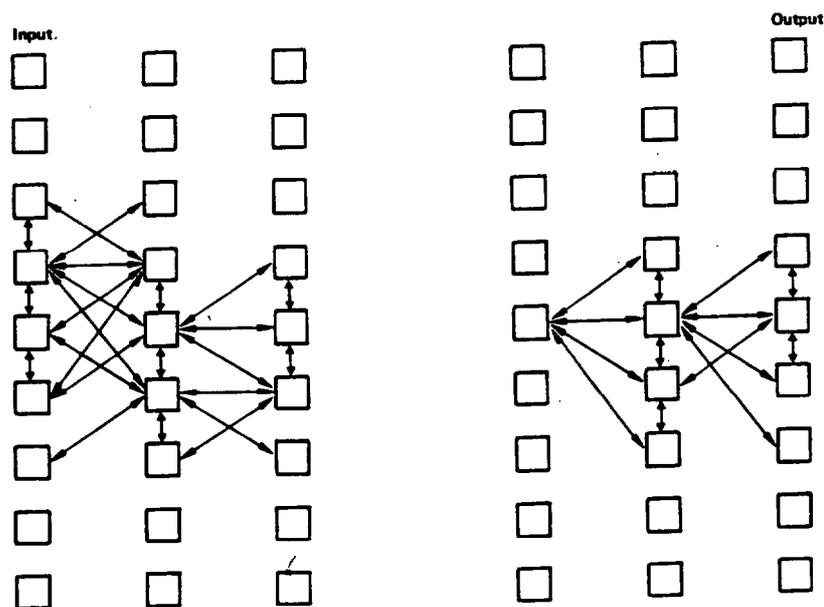
Por otro lado, las investigaciones en codificación de imágenes orientadas fundamentalmente por la teoría de la información no tienen en cuenta lo que el ojo ve, ni como lo ve. Los progresos recientes en el estudio de los mecanismos cerebrales de la visión, han abierto nuevas pistas en la codificación de imágenes. La sensibilidad direccional de las neuronas a lo largo del camino visual junto con los procesos separados de contornos y texturas, han dado lugar a una nueva clase de métodos de codificación, denominados de 2ª generación que alcanzan muy altas tasas de compresión.

## 1.2 EL SISTEMA VISUAL

EL Sistema Visual Humano (SVH) es una parte del Sistema Nervioso, el cual constituye, sin lugar a dudas, la red de comunicaciones más compleja que conocemos, la cual está a su vez controlada por el más potente computador: el Cerebro. La comunicación en esta red se realiza a través de células nerviosas denominadas neuronas. (El cerebro contiene alrededor de  $10^{11}$  neuronas que es aproximadamente el mismo número de estrellas de nuestra galaxia).

La transferencia de información de una a otra neurona, se realiza de forma electroquímica, denominándose a la neurona transmisora y receptora, presináptica y postsináptica respectivamente. La acción de una neurona puede ser de dos tipos: excitatoria e inhibitoria; la primera genera pulsos en la neurona postsináptica, mientras que la segunda inhibe los pulsos existentes.

Para tener una idea de la complejidad del Sistema Nervioso (SN) podemos imaginar una red de  $10^{12}$  neuronas, conectadas en cascada, en paralelo y con realimentación. En la siguiente figura se muestra una imagen correspondiente a este esquema organizativo.



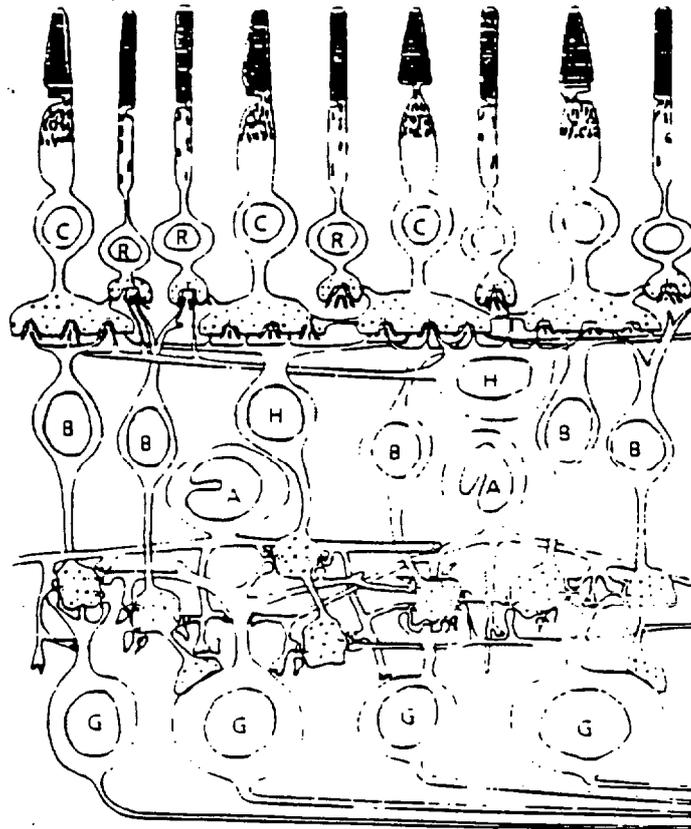
El análisis de este tipo de red podría parecer, a priori, si no imposible, si bastante difícil. Sin embargo, hay una serie de características que simplifican este estudio y que hacen posible actualmente analizar el SN sobre una base célula a célula. Estas características son:

En el SN sólo existen dos tipos de señales: una para distancias largas y otra para distancias cortas. Estas señales son casi idénticas en todas las neuronas, independientemente de la información que transporten, sea ésta visual, táctil, audible, etc. Más aún, su forma no varía de especie a especie. Lo que permite al cerebro distinguir entre dos señales idénticas, es la trayectoria seguida por cada una de las señales, en otras palabras, el cableado. Existe por tanto un entramado específico de neuronas correspondiente a cada tipo de excitación. Desde un punto de vista matemático existe un mapeo uno-a-uno entre las diferentes partes del cuerpo y el cerebro.

### **1.3 CODIFICACIÓN RETINAL**

El ojo es el sensor de las señales visuales, siendo la retina la capa neurosensorial del ojo. La anatomía de la retina muestra cinco tipos de células organizadas en capas, existiendo una correspondencia anatómico funcional entre ellas. Los fotorreceptores transforman la señal luminosa en una señal eléctrica cuya duración y amplitud dependen de la intensidad y longitud de onda de la luz incidente. Tras esta primera capa somática se sitúa la capa plexiforme externa donde se encuentran las conexiones sinápticas de fotorreceptores con células bipolares y horizontales y tras ella la capa somática interna con los cuerpos celulares de bipolares, horizontales e interplexiformes. Las células horizontales reciben información de los fotorreceptores y envían su salida tanto a fotorreceptores como a bipolares. Las bipolares por su parte, transforman la información aferente entre las dos capas plexiformes, es decir, sus salidas alimentan las conexiones con las células amacrinas

y las ganglionares (conexiones que forman la capa plexiforme interna), cuyos cuerpos celulares forman la capa somática mas externa de la retina. Los axones de las ganglionares forman el nervio óptico que abandona el ojo a través del punto ciego y lleva la salida de estas células al cuerpo lateral geniculado (CLG).



Este constituye la "primera parada" de la señal visual antes de llegar a la corteza visual, y morfológicamente está formado por seis capas de neuronas divididas en dos grupos: las capas parvocelulares (de células de pequeño tamaño) y las capas magnocelulares (de células más grandes). La neurofisiología relaciona la actividad del CLG con el procesamiento de color y codificación de movimiento, si bien por nuestro grupo se ha propuesto un mecanismo de cómputo invariante frente a cambios externos globales de iluminación que residiría en el CLG y que haría uso de una acción del tipo inhibición presináptica, tal y como se explicará, en su versión artificial, más adelante en

el capítulo 2.

En el SVH la imagen retiniana es muestreada por unos 125 millones de fotorreceptores, y el nervio óptico consta, por contra, de un millón de líneas [Lei72][Tru65].

Las retinas de muchas especies presentan una región central o fovea donde la agudeza visual presenta un máximo. En esta región, la población de fotorreceptores es mas densa que en otras zonas retinales y las células están desplazadas lateralmente, de forma que la luz incida lo mas directamente posible sobre los fotorreceptores.

Un concepto íntimamente ligado a la computación o proceso por capas es el de computación paralela. Dentro de cada capa existen multitud de células (procesadores) trabajando en paralelo sobre los datos de sus campos receptivos. Las células de una capa determinada no hacen todas ellas lo mismo ni son sensibles a exactamente los mismos tipos de estímulos sobre su campo receptivo. Este paralelismo en el proceso implica, pues, una extracción de características diferentes de la información de entrada. El proceso en paralelo dentro de una capa se complica conforme la información llega al cerebro, donde existen diversos caminos de bifurcación y la computación en paralelo ha de verse no solo dentro de una capa sino entre capas diferentes: áreas paralelas diferentes tienen funciones independientes. El análisis apropiado de los procesos retinales, pues, ha de basarse en los conceptos de computación por capas, contenidos en germen en los escritos de McCulloch y Pitts [McC90].

Intrínsecamente unido al concepto anterior de computación por capas, se encuentra el concepto de Campo Receptivo, hoy en día adoptado por ingenieros y matemáticos. El Campo Receptivo de una célula en una capa determinada es el conjunto de células de la capa anterior que le envían señales. El número de células que procesan la información y el

tipo y tamaño de los campos receptivos varía de capa a capa, existiendo una continua *contracción y expansión* de los lugares en donde es procesada la información. En cada capa se observa un altísimo grado de solape entre campos receptivos vecinos, lo que unido a la convergencia y divergencia de información entre capas, ha dado lugar a la formulación de diversos modelos de localización y detección de estímulos [Bol89][Grü73], preservación de información (y su relación con la complitud de las transformaciones llevadas a cabo por las neuronas vistas como procesadores [Bol89][Mor79]), percepción de movimiento [Grü73] y funcionamiento de ganglionares en retinas de anfibios [Mor65].

A través de todo lo anterior es presumible que, en este esquema de computación por capas, conforme la información progresa en su camino hacia zonas más centrales del cerebro, el significado de la actividad neuronal se incrementa dramáticamente. A medida que nos movemos desde los sensores, el grado de complejidad semántica aumenta y se hace más difícil descifrar la función neuronal atendiendo únicamente al registro de su actividad eléctrica. Es aquí donde resulta necesario un análisis tanto de la cooperatividad neuronal como de su morfogénesis, si se pretende llegar a un conocimiento de la función nerviosa. [Mat90][Mir] .

El uso de los dos conceptos ya expuestos, Campo Receptivo y Computación por Capas, hace ver la actividad neuronal global (tomando como entrada la salida de los fotorreceptores y como salida la señal en el nervio óptico), desde un punto de vista formal matemático, como la *implementación natural* de una operación convolutiva. Así, la capa de fotorreceptores de una retina daría el resultado del muestreo de una imagen a la siguiente capa de células, las cuales acceden a *trozos* de la imagen, los *trozos* que caen efectivamente sobre sus campos receptivos, que en conjunto *barren* toda una imagen digitalizada. Sin embargo, existen varias diferencias importantes: la actividad de las células ganglionares realizaría una convolución cuyos núcleos no tienen una funcionalidad constante ni espacial ni temporalmente.

### 1.3 CONVOLUCIONES GENERALIZADAS

Un Campo de Datos arbitrario, discreto y multidimensional, se puede definir por un espacio de entrada, de dimensión  $n$ , tal que la posición  $ij\dots n$  está ocupado por un número,  $f_{ij\dots n}$ . En el caso de que el Campo de Datos fuese una imagen,  $f_{ij\dots n}$  indica, por ejemplo, la intensidad luminosa (tonalidad de gris).

Una transformación sobre  $f_{ij\dots n}$  (espacio de entrada) es un conjunto de reglas  $R_{ij\dots n}$  (una regla para cada posible  $n$ -tupla de los subíndices  $ij\dots n$ ), tal que, aplicadas a todos los números  $f_{ij\dots n}$ , generan un nuevo espacio  $F_{ij\dots n}$  (denominado espacio de salida).

$$F_{ij\dots n} = R_{ij\dots n} \{ \forall_{ij\dots n} f_{ij\dots n} \} \quad \text{Ec. 1}$$

El marco general para la creación de transformaciones sobre la base de convoluciones generalizadas está basado en una propuesta de Moreno-Díaz y Rubio [Mor78] y modificaciones posteriores [Mor84], actualizada y adaptada a los objetivos del proceso visual. La hipótesis fundamental consiste en considerar la regla de decisión como una transformación, en lugar de considerarla como definitoria o determinante de un sistema.

Haremos el desarrollo teórico para una dimensión, ya que los resultados son fácilmente extensibles a más dimensiones.

Una **Convolución Generalizada** se basa en una única regla de transformación que se aplica a todos o parte de los valores  $f_i$  contenidos en los lugares  $i$ .

$$F_{i+k} = R_i \{ \forall_i f_{i+k} \} \quad \forall_k \quad \text{Ec. 2}$$

A partir de la regla del elemento en la posición cero, la regla para el elemento en

---

la posición  $k$  se obtiene sustituyendo cada  $f_i$  en la regla por  $f_{i-k}$

$$F_k = R_o\{\forall_i f_{i+k}\} \quad \text{Ec. 3}$$

Una transformación multidimensional en la que existe una convolución generalizada para una (o más) dimensiones, es **invariante** en esas dimensiones.

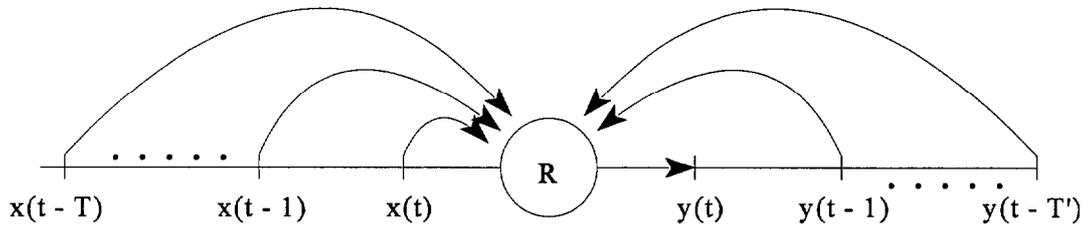
Al imponer restricciones sobre la regla de transformación, podemos obtener distintos tipos de convoluciones generalizadas, como son las algebraicas, las analíticas y las algorítmicas.

### 1.3.1 Convoluciones Algebraicas. Filtros Digitales

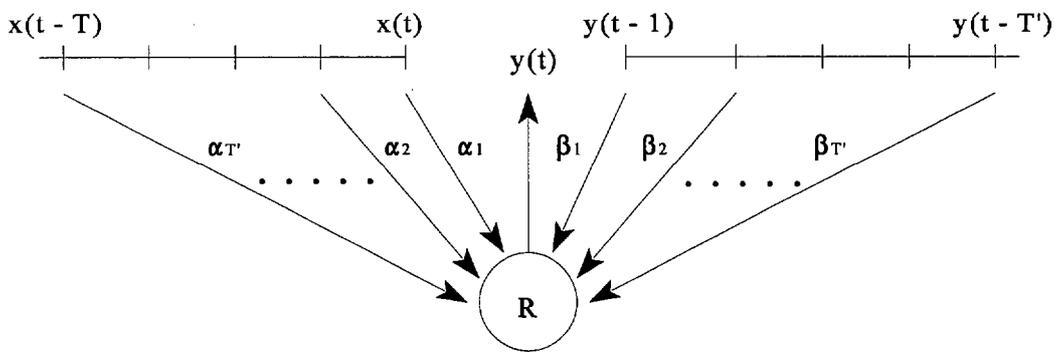
En el caso de señales temporales discretas y permitiendo que la regla de transformación sea una función arbitraria de los valores:

$$x(t), x(t-1), x(t-2), \dots, y(t-1), y(t-2)$$

se generan los filtros digitales arbitrarios:



Si la regla es lineal en la forma de factores de peso o núcleos, se obtienen los filtros digitales lineales, recursivos o no.



Si además de ser la regla  $R_i$  una expresión algebraica y lineal, el sistema es no recursivo, se tiene la convolución convencional. La regla:

$$F_i = \sum_i \alpha_{i,j} f_j \quad \text{EC. 4}$$

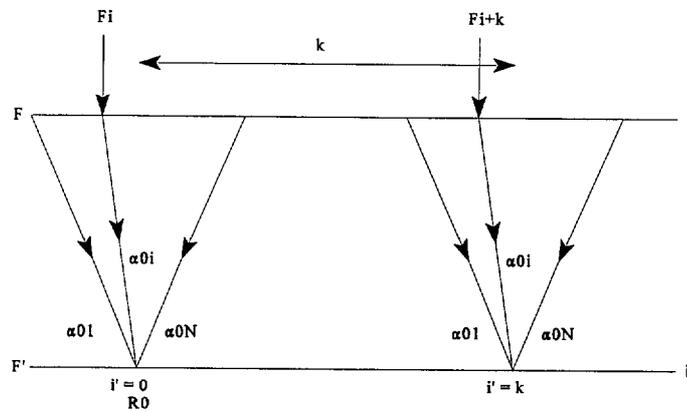
corresponde a una transformación lineal arbitraria. Con la condición de la ecuación (3), resulta la regla  $R_o$

$$R_o \rightarrow \sum_i \alpha_{oi} f_i$$

que con la condición (2):

$$F_k = \sum_i \alpha_{oi} f_{i+k} = \sum_i \alpha_i f_{i+k} = (\alpha_{oi} \alpha_i) \quad \text{Ec. 5}$$

El proceso se ilustra en la siguiente figura.



Los  $\alpha_i$  son los "factores de peso", o el núcleo de la convolución, referidos al origen.

### 1.3.2 Convoluciones Analíticas

Una familia importante de convoluciones generalizadas es la que caracteriza a los sistemas no lineales polinómicos, invariantes en espacio-tiempo, los cuales corresponden a la caracterización de Volterra-Wiener [Wie58]. En esta situación, la regla puede expresarse por una función analítica de los  $f_i$ . Es importante notar que estrictamente la regla  $R$  es un funcional ya que desde el punto de vista del análisis clásico, las variables no son los  $f_i$ , sino los índices, es decir las direcciones de memoria.

Al ser la regla una expresión analítica, puede ser desarrollada en series de potencias, correspondiendo en un caso unidimensional a un desarrollo de  $N$  variables.

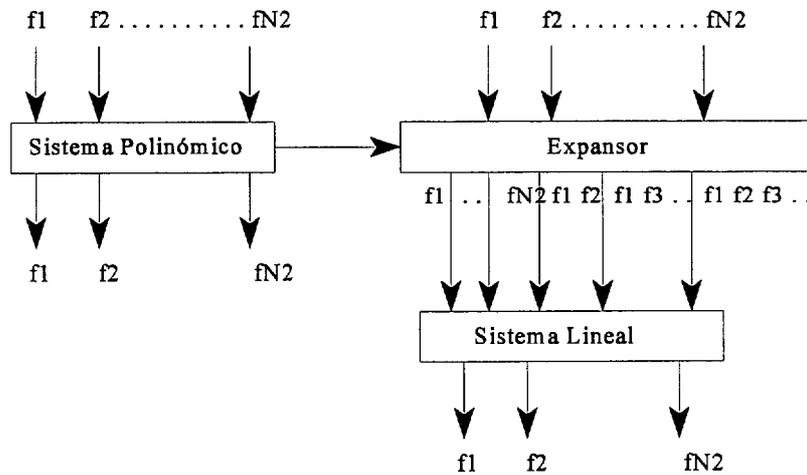
$$R(i,j,\dots,n) = R(0,\dots,0) + \sum_i (\partial R / \partial f_i)_{f_i=0} f_i + \sum_i \sum_j (\partial^2 R / \partial f_i \partial f_j)_{f_i=f_j=0} f_i f_j + \dots$$

siendo:

$$W_i = (\partial R / \partial f_i)_{f_i=0} \quad ; \quad W_{ij} = (\partial^2 R / \partial f_i \partial f_j)_{f_i=f_j=0}$$

donde  $W_i$ ,  $W_{ij}$ , son los núcleos de primer y segundo orden introducidos en la teoría de sistemas no lineales.

Moreno Díaz y Rubio [Mor78] demostraron que un sistema polinómico es equivalente a un expansor del número de funciones de entrada seguido de un sistema lineal, tal como se ilustra en la siguiente figura.



Consideraremos, como ilustración, el caso de un sistema de segundo orden, es decir, un sistema que vendría caracterizado hasta los núcleos de Wiener de orden dos.

Un sistema no lineal polinómico de orden dos comporta los coeficientes  $F_i$ :

$$F_i = \sum_j \alpha_{ij} f_j + \sum_j \sum_k \alpha_{ijk} f_j f_k$$

Introduciendo las nuevas variables funcionales  $\Gamma_{jk} = f_j f_k$  y red denominando los índices tales que, a cada pareja  $\mathbf{jk}$  le hacemos corresponder un nuevo índice  $\mathbf{l}$  ( $\mathbf{l}=1, \dots, \mathbf{N}^2$ ), resulta:

$$\Gamma_l = \Gamma_{jk} = f_j f_k$$

Es decir:

$$F_i = \sum_j \alpha_{ij} f_j + \sum_l \alpha_{il} \Gamma_l$$

un sistema lineal expandido y redundante.

Renombrando las variables, resulta:

$$F_i = \sum_j \alpha_{ij} \Gamma_j$$

donde  $\alpha_{ij}$  es una matriz de  $N^2 \times (N^2 + N^4)$ .

En principio, la solución inversa de la ecuación anterior, con más incógnitas funcionales  $\Gamma_j$  que variables funcionales  $F_i$  lleva a un problema diofantino en funcionales. Aunque ello pueda ser estrictamente el caso, el número de variables funcionales independientes no es  $N^2 + N^4$  ya que, obviamente, por encima del índice  $N^2$ , las nuevas variables funcionales no son independientes de las anteriores.

Las condiciones suficientes para la existencia del sistema inverso serían:

1º) Las variables funcionales presentes en la expansión han de ser independientes, para evitar redundancia. Por ejemplo, si  $f_i$  y  $f_j$  están presentes, no lo puede estar  $f_i f_j$ . Igualmente, si  $f_i$  está presente, no lo podrá estar  $f_i^2$  y viceversa.

2º) Lo anterior reduce a  $N^2$  (número original de grados de libertad) el número de los  $\alpha_{ij}$  que están presentes en la expansión. Además, si  $\alpha'_{ij}$  es la matriz reducida, debe cumplirse  $|\alpha'_{ij}| \neq 0$ .

3º) En general, ningún  $f_i$  debe ser cero. Este es un requisito fácil de conseguir en el preproceso de imágenes, por medio de una traslación del origen de intensidades.

Según esto, si se cumplen los requisitos anteriores, la transformación no lineal inversa existe tal como ha indicado Muñoz [Muñ87].

**Si una representación polinómica de segundo grado cumple 1º, 2º y 3º, existe la transformación inversa, que viene dada por las expansiones y reglas de solución expresadas a continuación :**

En efecto, partimos de la representación

$$F_i = \sum_j \alpha'_{ij} \Gamma_j \quad \text{con } |\alpha'_{ij}| \neq 0$$

computamos

$$\Gamma_i = \sum_j \beta_{ij} F_j \quad \text{con } \beta = \alpha^{-1}$$

Puesto que  $\Gamma_i$  contiene los  $f_k, f_l$  como  $f_k$  o bien  $f_k f_l$  y según las condiciones (1,2,3) resulta:

Si  $f_k$  está presente, así como  $f_k f_l$ , entonces

$\Gamma_k = f_k$  y  $\Gamma_{k'} = f_k f_l$ , de donde  $f_l = \Gamma_{k'}/f_k$ , de forma semejante se procede para otras combinaciones de productos de las funciones de origen.

Como ilustración práctica, consideremos una retina de cinco fotorreceptores ( $f_1, f_2, f_3, f_4, f_5$ ). Una forma sistemática de conseguir un conjunto  $\Gamma$  independiente de orden dos es el siguiente:

$$\Gamma_1 = f_1 ; \Gamma_2 = f_1 f_2 ; \Gamma_3 = f_2 f_3 ; \Gamma_4 = f_3 f_4 ; \Gamma_5 = f_4 f_5$$

luego, la transformación no lineal sería:

$$F_i = \alpha_{i1}f_1 + \alpha_{i2}f_1f_2 + \alpha_{i3}f_2f_3 + \alpha_{i4}f_3f_4 + \alpha_{i5}f_4f_5 \quad (i = 1, \dots, 5)$$

los  $\alpha_{ij}$  pueden ser los coeficientes de cualquier transformada lineal con inversa ( $|\alpha_{ij}| \neq 0$ ), por ejemplo, la de Fourier, Fourier-Bessel o los Momentos.

La transformación inversa será:

$$\Gamma_i = \sum \beta_{ij} F_j \quad \text{con } \beta = \alpha^{-1}$$

y

$$f_1 = \Gamma_1; f_2 = \Gamma_2/f_1; \dots; f_5 = \Gamma_5/f_4$$

La ampliación de estos conceptos a dos dimensiones no es única, y dependerá de la regla de definición. Por ejemplo, podríamos realizar una expansión por etapas, primero por filas y posteriormente por columnas. Dada una imagen de dimensión NXN la expansión será:

<b><u>Por Filas:</u></b>	<b>Si</b> $j = 1$	<b>entonces</b>	$\Gamma'_{ij} = I_{ij}$
		<b>si no</b>	$\Gamma'_{ij} = I_{ij} * I_{i,j-1}$

El resultado de la expansión por filas puede observarse en la siguiente figura.

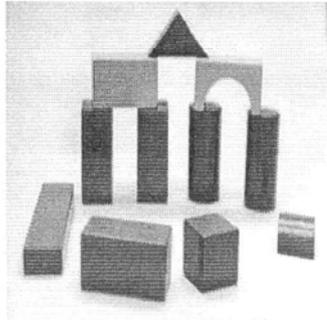


Imagen Original

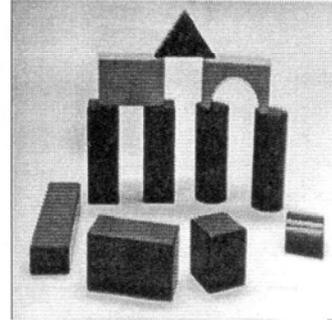


Imagen Expandida por Filas

Por columnas: Si  $i = 1$  entonces  $I''_{ij} = I'_{ij}$

si no  $I''_{ij} = I'_{ij} * I'_{i-1,j}$

de esta forma nos aparecerían los productos cuádruples:

$$I_{ij} * I_{i,j-1} * I_{i-1,j-1} * I_{i-1,j}$$

Esta transformación al ser multiplicativa tiene la ventaja de ser simétrica direccionalmente. En la siguiente figura se muestra el resultado de la expansión por columnas:

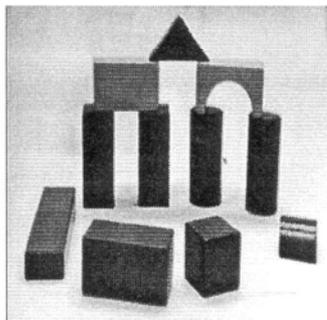


Imagen Expandida por Filas

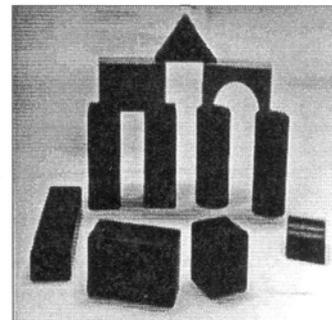


Imagen Expandida por Filas y Columnas

Otra posible definición sería tomar como referencia la diagonal y generalizar el ejemplo unidimensional del siguiente modo:

$$\begin{array}{ll} \text{Si } i=j & \text{ entonces } I'_{ij} = I_{ij} \\ \text{Si } i < j & \text{ " } I'_{ij} = I_{i,j-1} * I_{ij} \\ \text{Si } i > j & \text{ " } I'_{ij} = I_{i-1,j} * I_{ij} \end{array}$$

Observamos que en este caso sólo aparecen los dobles productos. En la siguiente figura se muestra el resultado de esta última expansión:

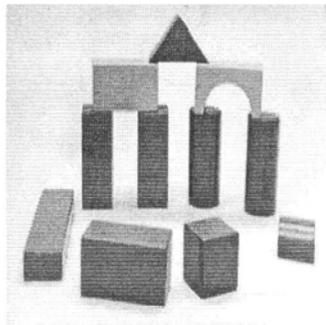


Imagen Original

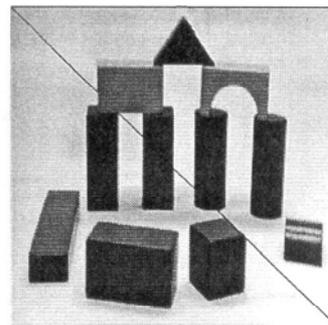


Imagen Expandida

### 1.3.3 Convoluciones Algorítmicas

En distintos trabajos anteriores, desarrollados por miembros de nuestro grupo de investigación, se han desarrollado ejemplos relacionados con esta convolución. Uno de esos ejemplos serían las transformaciones no lineales, cuyo objetivo es la normalización de las intensidades.

Según Muñoz [Muñ87], un método para alcanzar el objetivo anterior es mediante un filtro lineal paso alto, seguido de un procedimiento de escalado para estandarizar el rango de variación de las intensidades.

Una transformación sencilla que proporcione las invarianzas traslacionales y homotéticas se consigue por medio de la expresión local:

$$F'(x,y) = K \frac{I(x,y) - \text{Min}[I(x,y)]}{\text{Max}[I(x,y)] - \text{Min}[I(x,y)]}$$

Las invarianzas requeridas suponen la acción de un filtro no lineal. La naturaleza de la no linealidad viene dada por la homotecia en el dominio de las intensidades. En general, ésta es una propiedad que tiene como expresión el cociente de dos transformaciones lineales, cuyos núcleos cumplen la propiedad:

$$\int_C W(x,y,x',y') dx' dy' = 0$$

donde C es el campo receptivo. En el lenguaje de las estructuras computacionales paralelas, la expresión anterior requiere que el campo receptivo C de cada unidad computacional tenga partes excitadoras e inhibidoras que se cancelen.

Esto permitió establecer y demostrar a Muñoz [Muñ87], el siguiente teorema:

**Teorema.** Si  $T_1$  y  $T_2$  son dos transformaciones lineales, caracterizadas por los núcleos  $W_1$  y  $W_2$ , tales que:

$$\int_c W_1(x,y;x',y') dx' dy' = 0 \quad \text{y} \quad \int_c W_2(x,y;x',y') dx' dy' = 0$$

la transformación no lineal  $T$ , definida por:

$$T(f(x,y)) = \frac{T_1(f(x,y))}{T_2(f(x,y))} = f'(x,y)$$

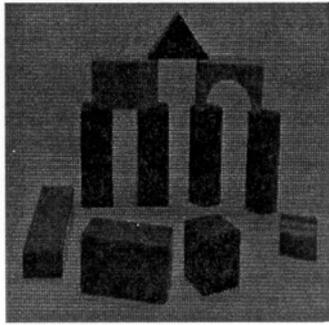
es tal que  $f'(x,y)$  es invariante frente a traslaciones y homotecias en intensidades.

A modo de ejemplo, presentamos una demostración práctica del teorema anterior. Convolucionamos una imagen con una máscara tipo Laplaciana. En la siguiente figura, presentamos:

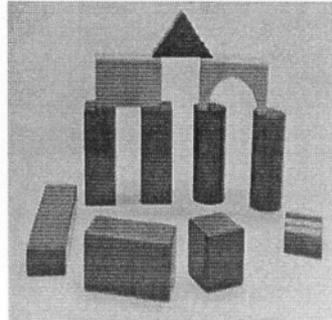
a) Imagen Original

b) Imagen resultante de aplicar a la imagen original una homotecia y una traslación.

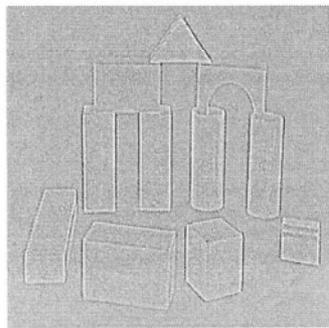
c) y d) Resultados de procesar las imágenes anteriores según la transformación no lineal, que corresponde a una Convolución Generalizada.



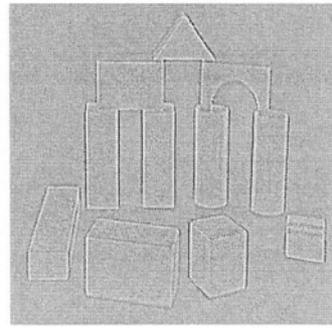
a)



b)



c)



d)

Se puede observar, y se ha constatado numéricamente, que las imágenes c) y d) son iguales.

## 1.4 FILTROS DE NEWTON.

Según un estudio realizado por Moreno Díaz [Mor93], un Filtro de Newton es el resultado en cascada de unos procesos  $S(+)$  y  $R(-)$  expresados por  $N(A_m, D_n, \dots)$  donde "m" es el orden del proceso aditivo (número de filas suma) y "n" es el orden de la diferencia (número de filas diferencia).

Por ejemplo, el filtro  $N(A_2, D_2)$  se obtiene por la aplicación de las reglas:

	1
A1	1 1
A2	1 2 1
D1	1 1 -1 -1
D2	1 0 -2 0 1

Una clase de transformaciones, son las denominadas transformaciones globales generadas a partir de Filtros de Newton de anchura fija.

Se define la Anchura de un Filtro de Newton,  $N(A_n, D_m)$ , como el número total de componentes del filtro:  $n+m+1$ .

Dado un número natural  $L \geq 1$ , existen  $L$  Filtros de Newton.

Veamos el conjunto de Filtros de Newton de una cierta anchura determinada,  $L$ , como un conjunto de vectores  $L$ -dimensionales:

$$\begin{aligned}
 N1 &= N(A_n) \\
 N2 &= N(A_{n-1}, D1) \\
 &\dots \\
 N_{L-1} &= N(A1, D_{n-1}) \\
 NL &= N(D_n)
 \end{aligned}$$

donde  $L=n+1$ . Las componentes de esos vectores se puede expresar como:

$$N1 = \begin{bmatrix} n \\ 0 \end{bmatrix} \begin{bmatrix} n \\ 1 \end{bmatrix} \begin{bmatrix} n \\ 2 \end{bmatrix} \dots \begin{bmatrix} n \\ n \end{bmatrix}$$

para el primer vector. Para el segundo, la componente k-ésima vendrá dada por:

$$\begin{bmatrix} n-1 \\ k \end{bmatrix} \dots \begin{bmatrix} n-1 \\ k-1 \end{bmatrix}$$

con  $k=0\dots n-1$ , obtenemos los primeros  $n$  elementos. El elemento  $n+1$  es  $(-1)^{2+1}=-1$ .

Convenimos que:

$$\begin{bmatrix} a \\ b \end{bmatrix} = 0$$

si  $b < 0$  o si  $b > a$ . De esta forma, para el tercer vector:

$$\begin{bmatrix} n-2 \\ k \end{bmatrix} - 2 \begin{bmatrix} n-2 \\ k-1 \end{bmatrix} + \begin{bmatrix} n-2 \\ k-2 \end{bmatrix}$$

donde  $k=0\dots n-1$  con las convenciones anteriores y el último elemento  $(-1)^{3+1}=1$ , y la  $k$ -ésima componente del vector  $i$  queda:

$$\sum (-1)^j \begin{bmatrix} n \\ j \end{bmatrix} \begin{bmatrix} L-i \\ k-j \end{bmatrix}$$

con  $k=0\dots n-1$ , donde la última componente del vector es  $(-1)^{i+1}$  e  $i \neq L$ . Para el último vector tendremos:

$$(-1)^k \begin{bmatrix} n \\ k \end{bmatrix}$$

con  $k=0\dots n$ .

Ejemplos:

1.- Para  $L=2$  formamos los filtros  $N(A1)$  y  $N(D1)$ :

$$N1 = N(A1) = (1,1)$$

$$N2 = N(D1) = (1,-1)$$

2.- Para  $L=3$ , tenemos:

$$N1 = N(A2) = (1,2,1)$$

$$N2 = N(A1,D1) = (1,0,-1)$$

$$N3 = N(D2) = (1,-2,1)$$

3.- Para  $L=4$  los vectores correspondientes a los filtros serán:

$$N1 = N(A3) = (1,3,3,1)$$

$$N2 = N(A2,D1) = (1,1,-1,-1)$$

$$N3 = N(A1,D2) = (1,-1,-1,1)$$

$$N4 = N(D3) = (1,-3,3,-1)$$

Con los ejemplos anteriores, se puede formar las matrices cuyas filas son los correspondientes vectores de pesos de los Filtros de Newton. Dichas matrices son cuadradas y las llamaremos Matrices de Newton. De esta forma, siguiendo con los ejemplo, la Matriz de Newton de orden 2 es:

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

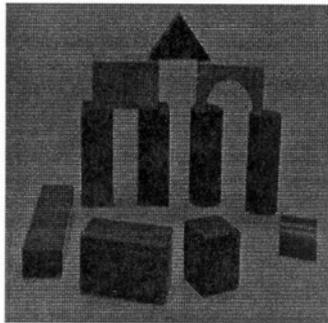
y la de orden 3:

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

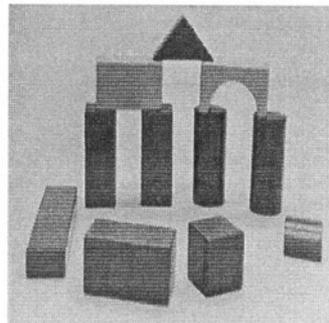
y así sucesivamente.

Dependiendo de la combinación de  $S(+)$  y  $R(-)$  que se aplique y existiendo al menos una de esta última, podemos conseguir fácilmente núcleos que cumplan las características presentadas en el teorema del apartado anterior. Por tanto, en este trabajo, hemos obtenido distintos núcleos a partir de los Filtros de Newton para aplicarlos al teorema del apartado anterior. En la siguiente figura se muestra una demostración práctica de dicho teorema con los núcleos obtenidos en esta sección. Presentamos:

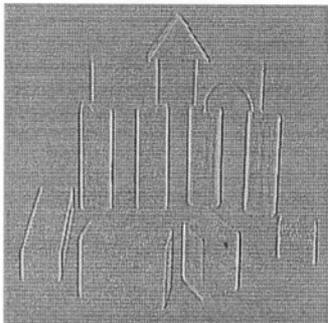
- a) Imagen Original
- b) Imagen resultante de aplicar a la imagen original una homotecia y una traslación.
- c) y d) Resultado de procesar las imágenes anteriores según la transformación no lineal, que corresponde a una convolución generalizada.



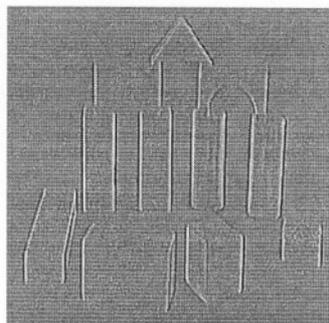
a)



b)



c)



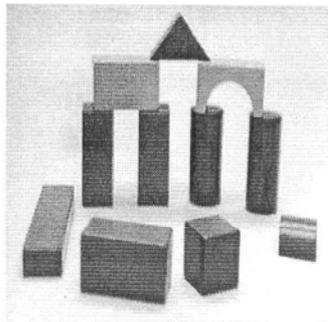
d)

Nuevamente, comprobamos que las imágenes c) y d) son iguales.

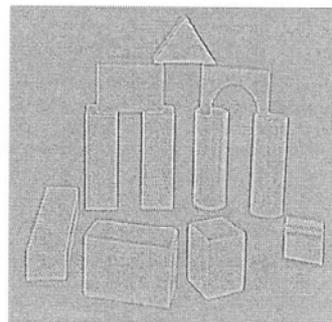
Como caso práctico, estos núcleos obtenidos a partir de los Filtros de Newton lo podemos aplicar para una operación de filtrado de una imagen para extraer los bordes de los objetos. En la siguiente figura presentamos:

a) Imagen Original

b) Imagen resultante de aplicar la operación especificada anteriormente.



a)



b)



a)



b)

## 1.5 TRANSFORMACIONES EN CAMPOS DE DATOS

A partir del estudio de las Transformaciones de Campo Receptivo Variable iniciado en 1.987 por Candela [Can87], han sido desarrollados una serie de teoremas, lemas y demostraciones (Bolívar[Bol89]), relacionados con conceptos de Complitud en transformaciones que combinan particiones con funcionales.

Un concepto básico en Transformaciones de Imágenes y en Visión Artificial es el de **Descripción Completa**. Una descripción es completa en su entorno visual, si la descripción contiene todos los datos y propiedades de éstas necesarias para cubrir unos objetivos. A nivel analítico, los objetivos que se establecen consisten en admitir que una descripción es completa cuando ella puede recuperarse del entorno visual.

Los conceptos de complitud de una transformación con relación al campo receptivo y a la función fueron introducidos por Candela en 1.987 [Can87]. Según dicho autor, 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 donde se toman éstos), para mantener una constancia de los grados de libertad impuestos por los objetivos del sistema.

Dicha constancia permitió establecer una especie de principio de conservación que viene determinado por los funcionales computados y por los campos receptivos que, además, reflejan una situación de dualidad, en el sentido de que la complitud requiere aumentar el número de ellas si el otro disminuye y viceversa.

Por otra parte, los conceptos de Campos de Datos y Particiones fueron introducidos por Moreno-Díaz y Bolívar en 1.989 [Bol89] como resultado de una generalización de la representación de imágenes en una, dos, o más dimensiones.

Un Campo de Datos Unidimensional de longitud  $N$  y resolución  $R$  es un conjunto ordenado de  $N$  lugares  $i$ . ( $i=1\dots N$ ), tal que a cada lugar puede asignársele un número  $l_i$  (real, complejo o de otra índole) con resolución  $R$ . Los lugares  $i$  deben entenderse como "hojas en archivo" mientras que  $R$  debe entenderse como el máximo número de huecos binarios necesarios para alojar los posibles datos  $l_i$ . El número de grados de libertad "extendido" de un campo de datos es, en general:

$$N * R$$

Dado un campo de datos  $D(N,R)$  una transformación de  $D$  es una regla  $\&$  que permite pasar de  $D$  a  $D'$

$$D(N,R) \text{ ----- } \& \text{ -----} \rightarrow D'(N',R')$$

Según dichos autores, y siguiendo a Moreno Díaz y Candela [Can87] una transformación  $\&$  proporciona una representación completa si y solo si existe la transformación inversa  $\&'$ , tal que

$$D'(N',R') \text{ ----- } \&' \text{ -----} \rightarrow D(N,R)$$

Dentro de la Teoría de Sistemas, una transformación corresponde a la acción de un sistema sobre el espacio de las entradas y quizás el de las salidas, para generar un nuevo

"slide" del espacio de salidas.

En tal estructura, se puede definir y visualizar lo que se entiende por "*Campo Receptivo*".

Según dichos autores, el sistema a considerar transforma  $f(x)$  en  $F(x')$ , donde  $x$  y  $x'$  son dominios que pueden coincidir, a través de una transformación arbitraria:

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

Se admite que existe un "dominio máximo" en la definición de  $f(x)$ , con respecto a la variable independiente  $x$ . Por ejemplo, si  $x=t$  (el tiempo),  $x$  va de  $-\infty$  (ó 0) a  $t$  (instante actual), igual que  $x'$ , analiza todos los valores posibles de  $f(t)$  antes de  $t$  para concretar el valor  $F(t)$ . Si la variable es espacial, el anterior "principio" de causalidad puede romperse, es decir,  $x'$  puede ser tal que  $x' > x$ .

Primeramente, se habla 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  $x_2$ ). Esto 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 estática, 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 ( $C_1$ ) como condición suficiente:

*$C_1$ : Establecido el dominio total de  $f(x)$ , una transformación  $F(x')$  es una Transformación de Campo Receptivo Variable (TCRV) 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 de Campo Receptivo Constante (TCRC).

Siguiendo con lo anterior, se puede llegar a una condición más restrictiva para las TCRC. Para ello, se define el concepto de "tamaño" del campo receptivo de una transformación  $F(x')$ . Admitamos 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 TCRC siguiente:

*$C_2$ : Dada una distancia  $d(x_i, x_j)$  medida en  $x$ , una transformación  $F(x')$  es TCRC si  $F(x')$  decide sobre el dominio comprendido por  $(x_i < x < x_j)$  tal que  $d(x_i, x_j) = cte$  para todo  $x'$ .*

Así, la generalización de lo anterior resulta en cierto modo trivial: 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  $(x_1, x_2, x_3, x_4, x_5, x_6, \dots, x_l, x_m)$ , 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, se llega a la condición  $C_3$ :

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

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

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  $f(x)$ ,  $N$ .

Unas TCR triviales son las  $TCR \in \{TRC\}$ , es decir, por ejemplo, la transformación base de identidad.

Para la situación de complitud 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:

*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.

# **CAPÍTULO 2**

**TRANSFORMACIONES FOVEALES**

**GENERALIZADAS**

## 2.1 ANTECEDENTES

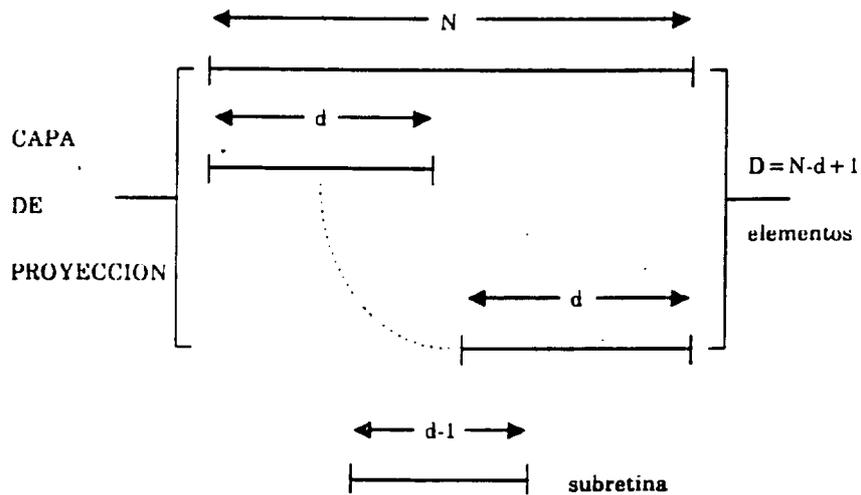
Como se indicó en el Capítulo 1, un dominio puede estar caracterizado por distancias  $d_{mn}$ . Consideremos los dominios tales que cada uno es caracterizable por una única distancia. Es decir, sea  $\{D\}$  un conjunto de  $M$  dominios  $D_i$ , con distancias  $d_i$ . Estas distancias pueden ordenarse de menor a mayor y en caso de igualdad, por la proximidad al origen del menor elemento de  $D$ . Llamaremos  $C_r$  a este criterio de ordenación. Según esto, una Transformación de Resolución Progresiva (TRP) es aquella en la que :

- a)  $\{D\}$  es independiente.
- b)  $\{D\}$  está ordenado según un cierto criterio  $C_r$ .
- c)  $\{D\}$  es completo.

Por ejemplo, la transformación de Haar es una TRP, y las propiedades relevantes de dicha transformación se deben, en esencia, a que es TRP y no la forma de las funciones (o núcleos) empleados.

Una clase particular de TRP's que están inspiradas en el "moving average", son las transformaciones denominadas "foveales", ya que para una retina, proporcionan la máxima resolución "directa" en la parte central o "fóvea".

La idea de TRP foveales parte del hecho de que en una partición completa global de una retina se puede 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. Para una dimensión dada,  $d$ , existen:



$D=N-d+1$  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:

*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.*

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

Seguidamente, 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:

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

Ya 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 considerarse como una subretina, al que nuevamente se le puede aplicar el procedimiento anterior.

Una TRP con las propiedades anteriores es lo que entendemos como Transformación Foveal.

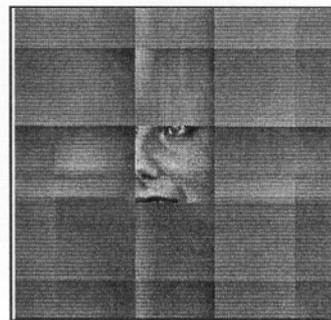
Como ilustración, consideremos una retina de 256 grados de libertad ( $N=256$ ), donde a la imagen original se le ha aplicado distintas transformaciones foveales.



Imagen Original



T.F.



T.F.

En este punto, proponemos una nueva transformación. de tal forma que la fovea se pueda dirigir a cualquier punto del Campo de Datos. Llamaremos a esta transformación, Transformación Foveal Dirigida (TFD).

## 2.2 TRANSFORMACIÓN FOVEAL DIRIGIDA

Al igual que la transformación foveal, la TFD está inspirada en el "moving average", con la diferencia que los dominios unitarios se pueden colocar en cualquier zona de la retina. Veamos un procedimiento para obtener la TFD.

Vamos a definir primeramente, 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 cada cardinal es:

$$d = (\text{ent}) \frac{N}{K}$$

donde N representa los grados de libertad.

Para la primera capa se tiene una dimensión (N-d+1)

$$d_1 = N - \frac{N}{K} + 1$$

e iterativamente, para la capa i

$$d_i = d_{i-1} - \frac{N}{K}$$

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

$$N = (K-1) \frac{N}{K}$$

Analicemos un ejemplo numérico. Consideremos una retina de  $N=128$ , con 4 capas de proyección ( $K=4$ ). Se tiene que el cardinal de cada capa es:

$$d = \frac{128}{4} = 32$$

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

La dimensión de los elementos de las capas son:

$$d_1 = 128 - 32 + 1 = 97$$

$$d_2 = 97 - 32 = 65$$

$$d_3 = 65 - 32 = 33$$

$$d_4 = 33 - 32 = 1$$

y las de las subretinas:

$$\text{SUB}_1 \quad 97 - 1 = 96$$

$$\text{SUB}_2 \quad 65 - 1 = 64$$

$$\text{SUB}_3 \quad 33 - 1 = 32$$

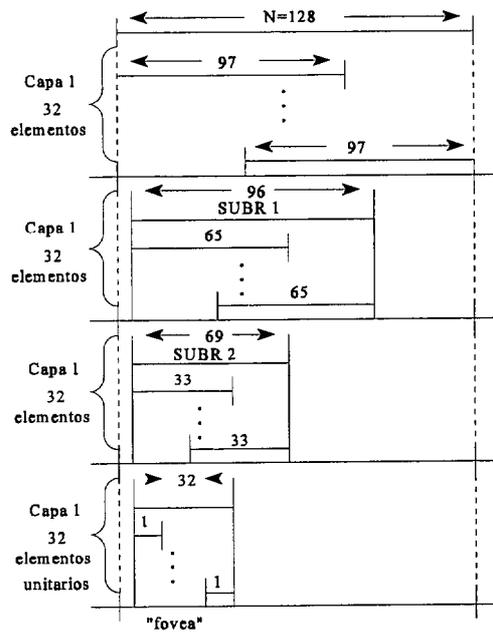
Si, por ejemplo, dirigimos la fovea a la zona izquierda de la retina, la posición de las subretinas serían:

$$\text{SUB}_1 \quad \text{--> posición 48}$$

$$\text{SUB}_2 \quad \text{--> posición 32}$$

SUB<sub>3</sub> --> posición 16

Gráficamente. observamos las 4 capas de proyección con la fóvea dirigida a la zona izquierda.



La restricción de la posición de las subretinas (especificado por I y J) viene dada por la expresión:

$$N/2 * K \leq I, J \leq N - N/2 * K$$

siendo N los grados de libertad y K el número de capas de proyección.

A continuación presentamos distintos ejemplos de la TFD aplicada a una misma imagen, donde hemos variado tanto la dirección de la fovea como el número de capas. Al

pie de cada imagen se indica el número de capas de la transformación (K) y la posición en fila y columna (I,J) donde se dirigió la fovea. El grado de libertad elegido en cada una de ellas es de 256 (N=256).



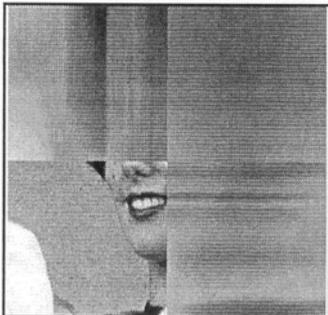
Imagen Original



K=2;I=64,J=64



K=2;I=64,J=128



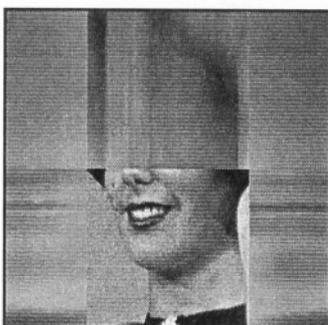
K=2;I=64,J=192



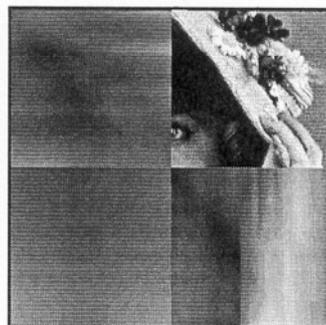
K=2;I=128,J=64



K=2;I=128,J=128



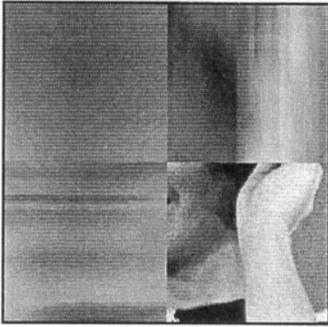
K=2;I=128,J=192



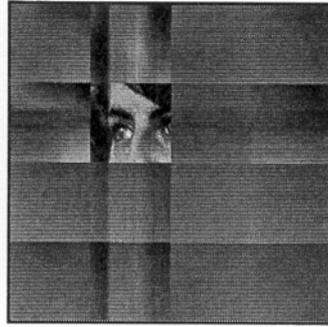
K=2;I=192,J=64



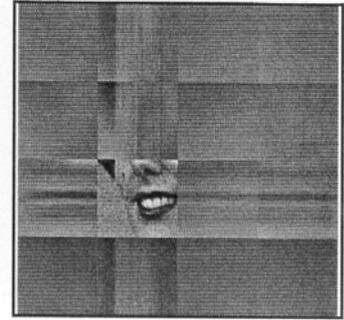
K=2;I=192,J=128



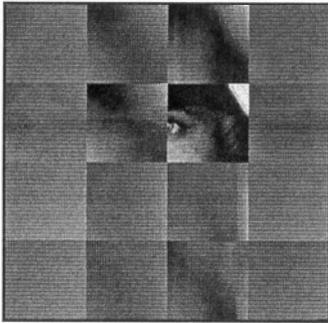
$K=2;I=192,J=192$



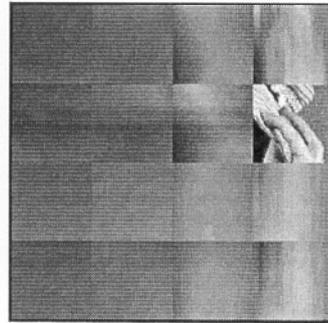
$K=4;I=96,J=96$



$K=4;I=96;J=160$



$K=4;I=160,J=96$



$K=4;I=224,J=96$

## **2.3 APLICACIONES DE LA TFD**

### **2.3.1 TFD como método Criptográfico**

La seguridad de la información surge de la necesidad de privatizar la transmisión de mensajes. Esta necesidad es tan antigua como la propia civilización. Por ejemplo, los antiguos Espartanos ya cifraban sus mensajes militares. Los primeros canales de comunicación eran muy simples y la seguridad se conseguía utilizando mensajeros de confianza. Por tanto, la seguridad dependía de la fiabilidad de los mensajeros.

Con el nacimiento de los sistemas computadores y la incorporación de las redes de computadores, ha cambiado drásticamente el rango de protección utilizado. En las primeras etapas del desarrollo de los sistemas computadores, la seguridad física utilizando un departamento policial, era suficiente para proporcionar seguridad. Pero esta comienza a ser insuficiente e inflexible después del nacimiento de los sistemas computadores de tiempo compartido con distintos terminales geográficamente dispersos.

La mayoría de la protección utilizada, tanto en sistemas computadores de tiempo compartido como en las redes de computadores, estaba estrictamente relacionada con la protección de los canales de comunicación. Debido a las características propias de los canales, tenemos un medio de comunicación que es fácilmente accesible y, por tanto, poco fiable. La única forma de proteger los canales de comunicación es aplicando la criptografía.

En general, la seguridad de los sistemas computadores de tiempo compartido y las redes de computadores consiste de tres componentes:

- a) Seguridad en los Centros de Computación.
- b) Seguridad de los Terminales.

### c) Seguridad de los Canales de Comunicación.

La protección de los centros de computación necesitan diferentes contramedidas de seguridad. En primer lugar, los centros deben protegerse de cualquier posible desastre natural, inundaciones, fuego, etc. También deben protegerse contra actividades exteriores, ataque terroristas, escuchas ocultas, etc. Todas estas contramedidas puede verse como seguridad externa. Sin embargo, la seguridad interna incluye medidas de protección utilizadas dentro del sistema computador (mecanismos de control de acceso, mecanismos de identificación etc.) y otras aplicadas fuera del computador como la protección física de los componentes del computador. La experiencia muestra que los terminales son los componentes más vulnerable de cualquier sistema computador. La mayoría de los intentos de acceso ilegales son realizados a través de los terminales.

La criptografía se utiliza para proteger la información a cualquier acceso ilegal y donde son insuficientes otras medidas de seguridad. Por tanto puede aplicarse para proteger los canales de comunicación, las bases de datos físicas, etc.

La operación base de la criptografía es el cifrado. Es una computación especial que opera sobre el mensaje convirtiéndolo en una representación que sólo tiene significado para el receptor de dicho mensaje.

Los algoritmos de cifrado se eligen de una familia de transformaciones invertibles conocidas como sistema general o criptosistema. El parámetro que selecciona la transformación particular de la familia se llama clave de cifrado o clave. El criptosistema puede ser un conjunto de instrucciones, un trozo de hardware o un programa sobre el cual se selecciona la clave de cifrado.

Formalmente, un criptosistema es una familia de parámetros de transformaciones

invertibles,

$$E_K; K \in \mathbf{K}$$

donde  $\mathbf{K}$  es el espacio de claves, el cual es de longitud finita. Sea  $\mathbf{M}$  el espacio de mensajes y  $\mathbf{C}$  el espacio de texto cifrado, entonces el sistema debe tener las siguientes propiedades:

\* Algoritmo de Cifrado:

$$E_K: \mathbf{M} \rightarrow \mathbf{C}$$

para cualquier clave de cifrado fija  $K \in \mathbf{K}$ , es una transformación invertible del espacio de mensajes en el espacio del criptograma, por ejemplo,  $E_K(M) = C$ , donde  $M \in \mathbf{M}$  y  $C \in \mathbf{C}$ ;

\* Existe un algoritmo inverso  $E_K^{-1} = D_K$  llamado Algoritmo de Descifrado:

$$D_K: \mathbf{C} \rightarrow \mathbf{M}$$

tal que  $D_K(C) = D_K[E_K(M)] = M$ ;

\* las claves son únicas:

$$E_{K_1}(M) \neq E_{K_2}(M) \text{ si } K_1 \neq K_2$$

La criptografía se ocupa del diseño y análisis de sistemas que suministran comunicaciones seguras o resisten al llamado criptoanálisis. Un sistema se dice que está comprometido vía el criptoanálisis si es posible recuperar el mensaje original (texto puro) del texto cifrado (texto puro cifrado) sin conocer la clave utilizada en el algoritmo de cifrado. El criptoanálisis es una faceta altamente especializada de la aplicación de varias disciplinas matemáticas como la teoría de la probabilidad, teoría de números, estadística, álgebra, etc. El criptoanalista experto debe conocer muy bien estos campos. También debe utilizar información secundaria sobre el sistema tales como la naturaleza de sus algoritmos, el lenguaje de comunicación, el contexto del mensaje y las propiedades estadísticas del lenguaje del texto puro (por ejemplo, su redundancia).

El criptoanálisis es un problema de identificación del sistema y el objetivo de la criptografía es crear sistemas que sean difícil de identificar. Las herramientas típicas de un oponente las podemos resumir como:

- a) Determinar el contenido del mensaje  $M$ .
- b) Alterar el mensaje  $M$  en  $M'$  y tener  $M'$  aceptado por el receptor como un mensaje del emisor de  $M$ .
- c) Iniciar una comunicación con el receptor como un emisor autorizado.

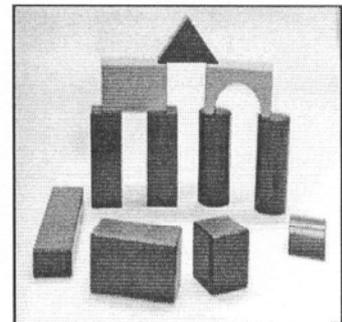
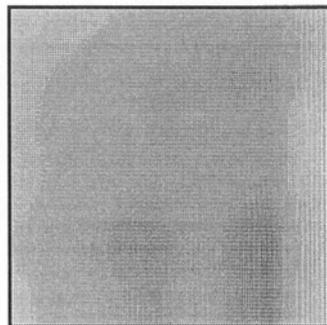
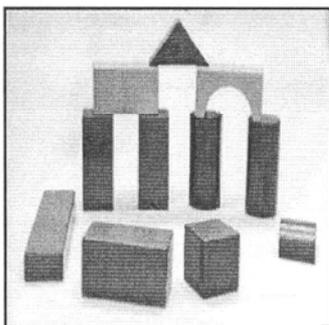
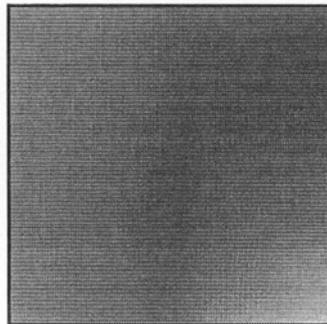
En este punto introducimos una propuesta sobre la utilización de la Transformación Foveal Dirigida como método criptográfico de imágenes. Proponemos utilizar la TFD para criptografiar las imágenes utilizando para ello una clave triple correspondiendo al número de capas de la transformación y la posición  $I, J$  donde está dirigida la fóvea.

A continuación, presentamos varios ejemplos de imágenes a la que le hemos aplicado una TFD. En los resultados podemos observar:

1.- A la izquierda, la Imagen Original.

2.- En el centro, las imágenes cifradas utilizando la Transformación Foveal Dirigida con una clave determinada.

3.- A la derecha, la recuperación de las imágenes a partir de la TFD Inversa, utilizando la misma clave que en el punto anterior.



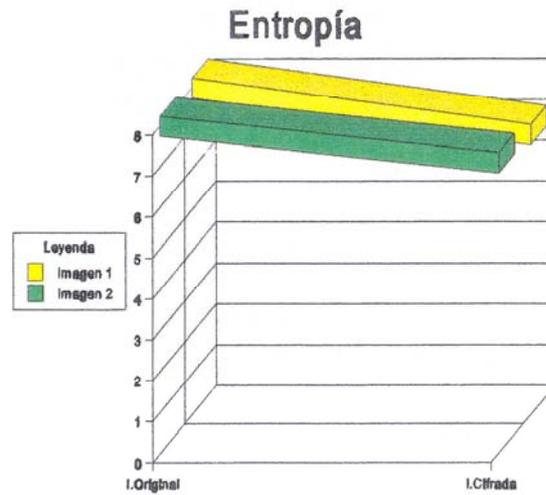
Como podemos observar, las imágenes resultantes de realizar el cifrado son inteligibles al observador. Para llevar a cabo una corta discusión sobre el método criptográfico, nos referiremos a la base fundamental de la teoría de la información. Claude Shannon, padre de esta disciplina, publicaba en 1949 un trabajo sobre Teoría de la Información [Sha49] el cual ha sido el fundamento teórico de la criptografía moderna. Las herramientas principales de la comunicación segura a través de un canal son los códigos y el cifrado. Un código es un "diccionario" predeterminado fijo donde para cada uno de los mensajes válidos, existe un mensaje codificado equivalente, llamado palabra clave. La teoría de la codificación direcciona por simisma el problema del "canal ruidoso", seleccionando un código en particular, si un mensaje  $M$  se distorciona a  $M'$  durante la transmisión, este error puede detectarse y corregirse. Por otro lado, los cifrados son métodos más universal de transformación de mensajes en un formato cuyo significado no es aparente.

Se define la Cantidad de Información en un mensaje como el número de bits por termino medio que se necesitan para codificar óptimamente todos los posibles mensajes. Una codificación óptima es aquella que minimiza el número de bits a transmitir sobre un canal. Más concretamente, la cantidad de información en un mensaje es medido por la Entropía del mensaje. Si  $M_1, M_2, \dots, M_n$  son los  $n$  posibles mensajes y  $P(M_1), P(M_2), \dots, P(M_n)$  son sus probabilidades respectivas de ocurrencias, entonces la entropía de un mensaje se define como:

$$H(M) = \sum_M P(M_i) \log_2 \frac{1}{P(M_i)}$$

La entropía de un mensaje,  $H(M)$ , también mide el número de bits de información que deben ser adquiridos para recuperar un mensaje distorsionado por un canal ruidoso o al realizarle un cifrado.

A continuación presentamos dos gráficas referentes a la entropía de las imágenes anteriores y su correspondientes imágenes cifradas.



Finalmente, proponemos como trabajo futuro, la comparación detallada del método propuesto con distintos métodos criptográficos existentes.

### 2.3.2 TFD como Mecanismo de Atención.

Otro campo de aplicabilidad de la TFD es utilizarla para "captar la atención" del observador sobre alguna zona de la imagen. Se pretende resaltar algún objeto de la imagen de tal forma que el resto de la información quede difuminada (sin ninguna información aparente). Con la TFD, dirigimos la fovea sobre el objeto a resaltar y aplicamos la transformación consiguiendo el objetivo marcado. Recuérdese que aunque la información parafoveal quede "distorsionada", con la TFD inversa recuperamos la imagen original en su totalidad.

El método consistirá en lo siguiente. En primer lugar, calculamos el centro de gravedad del objeto, que corresponderá al punto de atención o centro de la fovea. Para ello, utilizaremos como herramienta los Momentos, es decir, un conjunto de valores obtenibles de forma sistemática a partir de una función. Los momentos en el dominio discreto, que será nuestro caso, se definen como:

$$M_{pq} = \sum_X \sum_Y X^p Y^q f(x,y)$$

Una de las propiedades de los momentos de bajo orden es que nos dan las coordenadas del centro de gravedad del objeto, así que estas pueden ser calculadas mediante las expresiones:

$$X_G = M_{10} / M_{00} \quad Y_G = M_{01} / M_{00}$$

Por tanto, la fovea la centramos en el punto  $(X_G, Y_G)$ .

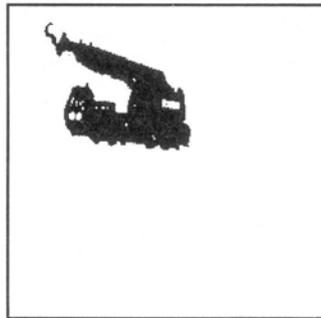
Un vez determinado el centro de la fovea, pasamos a calcular el tamaño de la

misma. Para ello calculamos el punto más alejado de la forma con respecto al centro de gravedad calculado y con este dato tendremos el límite exterior de la fóvea.

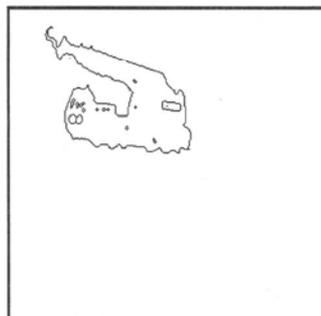
Para encontrar el punto más alejado del centro de gravedad, calculamos el contorno del objeto siguiendo, por ejemplo, la técnica del Código de Cadena, que es la codificación de los contornos en base a las coordenadas del primer punto del contorno  $(X_0, Y_0)$  y de las posiciones relativas de cada punto respecto al anterior del contorno. Una vez obtenido el contorno, se localiza el punto más alejado de este  $(X_i, Y_i)$  con respecto al centro de gravedad de la pieza  $(X_G, Y_G)$ .

A continuación presentamos un ejemplo práctico para ilustrar el proceso seguido.

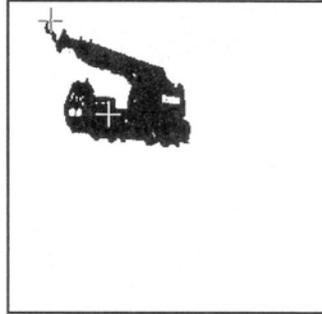
- 1) En primer lugar, seleccionamos una imagen-ejemplo.



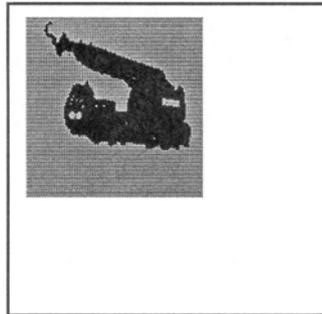
- 2) Calculamos el contorno de la pieza.



3) Localizamos el centro de gravedad y el punto más alejado, reflejado por una cruz en la figura:



4) Por último, presentamos en la siguiente figura, y en tono sombreado, la localización y tamaño de la fovea, calculada automáticamente.



---

## 2.4 INCIDENCIA EN TEORÍA RETINAL

La Teoría de la Visión persigue fundamentalmente construir un marco teórico y práctico para la explicación de la función visual en seres vivos y sus posibles contrapartidas artificiales. En este sentido cae dentro de la Cibernética, entendida esta en el sentido original que le imprimió Norbert Wiener [Wie59]: ciencia que trata de los fenómenos de comunicación, control y proceso de información en máquinas y seres vivos. El sentido de la vista es, en muchas especies, el que procesa en bruto y aporta una mayor cantidad de información con significado sobre el entorno del individuo. Siendo por tanto el que sirve en mayor medida de guía de conducta. La habilidad para discriminar parámetros o descriptores importantes y localizar en qué parte de la imagen se encuentran es esencial para un sistema visual, tanto si es natural como si es artificial.

Como queda desarrollado anteriormente, las Transformaciones Foveales Dirigidas, contempladas desde este punto de vista cibernético, se presentan como un primer paso en la modelización de los mecanismos de atención que rigen gran parte del proceso visual de los vertebrados superiores. El núcleo de la TFD es una variación del "moving average" que concentra la mayor resolución en una zona determinada de la imagen a la que se llama "fóvea" por razones obvias del paralelismo con el ojo. Originalmente, esta zona de mayor resolución se expandía a partir del centro geométrico de la imagen. En las TFD's, los dominios sobre los que la mayor resolución es guardada se pueden emplazar alrededor de cualquier punto de la imagen, a la vez que es posible elegir el tamaño de dicha zona foveal. En la zona parafoveal, es posible definir transformadas en las que el mantenimiento de resolución no sea un factor crítico. De esta manera, se entiende el papel de definir en zonas distintas de la imagen original estas transformaciones como el de una *atractor de interés*. Así, las zonas parafoveales extraerían información sobre un evento interesante, pero con pérdida de resolución, y una vez discriminado el evento, la fóvea se dirigiría y centraría en él para llevar a cabo un proceso mucho más detallado de ese fragmento de la imagen. La

idea biológica subyacente es una cierta economía de cómputo: ningún sistema visual conocido mantiene foveas grandes o que supongan un área mayor del 10% de la retina total [Lei91][Tru65][Bon69] ya que el volumen de información a procesar y transmitir crecería dramáticamente, con lo que se necesitarían nervios ópticos más complejos, más gruesos y muchas más conexiones neuronales en estadios intermedios del camino visual (dejando aparte otras consideraciones morfológicas que se escapan del objeto de esta discusión). Así pues, el objetivo natural sería optimizar el volumen de información detallada a transmitir, optimización que tiene también su importancia en el caso de querer implementar eficientes mimesis artificiales de dichos sistemas y que representaría una reducción de complejidad computacional (entendida como número de operaciones sobre datos) de hasta un 70%.

Otro aspecto interesante es el de la complitud. Utilizando una TFD, en principio, no se exigiría complitud a todo el conjunto de descriptores, sino sólo aquellos que actúan sobre la fovea. El grado de complitud de una transformada nos da idea (y la posibilidad) de la recuperabilidad posterior de los datos originales a partir de los descriptores de la imagen y en los sistemas visuales naturales es independientes, en principio, del grado de resolución del sistema (el ojo humano no discrimina a nivel fotorreceptor sino a nivel textural [Kol90][Jun73], que implica la integración de la información procedente de un conjunto de fotorreceptores).

# **CAPÍTULO 3**

## **REPRESENTACIONES DE IMÁGENES A PARTIR DE LAS TRANSFORMACIONES DE CAMPOS RECEPTIVOS**

### 3.1 CAMPOS RECEPTIVOS Y FUNCIONALES

En el capítulo 1, al presentar la codificación retinal, definimos el concepto de campo receptivo desde el punto de vista *natural* como:

**El Campo Receptivo de una célula en una capa determinada es el conjunto de células de la capa anterior que le envía señales.**

Patiendo de esto, consideremos en este apartado un Campo de Datos  $D(N)$  de resolución  $R$ , y consideremos un conjunto  $L$  de Campos Receptivos independientes de las direcciones  $i$  ( $i=1\dots N$ ). Este conjunto puede representarse por una matriz binaria  $P_{ik}$ , de  $N$  filas y  $L$  columnas tal que el elemento  $ik$  ( $k=1\dots L$ ) es 1 si la dirección  $i$  está en el campo receptivo  $k$ , y 0 en otro caso. Los vectores verticales  $P_{ik}$  (para cada  $k$ ) son independientes en el sentido de que ninguno de ellos puede obtenerse por unión de los otros. Por ejemplo, la matriz  $P_{ik}$  correspondiente a una transformación de resolución progresiva, para  $N=L=8$  es:

$$\begin{array}{l}
 1 \\
 2 \\
 3 \\
 \cdot \\
 \cdot \\
 \cdot \\
 \cdot \\
 \cdot \\
 N
 \end{array}
 \begin{bmatrix}
 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\
 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
 \end{bmatrix}$$

Una configuración de campos receptivos para  $L=4$ , es por ejemplo, representable por la matriz:

$$\begin{array}{l}
 1 \\
 2 \\
 3 \\
 \cdot \\
 \cdot \\
 \cdot \\
 \cdot \\
 \cdot \\
 N
 \end{array}
 \begin{bmatrix}
 1 & 1 & 1 & 0 \\
 1 & 1 & 1 & 0 \\
 1 & 1 & 0 & 0 \\
 1 & 1 & 0 & 0 \\
 1 & 0 & 0 & 1 \\
 1 & 0 & 0 & 1 \\
 1 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0
 \end{bmatrix}$$

Consideremos una subclase de campos receptivos de las direcciones  $i$ , donde cada dirección se toma, al menos, una vez. De esta forma, el campo receptivo cubre todo el campo de datos, es decir, no hay ninguna dirección desconsiderada. Esto puede observarse en las dos configuraciones de campos receptivos anteriores. En general, consideremos solamente configuraciones de este tipo, que llamaremos particiones  $P$ .

Se plantea, entonces, la siguiente cuestión. Dada una partición  $P$  de un campo de datos de dimensión  $N$ , ¿Cuál es un conjunto de funcionales que actuando sobre las direcciones de cada campo receptivo dentro de la partición, genera una representación completa?

Un teorema propuesto por Bolívar [Bol89], indica que dado un campo de datos de  $N$  direcciones y una partición  $P$  de  $L$  columnas tal que  $M = N/L$  sea entero, entonces el cómputo de  $M$  coeficientes de funcional, linealmente independientes y distintos de cero, en cada campo receptivo, proporciona una descripción completa del campo de datos.

Sean los  $M$  vectores de "factores de peso" del funcional representados en una matriz de  $m$  filas y  $d$  columnas  $F_{ij}$  ( $i=1, \dots, m; j=1, \dots, d$ ).

$$F = \begin{bmatrix} F_{11} & F_{12} & \dots & F_{1d} \\ F_{21} & F_{22} & \dots & F_{2d} \\ \dots & \dots & \dots & \dots \\ F_{m1} & F_{m2} & \dots & F_{md} \end{bmatrix}$$

Los vectores  $F_{ij}$  (para cada  $i$ ) son independientes por construcción.

Consideremos la matriz  $P_{ni}$  de campos receptivos:

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1l} \\ P_{21} & P_{22} & \dots & P_{2l} \\ \dots & \dots & \dots & \dots \\ P_{n1} & P_{n2} & \dots & P_{nl} \end{bmatrix}$$

Construyamos la matriz,  $M$ , de transformaciones de  $N \times N$ , por la "aplicación" de cada vector del funcional a cada campo receptivo:

$$\begin{bmatrix} F_{11} & \dots & F_{1d} \\ F_{21} & \dots & F_{2d} \\ \vdots & \vdots & \vdots \\ F_{m1} & \dots & F_{md} \end{bmatrix} \times \begin{bmatrix} P_{11} & \dots & P_{1l} \\ P_{21} & \dots & P_{2l} \\ \vdots & \vdots & \vdots \\ P_{n1} & \dots & P_{nl} \end{bmatrix} = N \begin{bmatrix} \text{---} & \text{---} & \text{---} & \text{---} \\ \text{---} & \text{---} & \text{---} & \text{---} \end{bmatrix} \begin{matrix} M_1 \\ M_2 \\ \vdots \\ M_k \end{matrix}$$

Como puede observarse, se han generado "grupos"  $M_1, M_2, \dots, M_k$  de  $M$  vectores, que tienen todos la misma extensión sobre el campo de datos. Dentro de cada grupo, los vectores son independientes entre sí, por construcción. Asimismo, los grupos entre sí son independientes, por corresponder a particiones independientes. Luego, los  $N$  vectores horizontales son independientes, y por consiguiente, la correspondiente transformación es completa. Es decir, la matriz  $M$  tiene inversa.

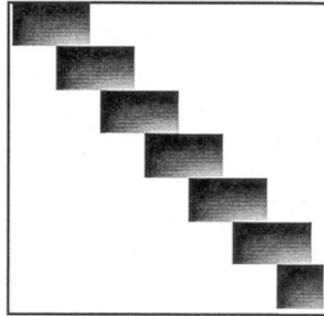
Como ejemplo numérico, consideremos el caso de una transformación de campo de datos de dimensión  $N=256$  (Grados de libertad o Resolución del Campo de Datos). Para la elección de los valores de  $L$  (número de campos receptivos) y  $d$  (grados de libertad del campo receptivo), la única restricción existente viene dada por la ecuación:

$$L = [(N-d)/z] + 1$$

siendo  $z$  el desplazamiento intercampo receptivo.

Esto quiere decir, que fijado el número de grados de libertad y el desplazamiento, obtenemos el número de campos receptivos necesarios.

En este caso, hemos elegido los valores  $d=64$  y  $z=32$ , ambos potencias de 2, lo cual simplifica apreciablemente el costo computacional para realizar la transformación. Los 7 campos receptivos que se generan son suficientes para cubrir todo el campo de datos, y por construcción, son linealmente independientes. Según el teorema anterior, la complitud requiere un número de funcionales por campo receptivo de  $M=N/L$  en media. En este caso al no ser entero, lo solucionamos eligiendo 36 funcionales para los campos receptivos 1,2,3,5,6,7, y 40 funcionales para el campo receptivo 4. En la siguiente figura, representamos el núcleo de la transformación utilizando como funcionales los momentos.



## **3.2 INFLUENCIA DE LOS PARÁMETROS CARACTERÍSTICOS DE LAS TRANSFORMACIONES DE CAMPOS RECEPTIVOS EN EL RECONOCIMIENTO DE FORMAS.**

### **3.2.1 Introducción**

Como ya es sabido, el Reconocimiento de Formas constituye la base de la Visión Artificial, una de las especialidades más importantes de la Inteligencia Artificial, y como consecuencia de ser un pilar de la Visión Computacional, el Reconocimiento de Formas también está vinculado a la Robótica.

Los objetos se engloban dentro de una misma clase si son más similares que los objetos pertenecientes a otra clase. En el sistema visual biológico, una base natural para la definición de similaridad puede derivarse del concepto de unidades de procesamiento con campos receptivos localizados.

En distintos trabajos realizados, se ha puesto de manifiesto que en Visión Artificial debe existir un compromiso básico entre los Funcionales a realizar y los Campos Receptivos, para mantener una "constancia" de los grados de libertad impuestos por los objetivos del sistema.

El estudio realizado consiste en la generación de espacios de representación para el reconocimiento visual. Dentro de este contexto vamos a utilizar transformaciones que combinan Campos Receptivos y Funcionales, y diseñar un sistema de clasificación en base a estas descripciones completas algebraico-analíticas. La eficacia del sistema clasificador depende en principio de factores como son la complejidad del campo de datos, la mayor o menor extensión de los campos receptivos, el número de descriptores en cada uno de éstos y la resolución. En trabajos realizados se ha puesto de manifiesto que cuando la resolución

es relativamente baja y la complejidad del campo de datos es pequeña, un número alto de campos receptivos y descriptores provoca una disminución de la eficacia en la clasificación, debido al ruido inherente y al aumento, no controlado por no ser detectable, de las redundancia en el campo de datos original.

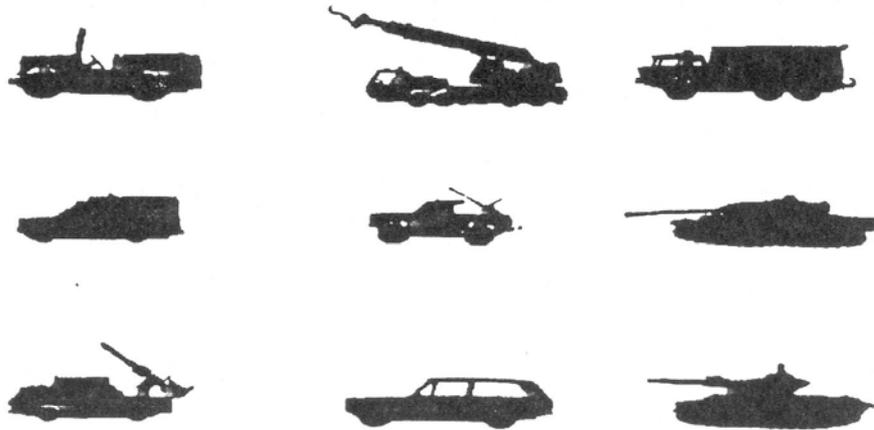
En este apartado, pretendemos alcanzar dos objetivos fundamentales:

- 1.- Estudiar la influencia de distintas configuraciones de Campos Receptivos y distintos Funcionales en la eficacia del sistema clasificador.
- 2.- Estudiar la dependencia de la eficacia del sistema en función de los distintos parámetros que actúan en las transformaciones de campos receptivos (Tamaño del Campo de Datos, Tipo de Funcionales, Tamaño del Campo Receptivo).

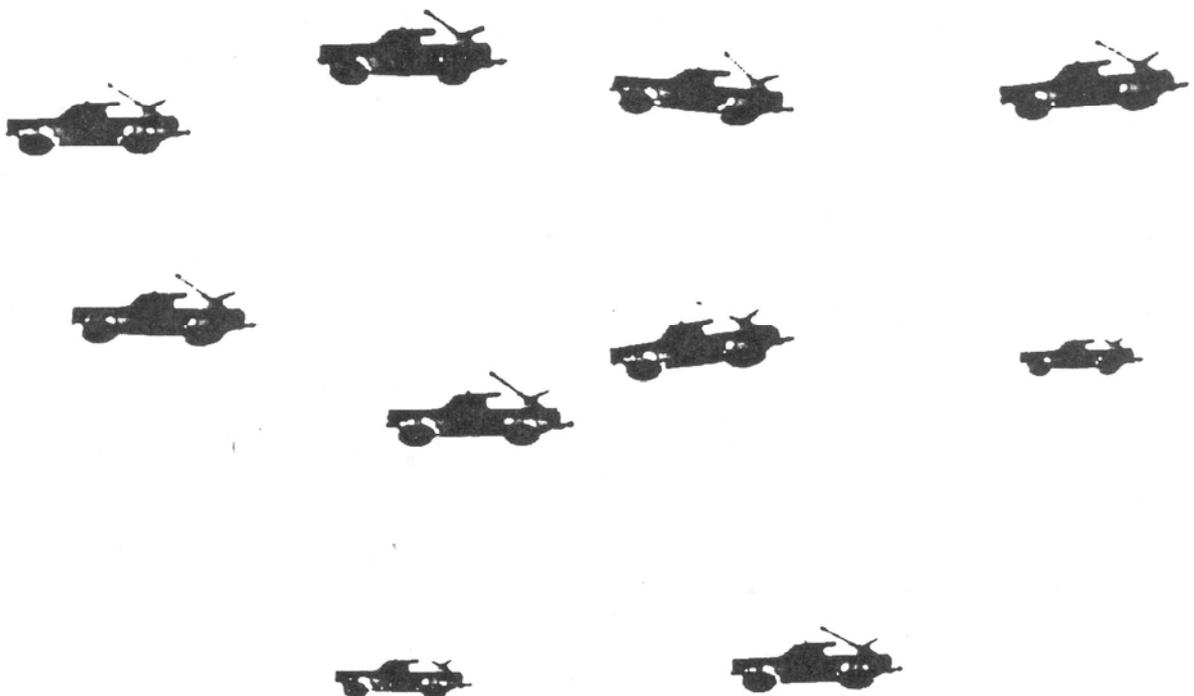
### **3.2.2 Banco de Imágenes**

Para realizar dicho estudio, se realizó la adquisición de un banco de 90 imágenes, utilizando para ello placas de adquisición FG-100 de Imaging Tec. Este consta de 10 tomas diferentes de cada una de las 9 formas distintas.

Las 9 formas diferentes de nuestro banco de datos se muestran en la siguiente figura:



En la siguiente figura se muestran las 10 tomas distintas realizadas de una de las nueve formas, donde se le han aplicado traslaciones, rotaciones y homotecias.



Analizando cada una de las imágenes adquiridas, podemos observar básicamente dos cosas:

- 1.- Objetos: normalmente etiquetables al satisfacer algún modelo de descripción.
- 2.- Regiones: zonas que poseen uniformidad/homogeneidad en alguna característica (color, nivel de gris, textura, etc.)

Esto nos conduce a un proceso de Segmentación, es decir, un proceso orientado a particionar la imagen digital en zonas disjuntas con significado propio. En nuestro caso, las imágenes contienen objetos claramente diferenciados del fondo, y en el caso particular en el que las frases descriptoras extraídas de las imágenes adquiridas van a depender sólo de la forma de las mismas, es posible generar un espacio de medida monodimensional donde sólo existan 2 regiones (objeto y fondo). Aplicando un método de segmentación como la umbralización basada en histograma, nos permite que todos los puntos menores que un cierto umbral sean asignados a una región y los que no verifiquen esta condición a otra. Por lo tanto, habrá que encontrar ese valor umbral y calcular las 2 regiones.

De los distintos métodos existentes para el cálculo del umbral, hemos seleccionado el ajuste a cuártica ya que este procedimiento tiene mucha más sencillez computacional que, por ejemplo, el ajuste a gaussiana y es resoluble de forma rápida empleando, por ejemplo, el método de Gauss-Seidel .

Una vez que las imágenes han sido segmentadas en dos regiones, objeto y fondo, le aplicamos otro proceso para normalizarlas y hacerlas invariantes frente a traslaciones, rotaciones y homotecias:

**a) Normalización frente a Traslaciones.**

Cuando se trabaja con imágenes que representan objetos aislados, una transformación que tiene especial interés es la traslación del objeto al origen de coordenadas. Para ello, se traslada el objeto hasta hacer coincidir su centro de gravedad con el centro de los ejes de coordenadas.

Para calcular el centro de gravedad del objeto utilizamos como herramienta los Momentos de igual forma que como fueron utilizados en el apartado 2.3.2 de este trabajo. Es decir, el centro de gravedad del objeto vendrá determinado por:

$$X_G = M_{10}/M_{00} \quad Y_G = M_{01}/M_{00}$$

Por lo tanto, aplicando este tipo de transformación conseguimos que las imágenes sean invariantes frente a cualquier tipo de traslación de coordenadas.

**b) Normalización frente a Rotaciones.**

Para obtener invarianza frente a rotaciones de las formas se puede girar los ejes de coordenadas hasta hacerlos coincidir con los Ejes de Inercia de la forma, que son una propiedad intrínseca de la misma y que, por tanto, la acompaña sea cual sea la situación relativa del objeto respecto al referencial de la cámara.

Otro método, que es el utilizado en este trabajo por su mayor exactitud y mayor velocidad de cómputo, consiste en calcular el ángulo de desfase como el formado por el vector que une el punto más alejado del contorno de la forma y el centro de gravedad de la

misma con respecto al eje horizontal. Este método sólo es válido es aquellos casos en que la forma no tenga ningún tipo de simetría respecto al centro de gravedad, ni respecto a los ejes cartesianos que pasan por él, como es nuestro caso.

Para ello, calculamos el contorno del objeto, por ejemplo, siguiendo la técnica del Código de Cadena, que es la codificación de los contornos en base a las coordenadas del primer punto del contorno  $(X_0, Y_0)$  y de las posiciones relativas de cada punto respecto al anterior. Una vez obtenido el contorno, se localiza el punto más alejado de este  $(X_i, Y_i)$  con respecto al centro de gravedad de la pieza  $(X_G, Y_G)$ . Por último, se gira el objeto hasta hacer coincidir la recta que une  $(X_i, Y_i)$  con  $(X_G, Y_G)$  con el eje X positivo de nuestro sistema de referencia y de esta forma la imagen queda normalizada frente a rotaciones.

### **c) Normalización frente a Homotecias.**

Para conseguir invarianzas frente a homotecias, es decir, variaciones del tamaño de la forma, fijamos que la imagen normalizada tenga un área predefinida constante. Esto nos lleva a realizar dos tratamientos distintos:

1.- Cuando el área predefinida es mayor que la de la imagen, existe una expansión del objeto apareciendo nuevos puntos. El procedimiento consiste en transformar la imagen de forma que a varios pixels de la imagen transformada le correspondan un único punto en la imagen original.

2.- Cuando el área predefinida es menor que la de la imagen, existe una compresión del objeto eliminándose puntos. El procedimiento consiste en realizar un enrejado en la imagen original y calcular para un punto transformado la media de los valores

de una cuadrícula del enrejado.

### 3.2.3 Frases Descriptoras: Campos Receptivos y Funcionales.

En este punto vamos a tratar la generación de las frases descriptoras que deben etiquetar a cada una de las clases que conforman las formas, tanto patrones como incógnitas, para su posterior reconocimiento.

Para ello, partimos del valor de N (Grados de libertad o resolución del Campo de Datos). Para la elección de los valores de L (número de particiones) y d (grados de libertad del Campo Receptivo), la única restricción existente viene dada por la ecuación:

$$L = [(N-d)/z] + 1$$

siendo z el desplazamiento interparticiones.

Esto quiere decir, que fijado el número de grados de libertad y el desplazamiento, obtenemos el número de particiones necesarias.

Por último, llevamos a cabo el proceso de clasificación. Para ello hemos utilizado, de los distintos métodos de clasificación existentes, una versión modificada de la Distancia Euclídea. Esta distancia se define de la siguiente forma:

Dada una frase descriptora incógnita  $D'(i)$  y un conjunto de frases descriptoras, correspondientes a cada una de las clases de la base de datos referencial,  $D(i,j)$ , se define la distancia Euclídea modificada como:

$$[Distancia(j)]^2 = \sum [(D'(i) - D(i,j)/10^{E(i)})^2]$$

donde:

$$E(j) = \min (\text{exponente } \{D'(i)\}, \text{exponente } \{D(i,j)\}) \quad (i=1,\dots,n)$$

siendo  $n$  el número de elementos de las frases descriptoras.

Una vez calculadas todas las distancias entre la frase descriptora incógnita y cada una de las frases descriptoras de las clases que componen la base de datos referencial, el sistema de inferencia asocia el patrón incógnita a aquella clase patrón cuya distancia al patrón incógnita sea mínima.

### 3.2.4 Estudio Experimental

#### Estudio Preliminar

Como punto de partida de este estudio, se realizó la adquisición de las 90 imágenes con una resolución de 128X128 pixels, teniendo, por tanto, un tamaño del Campo de Datos igual a 128. Se comenzó eligiendo los siguientes parámetros:

- Tamaño del Campo de Datos:  $N = 128$
- Grados de Libertad de cada Campo Receptivo:  $d = N/4 = 32$
- Desplazamiento Intercampo Receptivo:  $z = d/2 = 16$
- Nº de Campos Receptivos:  $L = 7$
- Nº de Funcionales por Campo Receptivo:

Campo Receptivo 1,2,3,5,6,7: 18

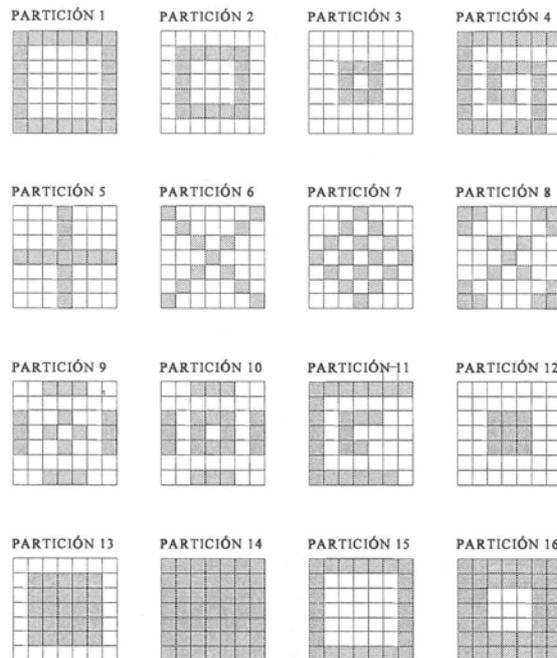
Campo Receptivo 4: 20

Una vez elegidos estos parámetros, pasamos a seleccionar un conjunto de

Funcionales y Campos Receptivos para analizar su influencia en el sistema clasificador.

Los funcionales elegidos primeramente en este estudio fueron los Momentos. Para la elección del conjunto de momentos, seguiremos las especificaciones de Alt [Alt62] que nos propone que los momentos que proporcionan mejores resultados para propósitos de clasificación son los de bajo orden. Por esta razón elegimos los momentos  $M_{0,3}$ ;  $M_{0,4}$ ;  $M_{1,2}$ ;  $M_{1,3}$ ; y otros distintos como  $M_{0,6}$ ;  $M_{0,8}$ ;  $M_{0,10}$ ;  $M_{1,4}$ ;  $M_{1,6}$ ;  $M_{1,8}$ ;  $M_{1,10}$ .

En cuanto a la distribución de los Campos Receptivos, es decir, dónde posicionamos los Campos Receptivos dentro del Campo de Datos, lo hacemos de forma aleatoria, ya que, a priori, no tenemos ninguna información que nos determine su posición. Por tanto, seleccionamos distintos Campos Receptivos Aleatorios resultando distintas configuraciones. Por claridad de representación, hemos etiquetado a las distintas configuraciones de campos receptivos seleccionados como Particiones. En la siguiente figura mostramos 16 configuraciones de campos receptivos etiquetadas como Partición 1 hasta Partición 16.



La zona sombreada de la figura representa la región de los Campos Receptivos y el número total de ellos pero no la localización física exacta de cada uno de ellos. Al estar tratando un problema bidimensional, la representación del desplazamiento intercampo receptivo exigiría dos dimensiones adicionales, con lo cual su representación no es posible. Por ello hemos optado por la representación anterior donde la zona sombreada refleja aproximadamente la localización de los campos receptivos elegidos.

Con los Funcionales y las Particiones señaladas, realizamos la última fase. Para ello, partimos del banco de imágenes tomando:

- 63 Imágenes Patrones.
- 27 Imágenes Incógnitas.

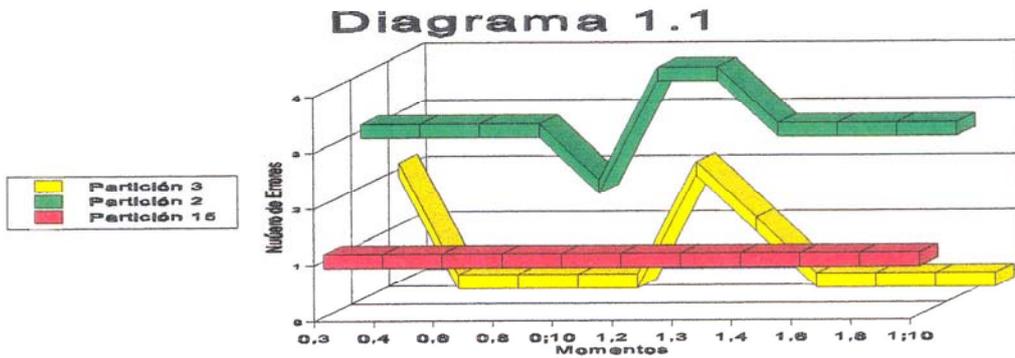
y arrancamos el proceso de clasificación. Los resultados obtenidos se muestran en la siguiente tabla donde se refleja los números de fallos producidos por cada una de las particiones y con cada funcional.

Representación de Imágenes a partir de las Transformaciones de Campos Receptivos

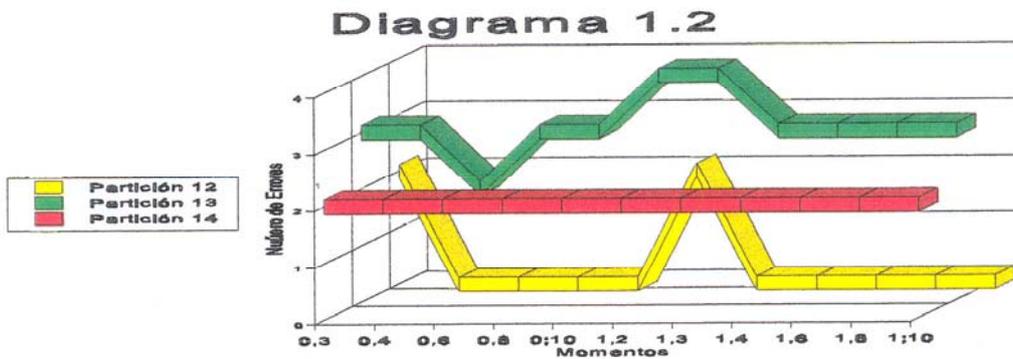
P/ M	M 0 3	M 0 4	M 0 6	M 0 8	M 0 10	M 1 2	M 1 3	M 1 4	M 1 6	M 1 8	M 1 10
1	1	1	1	1	1	1	1	1	1	1	1
2	3	3	3	3	2	4	4	3	3	3	3
3	2	0	0	0	0	2	1	0	0	0	0
4	3	2	3	3	3	2	2	2	2	3	3
5	1	1	1	1	1	2	1	1	1	1	1
6	4	4	4	4	3	3	0	0	1	1	2
7	2	2	2	2	2	2	2	2	2	2	2
8	1	1	3	3	2	0	0	0	3	2	2
9	1	1	2	4	2	1	2	1	1	2	1
10	1	1	1	3	2	1	1	1	1	2	1
11	2	1	1	1	3	1	2	1	1	3	3
12	2	0	0	0	0	2	0	0	0	0	0
13	3	3	2	3	3	4	4	3	3	3	3
14	2	2	2	2	2	2	2	2	2	2	2
15	1	1	1	1	1	1	1	1	1	1	1
16	2	2	2	2	2	2	2	2	2	2	2

Para analizar los resultados obtenidos, procedemos a reagruparlos en cuatro grupos distintos atendiendo a la configuración "progresiva" de los campos receptivos:

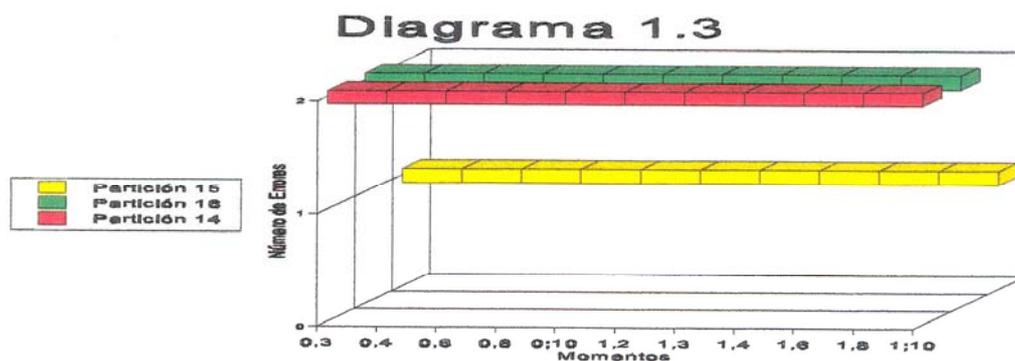
a) Campos Receptivos de Centro Creciente Simple. Corresponde a las particiones 3, 2 y 15.



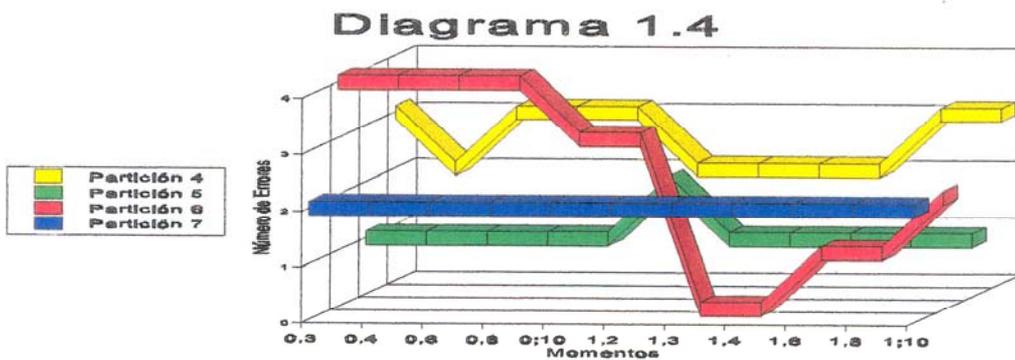
b) Campos Receptivos de Centro Creciente Compuesto. Corresponde a las particiones 12, 13 y 14.

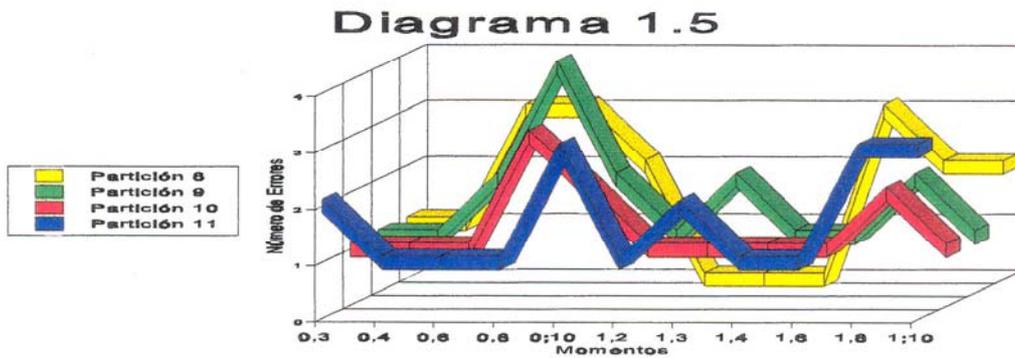


c) Campos Receptivos de Periferia Creciente Compuesto. Corresponde a las particiones 15, 16 y 14.



d) Campos Receptivos Variados. Corresponde al resto de las particiones.





Analizando los resultados, concluimos que:

1) Existe un conjunto de configuraciones o distribución de los Campos Receptivos que producen mejores resultados que otros para propósitos de clasificación. Esto enfatiza la relevancia de los campos receptivos frente a los funcionales clásicos. Esto puede comprobarse observando que la variabilidad de los campos receptivos versus funcionales es más pequeña que los funcionales versus campos receptivos.

Las configuraciones de los Campos Receptivos que producen mejores resultados son aquellas que engloban las zonas de la imagen que son más discriminantes, como puede ser el contorno de la misma.

2) Un aumento del número de descriptores produce una disminución en la eficiencia del sistema clasificador. Posiblemente esto es debido al ruido inherente y al aumento de redundancia los cuales no son controlados en el campo de datos original. Concluimos que el hecho importante no es el número de descriptores sino dónde y cómo son elegidos.

3) Corroboramos el hecho de que los momentos de bajo orden proporcionan, en general, mejores resultados.

Con los resultados obtenidos en este estudio preliminar, proseguimos analizando la influencia de distintos parámetros en el proceso de clasificación.

### **Influencia del Tamaño del Campo de Datos**

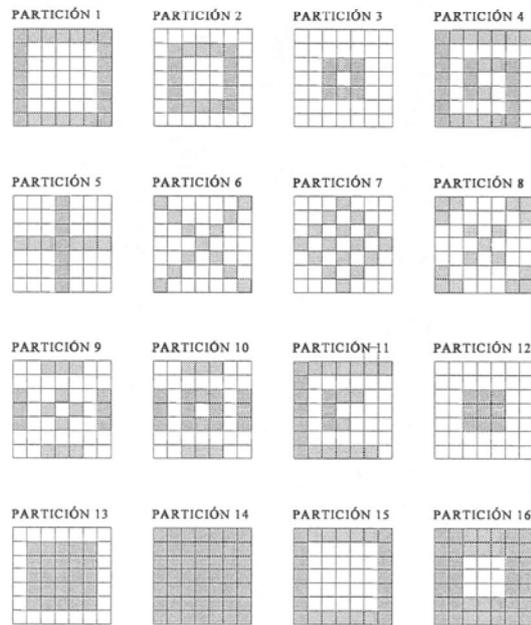
En primer lugar, nos planteamos como influiría en los resultados del sistema clasificador el aumentar el tamaño del Campo de Datos. es decir, pasar de imágenes adquiridas con una resolución de 128X128 pixels como era el caso anterior, a imágenes adquiridas con una resolución de 256X256 pixels. Por tanto los nuevos parámetros serían:

- Tamaño del Campo de Datos:  $N = 256$
- Grados de Libertad de cada Campo Receptivo:  $d = N/4 = 64$
- Desplazamiento Intercampos Receptivos:  $z = d/2 = 32$
- Nº de Campos Receptivos:  $L = 7$
- Nº Funcionales por Campo Receptivo:

Campo Receptivo 1,2,3,5,6,7: 36

Campo Receptivo 4: 40

Con estos nuevos parámetros, utilizando los mismos funcionales de antes (Momentos), y configuraciones de Campos Receptivos equivalentes al caso anterior etiquetadas como particiones como muestra la siguiente figura:

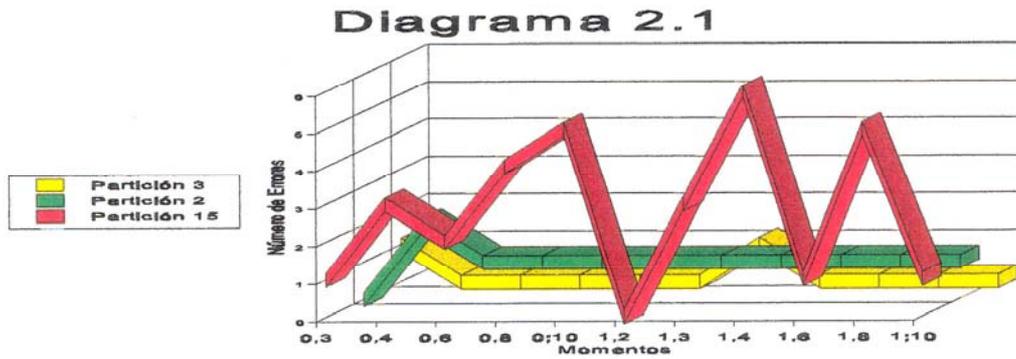


arrancamos el proceso de clasificación. Los resultados obtenidos se muestran en la siguiente tabla donde se refleja los números de fallos producidos por cada una de las particiones y con cada funcional.

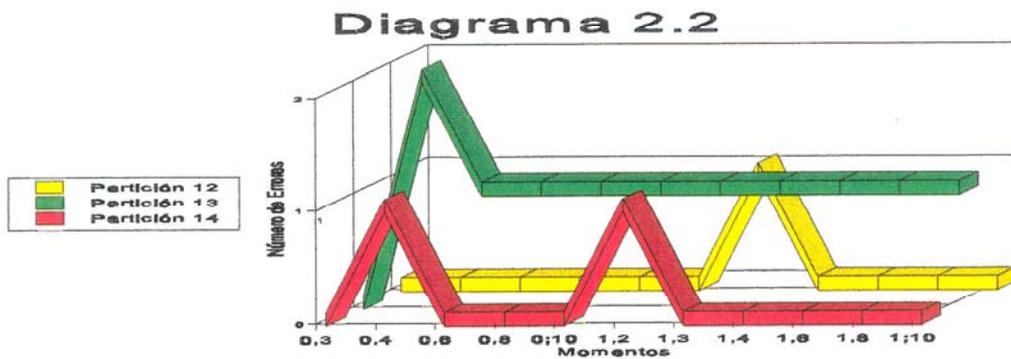
P/ M	M 0 3	M 0 4	M 0 6	M 0 8	M 0 10	M 1 2	M 1 3	M 1 4	M 1 6	M 1 8	M 1 10
1	1	3	1	2	5	0	1	2	1	3	1
2	0	2	1	1	1	1	1	1	1	1	1
3	1	0	0	0	0	0	1	0	0	0	0
4	2	1	2	4	2	0	3	1	3	2	3
5	0	0	0	0	0	0	0	0	0	0	0
6	2	2	2	1	1	2	2	2	2	1	1
7	1	1	1	2	2	1	1	1	1	2	2
8	1	1	0	3	0	0	0	0	0	0	0
9	1	4	1	2	3	1	2	2	1	3	1
10	1	4	1	2	3	0	2	2	1	3	1
11	0	0	1	1	0	0	0	0	0	0	0
12	0	0	0	0	0	0	1	0	0	0	0
13	0	2	1	1	1	1	1	1	1	1	1
14	0	1	0	0	0	1	0	0	0	0	0
15	1	3	2	4	5	0	3	6	1	5	1
16	0	1	0	0	0	1	0	0	0	0	0

Para analizar los resultados obtenidos, procedemos a reagruparlos nuevamente en cuatro grupos distintos atendiendo a la configuración de los campos receptivos:

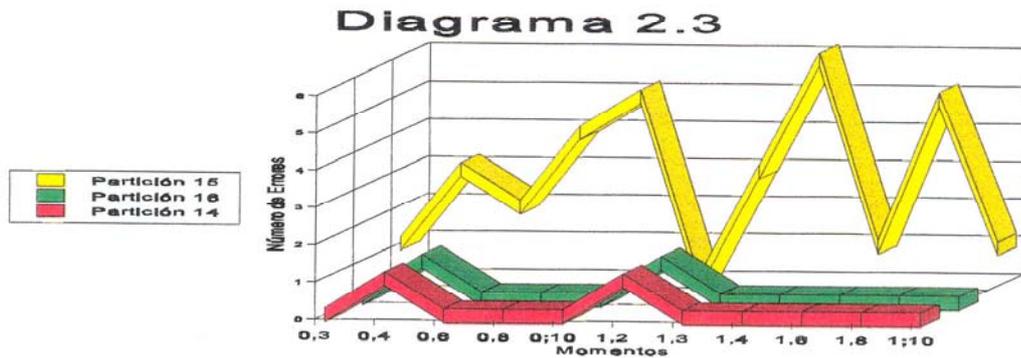
a) Campos Receptivos de Centro Creciente Simple. Corresponde a las particiones 3, 2 y 15.



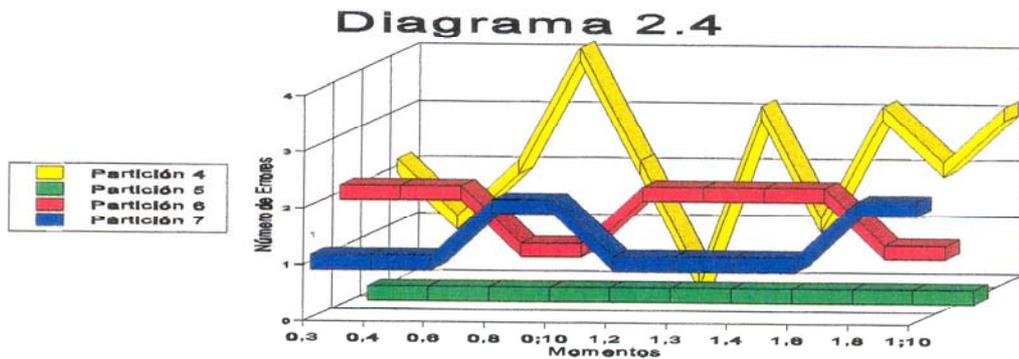
b) Campos Receptivos de Centro Creciente Compuesto. Corresponde a las particiones 12, 13 y 14.

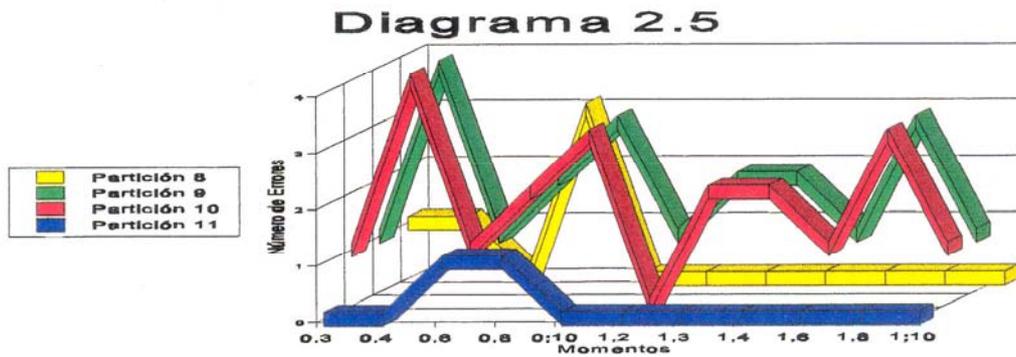


c) Campos Receptivos de Periferia Creciente Compuesto. Corresponde a las particiones 15, 16 y 14.



d) Campos Receptivos Variados. Corresponde al resto de las particiones.





Analizando los resultados, concluimos que:

- 1) En general se observa que un aumento del tamaño del Campo de Datos (Resolución) aumenta la eficiencia del sistema clasificador.

### Influencia del Tipo de Funcional

Al observar que, efectivamente, los resultados mejoraban aumentando la resolución, pasamos a estudiar como afectaría al sistema clasificador si utilizamos otros tipos de funcionales.

Para ello, elegimos como descriptores de Fourier, tres funcionales diferentes, que son:

#### 1.- Energía Total

$$E.T. = \sum_{1,N} \sum |F(U,V)|^2$$

2.- Energía del Centro

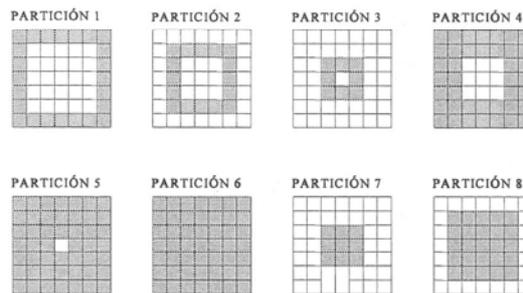
$$F.P.B. = \sum_i \sum_j |F(U,V)|^2$$

3.- Energía de la Periferia

$$F.P.A. = \sum_{1,i} \sum_{1,N} |F(U,V)|^2 + \sum_{j,N} \sum_{1,N} |F(U,V)|^2 +$$

$$\sum_{1,N} \sum_{1,i} |F(U,V)|^2 + \sum_{1,N} \sum_{j,N} |F(U,V)|^2$$

En cuanto a las configuraciones de los Campos Receptivos, seleccionamos un conjunto más reducido de los primeros, quedándonos con aquellos más determinantes en el sistema clasificador anterior, que corresponden a los siguientes:

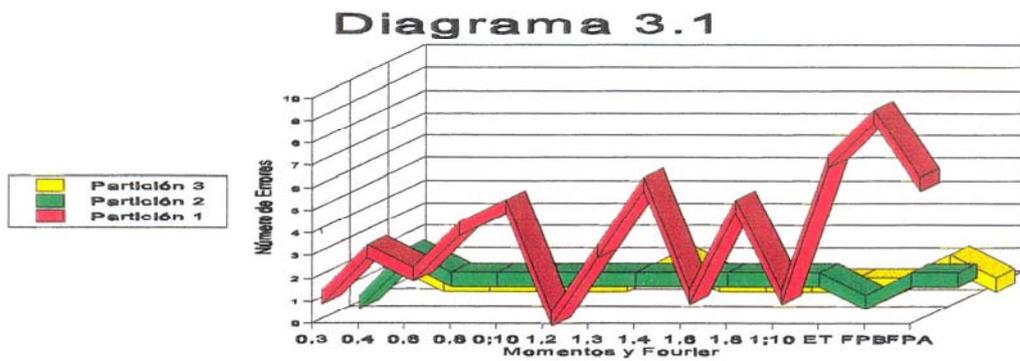


Seguidamente, realizamos el proceso de clasificación. Los resultados obtenidos se muestran en la siguiente tabla donde se refleja los números de fallos producidos por cada una de las particiones y con cada funcional.

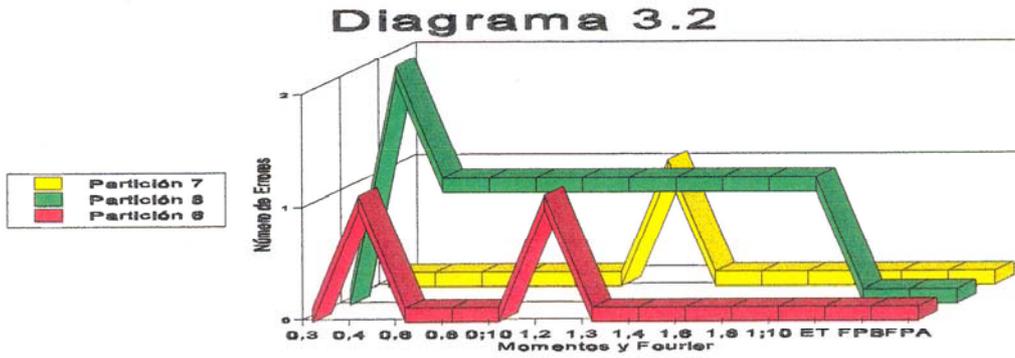
P/ M	M 0 3	M 0 4	M 0 6	M 0 8	M 0 10	M 1 2	M 1 3	M 1 4	M 1 6	M 1 8	M 1 10	E T	F P B	F P A
1	1	3	2	4	5	0	3	6	1	5	1	7	9	6
2	0	2	1	1	1	1	1	1	1	1	1	0	1	1
3	1	0	0	0	0	0	1	0	0	0	0	0	1	0
4	0	1	0	0	0	1	0	0	0	0	0	0	1	1
5	0	0	0	1	1	0	0	0	0	0	1	0	1	0
6	0	1	0	0	0	1	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0	0	0	0	0
8	0	2	1	1	1	1	1	1	1	1	1	0	0	0

Para analizar los resultados obtenidos, procedemos a reagruparlos nuevamente en diferentes grupos distintos atendiendo a la configuración de los campos receptivos:

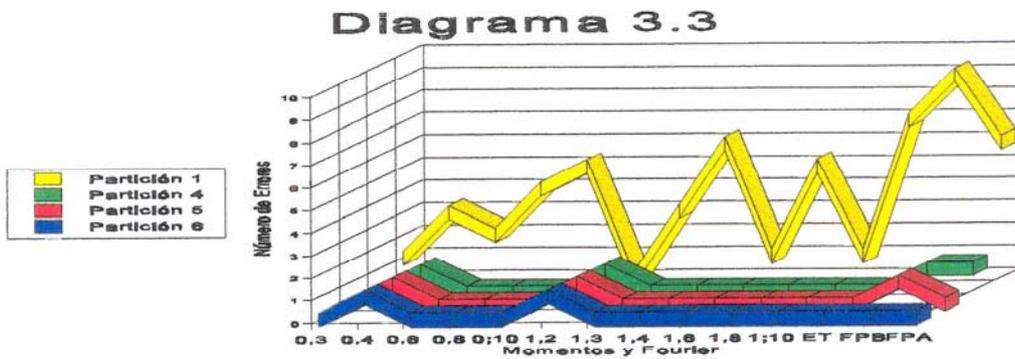
a) Campos Receptivos de Centro Creciente Simple. Corresponde a las particiones 3, 2 y 1.



b) Campos Receptivos de Centro Creciente Compuesto. Corresponde a las particiones 7, 8 y 6.



c) Campos Receptivos de Periferia Creciente Compuesto. Corresponde a las particiones 1, 4, 5 y 6.



Analizando los resultados, concluimos que:

- 1) En general se observa la poca significancia de los resultados con estos nuevos funcionales. Recordamos nuevamente que la máxima significancia se encuentra en la elección adecuada del lugar de los Campos Receptivos.

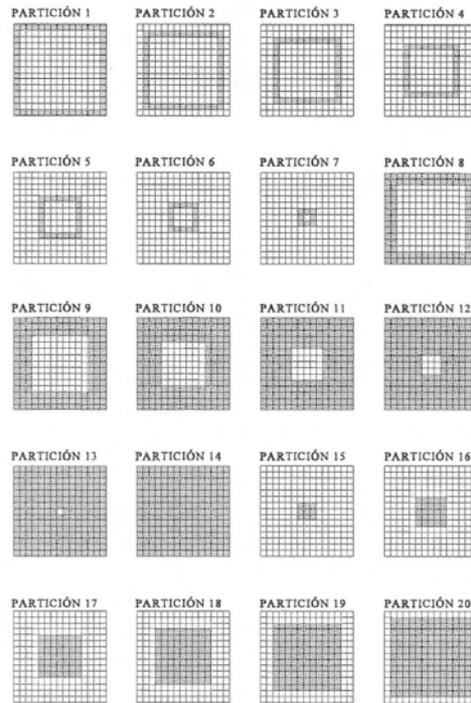
### **Influencia de la dimensión del Campo Receptivo**

Otro parámetro que no hemos modificado hasta ahora es la dimensión del Campo Receptivo. En este caso, vamos a estudiar que influencia produce la disminución del Campo Receptivo.

Para ello tomamos los siguientes parámetros:

- Tamaño del Campo de Datos:  $N = 256$
- Grados de Libertad de cada Campo Receptivo:  $d = N/8 = 32$
- Desplazamiento Intercampos Receptivos:  $z = d/2 = 16$
- N° de Campos Receptivos:  $L = 15$
- N° Funcionales por Campo Receptivo:
  - Campo Receptivo 8: 18
  - Resto de Campos Receptivos: 17

Los funcionales corresponden al caso anterior y los Campos Receptivos elegidos serían: Campos Receptivos Centro Creciente Simple, Centro Creciente Compuesto y Periferia Creciente Compuesta:

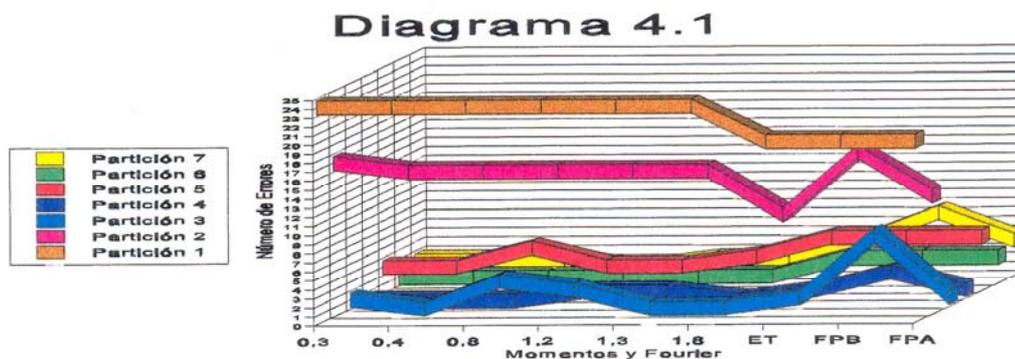


Seguidamente, realizamos el proceso de clasificación. Los resultados obtenidos se muestran en la siguiente tabla donde se refleja los números de fallos producidos por cada una de las particiones y con cada funcional.

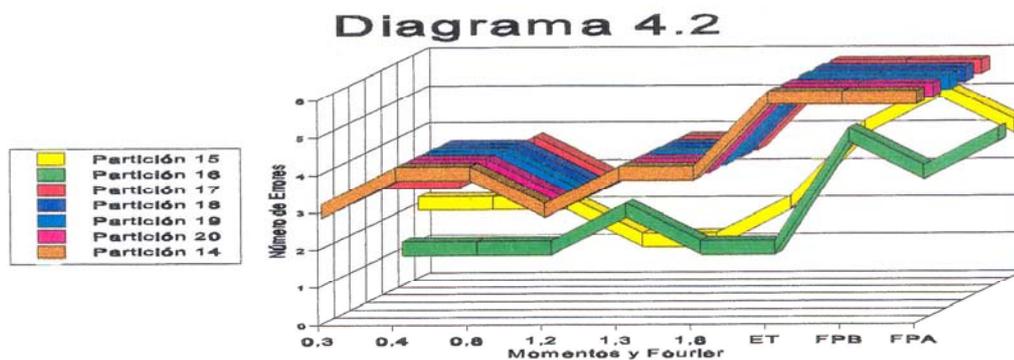
P/ M	M 0 3	M 0 4	M 0 8	M 1 2	M 1 3	M 1 8	E T	F P B	F P A
1	24	24	24	24	24	24	20	20	20
2	17	16	16	16	16	16	11	18	13
3	1	0	3	2	0	0	1	8	1
4	0	0	0	1	1	0	1	3	1
5	3	3	5	3	3	4	6	6	6
6	1	1	1	1	1	1	3	3	3
7	2	2	2	1	2	2	4	7	4
8	17	16	16	16	16	16	9	13	9
9	2	0	3	2	0	0	1	9	1
10	0	0	0	1	0	0	0	3	0
11	3	3	3	3	3	4	5	5	5
12	3	3	3	4	4	4	6	6	5
13	3	4	4	3	4	4	6	6	6
14	3	4	4	3	4	4	6	6	6
15	2	2	2	1	1	2	4	5	4
16	1	1	1	2	1	1	4	3	4
17	3	3	4	3	4	4	6	6	6
18	3	4	4	3	4	4	6	6	6
19	3	4	4	3	4	4	6	6	6
20	3	4	4	3	4	4	6	6	6

Para analizar los resultados obtenidos, procedemos a reagruparlos nuevamente en diferentes grupos distintos atendiendo a la configuración de los campos receptivos:

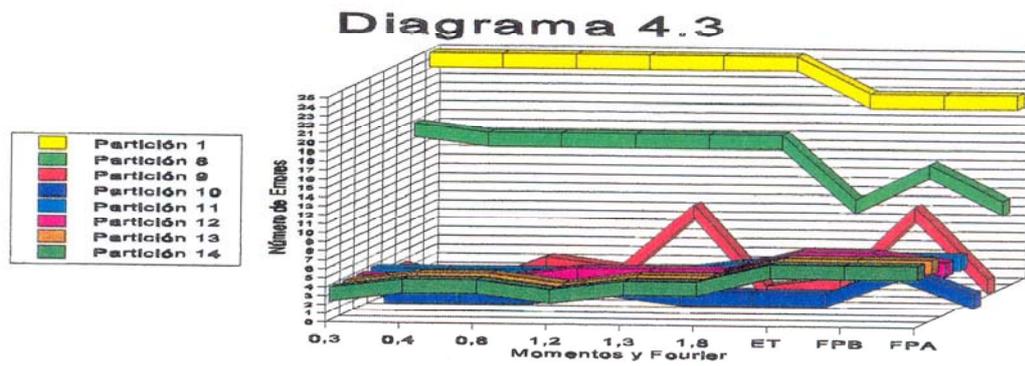
a) Campos Receptivos de Centro Creciente Simple. Corresponde a las particiones 7,6,5,4,3,2 y 1.



b) Campos Receptivos de Centro Creciente Compuesto. Corresponde a las Particiones 15,16,17,18,19,20,14.



c) Campos Receptivos de Periferia Creciente Compuesto. Corresponde a las Particiones 1,8,9,10,11,12,13,14.



Analizando los últimos resultados, resaltamos nuevamente la importancia del lugar donde elegimos los Campos Receptivos.

### **Conclusiones de este apartado**

- 1) Existe un conjunto de configuraciones o distribución de los Campos Receptivos que producen mejores resultados que otros para propósitos de clasificación. Esto enfatiza la relevancia de los campos receptivos frente a los funcionales clásicos. Esto puede comprobarse observando que la variabilidad de los campos receptivos versus funcionales es más pequeña que los funcionales versus campos receptivos. Dichas configuraciones de los Campos Receptivos que producen mejores resultados son aquellas que engloban las zonas de la imagen que son más discriminantes, como puede ser el contorno de la misma.
  
- 2) Corroboramos el hecho de que los momentos que proporcionan mejores resultados para propósitos de clasificación son los de bajo orden.
  
- 3) Al aumentar el tamaño del Campo de Datos, aumenta la eficiencia del sistema clasificador.
  
- 4) Un aumento del número de descriptores produce una disminución en la eficiencia del sistema clasificador. Posiblemente esto es debido al ruido inherente y al aumento de redundancia los cuales no son controlados en el campo de datos original. Concluimos que el hecho importante no es el número de descriptores sino dónde y cómo son elegidos.
  
- 5) En general, el utilizar un tipo de funcional u otro no influye significativamente

en los resultados de la clasificación.

6) El tamaño del Campo Receptivo tampoco influye significativamente en los resultados de la clasificación. Sin embargo, observando los resultados cuando se disminuyó el tamaño del campo receptivo, podemos enfatizar el hecho de la elección adecuada de los Campos Receptivos.

### 3.3 DESCRIPCIÓN AUTOMÁTICA DE FORMAS UTILIZANDO LOS CAMPOS RECEPTIVOS

Tal como concluimos en el apartado anterior, los campos receptivos que poseen mayor información discriminante son los situados en el contorno de las formas. Por tanto, en este apartado proponemos un método de describir una forma posicionando automáticamente los Campos Receptivos que la describen.

Como ilustración, partimos de una imagen que corresponde a una de las formas utilizadas en el estudio anterior.

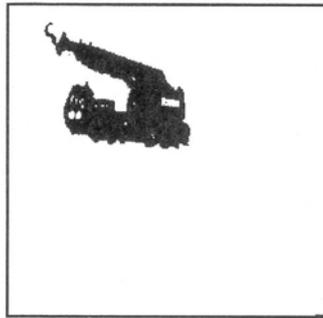
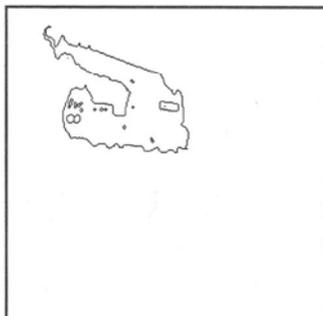


Imagen Original

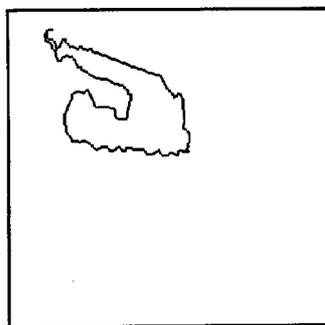
A dicha imagen le aplicamos un proceso de umbralizado y otro proceso para determinar el contorno de la misma.



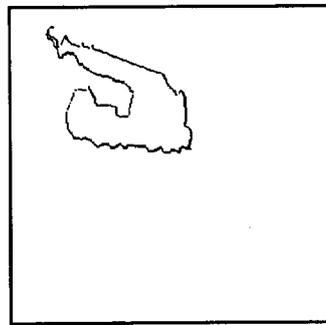
Contorno I. Original

Por último se determina la localización de los campos receptivos, dependiendo del tamaño del mismo.

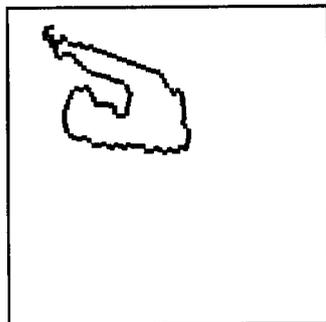
En las siguientes figuras, presentamos el resultado obtenido al aplicar el método de localización automática de los campos receptivos. En las figuras de la izquierda presentamos distintos resultados dependiendo del tamaño del campo receptivo (T.C.R.) seleccionado y a su derecha, la reconstrucción obtenida de la imagen original para dicho tamaño del campo receptivo.



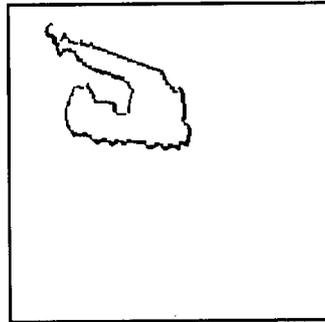
T.C.R. = 2



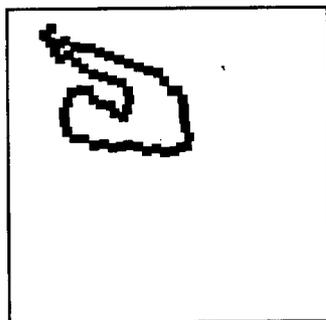
R.C.R. = 2



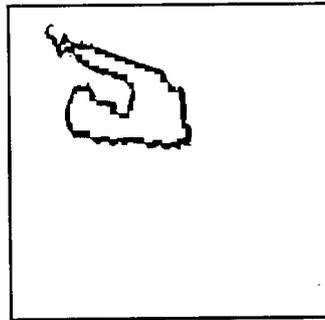
T.C.R. = 4



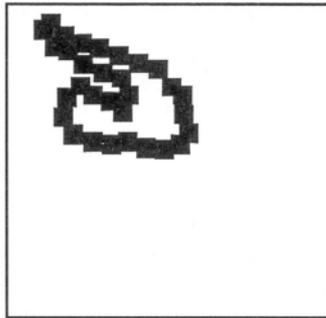
R.C.R. = 4



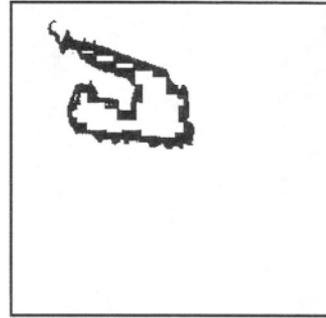
T.C.R. = 8



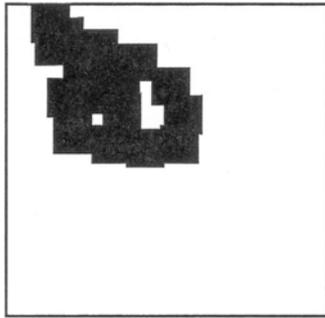
R.C.R. = 8



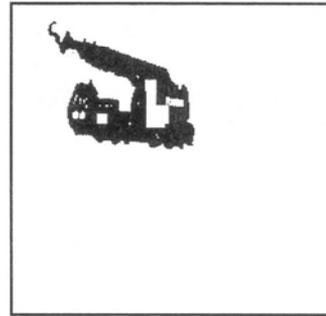
T.C.R. = 16



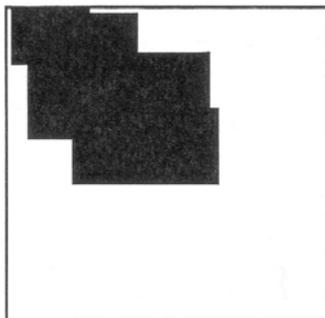
R.C.R. = 16



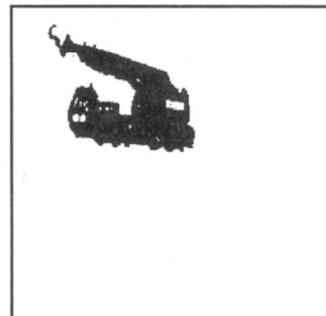
T.C.R. = 32



R.C.R. = 32

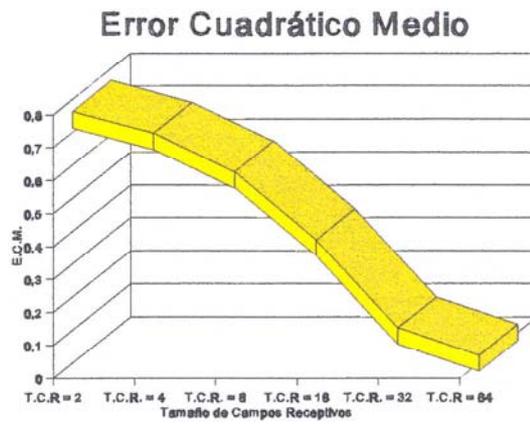


T.C.R. = 64



R.C.R. = 64

Por último, presentamos el Error Cuadrático Medio entre la imagen original y cada una de las reconstrucciones dependiendo del tamaño del campo receptivo.



# **CAPÍTULO 4**

**VARIABILIDAD DE LOS CAMPOS RECEPTIVOS.**

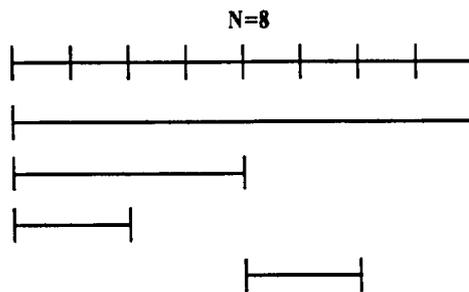
**COMPLITUD. GENERACIÓN DE FUNCIONALES**

## 4.1 INTRODUCCIÓN

Hasta ahora, hemos visto transformaciones en las que la partición del campo de datos era tal que la dimensión de los campos receptivos era constante, y en estas condiciones formulamos un teorema respecto a la complitud de los mismos. Pero ¿qué ocurre si siguiendo el paradigma neurofisiológico, consideramos campos receptivos de tamaño variable?. ¿Cómo se formularían los Teoremas de complitud correspondientes?.

En este capítulo, estudiaremos la variabilidad de los campos receptivos, es decir, conjuntos de campos receptivos de distintos tamaños. Se estudiará la complitud de la transformación con estos campos receptivos y distintas transformaciones para la generación de funcionales.

Primeramente, como ilustración gráfica, partamos de un campo de datos de longitud 8 ( $N=8$ ) y 4 campos receptivos ( $L=4$ ) de distinto tamaño:



Expresándolo en forma de matriz, tendríamos:

$$\begin{array}{l}
 1 \\
 2 \\
 3 \\
 . \\
 . \\
 . \\
 . \\
 8
 \end{array}
 \begin{bmatrix}
 1 & 1 & 1 & 0 \\
 1 & 1 & 1 & 0 \\
 1 & 1 & 0 & 0 \\
 1 & 1 & 0 & 0 \\
 1 & 0 & 0 & 1 \\
 1 & 0 & 0 & 1 \\
 1 & 0 & 0 & 0 \\
 1 & 0 & 0 & 0
 \end{bmatrix}$$

En este ejemplo, tenemos 8 grados de libertad y en media se ha de computar 8/4 funcionales por columna. En este caso, nos llevaría a una representación completa, si la elección de los funcionales es apropiada. Es decir, tenemos que garantizar la independencia de los funcionales cuando se aplican a cada columna, no bastando que lo sea para una columna determinada.

Por ejemplo, para la primera columna, podemos elegir distintos funcionales independientes:

$$F1 \rightarrow 1/8 (1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1)$$

$$F2 \rightarrow (1 \ 1 \ 1 \ 1 \ -1 \ -1 \ -1 \ -1)$$

que al aplicarlos a la matriz, nos daría el siguiente resultado:

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

Se observa que al aplicarlos a la segunda columna los números que se obtienen no son independientes. Sin embargo, si elegimos dos funcionales que son independientes para la columna de menor dimensión, por ejemplo los de pesos:

$$(1,1) \quad ; \quad (+1,-1)$$

y se generan dos funcionales nuevos de dimensión el de la columna más grande por yuxtaposición, es decir los:

$$\begin{aligned} &(1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1) \\ &(+1 \ -1 \ +1 \ -1 \ +1 \ -1 \ +1 \ -1) \end{aligned}$$

los números que se generan ya son independientes. La correspondiente matriz de transformación es:

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

tal que todos los vectores filas son independientes.

La solución encontrada en este ejemplo ha sido posible porque el número medio de funcionales a calcular por columna de la partición es igual (o menor) a la resolución de la columna de menor dimensión.

### TEOREMA

Sea un campo de datos de dimensión  $N$ . Sea una partición arbitraria  $P$ , de  $L$  columnas tal que  $a$  y  $b$  son las dimensiones de la columna de mayor y menor dimensión respectivamente, y se cumple además,  $a \geq N/L > 1$ . En estas condiciones, un conjunto de  $n=N/L$  funcionales independientes, de longitud  $a$  yuxtapuestos hasta la longitud  $b$ , proporciona una transformación completa.

La demostración del teorema se sigue con facilidad. En efecto, consideremos un conjunto  $\{a\}$  de  $n$  vectores linealmente independientes, por construcción, de dimensión  $a$ . Formemos la matriz de transformación  $T$  de la siguiente manera:

- 1) Se yuxtaponen las componentes de cada vector de  $\{a\}$  hasta lograr vectores de

dimensión  $\{b\}$ , en total  $n$ .

2) Se completan los vectores de  $\{a\}$  hasta la dimensión  $b$  ( $b-a$  lugares) con ceros.

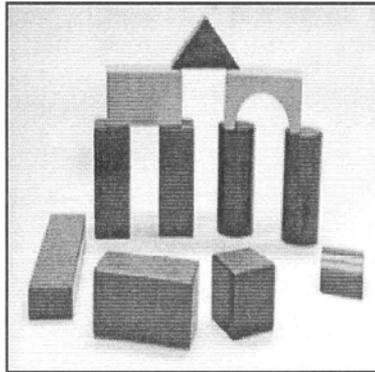
Es obvio que el conjunto de los  $2n$  vectores ( $2n < b$ ) son linealmente independientes.





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

En los siguientes apartados, se crearán distintas particiones, según lo planteado anteriormente, basándonos en distintas transformaciones. Haremos uso de la siguiente imagen que ha sido adquirida con una resolución de 256X256 ( $N=256$ ).



### 4.3 BASADOS EN LA TRANSFORMADA DE HAAR

En primer lugar, vamos a realizar un ejemplo ilustrativo suponiendo que el tamaño del campo de datos es  $N=8$ . Elegimos como funcional los de Haar y consideremos una partición (conjunto de campos receptivos) de tamaño 4 ( $L=4$ ). Según el teorema anterior, el número de funcionales a seleccionar será:

$$8/4 \leq 2$$

Elijamos como funcionales los basados en la transformada de Haar:

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

Generamos una partición siguiendo el método comentado anteriormente, cuyo resultado será:

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

El último paso sería obtener la matriz de transformación. Por ejemplo tendríamos la siguiente matriz de transformación:

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

Podríamos obtener distintas matrices de transformación, dependiendo del criterio de ordenación que se utilice.

Siguiendo con el estudio, elegimos como funcionales los de Haar de longitud 2 ( $dm=2$ ) y consideramos una partición de tamaño 128 ( $L=128$ )

$$256/128 \leq 2$$

Los funcionales elegidos fueron:

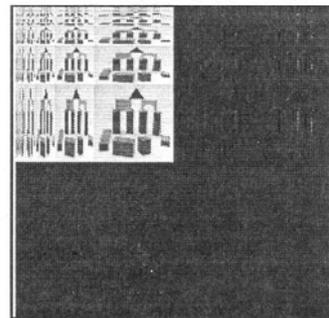
$$\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

La partición resultante tendrá la siguiente composición:

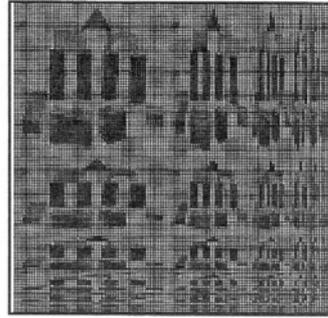
$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & \dots \\ \dots & \dots \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & \dots \\ \dots & \dots \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \dots \end{bmatrix}$$

Con la imagen original, realizamos distintas transformaciones con una matriz de partición ordenada de distintas formas. Ante la imposibilidad física de mostrar la matriz de transformación en su totalidad, hemos representado como guía ilustrativa a la izquierda de las imágenes, la matriz de transformación equivalente para el caso de un tamaño del campo de datos N=8, obteniendo los siguientes resultados:

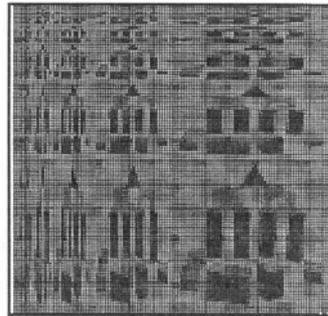
$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \end{bmatrix}$$



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



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



#### 4.4 BASADOS EN LA TRANSFORMADA DE HADAMARD

En este apartado, seleccionamos unos funcionales basados en la transformada de Hadamard. Elegimos unos de longitud 4 ( $dm=4$ ) y consideramos una partición de 64 ( $L=64$ )

$$256/64 \leq 4$$

Los funcionales elegidos fueron:

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

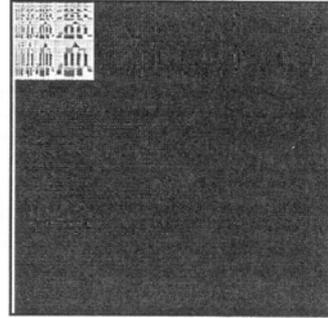
La partición resultante tendrá la siguiente composición:

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

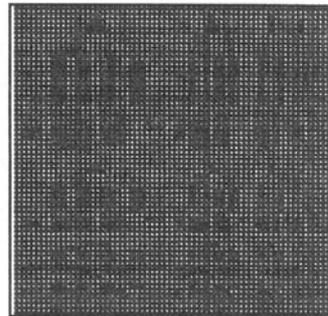
Con la imagen original, realizamos distintas transformaciones con una matriz de partición ordenada de distintas formas. De igual forma que el apartado anterior, hemos representado como guía ilustrativa a la izquierda de las imágenes, la matriz de

transformación equivalente para el caso de un tamaño del campo de datos N=8, obteniendo los siguientes resultados:

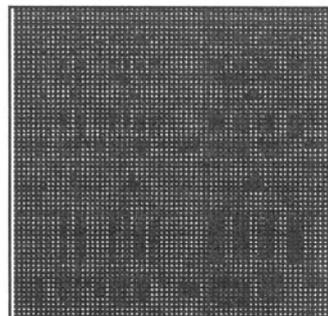
$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & -1 & 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & -1 & 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 & 0 & 0 & 0 & 0 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$



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



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

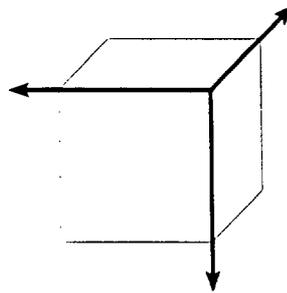


#### 4.5 BASADOS EN LAS TRANSFORMACIONES DIAGONALES.

Otro tipo de transformaciones son las denominadas Transformaciones Diagonales.

Los métodos de generación de coeficientes de peso para los funcionales, utilizados en el estudio anterior, sugieren otros procedimientos sistemáticos para generar transformaciones no ortogonales. Nos referiremos, en lo que sigue a un Campo de Datos donde no se realiza ninguna partición.

Consideremos el espacio de 3 dimensiones y el cubo de lado 2 centrado en el origen:



Sea el vector  $(1,1,1)$  (vértice del cubo). Una forma rápida de generar vectores no ortogonales pero independientes, es "movernos" en la dirección de las tres aristas confluyentes en  $(1,1,1)$ , para obtener los vectores:

$$(-1,1,1), (1,-1,1), (1,1,-1)$$

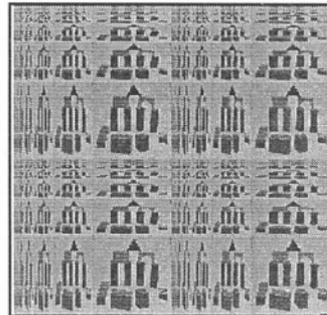
que forman una nueva base, pero no ortogonal. Su representación matricial es:

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

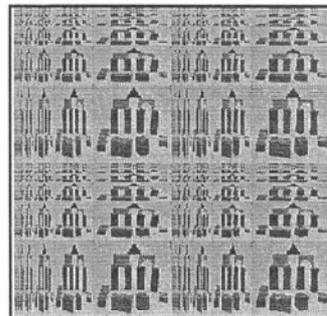
Una matriz de todos +1, excepto la diagonal. Obviamente, los 1's pueden ser números cualesquiera (distintos de cero).

A la imagen original le hemos aplicado distintas transformaciones diagonales obteniendo los siguientes resultados:

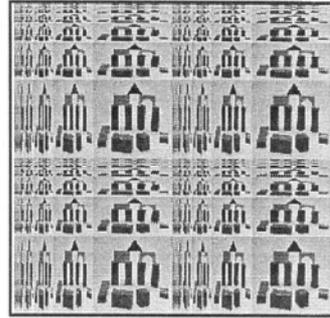
$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$



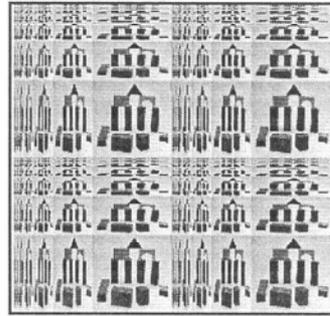
$$\begin{bmatrix} -1 & 2 \\ 2 & -1 \end{bmatrix}$$



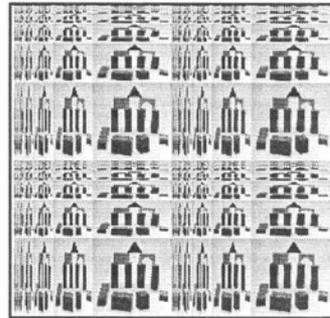
$$\begin{bmatrix} -1 & 4 \\ 4 & -1 \end{bmatrix}$$



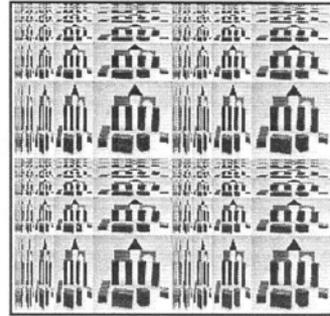
$$\begin{bmatrix} -1 & 8 \\ 8 & -1 \end{bmatrix}$$



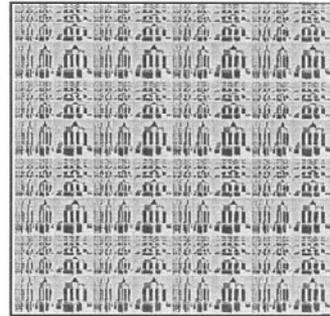
$$\begin{bmatrix} -1 & 16 \\ 16 & -1 \end{bmatrix}$$



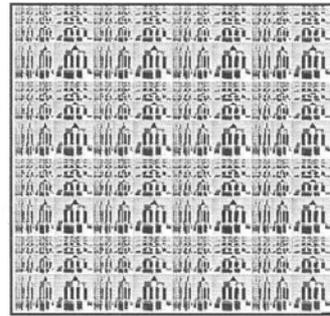
$$\begin{bmatrix} -1 & 32 \\ 32 & -1 \end{bmatrix}$$



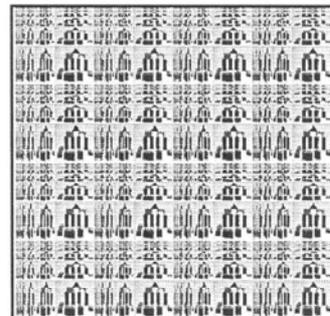
$$\begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix}$$



$$\begin{bmatrix} -1 & 2 & 2 & 2 \\ 2 & -1 & 2 & 2 \\ 2 & 2 & -1 & 2 \\ 2 & 2 & 2 & -1 \end{bmatrix}$$



$$\begin{bmatrix} -1 & 2 & 4 & 8 \\ 2 & -1 & 4 & 8 \\ 2 & 4 & -1 & 8 \\ 2 & 4 & 8 & -1 \end{bmatrix}$$



#### 4.6 BASADOS EN LOS FILTROS DE NEWTON

Una de las propiedades interesantes de los Filtros de Newton es que el conjunto de vectores formado por los Filtros de Newton de una anchura determinada  $L$  son linealmente independientes. Esto es:

Si  $A$  es la matriz de Newton de orden  $L$ , entonces:

$$A^*A = 2^{L-1} I, \text{ con } L \geq 1$$

donde  $I$  es la matriz de identidad. Sean los Filtros de Newton de anchura  $L$  denotados como vectores:  $N_1, N_2, \dots, N_L$  y formemos la correspondiente Matriz de Newton de dimensión  $L \times L$ :

$$A = \begin{bmatrix} N_1 \\ N_2 \\ \dots \\ N_L \end{bmatrix}$$

Por lo anterior sabemos que  $A^*A = 2^{L-1}I$  con lo que por propiedades de los determinantes:

$$\det(A^*A) = \det(A) \det(A) = (2^{L-1})^L$$

y de ahí:

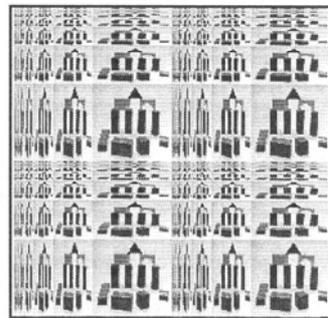
$$|\det(A)| = 2^{(L-1)L/2}$$

lo cual es distinto de cero para todos los posibles valores de  $L$ . De esta forma, si  $\det(A)$

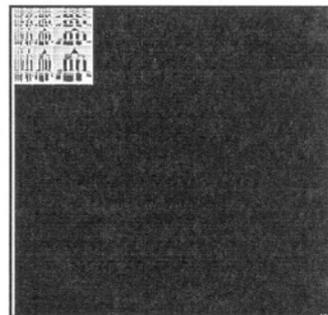
es distinto de cero, sus vectores componentes son linealmente independientes.

A continuación, hemos aplicado distintos Filtros de Newton como funcionales en la creación de distintas particiones. Los resultados obtenidos son:

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$



$$\begin{bmatrix} 1 & 3 & 3 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -3 & 3 & -1 \end{bmatrix}$$



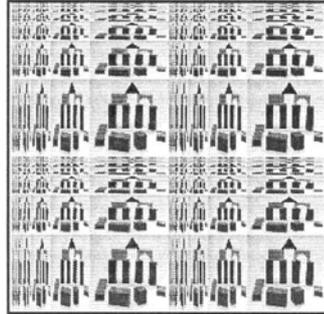
#### **4.7 INCIDENCIA EN LA COMPRESIÓN DE IMÁGENES.**

La compresión de los datos de una imagen es una de las grandes áreas de investigación en proceso de imágenes. Por lo general, las imágenes vienen definidas en un rango de 256X256 o 512X512 pixels. Suponiendo que cada pixel es representado por 8 bits (correspondiendo a las variaciones del nivel de gris entre 0 y 255), esto representa 64K o 256K bytes de memoria. Sería deseable representar la información de la imagen con un menor número de bits y al mismo tiempo ser capaz de reconstruir una imagen que se asemeje lo más posible a la imagen original. La compresión de datos es aplicada principalmente en la transmisión y almacenamiento de la información. En las aplicaciones de transmisión, las técnicas de compresión son obligadas por las consideraciones en tiempo real que limitan el tamaño y la complejidad del hardware. En las aplicaciones de almacenamiento, los requerimientos son menos estrictos, ya que la mayoría de los procesos se realizan offline.

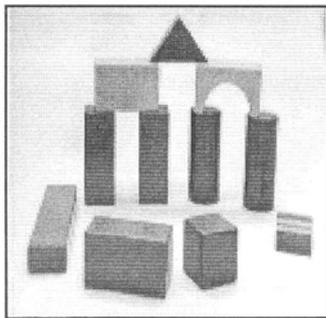
Existen múltiples técnicas que llevan a cabo distintas compresiones de imágenes. En este punto proponemos una nueva técnica de compresión de imágenes basada en lo expuesto anteriormente.

##### **Filtros de Newton.**

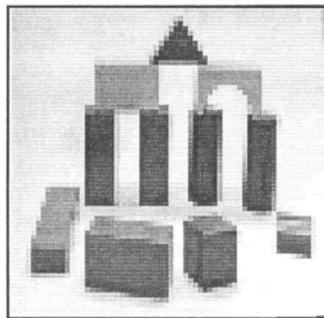
Partamos de la siguiente transformación:



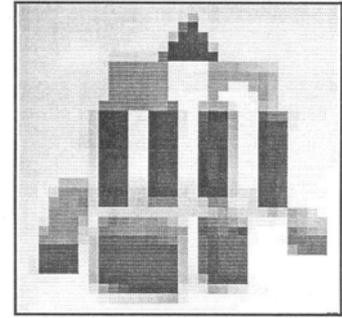
Si a dicha transformación le aplicásemos la inversa obtendríamos exactamente la imagen original. Para llevar a cabo la compresión nos quedaremos con alguna zona de la imagen. Por ejemplo, podemos almacenar 1/4 de imagen como puede ser alguno de los 4 cuadrantes de la transformación. A continuación presentamos distintas imágenes reconstruidas dependiendo del mayor o menor grado de compresión:



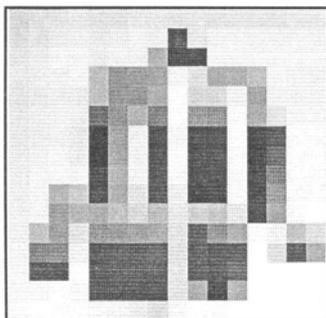
Compresión=1/4



Compresión=1/8



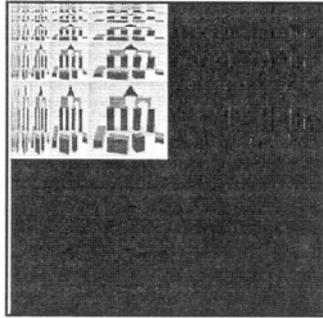
Compresión=1/16



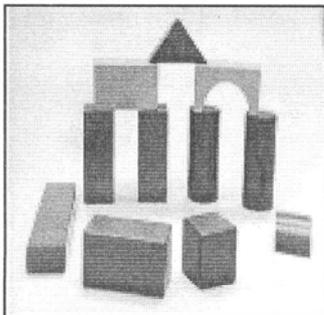
Compresión 1/32

## Haar

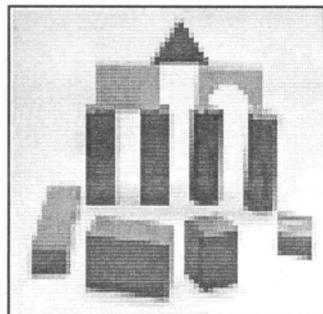
Partamos de la siguiente transformación:



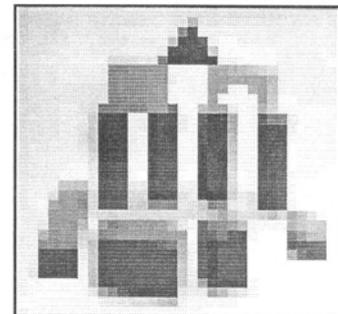
Si a dicha transformación le aplicásemos la inversa obtendríamos exactamente la imagen original. Para llevar a cabo la compresión nos quedaremos con alguna zona de la imagen. Por ejemplo, podemos almacenar 1/4 de imagen como puede ser el cuadrante superior izquierdo de la transformación. A continuación presentamos distintas imágenes reconstruidas dependiendo del mayor o menor grado de compresión:



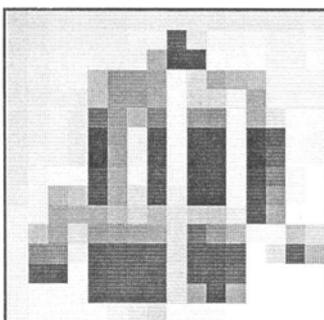
Compresión=1/4



Compresión=1/8



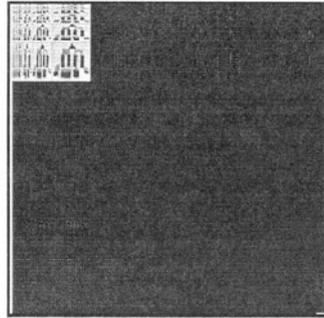
Compresión=1/16



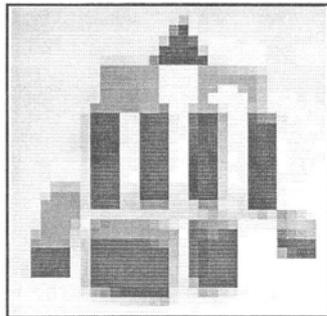
Compresión=1/32

## Hadamard

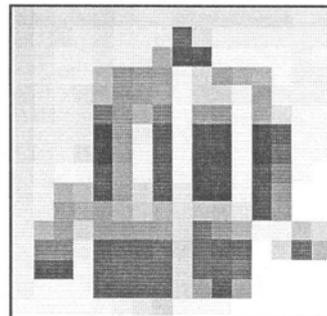
Partamos de la siguiente transformación:



Si a dicha transformación le aplicásemos la inversa obtendríamos exactamente la imagen original. Para llevar a cabo la compresión nos quedaremos con alguna zona de la imagen. Por ejemplo, podemos almacenar  $1/16$  de imagen como puede ser la esquina superior izquierda de la transformación. A continuación presentamos distintas imágenes reconstruidas dependiendo del mayor o menor grado de compresión:

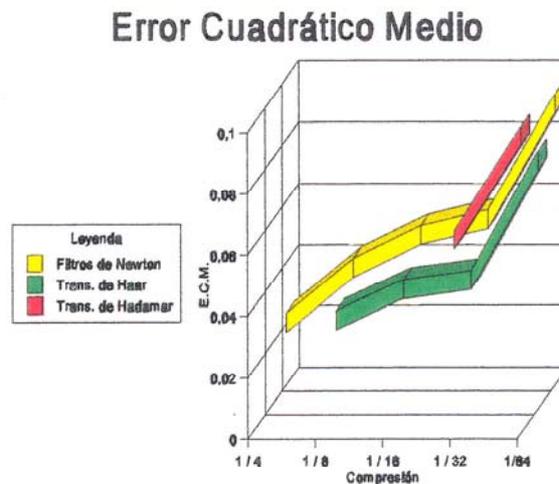


Compresión= $1/16$

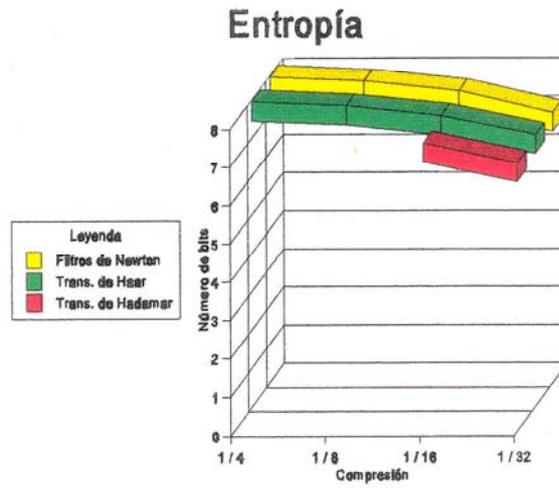


Compresión= $1/32$

Como puede observarse, al llevar a cabo una mayor compresión, se obtiene un peor resultado a la hora de realizar la reconstrucción de la imagen. Para medir la calidad de la imagen reconstruida, hemos calculado, en primer lugar, y de entre las distintas medidas que existen, el error cuadrático medio que existe entre la imagen original y las distintas imágenes reconstruidas. En el siguiente diagrama se muestra el resultado de estas medias:



En la década de los 40, Shannon publicó su famoso paper "A mathematical Theory of Information" [Sha49] que abrió la puerta al análisis cuantitativo de las fuentes de datos y a los sistemas de transmisión. La base de su teoría estriba en el concepto de entropía, que mide la cantidad de información de un mensaje y especifica el número mínimo de bits por pixel necesarios para codificar completamente una imagen. En el siguiente diagrama representamos la entropía calculada a las imágenes reconstruidas anteriormente:



**CONCLUSIONES Y PRINCIPALES**

**APORTACIONES**

La primera de las conclusiones se refiere a las Convoluciones Generalizadas que están basadas en una regla de transformación que se aplica a todos o parte de los valores de entrada. Aplicándole restricciones a la regla de transformación, alcanzamos distintos tipos de convoluciones. Dentro de las convoluciones algorítmicas, proponemos la utilización de los Filtros de Newton para crear núcleos que cumplan las propiedades presentadas en el apartado 1.3.3 y así utilizarlas en transformaciones cuyo objetivo es la normalización de las intensidades. Como aplicación, proponemos utilizarlas en la extracción de bordes de los distintos objetos que aparecen en una escena.

En el capítulo 2, se hizo referencia a la Transformación Foveal, resultado de una TRP con los dominios unitarios colocados en el centro de la retina. En este punto, se propone otra transformación, la Transformación Foveal Dirigida, donde la fovea se pueda dirigir a cualquier punto del campo de datos. Esto serviría para posicionar dicha fovea en la zona de la imagen donde se desee captar la máxima información, dejando para la perifovea la información menos relevante que quedaría más difusa. Como aplicaciones, proponemos utilizar esta nueva transformación como método criptográfico de imágenes y como mecanismo de atención.

En el Capítulo 3, presentamos un estudio relacionado con la representación de imágenes y campos receptivos. Las conclusiones aportadas en este trabajo referente a la influencia de los parámetros que rodean a las transformaciones de campos receptivos son las siguientes:

- 1) Existe un conjunto de configuraciones o distribución de los Campos Receptivos que producen mejores resultados que otros para propósitos de clasificación. Esto enfatiza la relevancia de los campos receptivos frente a los funcionales clásicos. Esto puede comprobarse observando que la variabilidad de los campos receptivos versus funcionales es más pequeña que los funcionales versus campos receptivos.

Dichas configuraciones de los Campos Receptivos que producen mejores resultados son aquellos que engloban las zonas de la imagen que son más discriminantes, como puede ser el contorno de la misma.

2) Corroboramos el hecho de que los momentos que proporcionan mejores resultados para propósitos de clasificación son los de bajo orden.

3) Al aumentar el tamaño del Campo de Datos, aumenta la eficiencia del sistema clasificador.

4) Un aumento del número de descriptores produce una disminución en la eficiencia del sistema clasificador. Posiblemente esto es debido al ruido inherente y al aumento de redundancia las cuales no son controladas en el campo de datos original. Concluimos que el hecho importante no es el número de descriptores sino dónde y cómo son elegidos.

5) En general, el utilizar un tipo de funcional u otro no influye significativamente en los resultados de la clasificación.

6) El tamaño del Campo Receptivo tampoco influye significativamente en los resultados de la clasificación. Sin embargo, observando los resultados cuando se disminuyó el tamaño del campo receptivo, podemos enfatizar el hecho de la elección adecuada de los Campos Receptivos.

Por último, dentro de este capítulo, resaltamos una forma de descripción automática de formas utilizando los campos receptivos.

En el Capítulo 4, estudiamos la variabilidad de los campos receptivos y su

conectividad con la complitud. Estudiamos un método para la generación de funcionales linealmente independientes y lo aplicamos basándonos, primeramente, en distintas transformadas clásicas. Además de éstas, desarrollamos otros tipos de transformadas como son las Diagonales y las basadas en los Filtros de Newton. Resaltamos la introducción de las transformaciones diagonales como puente entre el proceso de campo de datos y los formalismos de campos vectoriales. Concluimos que una forma sencilla de generar funcionales linealmente independientes se alcanza con el método de generación detallado en el apartado 4.2 utilizando distintas transformaciones (Haar, Hadamard, Diagonales y Filtros de Newton). Por último presentamos un marco de aplicabilidad exponiendo su incidencia en la compresión de imágenes, consiguiendo resultados aceptables hasta una reducción de 1/8 de su tamaño de almacenamiento.

Finalmente, exponemos a continuación algunos objetivos a alcanzar en futuras acciones:

- 1.- Realizar un estudio comparativo de distintos métodos criptográficos que actualmente se aplican en el cifrado de información con la propuesta realizada en este trabajo de utilizar la TFD como método criptográfico.
- 2.- Explorar el uso conjunto del mecanismo subyacente a las TFD's y a la descripción automática de formas utilizando los campos receptivos con vista a su aplicación práctica en sistemas de visión artificial para detección de movimiento y visión robótica.
- 3.- Considerar las Redes Neuronales y por Capas incorporando el parámetro tiempo en la formulación algebraico-analítica.

**REFERENCIAS  
BIBLIOGRÁFICAS**

- 
- [Alt62] Alt, F.L.: Digital pattern recognition by moments. ACM9, pp240-258 1962.
- [Bol89] Bolívar Toledo, Olga.: Hacia una Teoría de las Transformaciones en Campos Receptivos y Campos de Datos. Implicaciones en Teoría y Proceso de Imágenes. Tesis Doctoral. Universidad de Las Palmas de Gran Canaria 1989.
- [Bou69] Bourne GH: The structure and function of nervous tissue. Volms I-III Academic Press. Nueva York 1969.
- [Can87] Candela Solá, S.: Transformaciones de Campo Receptivo Variable en Proceso de Imágenes y Visión Artificial. Tesis Doctoral. Universidad Politécnica de Canarias 1987.
- [Cas85] Cassereau P.: A New Class of Optimal Unitary Transform for Image Processing. Master's thesis. MIT, Massachussets. 1985.
- [Gir77] N.C. Giri. Multivariate Statistical Inference, Chaper 10. Academic Press, Inc., Orlando, Fla. 1977.
- [Grü73] Grüsser OJ et al.: Neural mechanisms of visual movement perception and some psychophysical and behavioral correlations. En Hdbk of Sensory Physiol. Parte A. Springer Verlag. New York. 1973.
- [Jun73] Jung, R. :Visual perception and neurophysiology. Handbook of Sensory Physiology. Springer Verloag. New York. 1973.
- [Kol90] Kolb B., Wishaw IQ, Freeman and Co.:Human Neuropsychology. New York 1990.
- [Lei72] Leibovic KN.:Nervous Systems Theory: An Introductory Study. Academic Press. New York. 1972.
- [Lei91] Leibovic, K.N.: Science of Vision. Springer Verlag. Nueva York, 1991.

- 
- [Mat90] Maturana H., Varela F.: El árbol del conocimiento. Ed. Debate. Madrid 1990.
- [McC90] McCulloch WS. Pitts W.: How we know universals. En WS McCulloch. Embodiments of mind. MIT Press Cambridge Mass. USA. 1990.
- [Mir] Mira J.; Comunicación Personal.
- [Mor65] Moreno Díaz, R.: An Analytical Model of the Group 2 Ganglion Cell in the Frog's retina. Informe de Trabajo. Instrumentation Laboratory, MIT, Mass. USA. 1965
- [Mor78] Moreno Díaz, R.; Rubio Royo, E.: A generalized model for non-linear retinal processing. Current topics in Cybernetics and Systems. Ed. Rose J. Springer 1978.
- [Mor79] Moreno Díaz R. Rubio E.: A theoretical model for layered visual processing. Int. J. of Biomed. Computing. Vol 10:231-243. 1979
- [Mor84] Moreno Díaz, R.; Mira Mira, J.: Un marco teórico para interpretar la función neuronal a altos niveles. Biocibernética. Ed. Siglo XXI. Madrid 1984.
- [Mor93] Moreno Díaz, Roberto jr.: Computación Paralela y Distribuida: Relaciones Estructura-Función en Retinas. Redes de Newton. Transformadas de Hermite, Complejidad. Cooperación y Óptimos. Tesis Doctoral. Universidad de Las Palmas de Gran Canaria 1993.
- [Net80] Netravali A.N., Limb O.: Picture Coding: A review. Proc IEEE, vol 63, pp 366-406. 1980
- [Muñ87] Muñoz Blanco, J.A.: Jerarquización de Estructuras de nivel bajo y medio para reconocimiento visual. Aplicaciones a texturas y formas. Tesis Doctoral. Universidad de Las Palmas 1987.
- [Rao90] Rao K., Yip P. : Discrete Cosine Transform. Algorithms, Applications. Academic Press. San Diego, 1990.

- [Rov92] Rovaris Romero, Eduardo: Imágenes, Muestreo y Computación Distribuida de tipo Neuronal. Hacia un Enfoque Unificado. Tesis Doctoral. Universidad de Las Palmas de Gran Canaria, 1992.
- [Sha49] Shannon C.: A mathematical theory of information. 1949
- [Tru65] Truex. Carpenter: Human Neuroanatomy. Williams and Wilkins. New York. 1965
- [Wie58] Wiener N.: Nonlinear problems in random theory. MIT Press, Cambridge, Mass 1958.
- [Wie59] Wiener N.: Cybernetics. MIT Press. 1959.