



DEPARTAMENT INTERUNIVERSITARI D'ÒPTICA
FACULTAT DE FÍSICA
UNIVERSITAT DE VALÈNCIA

**CARACTERIZACIÓN NUMÉRICA DEL SISTEMA VISUAL
COMO FILTRO LINEAL EN UN DOMINIO DE GABOR**

Trabajo de Investigación presentado por:
Jesús Malo López

**CARACTERIZACIÓN NUMÉRICA DEL
SISTEMA VISUAL COMO FILTRO LINEAL
EN UN DOMINIO DE GABOR**

D. José María Artigas Verde, Prof. Titular del Departament d'Òptica de la Universitat de València CERTIFICA QUE:

El trabajo '*CARACTERIZACIÓN NUMÉRICA DEL SISTEMA VISUAL COMO FILTRO LINEAL EN UN DOMINIO DE GABOR*', ha sido realizado por D. **Jesús Malo López** bajo su dirección para optar a la consecución de los créditos correspondientes al Trabajo de Investigación del programa de doctorado 280 A.

Vº Bº Director del Departamento.

J.M. Artigas

ÍNDICE

1 INTRODUCCIÓN: Modelización numérica de la degradación de la información espacial. Planteamiento general y resultados.....	1
1.1 Planteamiento general del problema.....	2
1.1.1 Algunas ideas sobre el sistema visual	
1.1.2 Enfoque utilizado y problema a tratar	
1.1.3 Observaciones acerca del concepto de información	
1.1.4 Sistemas de codificación de bajo nivel	
1.2 Ámbito de nuestro estudio.....	12
1.2.1 Discusiones posibles asumiendo la caracterización del sistema visual como un codificador de bajo nivel.	
1.2.2 Hipótesis de partida	
1.2.3 Aportaciones originales	
2 HERRAMIENTA MATEMÁTICA: Funciones y Transformada de Gabor.....	17
2.1 Funciones de Gabor.....	17
2.1.1 Un poco de historia.	
2.1.2 Funciones de Gabor bidimensionales. Definición.	
2.1.3 Espectro de una función de Gabor 2D.	
2.1.4 Las funciones de Gabor como funciones de área mínima en el dominio conjunto espacio-frecuencial.	
2.1.5 Las funciones de Gabor como filtros pasa banda: control de las características del filtro mediante los parámetros variables en el dominio espacial.	
2.2 Transformada de Gabor.....	38
2.2.1 Descomposición de una función respecto de una base no ortogonal e incompleta.	
2.2.2 Solución general del problema: cálculo iterativo de los coeficientes de la proyección sobre cualquier base.	
2.2.3 Solución particular: cálculo de la transformada respecto de una base de funciones separables.	
3 DISCUSIÓN ACERCA DE LA TRANSFORMACIÓN EFECTUADA POR EL SISTEMA.....	49
3.1 Detectores, campos receptivos y transformaciones.....	50
3.1.1 Implementación fisiológica de la transformación T.	
3.1.2 Codificación en el dominio espacial o el frecuencial en función de la forma de los campos receptivos del sistema de detectores planteado.	
3.2 Ventajas de la utilización de detectores localizados espacial y frecuencialmente.....	53
3.2.1 Argumentos teóricos.	
3.2.2. Argumentos experimentales: datos fisiológicos y psicofísicos.	
3.3 Utilización de la formulación espacio-frecuencial en la actual investigación en visión.....	61
3.3.1 Situación general.	
3.3.2 ¿Funciones de Gabor o wavelets spline de orden n?.	

4 EL SISTEMA VISUAL COMO FILTRO LINEAL EN UN DOMINIO	
CONJUNTO: Función de sensibilidad a estímulos de Gabor (GSF)	65
4.1 <i>Filtrado lineal en un dominio de Gabor</i>	66
4.1.1 Dominios espacio-frecuenciales de Gabor.	
4.1.2 Filtrado lineal en un dominio espacio-frecuencial de Gabor 4D: posibilidades y limitaciones.	
4.2 <i>El sistema visual como filtro lineal en un dominio de Gabor</i>	70
4.2.1 El sistema visual como filtro lineal en el dominio de Fourier: la <i>CSF</i> .	
4.2.2 El sistema visual como filtro lineal en un dominio de Gabor: la <i>GSF</i> .	
4.2.3 Determinación experimental de la <i>GSF</i> .	
4.3 <i>Relaciones matemáticas entre la CSF y la GSF</i>	74
4.3.1 Relación entre dos funciones filtro R_1 y R_2 definidas en los dominios inducidos por las transformadas lineales T_1 y T_2 .	
4.3.2 Relaciones entre la <i>GSF</i> y la <i>CSF</i> .	
4.3.3 Limitaciones de la relación entre la <i>GSF</i> y la <i>CSF</i> .	
5 DETERMINACIÓN DE LA GSF 4D A PARTIR DE LA CSF 2D DE LA ZONA	
FOVEAL	80
5.1 <i>Requisitos para la determinación de la GSF 4D</i>	81
5.2 <i>Selección de la base: modelo de canales de Gabor</i>	82
5.2.1 Recubrimiento del dominio frecuencial.	
5.2.2 Recubrimiento del dominio espacial.	
5.3 <i>Transformada de Gabor respecto de nuestra base</i>	87
5.4 <i>Determinación experimental de los filtros en el dominio de Fourier</i>	93
5.4.1 Consideraciones generales.	
5.4.2 Práctica experimental.	
5.4.3 Filtros en el dominio de Fourier, respuestas impulsionales y simulaciones.	
5.4.4 Comentarios acerca del método. <i>CSFs</i> definitivas.	
5.5 <i>Cálculo de las GSFs equivalentes a las CSFs foveales</i>	103
5.5.1 Resultados.	
5.5.2 Equivalencia entre las <i>GSFs</i> calculadas y los filtros de Fourier iniciales.	
6 GSF DE UN SISTEMA ESPACIALMENTE VARIANTE	111
6.1 <i>Síntesis de un sistema espacialmente variante</i>	112
6.2 <i>Comportamiento del sistema sintetizado</i>	114
7 CONCLUSIONES	120
8 BIBLIOGRAFÍA	122

1. INTRODUCCIÓN.

Modelización numérica de la degradación de la información espacial: planteamiento general y resultados.

El inconveniente fundamental que presenta el estudio de la visión desde el punto de vista físico, es el escaso conocimiento que se tiene, de los procesos concretos de transmisión e intercambio de información entre los elementos del soporte biológico, donde tienen lugar. Para nosotros este inconveniente no es frustrante, en el sentido de que lo que se pretende no es llegar a comprender cómo funciona un sistema en particular, como por ejemplo el cerebro, sino establecer cuáles son los mecanismos que permiten la extracción eficiente de la información contenida en las imágenes naturales. El desconocimiento esencial de los procesos cerebrales relacionados con la visión, representa un problema porque implica que nuestro trabajo estará limitado exclusivamente por la fenomenología del problema, lo cual propicia la aparición de muchos modelos diferentes, difícilmente contrastables, para la explicación de los resultados psicofísicos.

El resultado es que en visión, tan lícito es discutir sobre los detalles de un modelo como sobre su misma base, planteando caracterizaciones alternativas esencialmente distintas. En este trabajo desarrollamos la idea clásica del filtrado lineal utilizando el reciente formalismo de las representaciones según bases de funciones localizadas, dando por sentada la ejecución de una

cierta transformación de las imágenes, la elección de una base, el funcionamiento del sistema como filtro, etc.

La amplitud del problema y el hecho de que la solución permanezca esencialmente abierta, hacen necesaria una justificación adecuada de las hipótesis de las que se parte. Por ello, los tres primeros capítulos están pensados como una introducción general que describa el problema, y justifique el planteamiento mediante el que lo hemos abordado.

El primer capítulo estará dedicado a delimitar cuál es el problema que pretendemos abordar y los diferentes aspectos sobre los que cabe discutir, señalando las hipótesis que asumimos, situando así nuestra contribución en el punto del problema que le corresponde.

En el segundo capítulo analizamos las propiedades de la herramienta matemática que vamos a utilizar: *las funciones de Gabor y la Transformación de Gabor*.

Basándonos en las propiedades de dichas funciones, en el tercer capítulo repasaremos algunos de los argumentos a favor de la utilización de funciones localizadas en el análisis de imágenes y en el modelizado del sistema visual.

1.1 PLANTEAMIENTO GENERAL DEL PROBLEMA.

1.1.1 Algunas ideas sobre el sistema visual.

En contra de lo que sucede en el resto de las ciencias, donde el objeto de estudio determina el método de trabajo, en física, es habitual que el enfoque utilizado, defina el objeto a estudiar. Así ocurre en el caso del análisis físico del problema de la visión. Considerando la visión como la adquisición de información respecto de un entorno, a partir de las imágenes que del mismo genera un sistema óptico, es evidente que al hablar del sistema visual no nos referiremos exclusivamente al sistema óptico formador de imágenes, sino que podremos incluir también el sistema de fotodetección y procesado posterior que permite la extracción y el análisis de la información, necesaria para la formación de conceptos acerca del entorno. Asumiendo esta

definición, determinar a qué llamamos sistema visual es una cuestión de convenio.

En general, se considera que el sistema visual es el que se encarga de la formación, adquisición y codificación de las imágenes. De esta manera, la información que pueda extraerse de la señal codificada estará limitada por la degradación que haya sufrido en los procesos previos efectuados por el sistema visual.

En el caso concreto de los mamíferos superiores, el concepto de sistema visual incluiría tanto al sistema óptico del ojo como a la retina y los canales neuronales que trasladan y transforman la señal procedente de los fotodetectores a través del núcleo lateral geniculado (LGN) hasta el córtex visual; mientras que, si se trata de un sistema de visión artificial, incluiría la cámara, el sistema de codificación de las señales RGB y los primeros algoritmos efectuados por el hardware de la tarjeta digitalizadora y el programa de tratamiento de imágenes.

1.1.2 Enfoque utilizado y problema a tratar.

De acuerdo con el modo en el que hemos establecido el concepto de sistema visual, el enfoque empleado será formal, y centrado en el planteamiento de algoritmos que permitan reproducir cuantitativamente el preprocesado de la información visual realizado por el sistema.

Aunque se acudirá a ciertos resultados fisiológicos para argumentar a favor o en contra de alguna herramienta matemática, o fijar alguno de sus parámetros libres, en ningún momento pretenderemos identificar las transformaciones matemáticas que se utilicen con las trayectorias neuronales del proceso visual y sus propiedades fisiológicas.

En definitiva, algunos resultados fisiológicos pueden servir de base a los algoritmos presentados, pero nuestro objetivo es que el modelo obtenido sea eficiente y explique la psicofísica, y no que sirva para explicar la función desempeñada por el LGN o el área V17 del córtex, por poner un ejemplo.

En este trabajo analizaremos exclusivamente la actuación del sistema respecto de imágenes acromáticas estáticas, obviando por tanto los efectos cromáticos y de movimiento, así como los efectos estereoscópicos y de interacción binocular.

Aquí consideraremos el sistema visual humano como un *codificador eficiente de bajo nivel*, que transforma la señal de entrada (una función $i(\mathbf{x})$ donde los puntos \mathbf{x} pertenecen al dominio bidimensional de posiciones que constituye el campo visual y, $i(\mathbf{x})$ es la distribución de intensidades objeto en dicho campo), en una señal codificada $f(\mathbf{p})$. Los diferentes valores del parámetro \mathbf{p} representan *características disjuntas* que pueden estar presentes en la imagen, de forma, que el valor de $f(\mathbf{p})$ en cada punto \mathbf{p} indica en qué medida está presente la característica representada por \mathbf{p} en la imagen $i(\mathbf{x})$.

Las operaciones realizadas por el SV sobre las imágenes incidentes se denominan *preprocesado* o *codificación de bajo nivel* porque las características respecto de las que se codifica, son características matemáticas sencillas de las funciones base según la que, como veremos, se descompone el estímulo $i(\mathbf{x})$. Después de este tipo de codificación, más allá de los límites de lo que hemos acordado llamar sistema visual, se sitúa el *procesado de alto nivel*, encargado de extraer información a partir de la señal codificada y realizar tareas de clasificación y reconocimiento elaborando descriptores complejos con contenido semántico.

Asumiendo el tratamiento del sistema visual como un codificador de bajo nivel, nuestro propósito es caracterizarlo cuantitativamente planteando una transformación \mathbf{T} tal que :

$$f(\mathbf{p}) = \mathbf{T} [i(\mathbf{x})] \quad (1.1)$$

La transformación \mathbf{T} representa el efecto conjunto de todos los procesos que realiza el sistema visual. En general, \mathbf{T} incluirá una transformación T sobre el dominio \mathbf{P} y una operación de reducción de información R en dicho dominio; es decir:

$$\mathbf{T} = R \circ T \quad (1.2)$$

Con una caracterización de este tipo, dada cualquier señal de entrada $i(\mathbf{x})$, podremos obtener una representación topográfica de la señal codificada:

$$\begin{aligned} f(\mathbf{p}) &= R(T[i(\mathbf{x})]) \\ i'(\mathbf{x}) &= T^{-1}[f(\mathbf{p})] \end{aligned} \quad (1.3)$$

que permita comprobar qué aspectos de la señal, se han perdido en el procesado. Lógicamente, la manipulación de $T [i(x)]$ mediante R implica que la imagen reconstruida a partir de la codificación diferirá de la inicial, es decir: $i'(x) \neq i(x)$

Antes de analizar con detalle las funciones de un codificador de bajo nivel, diremos unas palabras acerca del sentido en que utilizaremos los conceptos de: información, volumen de una señal, reducción de la información, etc., conceptos todos ellos de interés al definir la misión del sistema con el que tratamos.

1.1.3 Observaciones acerca del concepto de información.

Al hablar de sistemas caracterizados por funciones filtro, definidas en un dominio P , suele utilizarse el concepto de información en un sentido muy amplio que en ocasiones lleva a resultados contradictorios con lo obtenido aplicando estrictamente los conceptos de la teoría de la comunicación.

Por ejemplo, al estudiar un sistema como filtro, en el dominio de Fourier, si su función de transferencia es uniforme e igual a la unidad, se dice que la señal de salida, definida en el dominio frecuencial¹, contiene la misma información que la señal de entrada, en el sentido de que es posible recuperar la señal de entrada de forma perfecta. Si por el contrario, el sistema actúa como un filtro pasa baja, se dice que la señal de salida contiene menor información sobre las altas frecuencias y que, por tanto, aporta menor información que la imagen original.

Desde este punto de vista, la pérdida o el mantenimiento de la información se asocia con la distorsión sufrida por la imagen reconstruida, $i'(x)$, a partir de la señal codificada, $f(p)$, respecto de la imagen original, $i(x)$. Sin embargo, como veremos a continuación, aplicando las definiciones de la teoría de la comunicación, la información contenida en la transformada de Fourier de la imagen filtrada, es en general la misma que en la transformada de Fourier sin filtrar; y en cualquier caso, muy superior a la contenida en la imagen original.

¹ Recuérdese que consideramos que la salida del sistema es $f(p) = R[T[i(x)]](p)$ y no $i'(x) = T^{-1}[f(p)]$

Esta contradicción surge del hecho de que una pérdida de información, en el sentido estricto de los términos, efectivamente provoca la imposibilidad de obtener una reconstrucción perfecta de la señal original, pero lo contrario no es siempre cierto. Digamos que el concepto de *sistema que degrada la señal* incluye el concepto de *sistema que elimina información*, pero es más amplio que éste.

A continuación recordaremos, sin intención sistemática, los conceptos básicos sobre la medida de la información contenida en una señal, para clarificar las afirmaciones anteriores. En primer lugar, definiremos los conceptos de fuente de señales y entropía de la fuente para introducir el concepto de información de una señal como medida de la reducción de entropía de la fuente.

Fuente de señales.

Fenómeno físico que toma valores (continuos o discretos) en un determinado rango según una cierta distribución de probabilidad.

Cada punto de una escena 2D es una fuente de señales, ya que la intensidad i en cada punto x toma valores $i(x)$ según una densidad de probabilidad $P(x, i)$.

En el caso más general, las diferentes fuentes puntuales *no son independientes*.

Entropía de una fuente (H).

La entropía H de una fuente, es una medida de su aleatoriedad, de la incertidumbre existente respecto de la posibilidad de que tome uno u otro valor dentro de su rango de variación.

La formulación matemática de entropía debe satisfacer ciertas propiedades intuitivas:

- La entropía (aleatoriedad o indeterminación en el resultado) debe crecer con el incremento de posibilidades de elección, con lo cual, debe depender de:

- El número de valores posibles en el rango de variación de la fuente: A mayor número de valores posibles, mayor indeterminación y mayor H .
- La forma de la densidad de probabilidad del valor i en su rango de variación: H crecerá cuanto más uniforme sea la función densidad de probabilidad. La máxima indeterminación se tiene cuando todos los valores sean equiprobables.
- Si dos fuentes *independientes* se consideran como una sola fuente, la entropía de la fuente compuesta debe ser igual a la suma de las entropías fuentes simples.
- En los años cuarenta Shannon[1] propuso una expresión para H que cumple las propiedades citadas. Definió la entropía de una fuente F que puede tener N valores i_l con $l= 1, \dots, N$ según una distribución de probabilidad $P(i_l)$ dada por:

$$H(F) = - \sum_{l=1}^N P(i_l) \cdot \log_2 P(i_l) \quad (1.5)$$

En esta definición se toma un logaritmo de base 2 por convenio, porque así, la definición de la unidad de entropía resulta más intuitiva.

Unidades de medida de la entropía.

Tomando el logaritmo en base 2, conseguimos que la entropía sea la unidad para una fuente que posea dos estados equiparables: ($P(i_1)=P(i_2)=1/2$):

$$H(F) = - \sum_{l=1}^N P(i_l) \cdot \log_2 P(i_l) = - \frac{1}{2} \cdot \log_2 \left(\frac{1}{2} \right) - \frac{1}{2} \cdot \log_2 \left(\frac{1}{2} \right) = \frac{1}{2} + \frac{1}{2} = 1$$

Así, definimos la unidad de entropía, el bit, como la entropía de una fuente con dos estados equiprobables. Una fuente con una entropía de m bits es, por extensión, aquella que puede tomar 2^m valores, con igual probabilidad.

Información contenida en una señal S sobre la fuente F.

Una *señal* indica algunos o todos los valores que la fuente ha tomado, reduciendo así la indeterminación de la fuente.

La *información* contenida en una señal se define como la reducción de la indeterminación que provoca su conocimiento.

La *información que contiene la señal* se define como la diferencia entre la indeterminación inicial H_i , previa al conocimiento de la señal, y la indeterminación final H_f , que permanece una vez conocida la señal:

$$I(s) = H_i(F) - H_f(F) \quad (1.6)$$

Si la señal indica el valor de todas las variables de la fuente, la indeterminación final es nula, con lo cual :

$$I(s) = H_i(F) = H(F) \quad (1.7)$$

Esta definición implica que la información tiene dimensiones de entropía, y por lo tanto también se mide en bits.

Una señal de m bits, es aquella que puede contener toda la información procedente de una fuente cuya entropía es de m bits.

Una vez introducidos los conceptos de entropía e información veamos cómo la degradación o no degradación de la señal, no coincide exactamente con el concepto matemático de pérdida o mantenimiento de la información.

Lo primero que hay que hacer notar es que la entropía de la fuente cambia al pasar de una representación a otra[2]. Efectivamente, si consideramos una escena como fuente de señales en el dominio espacial, formada por $N \times N$ fuentes puntuales (píxeles) con una entropía de m bits/píxel¹, al pasar a otra representación mediante una transformación lineal de la forma:

$$f(p) = \sum_{\bar{x}} g(\bar{x}, p) \cdot i(\bar{x}) \quad (1.8)$$

¹ Su entropía global será $N^2 \cdot m$ bits sólo si los píxeles son independientes entre si.

el valor en cada punto depende de varios, o en general de todos los puntos de la imagen inicial, con lo cual, si cada píxel puede tomar 2^m valores diferentes, cada coeficiente $f(p)$ podrá tomar $(2^m)^{N \times N}$ valores diferentes, de forma que la entropía de la fuente en esta nueva representación será de: $m \cdot N \times N$ bits/coeficiente!!

Al efectuar un filtrado en el nuevo dominio multiplicando cada coeficiente $f(p)$ por un peso diferente, no se reducen las posibilidades de variación de la fuente, sólo se comprime o dilata el rango de variación de cada coeficiente $f(p)$; no se reduce el número de valores diferentes que puede tomar en dicho rango, con lo cual, no se modifica la entropía respecto de la señal sin filtrar, manteniéndose los $m \cdot N^2$ bits/coeficiente.

Este aumento espectacular en la entropía de la fuente al cambiar de representación, es estrictamente cierta cuando en la señal inicial $i(x)$ el valor de i en cada píxel no depende de su entorno. Si, por el contrario, la fuente no genera la intensidad aleatoriamente punto a punto, como es el caso de las escenas naturales con las que se enfrenta el sistema visual, la densidad de probabilidad de los valores $f(p)$ no será uniforme en su rango de variación, de forma que, aunque se incremente el número de valores posibles, pueda mantenerse o incluso reducirse la entropía de la fuente en dicha representación. Lo que sí es cierto, es que en general, la entropía varía al cambiar de representación y que esta variación depende de la interacción entre los coeficientes al efectuar la transformación y de cómo se modifique la densidad de probabilidad de los valores $f(p)$ en su rango de variación.

El objetivo de todas estas matizaciones es poner de manifiesto la ambigüedad que supone hablar de reducción de información al tratar con caracterizaciones del sistema como un filtro, si no se tiene en cuenta la estadística del tipo de imágenes a tratar en el dominio considerado.

De todos modos, el concepto intuitivo de reducción de la información, entendido como degradación de la imagen, tiene cierta relación con el concepto estricto de reducción de entropía cuando el proceso de degradación tiende a simplificar la descripción de la señal, de forma que, ésta puede quedar definida, por ejemplo, por un menor número de coeficientes. Desde este punto de vista, un filtrado pasa-baja puede considerarse como una reducción de información de la imagen en el sentido de que disminuye la importancia de ciertos

coeficientes, de forma que podrían obviarse, pudiéndose obtener así una reducción estricta del volumen de la señal.

1.1.4 Sistemas de codificación de bajo nivel.

El objetivo de la codificación de bajo nivel es facilitar las tareas de extracción de conceptos y clasificación que deben realizarse a posteriori, por ello, los criterios para valorar la eficiencia de un determinado sistema codificador serán básicamente dos:

- La capacidad que tenga de reducir el volumen de la señal codificada respecto del volumen de la señal incidente de forma que se facilite la transmisión y el almacenamiento de la señal codificada y se reduzca el coste computacional de las operaciones posteriores.
- La utilidad de la codificación para la extracción de determinada información a partir de la señal codificada. Así, por ejemplo, la clasificación de objetos en determinadas clases será más sencilla cuando, con la codificación utilizada, los objetos de la misma clase, se sitúen en regiones linealmente separables del espacio de coeficientes p , de forma que la condición matemática que diferencie las clases sea sencilla.

De acuerdo con lo dicho, González[2] señala que los procesos básicos que debe llevar a cabo un codificador de bajo nivel son:

- Transformar la función $i(x)$ descomponiéndola respecto de una base de funciones dependientes del parámetro p , obteniendo una representación $f(p)$ de la señal en el dominio de valores de p .
- Modificar la señal en el nuevo dominio para eliminar parte de la información superflua o resaltar los aspectos de interés.

Esquemáticamente:

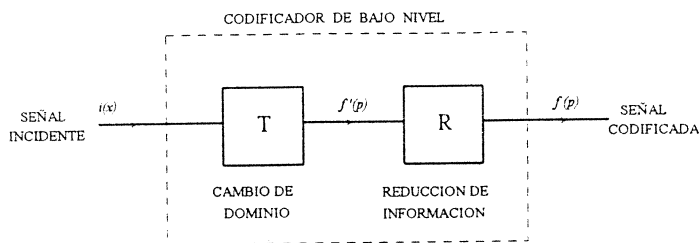


figura 1.-

Es conveniente representar la señal en otro dominio porque en el dominio espacial, la redundancia de las imágenes naturales es muy alta, en el sentido de que el valor de la función $i(x)$ en un punto x_0 está muy relacionado con el valor de la intensidad en su entorno, con lo cual, la información real que aporta cada píxel (la incertidumbre que elimina) es pequeña[3]. Esto es negativo, por dos razones: primero, almacenar la señal de esta manera es ineficiente debido a que cada píxel aporta poca información utilizándose muchos coeficientes para describir escenas de baja entropía, y segundo, porque en el dominio espacial es difícil establecer criterios fijos para la eliminación de la información superflua.

Será preferible representar la señal en un dominio donde la redundancia sea menor, cada coeficiente $f(p)$ contenga más información (permitiendo conservar lo substancial con una señal de menor volumen) y ésta tenga un significado que nos permita establecer un criterio fijo, para así centrar nuestra atención en unas regiones del dominio p en detrimento de otras menos interesantes. Ambas cuestiones: reducir el volumen de la señal y poder establecer criterios fijos para la eliminación de información tienen que ver con la estadística de las imágenes naturales en el dominio considerado. En el apartado anterior, hemos visto que el volumen de la señal que necesitamos para contener la información de una escena, depende de la representación utilizada ya que, la entropía depende de la uniformidad de la función de densidad de probabilidad de los valores de los coeficientes en cada representación

Por otra parte, si la estadística de las imágenes en el dominio considerado es muy asimétrica, será más sencillo decidir qué zonas debemos considerar fundamentalmente, y cuales despreciar o codificar con menor precisión.

El proceso de eliminación de información, en su sentido amplio ya comentado en el apartado anterior, puede iniciarse en la transformación misma: si la base de funciones sobre la que se proyecta $i(x)$ es incompleta, será imposible obtener una reconstrucción perfecta de la señal inicial a partir de $f(p)$, con la consiguiente pérdida de información. Sin embargo, si la transformación es completa $f(p)$ contiene toda la información presente en $i(x)$.

Los procesos de reducción de información efectuados sobre la señal en el dominio de los parámetros p pueden consistir en filtrados lineales, como por ejemplo, el efectuado por un sistema óptico en el dominio frecuencial; filtrados no lineales, como por ejemplo los que tienen lugar al implementar una cuantización vectorial¹ o en un sistema adaptativo de ganancia variable; o procesos en los que intervenga no sólo cada coeficiente $f(p)$ sino también su entorno, como por ejemplo, en las sumaciones y promediados de $f(p)$ sobre áreas del dominio P .

Así al tratar el sistema visual como un codificador de bajo nivel de imágenes acromáticas, el problema a resolver es determinar e implementar numéricamente la transformación T de la señal incidente y las operaciones R sobre la señal en el dominio P , lo cual es equivalente a proponer cuál debe ser la salida, $f(p)$, del sistema, de manera que se cumplan en lo posible los criterios de eficiencia y los resultados psicofísicos.

1.2 ÁMBITO DE NUESTRO ESTUDIO.

1.2.1 Discusiones posibles asumiendo la caracterización del sistema visual como un codificador de bajo nivel. Ámbito de nuestro estudio.

Habiendo centrado el problema en la determinación de la transformación T (transformación de dominio T y posterior transformación R) varias son las cuestiones básicas que pueden plantearse; cuestiones sobre las que aún se discute y que, como ya hemos dicho, no pretendemos resolver aquí.

¹ Una cuantización vectorial[4] representa una reducción de información en el sentido estricto de los términos, porque implica reducir el número de niveles posibles de la señal, reduciendo así su aleatoriedad y su contenido informativo.

El debate central acerca del modelo a utilizar, tiene dos aspectos fundamentales:

- Plantear de que tipo es la transformación T , es decir, discutir sobre cuál es el dominio P sobre el que debe proyectarse $i(x)$. En este punto se incluye además discutir acerca de si debe utilizarse una transformación completa o no, y en el caso de que existan varias bases de funciones posibles, elegir cuál es la que debe utilizarse.
- Decidir cómo se caracteriza la acción del codificador para reproducir la distorsión que introduce en la señal original. Si la transformación T es completa, esta cuestión se reduce a discutir sobre la naturaleza de R . Antes se han nombrado varias posibilidades para implementar los algoritmos de reducción de la información sobre la señal transformada. Si se acepta la opción del filtrado lineal habrá que definir los coeficientes de la función filtro en el dominio P . Si se opta por un modelo no lineal, como una cuantización vectorial o un filtro de ganancia variable, habrá que definir cuáles son los intervalos de cuantificación [4,5], o el mecanismo de ajuste de ganancia respectivamente.

Todas estas cuestiones, pueden abordarse desde dos puntos de vista: por una parte, puede analizarse la eficiencia objetiva de cada modelo decidiendo cuál debe ser la transformación o las operaciones posteriores para que, según la estadística de las imágenes tratadas, la codificación sea óptima según algún criterio de reducción de entropía sin eliminación de *información esencial*. Por otro lado, puede investigarse en las bases fisiológicas de los diferentes modelos relacionando por ejemplo, las funciones base respecto de las cuales se realiza la transformación, con los campos receptivos de las células del córtex y seleccionando la base empleada de acuerdo con la distribución en el dominio P de los campos receptivos.

1.2.2 Hipótesis de partida.

Cualquiera de las opciones planteadas representa, con sus ventajas y sus inconvenientes, una forma diferente de abordar el problema desde su misma raíz. Nuestro objetivo aquí no es

plantear un modelo alternativo, ni demostrar rigurosamente que el modelo elegido es el más eficiente o al que efectivamente responde nuestro sistema visual.

Recopilando las ventajas expuestas en la bibliografía a favor de una transformación cuyas funciones base estén simultáneamente localizadas en el dominio espacial y frecuencial (véase el apdo. 3.2), partiremos de la hipótesis de que la transformación T es una descomposición de la imagen original respecto de una base completa de funciones de Gabor.

Para implementar la transformación empleamos *ad hoc* una base de funciones inspirada en un modelo ya utilizado adaptándola convenientemente a nuestras necesidades[6].

Nuestra contribución se centra en proponer que el proceso posterior de reducción de información en el dominio espacio-frecuencial asumido, puede caracterizarse mediante un filtrado lineal análogo al que se plantea en otros modelos que asumen otra transformación (la de Fourier, por ejemplo), y calcular la función filtro correspondiente.

Hemos aceptado las limitaciones esenciales que implica un modelo de filtro lineal (véase el apartado 4.2), porque la relación de la función filtro en este dominio, con las funciones análogas en otros dominios es sencilla (véase el apartado 4.3) y disponemos del material experimental para la determinación de la función filtro en el dominio de Fourier, con lo cual, su obtención será más o menos directa.

Para centrar ideas, podemos esquematizar todo lo dicho hasta ahora en tres hipótesis básicas, que asumimos como punto de partida para nuestro estudio:

- El preprocesado de las imágenes acromáticas estáticas que efectúa el sistema visual, puede modelizarse como un codificador de bajo nivel; es decir, su comportamiento quedará totalmente descrito por la transformación T que aplica sobre la señal de entrada. La transformación T se compone de un cambio de dominio T y unas operaciones R de distorsión (y simplificación) de la señal en el nuevo dominio.
- La transformación T aplicada es una descomposición de la señal respecto de una base completa de funciones de Gabor; es decir, el dominio P donde queda definida la

señal de salida, es un dominio espacio-frecuencial de dimensión 4 (véase el apartado 4.1.1) donde cada punto p está definido por dos coordenadas espaciales y dos frecuenciales.

- El proceso R de reducción de información en el dominio p es un filtrado lineal de los coeficientes de la transformación T .

1.2.3 Aportaciones originales.

Sobre la base de estas hipótesis, nuestro trabajo ha tenido dos tipos de resultados:

Aportaciones teóricas:

- Caracterización del sistema visual como un filtro lineal en un dominio de Gabor, introduciendo el concepto de Función de Sensibilidad a estímulos de Gabor (GSF), poniendo de manifiesto las ventajas de tal representación en la caracterización de sistemas espacialmente inhomogéneos.
- Establecimiento de la relación matemática entre la función filtro en un dominio de Gabor (GSF), y la función filtro en el dominio de Fourier (Función de Sensibilidad al Contraste: CSF).
- Estudio de las limitaciones de la relación entre la GSF y la CSF con propuestas para superar dichas limitaciones.

Resultados experimentales y simulaciones:

- Establecimiento de la caracterización de Fourier.
 - Medida experimental de la función filtro (CSF), en todo su dominio bidimensional para obtener una modelización cuantitativa del funcionamiento del sistema.
 - Simulación de la respuesta impulsional y del procesamiento de imágenes complejas mediante las CSF s experimentales.
- Establecimiento de la caracterización de Gabor.

- Determinación de la función filtro en el dominio 4D de Gabor a partir de la *CSF 2D* experimental y de la base de funciones elegida.
- Demostración de la equivalencia de ambas caracterizaciones para sistemas espacialmente invariantes (zona foveal).
- Introducción de inhomogeneidades espaciales en la *GSF*.

Como se ve, lejos de plantear una alternativa radical, o tratar de justificar rigurosamente un modelo concreto de distribución de campos receptivos, nuestro objetivo ha sido simplemente destacar las potencialidades de las representaciones espacio-frecuenciales para caracterizar el comportamiento del sistema visual. En este sentido, nuestra contribución se ha centrado en la modelización simple (como filtro lineal) del proceso de reducción de información R , asumiendo que la transformación de codificación es una TG .

En este trabajo se han dado unos primeros pasos en el manejo de los dominios conjuntos estudiando una transformada conjunta general, y estableciendo una relación, también general, entre funciones filtro definidas en dominios diferentes.

El trabajo futuro (previo al inicio de experiencias psicofísicas con este tipo de funciones) debe estar orientado a refinar la formulación matemática para superar las dificultades (apdo. 6.2) asociadas a la base que hemos utilizado en nuestro caso particular (apdo. 5.2).

Una vez resueltos estos problemas (encontrada una base más estable), podremos aplicar las ideas aquí expuestas y plantearnos modelizaciones de R que vayan más allá del simple filtrado lineal.

2. HERRAMIENTA MATEMÁTICA.

Funciones y Transformada de Gabor.

Según lo dicho en el capítulo anterior, el objetivo de una caracterización numérica del comportamiento del sistema visual, es la determinación de las transformaciones T y R de codificación y reducción de información respectivamente. Como también se ha dicho antes, asumimos que la transformación T es una descomposición de la imagen incidente respecto de un conjunto de funciones $g(x,p)$ de Gabor, es decir, se trata de una transformada de Gabor.

En este capítulo, estudiaremos las propiedades matemáticas de este tipo de funciones y analizaremos el problema del cálculo de esta transformación T . En el siguiente capítulo utilizaremos dichas propiedades para justificar la elección de este tipo de transformada.

2.1 FUNCIONES DE GABOR.

2.1.1 Un poco de historia.

En el ámbito del procesado de señal, las funciones armónicas moduladas por gaussianas se conocen como funciones de Gabor en honor a D. Gabor que en su *Teoría de la Comunicación*[7] (1945) propuso este tipo de funciones como base óptima para la

representación de cualquier señal en lo que él llamó *diagrama de información*: un domino conjunto tiempo-frecuencia.

Gabor introdujo la descripción de las señales en un dominio tiempo-frecuencia porque, hacía poco que Nyquist[8] había justificado matemáticamente la limitación en la transmisión de datos mediante una señal de ancho de banda y duración limitadas: efectivamente, como, dada una señal $s(t)$ periódica con periodo Δt (véase la figura 2.a)

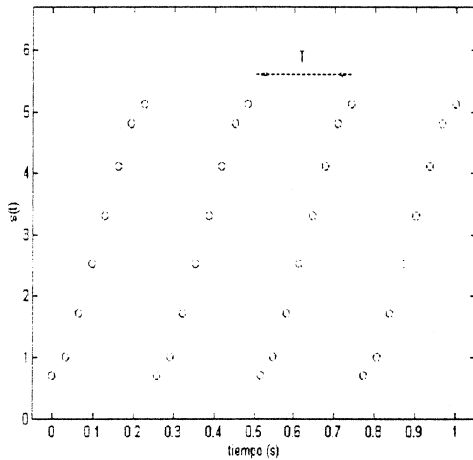


Fig 2.a: Señal $s(t)$ con periodo T

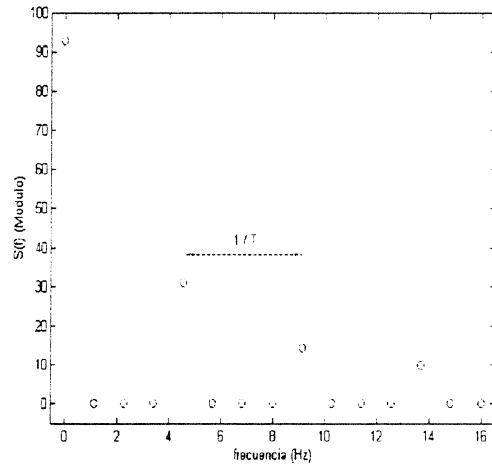


Fig 2.b: Espectro de $s(t)$

su espectro será discreto (fig. 2.b) con separación entre líneas igual a $1/\Delta t$ y cada coeficiente de $S(f)$ da dos datos independientes (módulo y fase) no contenidos en el resto del espectro, la cantidad de datos independientes que se pueden enviar mediante la señal $s(t)$ en cada intervalo Δt será proporcional al área que la señal ocupa en el dominio conjunto:

$$N = 2 \frac{f_M - f_m}{\frac{1}{\Delta t}} = 2 \Delta f \Delta t \propto \Delta f \Delta t \quad (2.1)$$

ya que el número de coeficientes por unidad de frecuencia es de $1/\Delta t$.

La limitación del número de datos independientes que puede aportar una señal en función de su ancho de banda y duración temporal sugirió a Gabor que el área que una señal ocupa en un dominio conjunto tiempo-frecuencia podría considerarse una medida de la cantidad de información¹ que contiene la señal.

Por otra parte, apoyándose explícitamente en la formulación de la mecánica cuántica, Gabor recuerda la imposibilidad de conseguir señales con un área arbitrariamente pequeña en el dominio conjunto debido a la relación de incertidumbre:

$$\Delta f \Delta t \geq \text{cte} \quad (2.2)$$

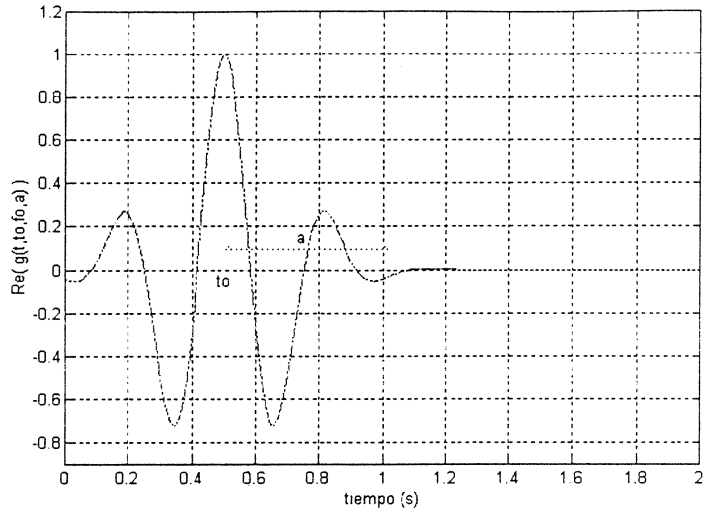
con lo cual, deduce que la base apropiada respecto de la cual descomponer las señales para su análisis en este tipo de dominios conjuntos debía estar formada por funciones que minimizasen su área $\Delta f \Delta t$ y permitiesen recubrir todo el dominio con mínimo solapamiento, constituyendo así una base completa y mínimamente redundante. Como, según Gabor, el área que ocupa una función en el dominio tiempo-frecuencia es proporcional a información que puede contener, llama, un tanto grandilocuentemente, señales elementales o *Logones* (cuantos de información o unidades indivisibles del diagrama de información), a aquellas funciones con área $\Delta f \Delta t$ mínima.

Gabor recoge de nuevo ideas de la mecánica cuántica al recordar que las funciones que satisfacen la igualdad para en la relación (2.2) son los *estados de mínima incertidumbre* o *paquetes gaussianos*: ondas planas moduladas por una envolvente gaussiana:

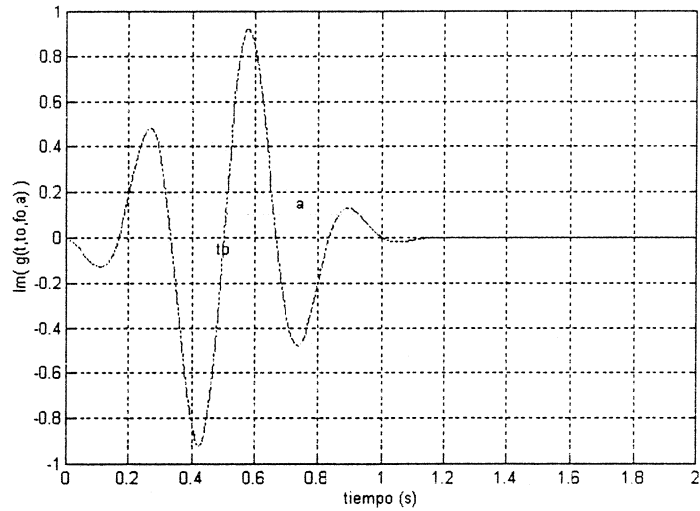
$$g(t, t_0, f_0, a) = e^{-\pi \left(\frac{t-t_0}{a} \right)^2} e^{i(2\pi f_0 t + \phi)} \quad (2.3)$$

En 1988, en pleno renacimiento del interés por el procesado de señal en dominios conjuntos, Daugman[9] demuestra que esta propiedad es generalizable al caso bidimensional y esgrime este hecho como argumento a favor del análisis de imágenes mediante una base de funciones de Gabor 2D.

¹ Gabor no estaba utilizando el concepto estricto de información definido por Shannon poco después.



(a)



(b)

Fig. 3: Parte real (a) e imaginaria (b) de una función de Gabor unidimensional. El parámetro t_0 controla la posición central, f_0 la frecuencia de oscilación, y a la anchura de la gaussiana. En este caso, $t_0 = 0.5$ s, $f_0 = 2$ Hz, y $a = 0.5$ s.

2.1.2 Funciones de Gabor bidimensionales: definición.

En el análisis de imágenes tratamos con señales definidas en el dominio bidimensional de posiciones espaciales formado por los puntos $\mathbf{x}=(x, y)$. De la misma manera que en el caso 1D, una función de Gabor general definida en este dominio vendrá dada por el producto de dos funciones:

- Una función armónica bidimensional cuya oscilación se extiende perpendicularmente al vector $\mathbf{f}_0 = (f_{0x}, f_{0y})$ con frecuencia $|f_0|$, con origen en un determinado $\mathbf{x}_0 = (x_0, y_0)$ y fase ϕ .
- Una gaussiana de base elíptica centrada en el punto \mathbf{x}_0 y cuyo eje mayor forma un ángulo α con el eje x .

$$g(\mathbf{x}, \mathbf{x}_0, \mathbf{f}_0, a, b, \alpha, \phi) = e^{-\pi \left(\frac{x'^2}{a^2} + \frac{y'^2}{b^2} \right)} e^{i(2\pi \mathbf{f}_0 \cdot (\mathbf{x} - \mathbf{x}_0) + \phi)} \quad (2.4)$$

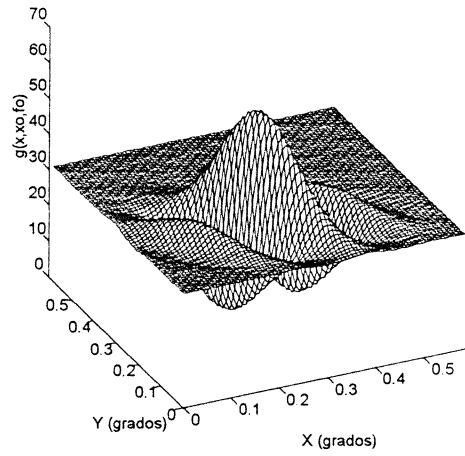
donde $\mathbf{x}' = (x', y')$ resulta de una traslación y un giro de ejes:

$$\mathbf{x}' = G_{-\alpha} [T_{-\mathbf{x}_0}(\mathbf{x})] \quad (2.5)$$

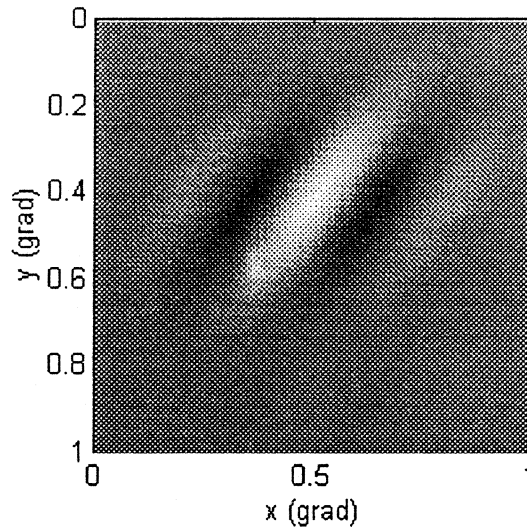
explícitamente:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(-\alpha) & \text{sen}(-\alpha) \\ -\text{sen}(-\alpha) & \cos(-\alpha) \end{bmatrix} \cdot \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix} \quad (2.6)$$

Para describir el dominio espacial utilizaremos las unidades en las que se mide el campo visual. La posición de un punto se da en grados respecto del eje óptico que subtiende dicho punto. La medida de longitudes en grados implica que las frecuencias espaciales (ciclos por unidad de longitud) se expresen en ciclos por grado. En la figura 4.a se muestran los valores de la parte real de la función para unos valores concretos de los parámetros, y en 4.b, el patrón de luminancias que corresponde a dicha función.



(4.a)



(4.b)

Fig. 4: Parte real de una función de Gabor bidimensional centrada en $x_0=(0.5,0.4)$ deg, con una frecuencia $f_0=(3,2)$ c/deg, anchuras $a=0.5$ deg y $b=0.3$ deg, fase $\phi=0^\circ$, e inclinación $\alpha=25^\circ$

2.1.3 Espectro de una función de Gabor 2D.

Para calcular la descomposición frecuencial de una función de Gabor aplicaremos las propiedades de la transformada de Fourier 2D [10]. Si una función $f(x)$ tiene como transformada de Fourier la función $F(f)$, se cumple que:

1. La introducción de una oscilación de frecuencia f_0 en el dominio espacial provoca un desplazamiento f_0 en el dominio frecuencial:

$$e^{i 2\pi f_0 x} f(x) \leftarrow TF \rightarrow F(f - f_0)$$

2. Un giro de la función entorno al origen un ángulo α en el dominio espacial induce un giro análogo en el espectro:

$$f(G_{-\alpha}(x)) \leftarrow TF \rightarrow F(G_{-\alpha}(f))$$

3. Un desplazamiento x_0 en el dominio espacial induce una oscilación armónica en el dominio frecuencial con máximos separados una distancia $1/|x_0|$.

$$f(x - x_0) \leftarrow TF \rightarrow e^{-i2\pi f x_0} F(f)$$

Como nuestra función problema puede expresarse como:

$$g(x, x_0, f_0, a, b, \alpha, \phi) = f(G_{-\alpha}(x - x_0)) e^{i(2\pi f_0(x - x_0) + \phi)}$$

donde:

$$f(x) = e^{-\pi\left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)} \quad (2.8)$$

si aplicamos las propiedades [1 - 3] se cumplirá que:

$$TF \left[f(G_{-\alpha}(x - x_0)) e^{i(2\pi f_0(x - x_0) + \phi)} \right] (f) =$$

$$\text{Prop. 1} \Rightarrow = e^{i\phi} TF \left[f(G_{-\alpha}(x - x_0)) \right] (f - f_0)$$

$$\text{Prop. 2} \Rightarrow = e^{i\phi} TF \left[f(x - x_0) \right] (G_{-\alpha}(f) - f_0) \quad (2.9)$$

$$\text{Prop.3} \Rightarrow = e^{-i (2 \pi x_0 (G_{-\alpha}(f) - f_0) - \phi)} TF[f(x)] (G_{-\alpha}(f) - f_0)$$

y como la transformada de $f(x)$ es[10]:

$$TF[f(x)] (f) = e^{-\pi (a^2 f_x^2 + b^2 f_y^2)} \quad (2.10)$$

resulta que el espectro de $g(x, x_0, f_0, a, b, \alpha, \phi)$ será otra función de Gabor definida en el dominio frecuencial:

$$TF[g(x, x_0, f_0, a, b, \alpha, \phi)] = G(f, f_0, -x_0, 1/a, 1/b, \alpha, \phi) \quad (2.11)$$

En la figura 5.a podemos observar la parte imaginaria de la función de Gabor definida en la figura 4 y la parte imaginaria de su espectro.

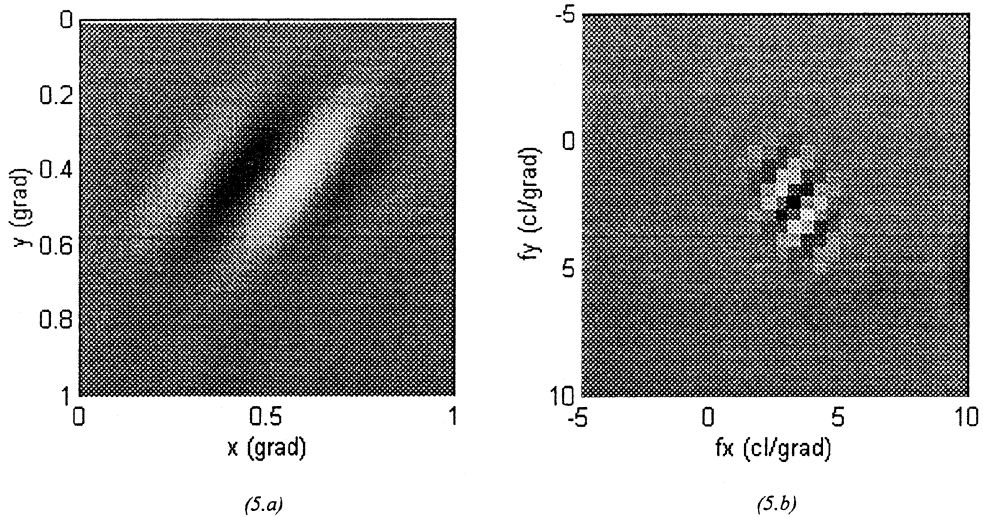


Fig.5: Patrón de intensidades -parte imaginaria- (a) y espectro (b) de la función de Gabor con los parámetros especificados en la figura 4. Nótese que el espectro es también una función de Gabor: una oscilación con vector de ondas $-x_0$ modulada por una gaussiana centrada en f_0 inclinada un ángulo α y con anchuras $1/a$ y $1/b$.

En el apartado 2.1.5 se analizan con detalle las implicaciones de la relación (2.11).

$$g(x) = e^{-\frac{1}{2}(x-x_0)^T \Sigma_x^{-1} (x-x_0)} e^{i 2\pi f x}$$

2.1.4 Las funciones de Gabor como funciones de área mínima en el dominio conjunto espacio-frecuencia espacial.

En los apartados anteriores hemos visto que una función de Gabor está *simultáneamente localizada* en el dominio espacial y en el dominio frecuencial, debido a las envolventes gaussianas presentes en ambos dominios, como queda de manifiesto en el ejemplo de la figura 5. En los apartados sucesivos veremos que esta propiedad es importante porque permite la utilización de las funciones de Gabor como filtros pasa-banda adaptables, capaces de detectar la presencia de cierta información espectral en una zona concreta del espacio.

En este apartado definiremos cuantitativamente lo que entendemos por localización de una señal y analizaremos las propiedades de localización de las funciones de Gabor, propiedades que tienen gran trascendencia a la hora de considerar su utilidad como filtros o detectores.

Anchura de una función en los dominios espacial y frecuencial: Relación de incertidumbre[9].

La forma standard de cuantificar la dispersión de una función $f(x)$ respecto de un determinado punto x_0 consiste en calcular los momentos de orden 2 de la distribución de energía $f(x)f(x)^*$ respecto de dicho punto.

Para señales 2D existen tres posibles momentos de orden 2: respecto del eje x , respecto del eje y , y el momento conjunto xy . Refiriendo la distribución respecto de sus ejes principales, el momento cruzado se anula, y los otros dos dan las anchuras efectivas de la distribución en sus ejes de simetría. Teniendo esto en cuenta, definiremos las anchuras de una función $f(x)$, referida a sus ejes principales, respecto de un punto $x_0 = (x_0, y_0)$ como:

$$\Delta x = \left[\frac{\int (x-x_0)^2 |f(x)|^2 dx}{\int |f(x)|^2 dx} \right]^{1/2} \quad (2.12)$$

$$\Sigma_x = \frac{\int (x-x_0)(x-x_0)^T |g(x)|^2 dx}{\int |g(x)|^2 dx} = \frac{\int (x-x_0)(x-x_0)^T e^{-2(x-x_0)^T \Sigma_x^{-1} (x-x_0)} dx}{\int e^{-2(x-x_0)^T \Sigma_x^{-1} (x-x_0)} dx}$$

$$|g(x)|^2 = e^{-2(x-x_0)^T \Sigma_x^{-1} (x-x_0)}$$

$$\int e^{f(x)} dx = \dots$$

$$\int f(x) e^{f(x)} dx = e^{f(x)}$$

$$\begin{pmatrix} x-x_0 & | & x-x_0 \\ & & \end{pmatrix}$$

$$\begin{pmatrix} & & \\ & D & \\ (x-x_0) & | & x-x_0 \end{pmatrix}$$

$$e^{(\dots) \Sigma_x^{-1} (\dots)}$$

la definición de Δy es análoga. De la misma manera podemos definir las anchuras de $F(f)$ en el dominio frecuencial respecto de una f_0 :

$$\Delta f_x = \left[\frac{\int (f_x - f_{ox})^2 |F(f)|^2 df}{\int |F(f)|^2 df} \right]^{1/2} \quad (2.13)$$

la definición de Δy es también análoga. Con estas definiciones se puede demostrar[11] que el producto de estas anchuras unidimensionales satisface la relación de incertidumbre usual:

$$\Delta f_x \Delta x \geq 1/4\pi ; \quad \Delta f_y \Delta y \geq 1/4\pi \quad (2.14)$$

con lo cual, como al girar la función respecto de sus ejes principales crecen los productos de las anchuras efectivas, puede asegurarse que en todo caso el área de cualquier señal en el dominio conjunto espacio-frecuencia siempre cumplirá:

$$\Delta x \Delta y * \Delta f_x \Delta f_y \geq 1/16\pi^2 \quad (2.15)$$

Área de una función de Gabor en el dominio conjunto.

Para utilizar las expresiones 2.12 y 2.13 la función debe estar definida respecto de sus ejes principales: $\alpha=0$ en las expresiones 2.4 y 2.11. Aplicando las definiciones de las anchuras a la función g y a su espectro G , se obtiene de forma inmediata:

$$\begin{aligned} \Delta x &= \frac{a}{2\sqrt{\pi}} ; \quad \Delta y = \frac{b}{2\sqrt{\pi}} \\ \Delta f_x &= \frac{1}{2a\sqrt{\pi}} ; \quad \Delta f_y = \frac{1}{2b\sqrt{\pi}} \end{aligned} \quad (2.17)$$

con lo que el área de una función de Gabor en el dominio conjunto minimiza la expresión 2.15.

El hecho de que las funciones de Gabor tengan área mínima en el dominio conjunto tiene varias consecuencias:

- Atendiendo al concepto de información definido por Gabor¹, estas funciones son las que contienen una cantidad de información mínima². Como para analizar una señal compleja, portadora de gran cantidad de información, es deseable descomponerla respecto de una base de funciones elementales, portadoras de una cantidad de información mínima, las funciones de Gabor constituyen una base adecuada (óptima) para representar las señales complejas.
- La relación de incertidumbre implica que para un grado de localización dado en uno de los dominios (espacio o frecuencia), la deslocalización en el otro dominio será tanto mayor cuanto más lejos se esté la igualdad en 2.15. Como las funciones de Gabor satisfacen la igualdad, la deslocalización en el dominio recíproco es mínima, por lo tanto, si utilizamos dichas funciones como filtros o detectores, tenemos la máxima resolución simultánea posible en ambos dominios.

A pesar de lo atractivo del concepto de dominio conjunto como diagrama de información divisible en cuantos elementales o Logones, lo cierto es que esta formulación de la idea de información resulta poco práctica y rápidamente fue eclipsada por las ideas de Shannon[1] expuestas anteriormente. El interés de las funciones de Gabor en el análisis de señales reside en el segundo de los aspectos citados: el hecho de tener la máxima resolución en ambos dominios implica que la aplicación de un filtro de Gabor sobre una señal nos da información precisa sobre la misma en ambos dominios, con lo cual, un conjunto de este tipo de filtros constituye un potente método de análisis de las señales.

Como curiosidad citaremos que existen opiniones aisladas contrarias a considerar que los momentos de orden dos sean una medida adecuada del área de una función en aplicaciones de modelizado de sistemas biológicos, así como contra la utilización de funciones complejas en dichos ámbitos[12]. Wilson argumenta que planteando otras métricas (por ejemplo momentos pares de mayor orden), otras formas de computar la

¹ Para Gabor la información contenida en una señal es proporcional al número de datos independientes que facilita.

² El número de datos independientes contenidos en una señal es proporcional al producto de su área espacial por su ancho de banda (ec. 2.1)

indeterminación o exigiendo que las funciones sean reales, las funciones de Gabor (o sus partes reales o imaginarias aisladas), no satisfarían la igualdad en la relación de incertidumbre.

De todos modos, según los cálculos del propio Wilson, la localización simultánea de las funciones de Gabor es, si no máxima, si muy alta, independientemente del criterio utilizado en la medida de las anchuras.

Lo cualitativamente importante en nuestro caso es tener una alta localización simultánea, para tener una alta resolución espacial y frecuencial, pudiéndose tolerar formulaciones sub-óptimas a cambio de obtener ventajas matemáticas. Este mismo razonamiento puede aducirse a favor de otros tipos de funciones localizadas en dominios conjuntos (wavelets o bases de soporte compacto[13]) ya que también presentan las mismas ventajas cualitativas. La discusión no debe plantearse pues, entorno a si unas u otras funciones minimizan unos u otros criterios de localización, sino en ver que modelo resulta más manejable y potente.

En cuanto a la exigencia de que las funciones sean reales, hay que decir que está demostrada la posibilidad de una representación biológica de cantidades complejas mediante canales sensibles a estímulos desfasados $\pi/2$ [14]. En cualquier caso, debido a la fragmentariedad de los datos fisiológicos, es complicado decidir entre diferentes modelos que coincidan cualitativamente (funciones de Gabor o spline wavelets?). Así pues, no parece razonable rechazar una determinada formulación mientras no exista una imposibilidad fisiológica manifiesta.

2.1.5 Las funciones de Gabor como filtros pasa-banda: control de las características del filtro mediante los parámetros variables en el dominio espacial.

Las funciones de Gabor pueden utilizarse como analizadores espectrales locales. En efecto, su localización espectral y la facilidad para controlar su posición y anchura en el dominio frecuencial, permiten su utilización como filtro pasa-banda fácilmente sintonizable. Por otro lado, su localización espacial hace que el producto escalar de una imagen con una determinada función de Gabor, de idea de la magnitud de la

contribución de la banda frecuencial considerada en la zona espacial definida por la ventana gaussiana.

Para elegir las funciones adecuadas para un determinado tipo de análisis hay que conocer como influye en el espectro de $g(x,p)$ la elección un determinado conjunto de los parámetros p . Aunque esta cuestión queda resuelta por la expresión 2.11, en este apartado estudiaremos con ejemplos concretos cuales son los grados de libertad que podemos fijar en una función de Gabor y como influyen en su espectro, obteniendo así un conocimiento intuitivo de sus características como filtro (banda de paso, selectividad a orientaciones, desfase introducido...) que nos permitirá establecer criterios para seleccionar un conjunto de estas funciones como base.

En una función de Gabor 2D tenemos 8 grados de libertad:

$x_0 = (x_0, y_0)$	Posición central en el dominio espacial.
$f_0 = (f_{0x}, f_{0y})$	Frecuencia de oscilación de la senoide.
a, b	Parámetros que controlan la anchura de la envolvente gaussiana (longitudes de los ejes principales de la elipse).
α	Ángulo que forma el eje mayor de la elipse con el eje x .
ϕ	Fase de la oscilación en el punto x_0

La modificación de la forma de la función en el dominio espacial provoca los efectos usuales[9-11] sobre su espectro. Las propiedades fundamentales son las siguientes:

<i>DOMINIO ESPACIAL</i>	<i>DOMINIO FRECUENCIAL</i>
Oscilación de frecuencia f_0	Filtro sintonizado entorno a f_0
Posición central x_0	Oscilación en el dominio de la espectral con frecuencia $-x_0$
Incremento del área	Disminución de la banda de paso
Inclinación de la elipse un ángulo α	Inclinación de la envolvente espectral un ángulo α

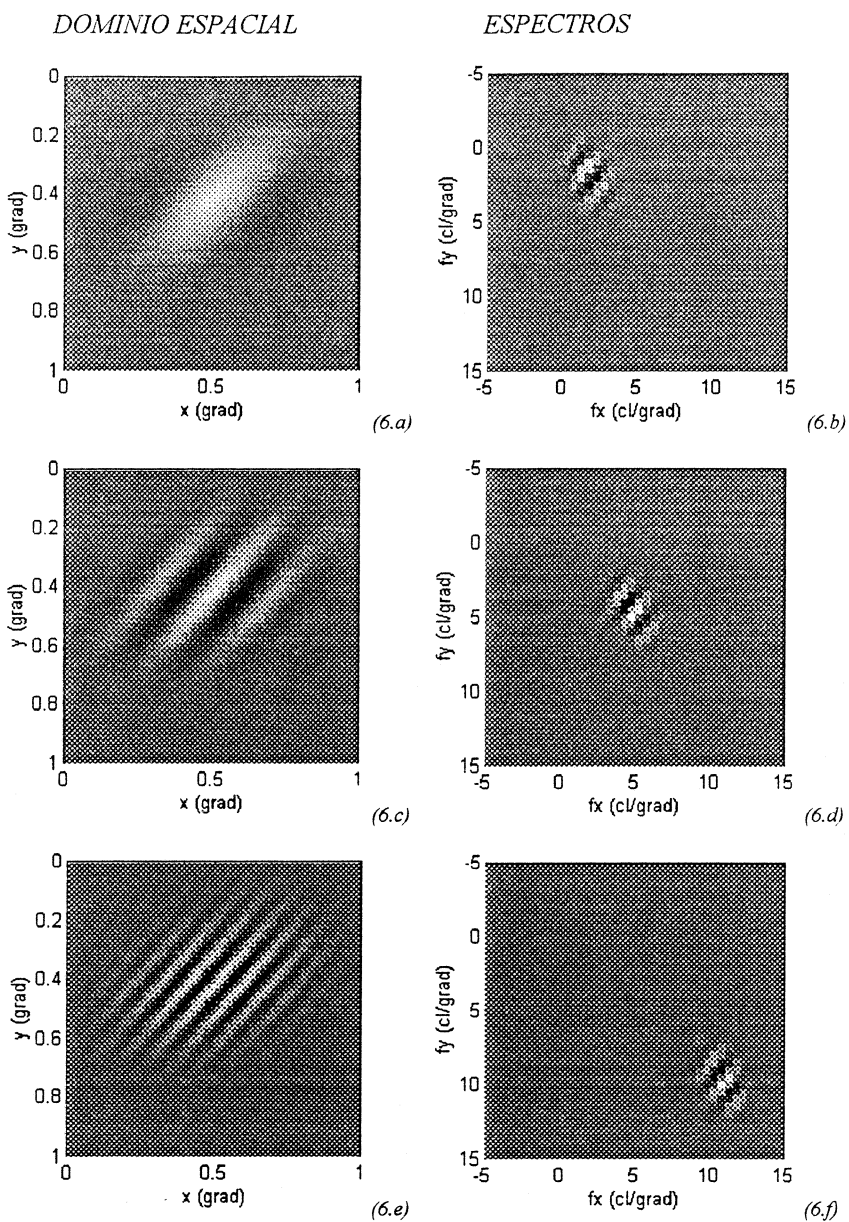


Fig. 6: La oscilación espacial determina la posición en el dominio frecuencial (determina el centro de la banda de paso). En las figuras 6.a-6.f se muestra como la frecuencia de la oscilación determina la distancia al origen del dominio de Fourier.

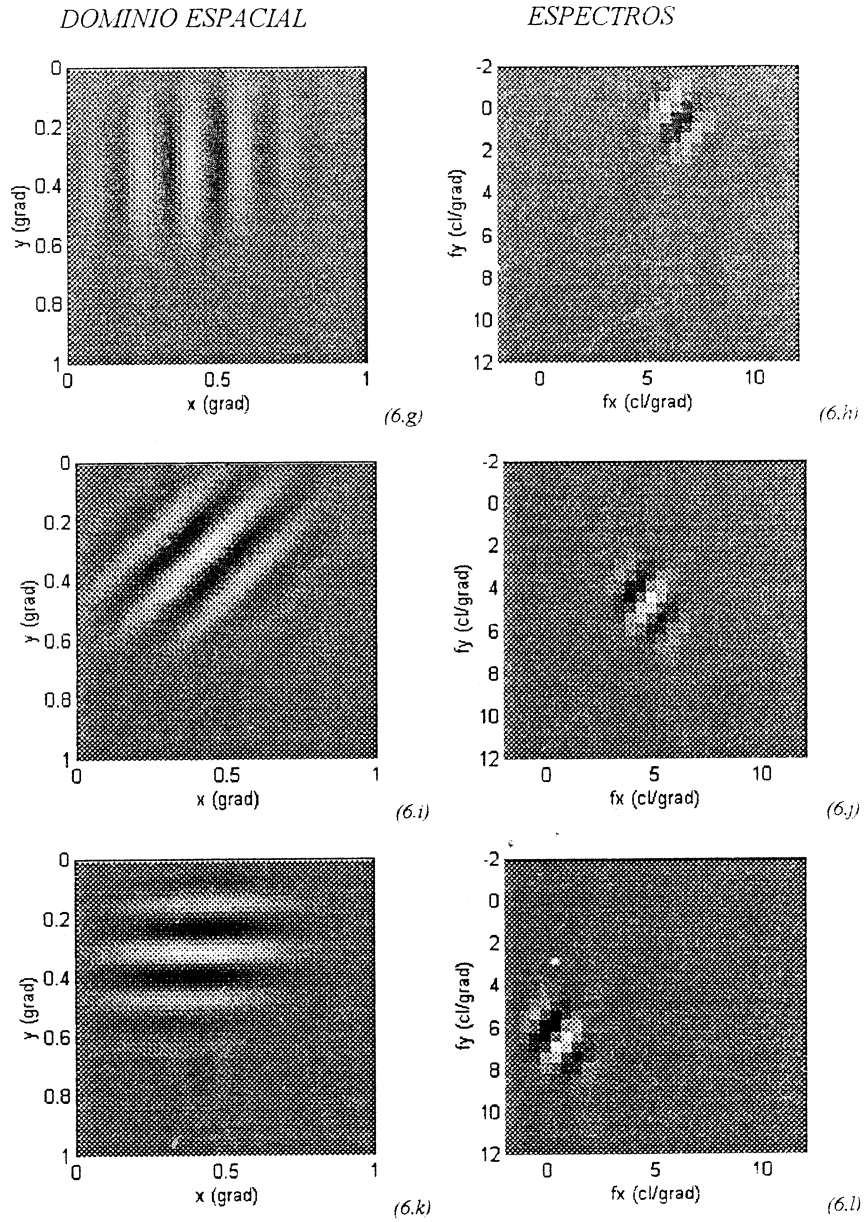


Fig. 6: La oscilación espacial determina la posición en el dominio frecuencial (determina el centro de la banda de paso). En las figuras 6.g-6.l se muestra como la orientación de la oscilación determina la dirección a la que es sensible el filtro.

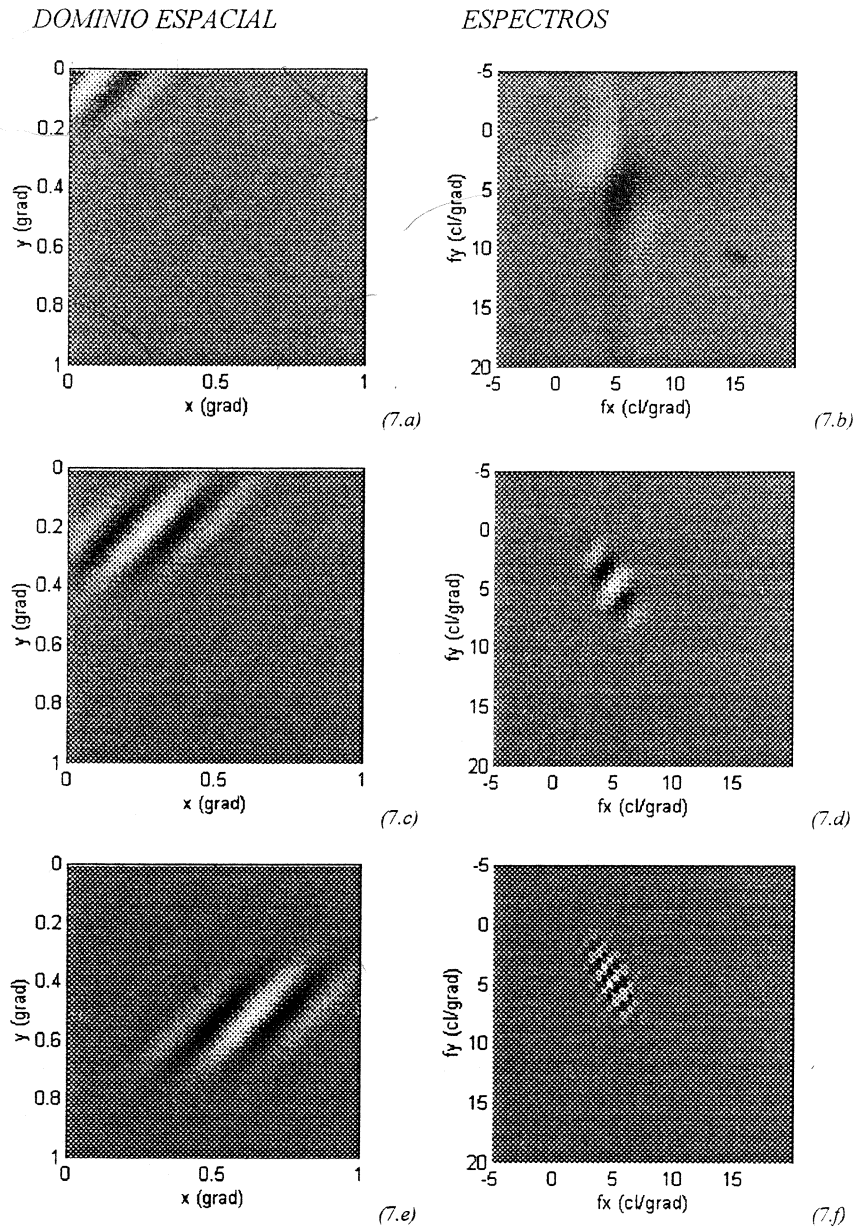


Fig. 7: La posición espacial determina la oscilación en el dominio frecuencial. En las figuras 7.a-7.f se muestra como el incremento de la distancia al origen en el dominio espacial provoca un aumento de la frecuencia de oscilación en el dominio frecuencial.

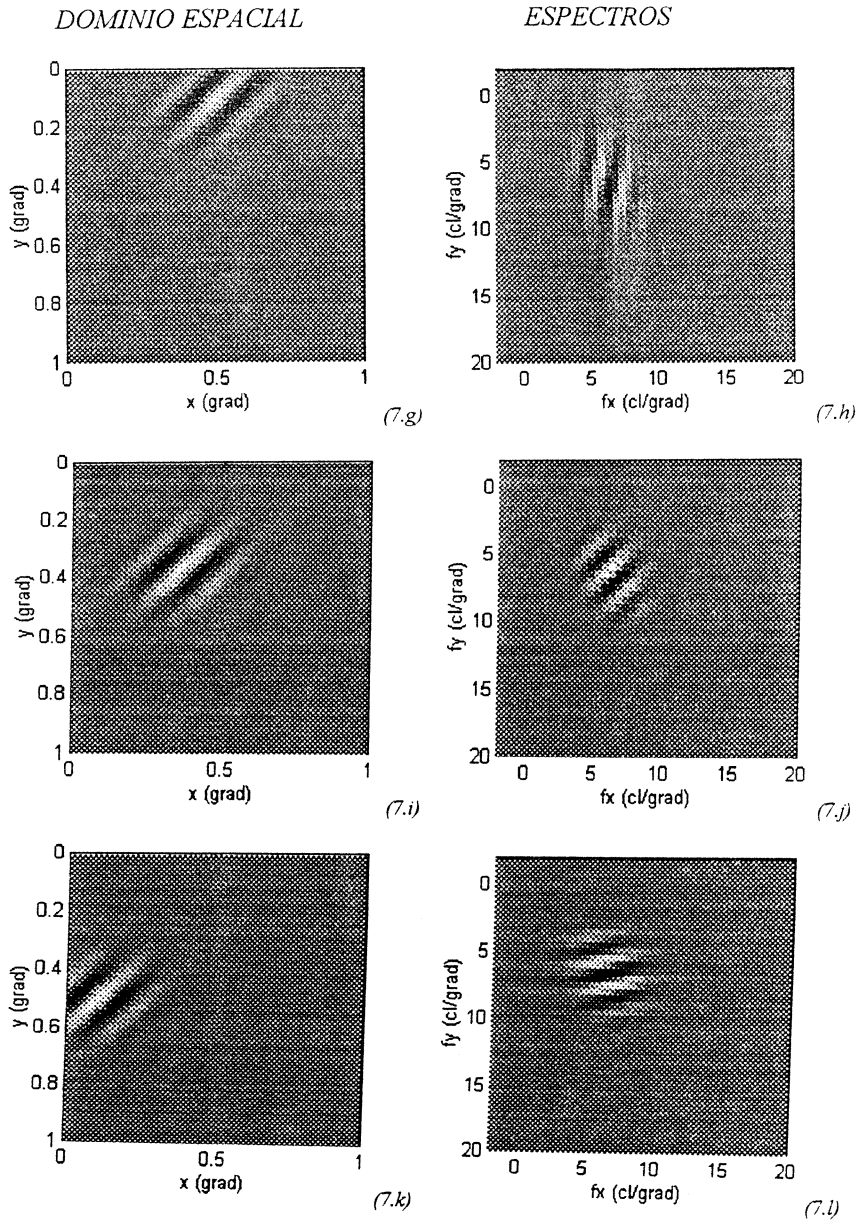


Fig. 7: La posición espacial determina la oscilación en el dominio frecuencial. En las figuras 7.g-7.l se muestra como la variación en la orientación del vector de posición de la envolvente gaussiana provoca una reorientación del vector de onda de la oscilación en el dominio frecuencial.

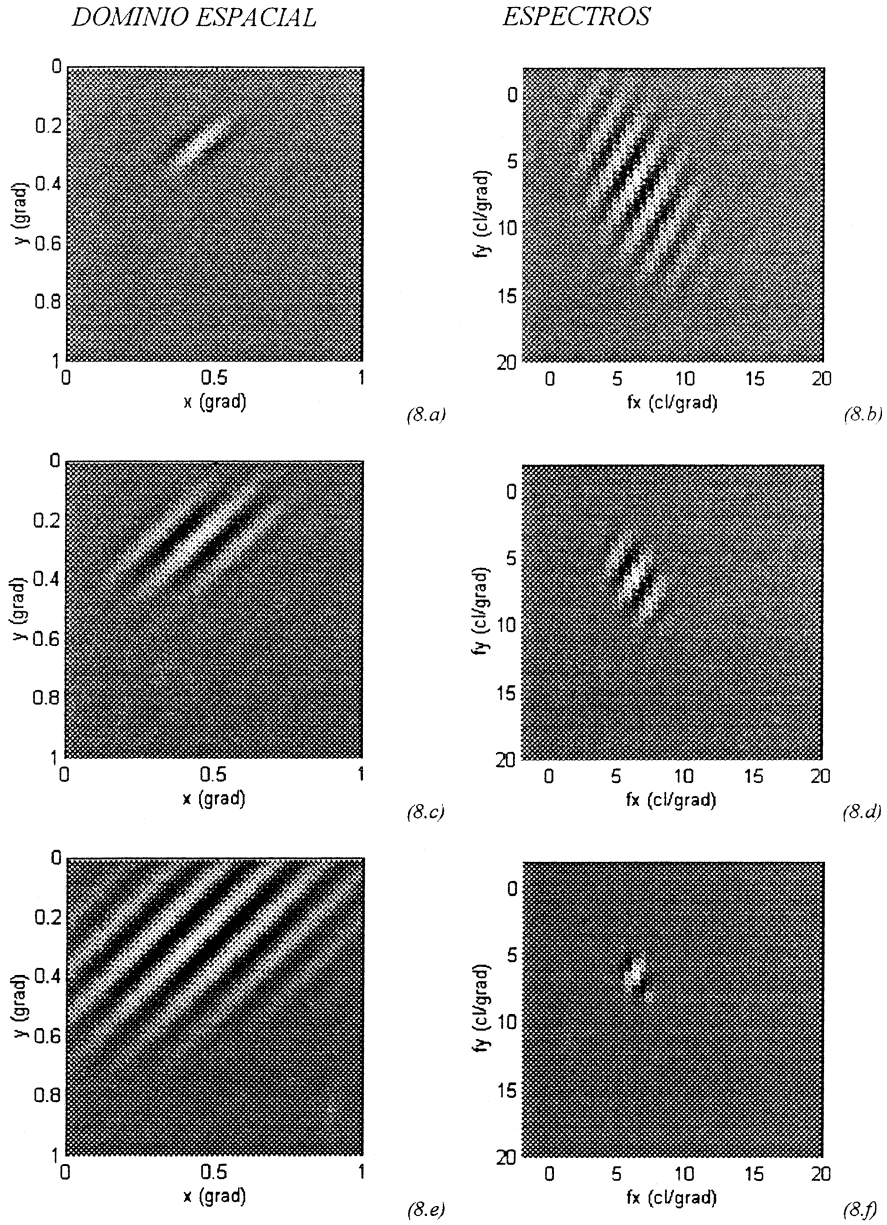


Fig. 8: Una expansión o compresión en uno de los dominios provoca el efecto contrario en el otro dominio. Esto implica que la anchura frecuencial y la selectividad a orientaciones pueden controlarse fácilmente mediante las anchuras espaciales de la envolvente gaussiana.

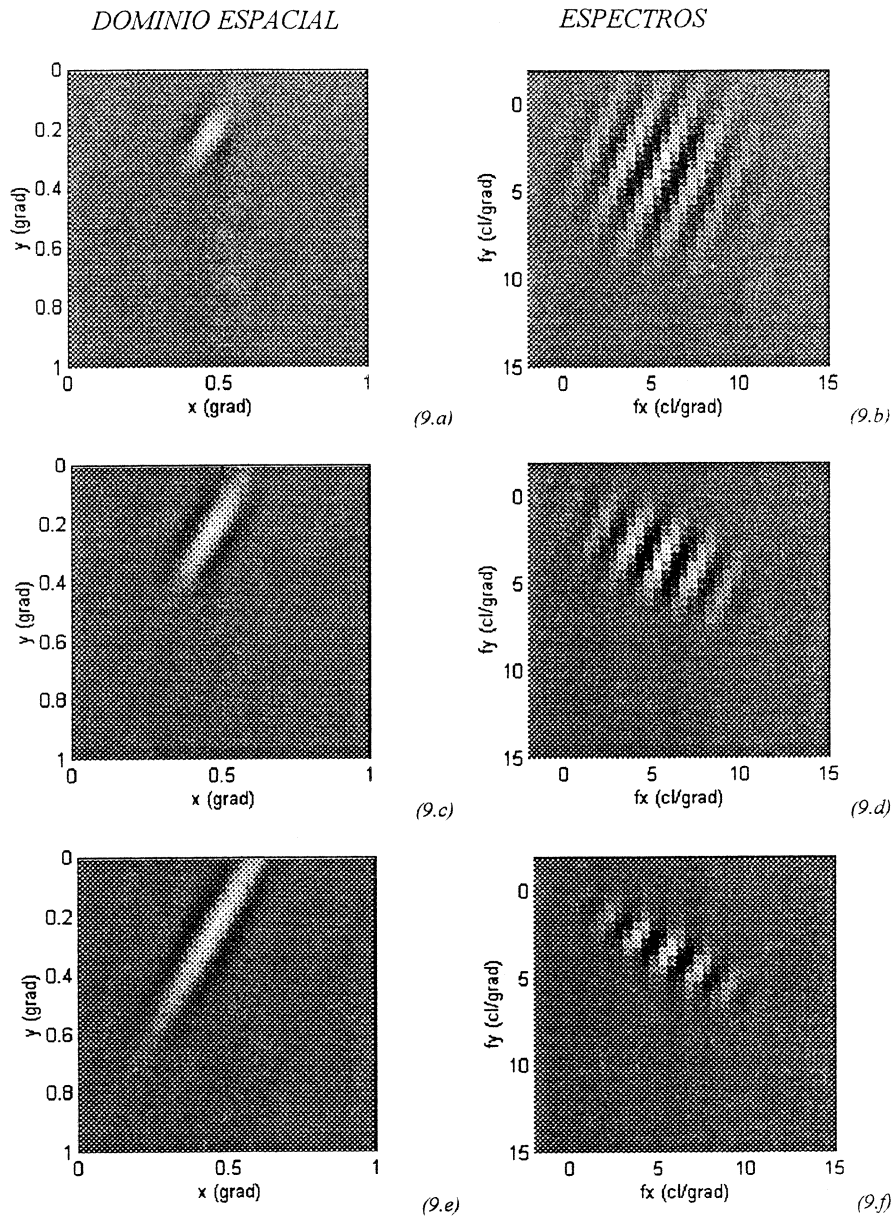


Fig. 9: La banda de paso y la selectividad a orientaciones puede variarse independientemente. En las figuras 9.a-9.f se muestra como es posible aumentar la selectividad a una orientación sin variar la anchura de la banda de paso.

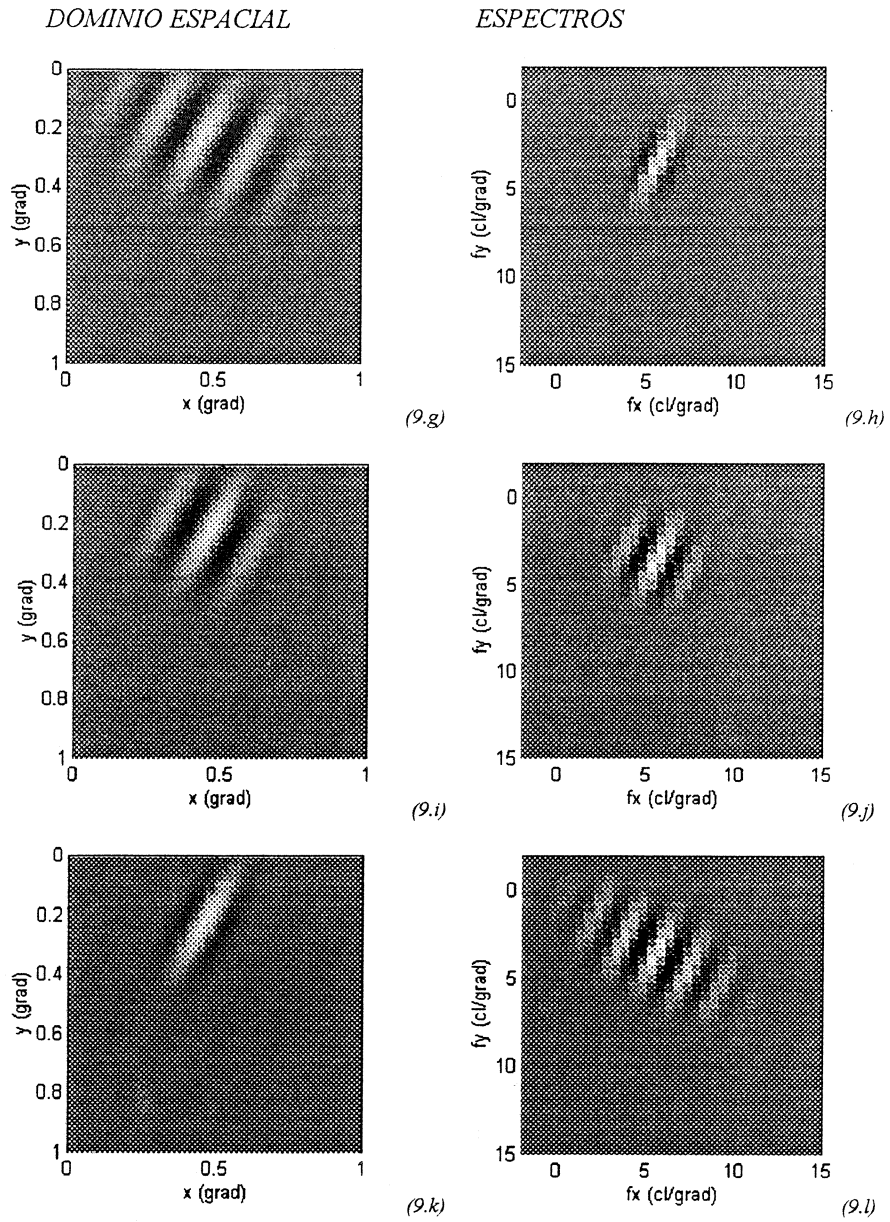


Fig. 9: La banda de paso y la selectividad a orientación pueden variarse independientemente. En las figuras 9.g-9.l se muestra la variación de la banda de paso dejando fija la selectividad a orientaciones.

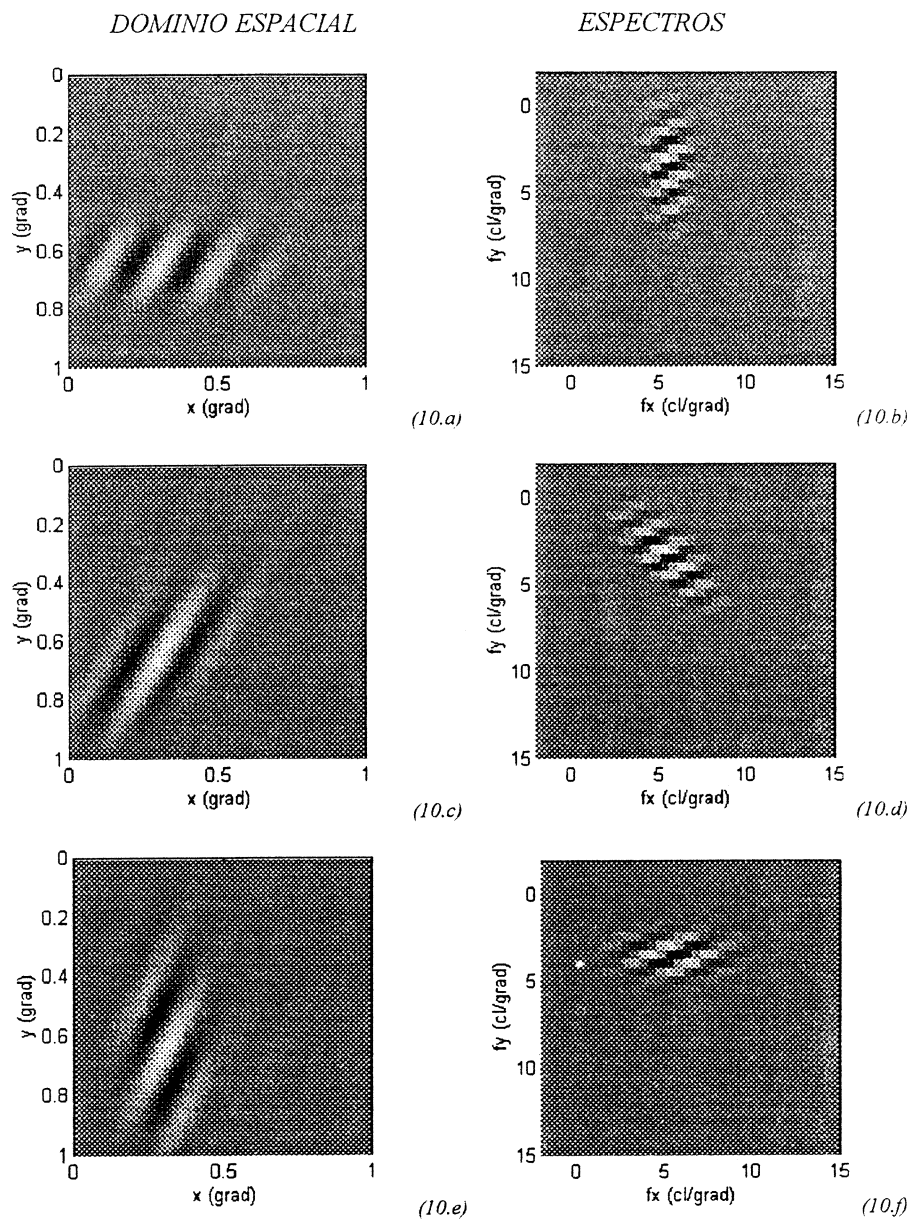


Fig. 10: Un giro de la envolvente en el espacio de posiciones provoca un giro análogo de la envolvente en el dominio frecuencial.

2.2 TRANSFORMADA DE GABOR.

La transformada de Gabor es la proyección de una función cualquiera, $i(x)$ definida en el dominio de posiciones, respecto de una base de funciones $\{g(x,p)\}$ donde las funciones g son funciones de Gabor. Los razonamientos que utilizaremos para el cálculo de los coeficientes de la proyección son válidos para cualquier base de funciones, con lo cual, en lo sucesivo no haremos referencia explícita a la forma de las funciones base, y todas las expresiones que se obtengan serán válidas para cualquier otro tipo de funciones de soporte compacto.

2.2.1 Descomposición de una función respecto de una base no ortogonal e incompleta.

La codificación de la señal $i(x)$ en un dominio de parámetros p supone una descomposición de la misma respecto de una base $\{g(x,p)\}_{p \in P}$ de manera que conocidos los coeficientes $f(p)$ de la transformada respecto de dicha base, sea posible reconstruir la señal original sumando las funciones base pesadas por el coeficiente correspondiente¹. Como decíamos en el apartado 1.1.2, cada coeficiente $f(p)$ indica en que medida están presentes las características de cada $g(x,p)$ en la imagen considerada.

Si la base no es completa, la señal reconstruida diferirá en general de la señal original, es decir, conocida la transformada

$$f(p) = T [i(x)](p) \quad (2.18)$$

podemos reconstruir:

$$h(x) = \int_P f(p) g(x, p) dp \quad (2.19)$$

¹ En el problema que nos ocupa, lo realmente importante es que el sistema mantenga toda la *información esencial* (aquella que es útil para el reconocimiento de formas), y no tanto que, a partir de la codificación generada, se pueda obtener una buena reconstrucción de la señal original. De hecho, las transformaciones óptimas para conseguir codificaciones con mínima distorsión o máxima discriminación son en general diferentes[15]. A pesar de esta matización, en lo sucesivo utilizaremos como criterio la señal reconstruida a partir del código generado.

pero en general se tendrá que $h(x) \neq i(x)$. Así pues, elegida una cierta base $\{g(x,p)\}_p \in P$, el cálculo de la transformación T se reduce al cálculo de los coeficientes $f(p)$ tales que la señal reconstruida se parezca al máximo a la original, en el sentido de minimizar, por ejemplo, el error cuadrático medio. Esto implica que *el problema de la proyección de una imagen respecto de una base general se trata de un problema de mínimos cuadrados.*

En los siguientes apartados abórdaremos dos soluciones diferentes del problema: una general propuesta por Daugman[16], y otra particular propuesta por Ebrahimi & Kunt [6]. Hemos tratado en detalle ambos algoritmos por dos razones:

- Para hacer más explícitas ciertas cuestiones que no quedaban claras en la bibliografía, como por ejemplo la implementación concreta de la red neuronal de Daugman o el significado espacio-frecuencial de los coeficientes de la transformada de Ebrahimi & Kunt.
- Para demostrar que la transformada más geneal puede resolverse mediante procesos de sumación pesados, que son los procesos neurofisiológicos típicos, de forma que puede establecerse una relación directa entre las características fisiológicas, *campos receptivos* de las células del córtex, y la transformación aplicada (apdo. 3.1.1)

2.2.2 Solución general del problema: cálculo iterativo de los coeficientes de la proyección sobre cualquier base $\{g(x,p)\}_p \in P$

Sin pérdida de generalidad podemos considerar un dominio de características P discreto y finito, de forma que el número N de funciones base sea también finito (e igual al número de parámetros diferentes p de P).

El cálculo de la transformación sobre esa base consistirá en hallar los coeficientes $f_p=f(p)$ tales que el error cuadrático medio:

$$\|\varepsilon(x)\|^2 = \left\| i(x) - \sum_p f_p g_p(x) \right\|^2 \quad (2.20)$$

sea mínimo.

Los coeficientes vendrán definidos por el conjunto de condiciones:

$$\frac{\partial \|\varepsilon\|^2}{\partial f_p} = 0 ; \quad \forall p \in P \quad (2.21)$$

Estas igualdades permiten obtener un sistema de ecuaciones de donde pueden obtenerse los f_p . Efectivamente, asumiendo un dominio espacial discreto, lo cual tampoco resta generalidad a los razonamientos y permite una implementación numérica directa, se tiene:

$$\frac{\partial \|\varepsilon\|^2}{\partial f_p} = 2 \sum_x (h(x) - i(x)) \frac{\partial h(x)}{\partial f_p} = 2 \sum_x (h(x) - i(x)) g(x, p) = 0 \quad (2.22)$$

con lo cual, aplicando la expresión discreta equivalente a 2.19:

$$-2 \sum_x i(x) g(x, p) + 2 \sum_{p'} f_{p'} \sum_x g(x, p') g(x, p) = 0 \quad (2.23)$$

y por lo tanto, sustituyendo el sumatorio del producto sobre el dominio de definición por el producto escalar, la condición 2.21 para hallar los coeficientes queda como:

$$\langle i(x), g(x, p) \rangle = \sum_{p'} f_{p'} \langle g(x, p'), g(x, p) \rangle ; \quad \forall p \in P \quad (2.24)$$

El conjunto de ecuaciones 2.24 es un sistema de N ecuaciones (una expresión para cada p), y N incógnitas (en cada ecuación intervienen todos los coeficientes f_p), con lo cual es compatible y determinado.

Nótese que si la base empleada es ortonormal

$$\langle g(\mathbf{x}, \mathbf{p}'), g(\mathbf{x}, \mathbf{p}) \rangle = \delta_{pp'} \quad (2.25)$$

en cada ecuación de 2.24 nos quedamos con un solo término, y entonces la obtención de cada coeficiente es directa:

$$f_p = \langle i(\mathbf{x}), g(\mathbf{x}, \mathbf{p}) \rangle \quad \forall \mathbf{p} \in \mathbf{P} \quad (2.26)$$

que es la manera habitual de calcular los coeficientes de una transformada respecto de una base ortogonal, como por ejemplo en el caso de una transformada de Fourier. Como en general la base puede no ser ortogonal, habrá que resolver el sistema, lo cual es computacionalmente muy costoso.

Descartada la resolución directa del sistema de ecuaciones, es posible plantear una solución iterativa del problema porque la función a minimizar ($\|\varepsilon\|^2$) es una función cuadrática de los coeficientes f_p , con lo cual existe un único mínimo y en ese caso, un método de búsqueda de descenso por gradiente alcanza con seguridad el mínimo absoluto.

Los métodos de descenso por gradiente parten de unos f_p iniciales que posteriormente se van modificando en caso de que dichos coeficientes no satisfagan la condición 2.21:

$$f_{p_{i+1}} = f_{p_i} + \Delta f_{p_i} \quad \text{con} \quad \Delta f_{p_i} = -\frac{\beta}{2} \frac{\partial \|\varepsilon\|^2}{\partial f_p} \Bigg|_{f_p = f_{p_i}} \quad (2.27)$$

donde i indica la iteración y β es el parámetro que controla la velocidad de convergencia del algoritmo (en general puede plantearse un β que dependa del coeficiente, facilitando la convergencia de los coeficientes para los cuales se tenga menos gradiente, y de la iteración, propiciando una rápida modificación al principio, y modificaciones más precisas cuando se está cerca de la solución).

Tomando como valores iniciales del algoritmo los dados por los productos escalares 2.26 (que serán una buena aproximación cuando las funciones base no se solapen mucho) y teniendo en cuenta que el factor corrector de 2.27 queda reducido al cálculo de productos escalares:

$$\Delta f_{p_i} = \beta \left[\langle i(\mathbf{x}), g(\mathbf{x}, p) \rangle - \sum_{p'} f_{p'} \langle g(\mathbf{x}, p'), g(\mathbf{x}, p) \rangle \right] \quad (2.28)$$

podemos implementar el algoritmo en un sistema de procesamiento masivamente paralelo como el de la figura 11 (Véase nota al pie)¹.

En esta figura se muestra un ejemplo explícito de la red neuronal empleada por Daugman para el cálculo de la transformada de Gabor 2D[16]. Esta red calcula los coeficientes de la transformada de una imagen $i(\mathbf{x})$ definida en un dominio discreto de cuatro puntos ($\mathbf{x}=1, 2, 3, 4$) proyectándola sobre un dominio P discreto de tres puntos ($p=1, 2, 3$).

2.2.3 Solución particular del problema: cálculo de la transformada respecto de una base de funciones separables

En primer lugar demostraremos que la combinación lineal utilizada en la reconstrucción de $h(\mathbf{x})$

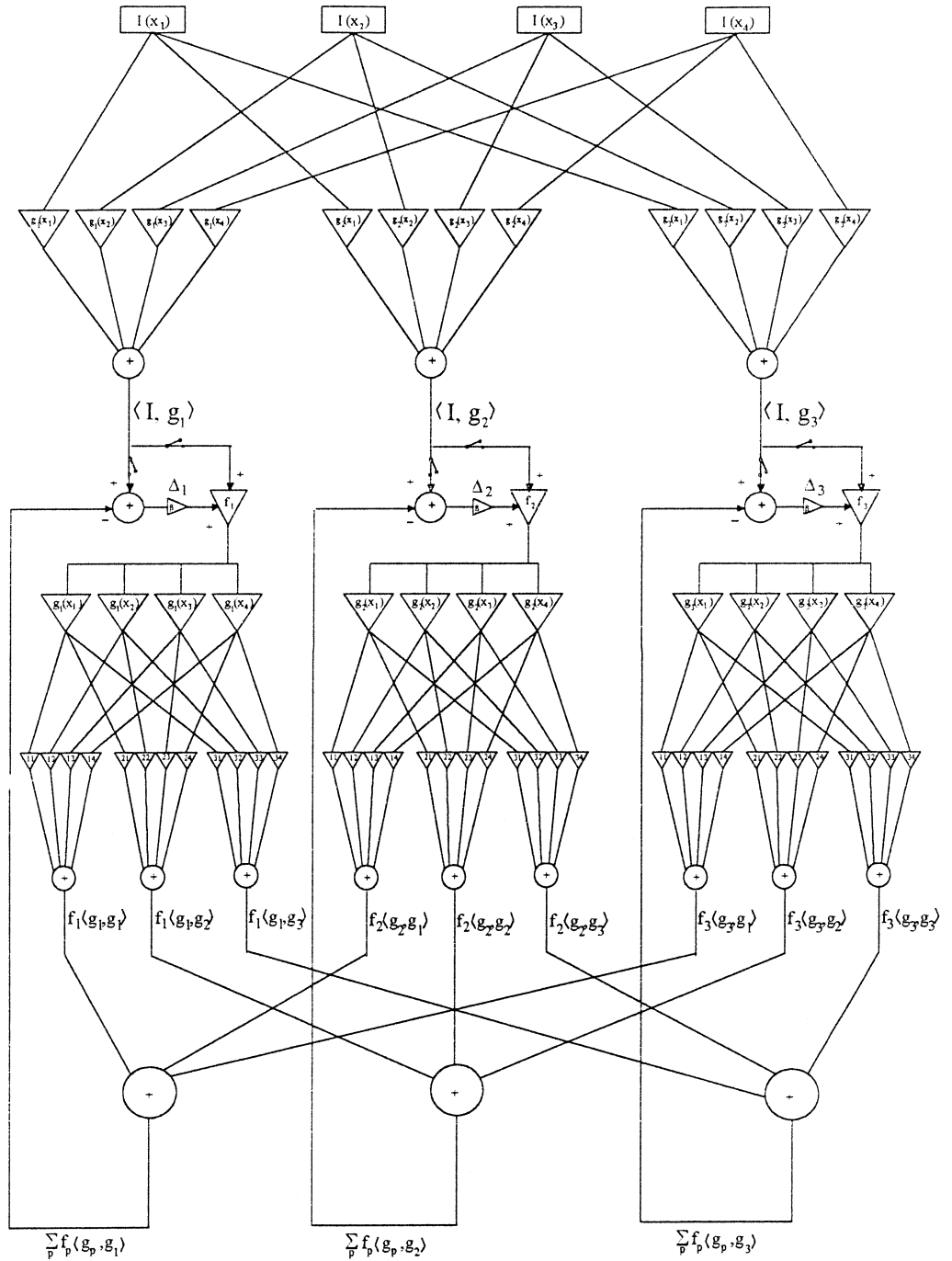
$$h(\mathbf{x}) = \sum_p f_p g_p(\mathbf{x}) \quad (2.29)$$

puede expresarse, como un producto de matrices:

$$H = G_1 F G_2 \quad (2.30)$$

si la base de funciones es separable. En la expresión 2.30 los elementos de la matriz H son los valores de $h(\mathbf{x})$ definida en un dominio espacial discreto, G_1 y G_2 están relacionadas con los valores de las funciones base $g_p(\mathbf{x})$, y los elementos de la matriz F son los coeficientes f_p .

Fig. 11: Red neuronal de Daugman para la resolución del sistema (2.24). El algoritmo comienza a partir del estímulo representado en el dominio de posiciones: $I(x_j)$ es la intensidad captada por el fotodetector situado en la posición j . Los símbolos $g_i(x) = g_{ij}$ representan la función campo receptivo del detector i (Véase el apdo. 3.1). $\langle I, g_i \rangle$ es la respuesta de cada detector i al estímulo $I(x)$. Las diferentes soluciones aproximadas que da de proceso iterativo se van obteniendo en la capa de las f_i . En la primera iteración f_i coincide con $\langle I, g_i \rangle$ pero posteriormente este valor va modificándose según (2.27).



La ventaja de la expresión 2.30 es que puede demostrarse[17-18] que en tal situación, la solución que minimiza el error cuadrático medio viene dada por el producto matricial:

$$F=(G_1^T G_1)^{-1} G_1^T I G_2 (G_2^T G_2)^{-1} \quad (2.31)$$

Donde I es la matriz que contiene los valores $i(x)$, con lo que el cálculo de la solución óptima se obtiene de forma directa, no iterativa.

Expresión de la reconstrucción discreta como producto matricial: significado de los elementos de la matriz F .

En el caso particular de que las funciones base sean separables, podemos expresarlas como:

$$g_p(x) = g_{kl}(x) = g_k(x) g_l(y) \quad (2.32)$$

donde k es el vector de parámetros que controla las propiedades (posición y frecuencia) de la función en el eje x y l es el vector análogo correspondiente al eje y .

Si $k = 1 \dots A$, y $l = 1 \dots B$, es decir, a partir de un número A de funciones 1D $g_k(x)$ y de un número B de funciones 1D $g_l(y)$ construimos AB funciones 2D $g_{kl}(x)$. De esta manera, las AB funciones bidimensionales definidas en un dominio discreto de $N \times N$ puntos ($x=1 \dots N$, $y=1 \dots N$), pueden ponerse como producto exterior de las funciones base 1D (vectores fila $1 \times N$ y columna $N \times 1$):

$$\begin{bmatrix} g_{kl}(x) \end{bmatrix} = \begin{bmatrix} g_k(1)g_l(N) & \dots & g_k(N)g_l(1) \\ \vdots & & \vdots \\ g_k(1)g_l(1) & \dots & g_k(N)g_l(1) \end{bmatrix} = \begin{bmatrix} g_l(N) \\ \vdots \\ g_l(1) \end{bmatrix} \begin{bmatrix} g_k(1) & \dots & g_k(N) \end{bmatrix} \quad (2.33)$$

la relación de los índices del dominio espacial discreto con los índices de la matriz $g_{kl}(x)$ será:

$$[g_{kl}(x)]_{ij} = g_k(i)g_l(N+1-j) \quad (2.34)$$

El truco que nos permite llegar a obtener la relación 2.30 consiste en expresar cada función base 2.33 como producto de

tres matrices G_1 , F y G_2 . Por ejemplo, tomando $N=3$, $A=4$, $B=5$, $k=2$ y $l=4$, la función $g_{2,4}(x)$ puede expresarse como:

$$\begin{bmatrix} g_H(1,3) & g_H(2,3) & g_H(3,3) \\ g_H(1,2) & g_H(2,2) & g_H(3,2) \\ g_H(1,1) & g_H(2,1) & g_H(3,1) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & g_l(3) & 0 \\ 0 & 0 & 0 & g_l(2) & 0 \\ 0 & 0 & 0 & g_l(1) & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ g_t(1) & g_t(2) & g_t(3) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

donde:

G_1 Matriz de dimensiones $N*B$ que contiene al vector $g_l(y)$ en la columna l

F Matriz de dimensiones $B*A$ cuyo único elemento no nulo es $F_{lk}=1$

G_2 Matriz de dimensiones $A*N$ que contiene al vector $g_k(x)$ en la fila k

Esto quiere decir que el coeficiente F_{lk} controla el peso de la función base $g_{kl}(x)$.

Si construimos la matriz G_1 de manera que cada una de las B columnas esté formada por una de las funciones base unidimensionales $g_l(y)$ con $l=1....B$, y análogamente construimos las A filas de G_2 a partir de las funciones 1D $g_k(x)$ ($k=1....A$):

$$G_1 = \left[\begin{pmatrix} g_1(y) \end{pmatrix} \quad \begin{pmatrix} g_2(y) \end{pmatrix} \quad \dots \quad \begin{pmatrix} g_B(y) \end{pmatrix} \right]$$

$$G_2 = \begin{bmatrix} (g_1(x)) \\ (g_2(x)) \\ \cdot \\ \cdot \\ (g_A(x)) \end{bmatrix} \tag{2.35}$$

tendremos que la combinación lineal análoga a la 2.29:

$$h(\mathbf{x}) = \sum_{kl} f_{kl} g_{kl}(\mathbf{x}) \quad (2.36)$$

puede ponerse como $H = G_1 F G_2$. Donde los coeficientes f_{kl} de 2.36 que controlan la contribución de cada $g_{kl}(\mathbf{x})$ son ahora los elementos de la matriz F , es decir:

$$f_{kl} = F_{lk} \quad (2.37)$$

Con esto, hemos demostrado que si las funciones base son separables, es posible poner la combinación lineal 2.29 (o la 2.36) como un producto matricial de la forma 2.30 (de manera que es posible aplicar la solución 2.31). Además, hemos analizado el sentido de cada uno de los elementos de F : cada elemento F_{lk} determina cual es la contribución de la función que resulta del producto exterior de la columna l de G_1 y la fila k de G_2 , pero cual sea esa $g_{kl}(\mathbf{x})$ depende de como se hayan ordenado las funciones unidimensionales en las matrices G_1 y G_2 . En el apartado 5.3 describiremos la ordenación concreta utilizada en nuestro caso, y el correspondiente significado espacio-frecuencial de los elementos de F .

Proyección sobre una base separable: obtención de la matriz de coeficientes.

La no ortogonalidad de las funciones base 1D provoca un inconveniente computacional a la hora de calcular la transformada mediante 2.31.

La redundancia en las columnas (o filas) de las matrices G_i hace que estas sean matrices *aproximadamente singulares*, con lo que no es posible efectuar directamente las inversiones que se precisan en la expresión 2.31. Para evitar estas singularidades utilizamos la descomposición en valores singulares (SVD, singular value decomposition)[18].

Mediante la SVD, cualquier matriz G_i (aunque no sea cuadrada) puede expresarse como producto de otras tres: U , S y V , siendo U y V unitarias (con traspuesta igual a la inversa), y S diagonal. De esta manera:

$$\begin{aligned}
 G_{1(N*B)} &= U_{1(N*N)} S_{1(N*B)} V_1^T \quad (B*B) \\
 G_{2(N*A)} &= U_{2(N*N)} S_{2(N*A)} V_2^T \quad (A*A)
 \end{aligned}
 \tag{2.38}$$

Los valores presentes en las diagonales de S_i son los llamados valores singulares de G_i . Las redundancias dan lugar a valores singulares muy pequeños[17], de forma que al calcular la inversa de S_i existe el peligro de obtener valores divergentes:

$$S_i^{-1} = \begin{bmatrix} s_{11} & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & s_{22} & 0 & & & & \\ 0 & 0 & s_{33} & & & & \\ \dots & & & & & & \\ \dots & & & & 0 & & \\ 0 & \dots & \dots & \dots & 0 & s_{NN} & 0 & \dots & 0 \end{bmatrix}^{-1} = \begin{bmatrix} 1/s_{11} & 0 & 0 & \dots & 0 \\ 0 & 1/s_{22} & 0 & & & & \\ 0 & 0 & 1/s_{33} & & & & \\ \dots & & & & & & \\ \dots & & & & & & 0 \\ 0 & \dots & \dots & \dots & 0 & 1/s_{NN} \\ \dots & & & & & & 0 \\ \dots & & & & & & \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 \end{bmatrix}$$

(2.39)

La solución a este problema consiste en asignar cero a la inversa de los valores singulares muy pequeños[18] (del orden de la precisión con la que se esté trabajando).

Con todo esto, aplicando las descomposiciones 2.48 sobre la igualdad 2.31, llegamos a la expresión:

$$F = V_1 (S_1^T)^{-1} U_1^T I V_2 (S_2^T)^{-1} U_2^T \tag{2.40}$$

mediante la cual, asumida cierta una base de funciones, podemos calcular la matriz de pesos F (los coeficientes de la transformada), para cualquier imagen discreta I .

Aunque tanto el algoritmo de Daugman como el de Ebrahimi & Kunt se han desarrollado de manera que a partir de los coeficientes de la transformada se obtiene una reconstrucción óptima, el hecho de que se pueda obtener una reconstrucción perfecta depende de la base que se elija. Como hemos anticipado antes (apartado 1.1.4), y comprobaremos después con ejemplos

concretos (apartado 5.2.2), una incorrecta elección de la base puede implicar pérdidas importantes de información.

3. DISCUSIÓN ACERCA DE LA TRANSFORMACIÓN EFECTUADA POR EL SISTEMA.

Aceptada la caracterización del sistema visual como un codificador que efectúa las transformaciones T y R , y conocidas las propiedades de las funciones de Gabor, en este apartado expondremos las razones que apoyan la idea de que la transformación T es una representación de la señal en un dominio conjunto espacio-frecuencial.

En primer lugar introduciremos el concepto de detector lineal y describiremos matemáticamente su funcionamiento. Veremos que la respuesta de un detector de este tipo ante cualquier estímulo queda caracterizada por su función de campo receptivo, con lo cual, la discusión entorno a la transformación T se reduce a la discusión entorno a la forma de los campos receptivos de los detectores empleados.

A continuación analizaremos las ventajas que supone la utilización de detectores con campos receptivos localizados espacial y frecuencialmente, y por lo tanto, la oportunidad de plantear una transformación T como descomposición de la señal incidente respecto de una base de funciones localizadas como por ejemplo una base de Gabor.

3.1 DETECTORES, CAMPOS RECEPTIVOS Y TRANSFORMACIONES.

3.1.1 Implementación fisiológica de la transformación T : Concepto de detector y campo receptivo.

La modelización de la parte del procesado visual que sigue a la adquisición de los datos por parte de los fotorreceptores, suele hacerse suponiendo un conjunto de sensores o detectores independientes cuya respuesta es una combinación lineal de las señales dadas por los fotodetectores[19-21].

Estos detectores abstractos son sensibles a una determinada característica p de la señal de entrada, de manera que la imagen incidente no queda codificada mediante señales proporcionales a la intensidad luminosa en cada punto $i(\mathbf{x})$, sino por el conjunto de respuestas $r(p)$ dadas por estos detectores lineales; lo que Watson llama vector de características[19].

Como dijimos en el apartado 1.1.2, el valor de $f(p)$ ($r(p)$ en el caso de las respuestas de los detectores citados) indicará en que medida está presente la característica p en la imagen dada. De esta forma, el procesado de alto nivel se realiza ya sobre estos vectores de características.

Matemáticamente, un detector p queda definido por la función que da cual es la influencia en el detector de la intensidad luminosa en cada punto. Esta es la *función peso* o *campo receptivo* del detector p . En general el campo receptivo se expresa como una función definida en el dominio de posiciones y que depende del vector de parámetros p : $g(\mathbf{x}, p)$.

De esta manera, la respuesta que da un detector p ante un estímulo $i(\mathbf{x})$ es el resultado de sumar todas las contribuciones del estímulo en el área sensible del detector, pesando cada valor de la intensidad por el factor de influencia $g(\mathbf{x}, p)$ correspondiente:

$$r(p) = \int i(\mathbf{x}) g(\mathbf{x}, p) d\mathbf{x} = \langle i(\mathbf{x}), g(\mathbf{x}, p) \rangle \quad (3.1)$$

es decir, el producto escalar de la señal incidente por la función campo receptivo.

El sistema de detectores $\{ g(x,p) \}$ descrito constituye una implementación posible de la transformación T postulada en la introducción. Como dijimos en el apartado 2.2.1, la transformación T (proyección de $i(x)$ sobre el dominio P respecto de una base $\{ g(x,p) \}$) consiste en hallar los coeficientes $f(p)$ tales que la diferencia entre la señal reconstruida $h(x)$:

$$h(x) = \int_P f(p) g(x, p) dp \quad (3.2)$$

y la original $i(x)$ sea mínima. En el apartado 2.2.2, vimos que el cómputo de los productos escalares $\langle i(x), g(x,p) \rangle$ es requisito imprescindible para el cálculo de los coeficientes $f(p)$, y en ocasiones (cuando la base es ortogonal), los coeficientes $f(p)$ coinciden con dichos productos escalares.

Por lo tanto, según lo expuesto en 2.2.2, a partir de las respuestas de un sistema de detectores lineales definidos por las funciones de peso $\{ g(x,p) \}$ pueden obtenerse los coeficientes $f(p)$ de la transformación T postulada.

En resumen: la discusión acerca de la transformación T se reduce a proponer las funciones campo receptivo apropiadas. Estas funciones determinan la transformación realizada y las características de la señal incidente que son tenidas en cuenta.

Enunciar el problema de la transformación T en términos de *detectores* y *campos receptivos* tiene sentido fisiológico porque este planteamiento coincide con la modelización matemática de la respuesta de las células del córtex visual[22]. En efecto, la investigación neurofisiológica en ese área está orientada a la determinación de las funciones campo receptivo mediante la medida directa de la respuesta de las células corticales a diferentes estímulos $i(x)$ [23-25].

3.1.2 Codificación en el dominio espacial o el frecuencial en función de la forma de los campos receptivos del sistema de detectores planteado.

Para comprobar la gama de transformaciones que puede englobar la formulación propuesta, estudiaremos dos casos concretos extremos:

- Un sistema de detectores cuyos campos receptivos están perfectamente localizados en el espacio, con lo cual cada detector está caracterizado por la posición espacial a la que es sensible ($p = x_0$):

$$g(x, p) = \delta(x - x_0) \quad (3.3)$$

- Un sistema de detectores cuyos campos receptivos están perfectamente localizados en el dominio frecuencial, con lo cual quedan caracterizados por la frecuencia a la que están sintonizados ($p = f_0$):

$$g(x, p) = e^{-i2\pi f_0 x} = \delta(f - f_0) \quad (3.4)$$

En dichos casos, la respuesta de cada detector será respectivamente:

$$r(x_0) = \int i(x) \delta(x - x_0) dx = i(x_0) \quad (3.5)$$

$$r(f_0) = \int i(x) e^{-i2\pi f_0 x} dx = TF[i(x)](f_0) \quad (3.6)$$

Como se ve, la elección de un determinado conjunto de detectores es realmente la elección de una determinada base respecto de la que proyectar, y por lo tanto, la elección de una determinada transformación. Así, si elegimos una base de funciones localizadas en el dominio espacial, tendremos una codificación de la imagen en función de los valores de la luminancia en cada punto, mientras que si elegimos una base de funciones armónicas, estaremos codificando la señal según los coeficientes de su transformada de Fourier. Una formulación que utilice funciones base localizadas en ambos dominios constituye un puente entre estos dos planteamientos extremos. Variando los parámetros a y b (expresión 2.11) que controlan la anchura de las funciones de Gabor en los dominios espacial y frecuencial conseguimos una transición gradual entre $\delta(x - x_0)$ y $\delta(f - f_0)$.

De estos ejemplos se deduce también que la elección de uno u otro tipo de campo receptivo facilita la extracción de uno u otro tipo de información: campos receptivos muy localizados espacialmente dan lugar a transformaciones cuyos coeficientes tienen un fuerte significado espacial (3.5), por el contrario,

campos receptivos de frecuencia muy bien definida implican que los coeficientes de la transformación una información frecuencial muy precisa, pero muy poca información espacial (3.6). Si la base que utilizemos contiene a la vez funciones con diferente grado de localización en ambos dominios, tendremos información espacial o frecuencial en el grado que se desee. Esto puede conseguirse fácilmente proponiendo una base de funciones de Gabor con un amplio rango de valores de a y b .

3.2 VENTAJAS DE LA UTILIZACIÓN DE DETECTORES LOCALIZADOS ESPACIAL Y FRECUENCIALMENTE.

Desde la publicación de los trabajos clásicos de Campbell, Green y Robson [26,27] sobre la sensibilidad del sistema visual a redes sinusoidales, la caracterización habitual de la visión espacial [28,29] asume que la transformación de codificación T es una transformada de Fourier y que los procesos de reducción de información R pueden modelizarse, al menos a nivel umbral, como un filtrado lineal en el dominio frecuencial, aplicando una función de pesos, la Función de Sensibilidad al Contraste (CSF), sobre la transformada de Fourier de la imagen.

Sin embargo, desde hace ya bastantes años, es conocido el hecho de que los mecanismos de detección de sinusoides tienen banda ancha [30], y por lo tanto, no es estrictamente correcta la suposición de una codificación de las imágenes en el dominio de Fourier. Desde los años 80, numerosos datos fisiológicos y psicofísicos, así como análisis puramente matemáticos, han propiciado el planteamiento de modelos alternativos basados en conjuntos de detectores cuyos campos receptivos estén localizados espacial y frecuencialmente. Las razones para apoyar este tipo de análisis son, pues, de dos tipos: experimentales y teóricas.

3.2.1 Argumentos teóricos.

Los argumentos teóricos más utilizados a favor del análisis espacio-frecuencial pueden resumirse en estos cinco puntos:

1.- *El análisis espacio-frecuencial facilita el procesado de alto nivel*, entendido como extracción de características complejas, reconocimiento de objetos y establecimiento de relaciones entre ellos. Como se ha dicho previamente, los coeficientes de una transformada respecto de una base de funciones localizadas en los dominios frecuencial y espacial, contienen simultáneamente ambas informaciones. Esto significa que un sistema adecuado de detectores de este tipo permite la extracción local de las características frecuenciales de la imagen incidente, lo cual es una ventaja frente a los esquemas extremos analizados en 3.1.2.

Los modelos que utilizan un sistema de detectores localizados en el dominio espacial no son eficientes a la hora de extraer características globales de la imagen, requiriéndose complicados mecanismos de interacción entre las señales de detectores muy alejados espacialmente para poder dar cuenta de las relaciones de largo alcance presentes en la escena.

Por otro lado, los modelos que utilizan detectores con una anchura de banda muy limitada, de forma que la escena es analizada en base a sus componentes de Fourier, consiguen extraer de forma sencilla las relaciones de largo alcance, porque las funciones base de la representación están extendidas en todo el espacio, pero no son eficientes en la extracción de información muy localizada espacialmente. Los detalles aislados, provocan la excitación de un gran número de detectores frecuenciales, de forma que se requiere un complejo esquema de interacción entre dichos canales para identificar la posición espacial de los estímulos aislados que generan las contribuciones sobre los canales frecuenciales.

La formulación espacio-frecuencial permite la extracción simultánea de características locales y de largo alcance, con lo cual, es adecuada cuando, para extraer información de las imágenes problema, se requieren ambos tipos de características.

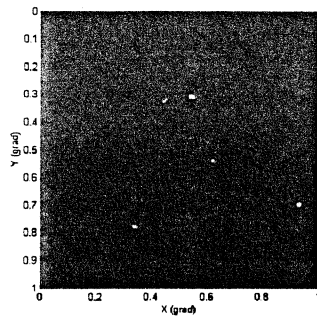
2.- *La representación espacio-frecuencial es adecuada para codificar imágenes naturales*. Como acabamos de decir, la capacidad para extraer simultáneamente información local y global, es una ventaja cuando se requiere de ambas informaciones para la clasificación y el reconocimiento. Esto

es así en el caso de las imágenes naturales. Efectivamente, los análisis de la estadística de las imágenes naturales[31-33] revelan que la energía y la varianza se distribuyen según $1/f$ en el dominio frecuencial. Esto implica que para plantear un sistema óptimo de detectores, en el que todos ellos estén igualmente utilizados (un sistema mínimamente redundante), habrá que distribuir no uniformemente los máximos de sensibilidad (y las anchuras de banda) en el dominio frecuencial[31], acumulando más detectores (con anchura de banda estrecha) en la zona de bajas frecuencias, y utilizando una densidad de detectores menor (y con mayor banda sensible) en la zona de alta frecuencia. Esa distribución de detectores solo es posible mediante unas funciones campo receptivo a medio camino entre los detectores puntuales y los detectores frecuenciales. En la figura 12 se muestran unos ejemplos muy intuitivos acerca del tipo de detectores a utilizar en función del tipo de imágenes problema con las que nos enfrentemos.

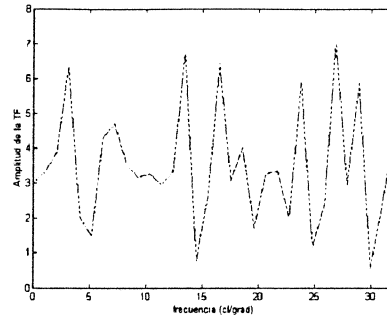
La adecuación a la estadística de las imágenes naturales ya ha sido utilizada para obtener fuertes tasas de compresión en algoritmos de codificación de imágenes ordinarias mediante diferentes transformaciones espacio-frecuenciales[6,16,34-36]. Además de eso, hay que notar que la localidad de los coeficientes de las transformadas espacio-frecuenciales implica que la entropía de la señal transformada no crece tanto como en el caso de las transformaciones respecto de bases extendidas en todo el dominio espacial, recuérdese como la entropía de la transformada depende del número de pixels que influyen en cada coeficiente (apdo. 1.1.3).

Una transformación local y piramidal de este tipo permite la uniformización de las propiedades estadísticas en cada zona, propiedades que en principio son espacialmente variantes (las imágenes naturales son no estacionarias). Conseguida la estacionariedad local, ya son aplicables los razonamientos que conducen a transformaciones óptimas como la KLT (Karhunen Loeve Transform) y la DCT (Discrete Cosine Transform) [36].

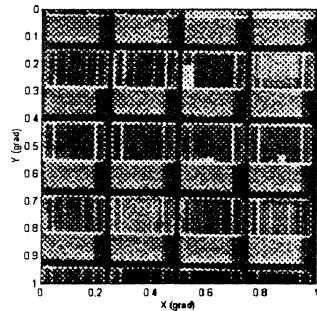
- 3.- *La representación espacio-frecuencial es adecuada para modelizar sistemas espacialmente variantes.* Efectivamente, como los coeficientes de las transformaciones respecto de bases



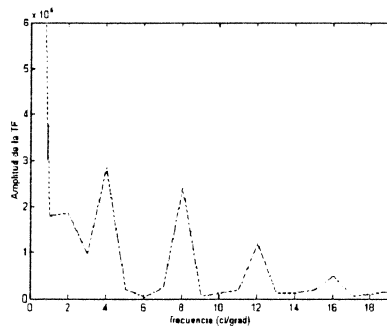
(12.a)



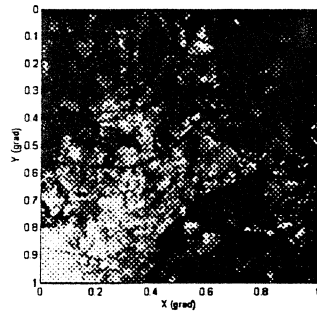
(12.b)



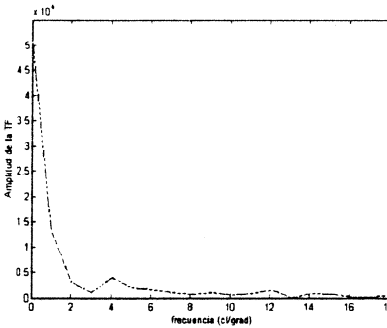
(12.c)



(12.d)



(12.e)



(12.f)

Fig. 12: En cualquier sistema de análisis de imágenes el dominio de codificación y la organización del conjunto de detectores en dicho dominio deben estar adaptados a las características de las señales con las que se enfrenta. En esta figura presentamos tres ejemplos extremos que inducirían diferentes organizaciones de los detectores: si las imágenes a tratar están formadas por ruido impulsional de espectro plano (fig. 12.a, 12.b), lo más eficiente es codificar las señales en el dominio espacial, recubriéndolo mediante detectores uniformemente espaciados cuyo campo receptivo sea una delta de Dirac. Si las señales a analizar son fuertemente periódicas, formadas por muy pocas componentes del dominio de Fourier (figs. 12.c, 12.d), lo más eficiente será codificarlas en función de su espectro, utilizando campos receptivos sinusoidales. En el caso de las imágenes naturales, cuyo espectro es de la forma $1/f$ (fig. 12.e y 12.f) no es apropiada ninguna de estas aproximaciones. En este caso habrá que utilizar un sistema de detectores con campos receptivos simultáneamente localizados en ambos dominios y que recubran no uniformemente el dominio frecuencial, (según una densidad $1/f$).

tipo wavelet, contienen información local, es sencillo plantear una modelización de la reducción de información espacialmente variante[37-39], tratando de forma diferente a los coeficientes con igual significado frecuencial pero diferente localización espacial. (Véase el apartado 4.1.2). Esto es una ventaja respecto del análisis de Fourier, donde cada coeficiente contiene información de todos los puntos de la imagen, y por lo tanto es complicado plantear procesos *R* espacialmente variantes. La formulación de Fourier solo será válida en zonas donde el sistema sea homogéneo, como es la fóvea en el caso del sistema visual. Para conseguir una descripción correcta del comportamiento extrafoveal, es necesario un modelo no homogéneo, modelo fácilmente formulable en un dominio conjunto espacio-frecuencia espacial.

- 4.- *No es posible una detección independiente de frecuencias y orientaciones mediante unos campos receptivos centro-periferia.* Una propiedad deseable en los detectores encargados de realizar el análisis de las señales visuales (y presente en el caso del sistema visual[9,23-25,40-42]) es la sintonización simultánea de frecuencias y orientaciones, es decir, detectores que sean sensibles exclusivamente a ciertas frecuencias con ciertas orientaciones. En un contundente trabajo de 1983, Daugman[42] demuestra que los detectores centro periferia no pueden presentar dicha propiedad aunque se deformen o se dispongan repetitivamente. Eso implica que la misión del filtrado pasa-banda representado por los campos receptivos centro-periferia de las primeras etapas del procesado visual es un realce de bordes y una eliminación de ruido de alta frecuencia, pero necesariamente debe ir seguido por una reorganización de las señales, de forma que a un nivel superior, tengamos una transformada cuyos *detectores* estén localizados tanto en frecuencia como en orientación en el dominio de Fourier, es decir, los campos receptivos de estos detectores deben ser funciones oscilantes (cuya oscilación tiene una orientación bien definida) y simultáneamente localizadas en el dominio espacial y el frecuencial.

- 5.- *La implementación directa de transformaciones espacio-frecuenciales requiere, en principio, menos conexiones entre pixels* debido a que se minimiza la porción de imagen

empleada en el cálculo de cada coeficiente[13]. Además de simplificarse la red de conexiones físicas entre los fotorreceptores, este hecho robustece el sistema en su conjunto, porque evita que un ruido localizado en la imagen, o un defecto local en la red de conexiones perturbe todos los coeficientes de la transformada.

3.2.2 Argumentos experimentales: datos fisiológicos y psicofísicos.

Los datos obtenidos electro-fisiológicamente acerca de la forma de los campos receptivos de las células del córtex de gatos y macacos[23-25] muestran que dichas células presentan máximos de sensibilidad frecuencial separados por octavas y que su anchura de banda crece con la frecuencia, es decir, recubren más densamente la zona de bajas frecuencias, y menos densamente la zona de altas frecuencias. Las oscilaciones detectadas en los campos receptivos y sus propiedades en el dominio de Fourier 2D han sido ajustadas por varios autores[9,44] a funciones de Gabor.

Además de estos estudios fisiológicos y de las modelizaciones correspondientes, también se han realizado estudios psicofísicos en humanos para determinar la extensión y características de los mecanismos de detección en el dominio frecuencial[40-42]. Las más modernas de estas experiencias[41,42] se basan en masking de frecuencias: la determinación de las diferencias en la sensibilidad del sistema visual a diferentes sinusoides en presencia de un estímulo sinusoidal enmascarador. En estos casos, el estímulo enmascarador de alto contraste fija el mecanismo de detección que es utilizado. Midiendo la sensibilidad del sistema a otras componentes frecuenciales determinamos cuáles son las características del canal seleccionado mediante el estímulo enmascarador. Los resultados obtenidos son consistentes con las medidas fisiológicas. En concreto se confirma la sintonización simultánea en frecuencia y orientación, las anchuras de banda y anchura angular máxima de los mecanismos de detección. Con todo esto, se reafirma la idea de que la operación básica de codificación se realiza en función de unos detectores localizados espacial y frecuencialmente.

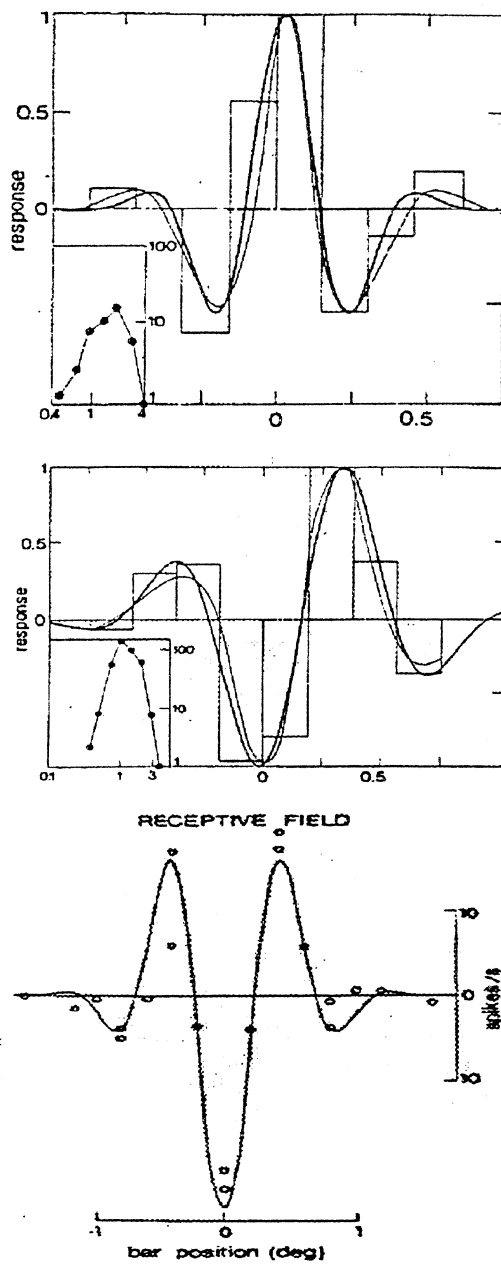


Fig. 13: Ajuste de funciones de Gabor a los resultados de campos receptivos corticales medidos electrofisiológicamente. (Marcelja [44])

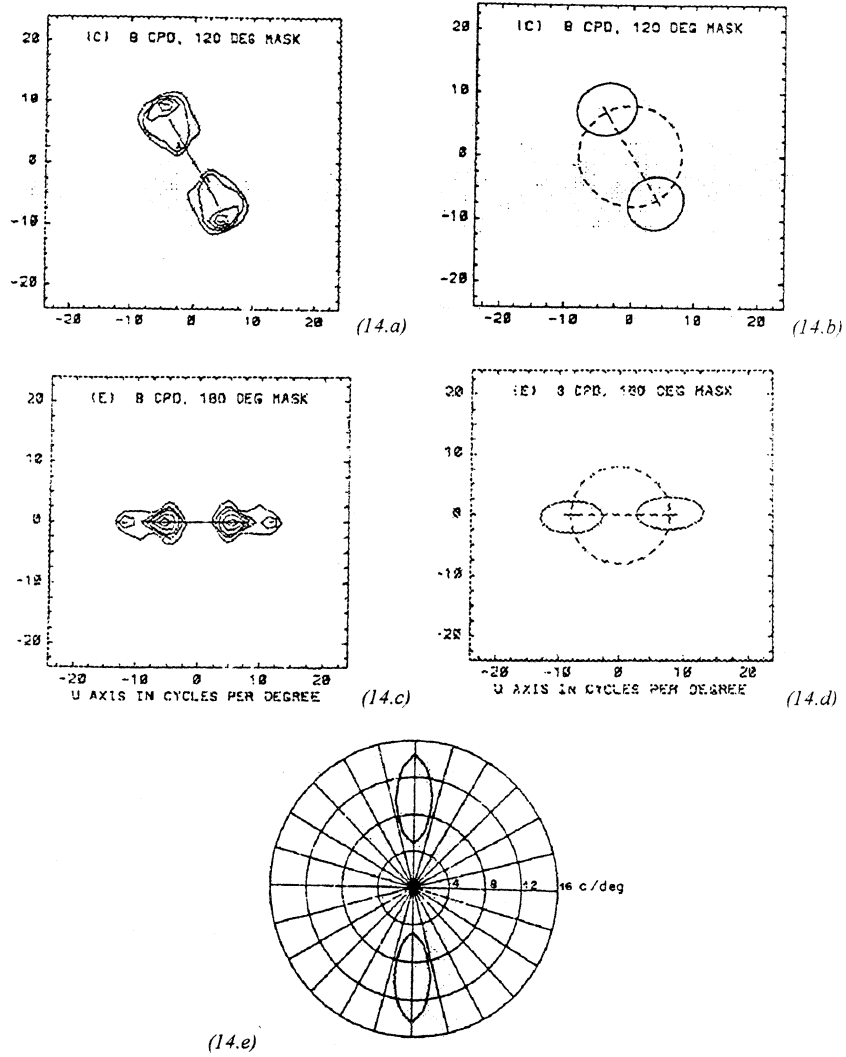


Fig. 14: Datos psicofísicos respecto a la sensibilidad frecuencial de los mecanismos de detección obtenidos mediante masking de frecuencias. Figs. 14.a y 14.c medidas experimentales de Harvey et al. [42], figs. 14.b y 14.d espectros de las funciones de Gabor correspondientes a dichos mecanismos. Fig. 14.e: espectro ajustado por Daugman[41]

Además, los experimentos de masking estaban orientados a poner de manifiesto la separabilidad o no separabilidad polar-radial de los mecanismos en el dominio frecuencial. El resultado del estudio sistemático de esta cuestión puso de manifiesto que no existe tal separabilidad y que la banda de paso de los mecanismos en el dominio frecuencial ajusta mejor a una gaussiana bidimensional (no separable) que a otros filtros localizados, como los laplacianos[45] o los sectores circulares[46,47].

3.3 UTILIZACIÓN DE LA FORMULACIÓN ESPACIO-FRECUENCIAL EN LA ACTUAL INVESTIGACIÓN EN VISIÓN.

3.3.1 Situación general.

Con la contribución de los avances en el campo de la matemática aplicada y el tratamiento de señal, se han establecido los fundamentos matemáticos de una representación completa y eficiente de las señales 1D y 2D en función de bases de soporte compacto, y se han demostrado las ventajas, ya comentadas, de dicha formulación[7,25,48-53].

Estas ideas han sido rápidamente aplicadas en el ámbito de la visión artificial, fundamentalmente para codificación óptima, compresión de imágenes y cancelación de ruidos[34,35,54].

Por otro lado, en el ámbito del estudio y la modelización del sistema visual, la introducción de este tipo de análisis se ha debido tanto al reconocimiento de sus ventajas respecto del análisis de Fourier, como a su adecuación a los datos experimentales citados. Desde hace ya algunos años, se proponen modelos de visión espacial que incluyen en su primera etapa una transformación T diferente de la transformada de Fourier, basada en una descomposición respecto de funciones de soporte compacto[9,15,19,32,46,47,55-58]. Esta formulación está resultando muy popular en el análisis y percepción de texturas[58-60], y además, en los experimentos psicofísicos, comienzan a utilizarse estímulos localizados espacial y frecuencialmente para estudio de la interacción entre los diferentes mecanismos de detección en el reconocimiento de patrones complejos[56], o los estudios de inducción y enmascaramiento[61,62].

3.3.2 ¿Funciones de Gabor o wavelets spline de orden n ?

Aceptado este tipo de análisis, queda por resolver la cuestión de elegir alguna de las diferentes formulaciones espacio-frecuenciales posibles. Como explicaremos a continuación, esta es una cuestión menor desde el punto de vista de cual va a ser nuestra aportación.

Cualquiera de las diferentes aproximaciones al análisis espacio-frecuencial: análisis mediante wavelets[13, 51-53,63,64], Short Time Fourier Transform (STFT) [65,66], análisis de Gabor[7,48,49], etc., posee la propiedad que está detrás de todas las ventajas citadas en 3.2.1: todas ellas se fundamentan en funciones que están simultáneamente localizadas en ambos dominios y pueden constituir una base completa integrada por funciones con un amplio rango de anchuras espaciales y anchuras de banda.

La diferencia entre las formulaciones posibles reside en la forma de obtener las funciones base y la expresión analítica final de las mismas:

- En el análisis mediante wavelets el conjunto de funciones base se obtiene a partir de dilataciones, compresiones, traslaciones y rotaciones (todas ellas continuas o discretas) de una única función generatriz[51,53]. La elegancia de este método tiene el inconveniente de que no existe una expresión analítica concreta para el cálculo de esa función generatriz, sino que puede ser cualquier función que cumpla una serie de teoremas[13].
- En el análisis mediante STFT la base está formada por funciones armónicas de frecuencia f , que están moduladas por una función ventana $w(x, x_0)$, centrada entorno a x_0 , que restringe la extensión espacial de la función:

$$g(x, p) = g(x, x_0, f) = w(x, x_0) e^{-i2\pi fx} \quad (3.7)$$

- En el caso del análisis de Gabor, la ventana utilizada es una función gaussiana. Eso permite un control intuitivo de la extensión espacial y frecuencial de la función (expresiones 2.4 y 2.11), además de un tratamiento paralelo de la transformada en ambos dominios[67].

También existen otras diferencias menores entre las diferentes formulaciones, como por ejemplo la separabilidad del espectro de las funciones base en frecuencia y orientación, o el grado en que minimizan la relación de incertidumbre. Es cierto que existen resultados psicofísicos y teóricos que restringen un poco el tipo de funciones que podemos utilizar: por ejemplo, la no separabilidad experimental en frecuencia y orientación de las funciones de sintonizado en el plano de Fourier[41,42], excluiría la posibilidad de utilizar algunas funciones de soporte compacto como por ejemplo las de Watson[46,47] (que son si separables), o como dijimos en 2.1.4, la exigencia estricta de área espacio-frecuencial mínima, solo permitiría la utilización de funciones de Gabor. Sin embargo, también dijimos en 1.1.2 y en 2.1.4 que si las diferencias cuantitativas entre dos caracterizaciones posibles son pequeñas (véanse las figuras 15 y 16), pueden admitirse pequeñas inconsistencias con los datos experimentales, o comportamientos teóricos sub-óptimos, si las ventajas matemáticas de elegir esa caracterización son grandes. De hecho, es difícil afirmar, por ejemplo, que un experimento psicofísico hecho con estímulos como los de 15.a o los de 16.a vaya a dar resultados sustancialmente diferentes.

En nuestro caso particular, la idea que proponemos, aunque asume una transformación espacio-frecuencial, se refiere fundamentalmente a R , con lo cual la elección de uno u otro tipo de función no es relevante¹.

En resumen, nosotros hemos elegido la formulación de Gabor porque reunía las ventajas propias de cualquier función de soporte compacto (apdos. 3.2.1 y 3.2.2), y se mostraba matemáticamente tratable. A la hora de decidir un algoritmo concreto de transformada (apdos. 2.2.3 y 5.3), o una base concreta respecto de la que transformar (apdo. 5.2), aplicaremos el mismo razonamiento: alcanzar un compromiso entre la comodidad matemática, la eficiencia teórica y la adecuación a los resultados experimentales.

¹ Recordemos que una transformación T completa, centra en R el peso de todos los procesos de reducción de información. Además, como veremos en 4.1.2, los métodos presentados son válidos para cualquier T completa.

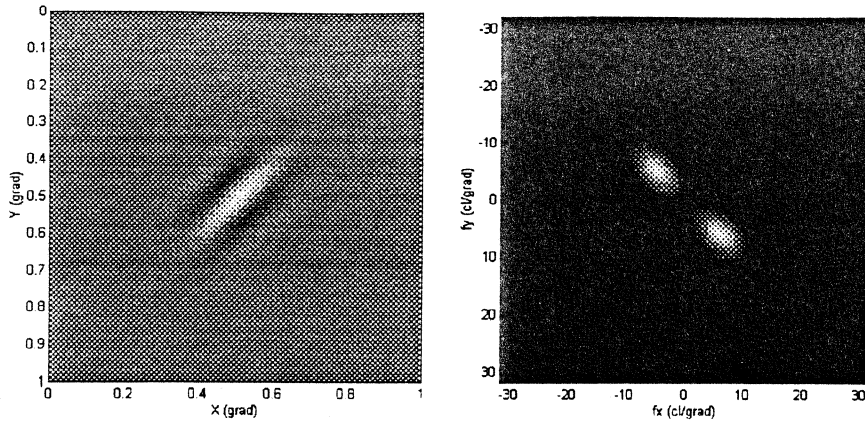


Fig. 15: (a) Parte real de una función de Gabor centrada en $x=(0.5, 0.4)$ deg y con frecuencia $f=(5.5, 5.5)$ cl/deg.
(b) Módulo del espectro de la función de 15.a.

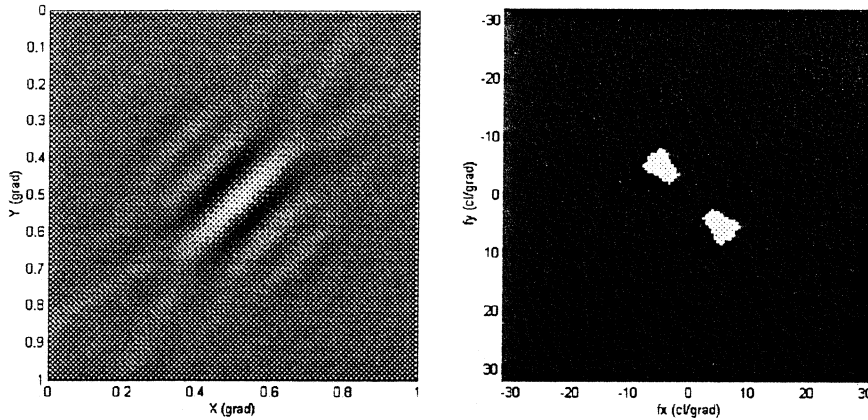


Fig. 16: (a) Parte real de una función base de la Cortex Transform de Watson [46,47]. Estas funciones están definidas por los parámetros de su espectro.
(b) Módulo del espectro de la función de 16.a. Banda de paso: entre 4 y 10 cl/deg, orientación: 45° , anchura angular: 30° .

4. EL SISTEMA VISUAL COMO FILTRO LINEAL EN UN DOMINIO CONJUNTO.

Función de Sensibilidad a estímulos de Gabor: GSF.

Después de reducir nuestro problema a la determinación de T y R (capítulo 1), estudiar las propiedades de las funciones y la Transformada de Gabor (capítulo 2), y asumir, conocidos los argumentos a su favor, que la transformación T es una transformación espacio-frecuencial, y en particular una Transformada de Gabor completa (capítulo 3), es en este capítulo donde hacemos nuestra aportación original más importante: proponemos que los procesos de reducción de información R pueden modelizarse, al menos a nivel umbral, através de un filtrado lineal en un dominio de Gabor 4D. Llamamos Función de Sensibilidad a estímulos de Gabor (GSF) a la función de pesos definida en el dominio conjunto espacio-frecuencial que da cual es la atenuación relativa que el sistema impone sobre cada una de las componentes de la Transformada de Gabor (TG) de la imagen incidente.

En este capítulo deducimos las relaciones generales entre dos funciones filtro lineal R_1 y R_2 que representan la actuación de *un mismo sistema* en los dominios de *dos transformadas diferentes* T_1 y T_2 . Aplicando estas relaciones generales, encontramos la expresión que liga la GSF, filtro lineal en un dominio de Gabor, con la Función de Sensibilidad al Contraste (CSF), filtro lineal en el dominio de Fourier. De este modo se hace posible el cálculo de la GSF a partir de las medidas experimentales disponibles sobre CSFs.

4.1 FILTRADO LINEAL EN UN DOMINIO DE GABOR.

Antes de lanzarnos al cálculo de la función filtro que caracteriza el comportamiento del sistema visual en un dominio de Gabor, estudiaremos las posibilidades que aporta dicho dominio, y las precauciones que hay que tomar a la hora de definir, en el mismo, una función filtro cualquiera.

4.1.1 Dominios espacio-frecuenciales de Gabor.

Como vimos en 2.2, la Transformada de Gabor de una señal $i(x)$ consiste en hallar su proyección respecto del conjunto de funciones de Gabor $\{g(x,p)\}_{p \in P}$, es decir, obtener su representación en el dominio de parámetros P . En el apartado 2.1.5 vimos que las funciones de Gabor bidimensionales dependen de 8 parámetros: $P = \{x_0, y_0, f_0, a, b, \alpha, \phi\}$, con lo cual, en principio, el dominio de la transformada tendría que ser de dimensión 8, sin embargo, como ya anunciábamos en 2.2, y veremos en detalle en 5.2, debido a factores estadísticos y fisiológicos, conviene ligar las anchuras y la orientación de la envolvente gaussiana de las funciones base con la frecuencia de su oscilación. Si además fijamos en $\phi=0$ el valor de la fase inicial de las oscilaciones, solo quedan cuatro grados de libertad: dos valores fijan la posición central de la función y otros dos la frecuencia de la oscilación, es decir $p = (x_0, f_0)$. Esto implica que al efectuar una transformada de Gabor pasamos de un dominio espacial 2D a uno espacio-frecuencial de doble dimensión.

Cada *posición* $p = (x_0, f_0)$ en el dominio 4D de la transformada tendrá un significado espacial y frecuencial concreto: el valor f_p de la transformada en ese punto indicará que en un entorno de x_0 la señal $i(x)$ tiene una contribución de la banda f_0 proporcional a f_p . La posición, la frecuencia, el tamaño de ese entorno y la anchura de la banda f_0 dependerán del conjunto de funciones de Gabor que hayamos elegido como base. Por eso debe hablarse de *un* dominio de Gabor y no de *el* dominio de Gabor.

Este significado espacio-frecuencial tan accesible es el que facilita la caracterización de sistemas espacialmente variantes, porque permite la actuación diferenciada sobre contribuciones frecuenciales en diferentes posiciones espaciales.

4.1.2 Filtrado lineal en un dominio de Gabor 4D: posibilidades y limitaciones.

Un determinado dominio será muy adecuado para efectuar operaciones de tratamiento de señales si el resultado de la actuación sobre sus coeficientes es potente y fácilmente interpretable.

El mejor de los casos se presenta cuando la transformación T es una proyección sobre una base ortogonal. En ese caso, el coeficiente de la transformada en cada punto del dominio P solo depende de la relación entre $i(\mathbf{x})$ y $g_p(\mathbf{x})$, con lo cual, los efectos de la modificación de cada coeficiente solo involucran a la función base correspondiente. En ese caso, la interpretación de la modificación aplicada es sencilla.

La cosa se complica cuando la base es redundante. En ese caso, el coeficiente en cada punto p también depende de la relación de la función $g_p(\mathbf{x})$ con el resto de funciones base $g_{p'}(\mathbf{x})$. Si esto es así, la interpretación del efecto de la modificación de un coeficiente particular es más compleja, porque el valor que posee depende también del valor de los coeficientes de su entorno¹.

En el nuestro caso particular², tenemos la ventaja de que se duplica la dimensionalidad del dominio de la transformada, con lo cual podemos realizar un filtrado más selectivo que el filtrado de Fourier, pero como veremos, existe el inconveniente de la no ortogonalidad de la base, que hace más delicado el diseño de filtros en ese dominio.

Como es conocido, un procesado espacialmente variante no puede ser formulado de forma sencilla en un dominio de Fourier porque en el caso de que la respuesta impulsional del sistema varíe con la posición, la integral de superposición no puede reducirse a una operación de convolución[68]. En estos casos, la única posibilidad es definir múltiples funciones filtro de validez local. Esta dificultad es lógica considerando que las funciones base asociadas a la representación de Fourier están extendidas en todo el espacio, con lo cual, la modificación de los coeficientes de la transformada de Fourier tiene un efecto global.

¹ Entorno formado por los puntos de P cuyas $g_{p'}(\mathbf{x})$ no sean ortogonales a $g_p(\mathbf{x})$.

² Siendo T una TG completa respecto de una cierta base separable. Véase 2.2.3 para el algoritmo general y 5.2 y 5.3 para los detalles a cerca de la base empleada.

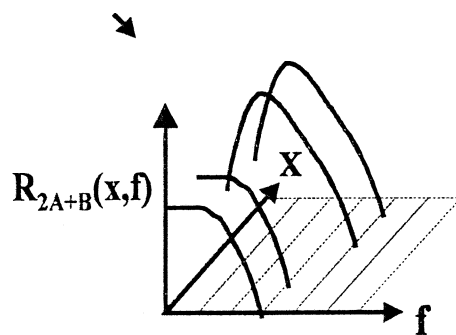
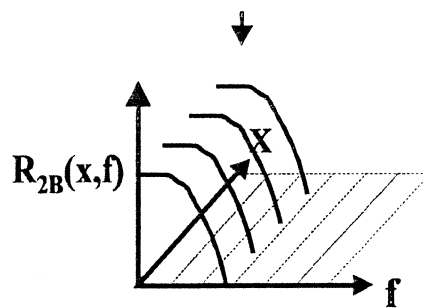
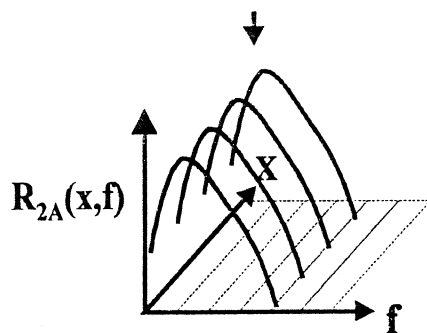
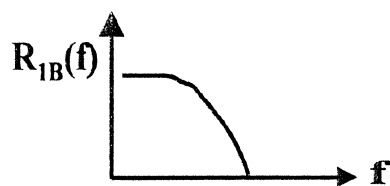
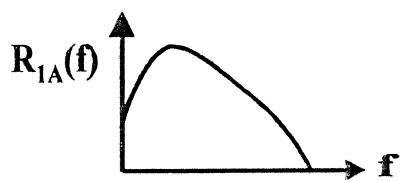
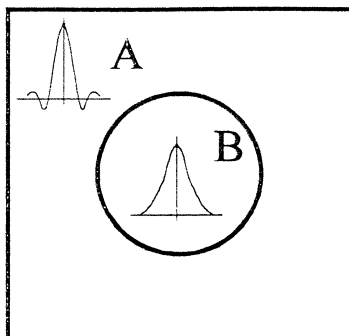
Las dificultades con los sistemas no homogéneos podrían superarse operando en un dominio tal que las funciones base de la representación estén localizadas en el espacio además de poseer cierto significado frecuencial. Esa es precisamente la propiedad común a todas las funciones de soporte compacto. Apoyándonos en esa idea fundamental, el problema del procesado espacialmente variante queda resuelto conociendo la forma de relacionar un filtro en el dominio de Fourier con su equivalente en el dominio espacio-frecuencial. La síntesis, o la caracterización, de un sistema espacialmente variante podría realizarse según un proceso como este:

1. Partimos del conocimiento de las respuestas impulsionales (o las funciones de transferencia en el dominio de Fourier R_{1A} , R_{1B}), que deseamos aplicar localmente.
2. Calculamos los filtros R_{2A} y R_{2B} , en el dominio conjunto espacio-frecuencia, equivalentes a los filtros de Fourier dados. (Véase el apartado 4.3)
3. Combinamos, en diferentes posiciones espaciales del dominio conjunto, los valores de los filtros 4D calculados, R_{2A} y R_{2B} , para generar R_{2A+B} .

La figura 17 muestra un ejemplo del proceso general descrito. En el apartado 4.3 obtendremos las relaciones necesarias para pasar de las funciones filtro R_1 a las funciones filtro R_2 .

Cuando, como en nuestro caso, la base no sea ortogonal este proceso estará sujeto a ciertas restricciones: al combinar en el dominio conjunto los diferentes filtros R_2 habrá que evitar grandes diferencias de atenuación sobre coeficientes *muy próximos*, es decir, sobre coeficientes cuyas funciones base tengan un alto grado de solapamiento. Volveremos sobre este tema al tratar ejemplos concretos con el algoritmo y la base particular utilizada (Véanse los apartados 6.1 y 6.2).

Fig. 17 (Pag. sig.): Ejemplo de síntesis de un sistema espacialmente variante. Se desea sintetizar un sistema que realice un promediado en el centro de la imagen (zona B) y un realce de bordes en la periferia (zona A), es decir, han de aplicarse en cada zona las respuestas impulsionales correspondientes. Para sintetizar el filtro, se parte de los filtros de Fourier correspondientes a las respuestas impulsionales deseadas (pasa baja y pasa banda) y se calculan los filtros equivalentes en el dominio conjunto (que son espacialmente invariantes). Por último se combinan los valores de los filtros conjuntos en diferentes posiciones espaciales (el pasa baja en la zona cercana al origen y el pasa banda en la zona más externa)



4.2 EL SISTEMA VISUAL COMO FILTRO LINEAL EN UN DOMINIO DE GABOR.

4.2.1 El sistema visual como filtro lineal en el dominio de Fourier: la CSF.

La caracterización habitual de los procesos de reducción de información dados en las primeras etapas del procesado visual, suelen modelizarse proponiendo un filtrado lineal en el dominio de Fourier[26-29]. En ese caso, se supone que el sistema actúa como un banco de filtros pasa-banda, cuya banda de paso es una función delta en el dominio frecuencial, de modo que cada detector solo es sensible a la amplitud (o el contraste) de las sinusoides de una frecuencia y orientación determinadas. Así pues, la señal queda descompuesta en sus componentes de Fourier.

La actuación conjunta del sistema está determinada por la sensibilidad individual de los citados detectores. Si los detectores sensibles a frecuencias de una cierta banda son más sensibles que el resto, dicha banda resulta resaltada respecto del resto, aumentándose la importancia relativa de ciertas características de la señal. Generalmente se admite que estos detectores de sinusoides, responden linealmente al contraste al menos a nivel umbral[28,29,69]. De esa forma, en la zona de bajos contrastes, la actuación del sistema estará totalmente determinada mediante un conjunto de detectores de sinusoides y por una función de pesos que atenúe linealmente la salida de cada uno de los detectores. Admitiendo, pues, que la transformación T es una transformada de Fourier, las modificaciones R que introduce el sistema pueden caracterizarse a nivel umbral mediante un filtrado lineal en el dominio de Fourier.

La Función de Sensibilidad al Contraste (CSF) es la función filtro que define la importancia relativa concedida por el sistema a cada uno de los coeficientes de la transformada de Fourier. (fig. 18).

Según esto, la codificación y distorsión efectuada por el sistema visual viene dada por:

$$f(p) = \mathbf{T}[i(x)](p) = R [T [i(x)](p)](p) = CSF(f) * \mathbf{TF}[i(x)](f)$$

(4.1)

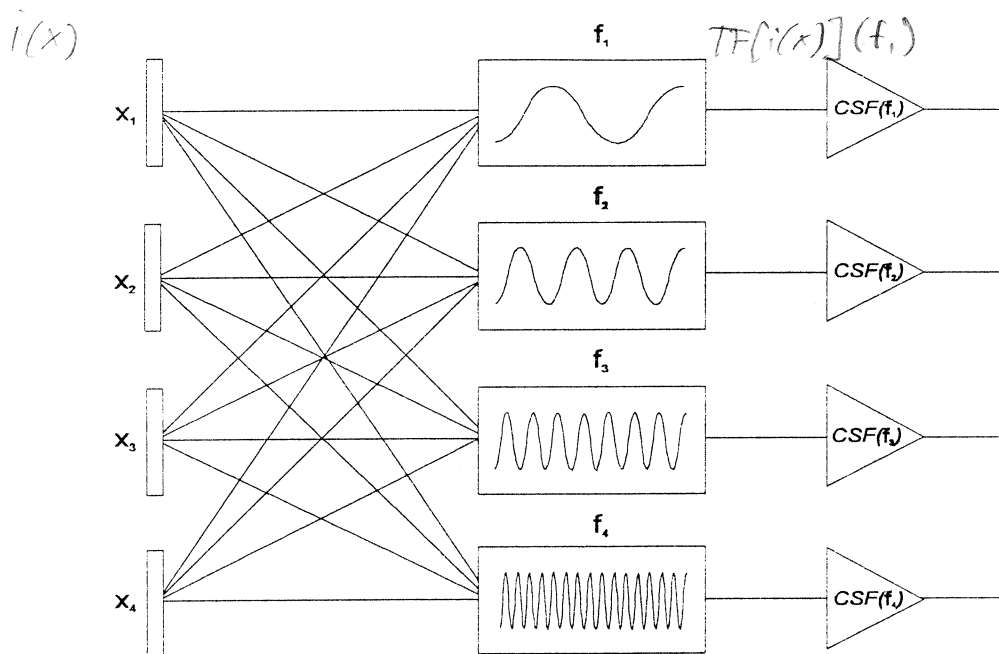


Fig. 18: Caracterización de los procesos T y R en el dominio de Fourier. La transformación T se produce mediante la actuación de un sistema de detectores sensibles a cada función base de la representación de Fourier. La eliminación de información R se modeliza mediante una atenuación lineal de la señal procedente de cada detector frecuencial.

4.2.2 El sistema visual como filtro lineal en un dominio de Gabor: la Función de Sensibilidad a estímulos de Gabor.

Como hemos visto en el capítulo 3, numerosas razones teóricas y algunas evidencias experimentales, sugieren que la transformación T no es una transformación de Fourier: sino algún tipo de transformación espacio-frecuencial como por ejemplo una transformada de Gabor. Lo que proponemos en este trabajo, es caracterizar el comportamiento del sistema en los dominios espacio-frecuenciales de forma análoga a como se venía haciendo en el dominio de Fourier: mediante una función filtro lineal.

Asumimos que los campos receptivos del sistema de detectores son funciones de soporte compacto tipo Gabor, de forma que el resultado de su aplicación es una transformada de Gabor. Proponemos que la importancia concedida por el sistema visual

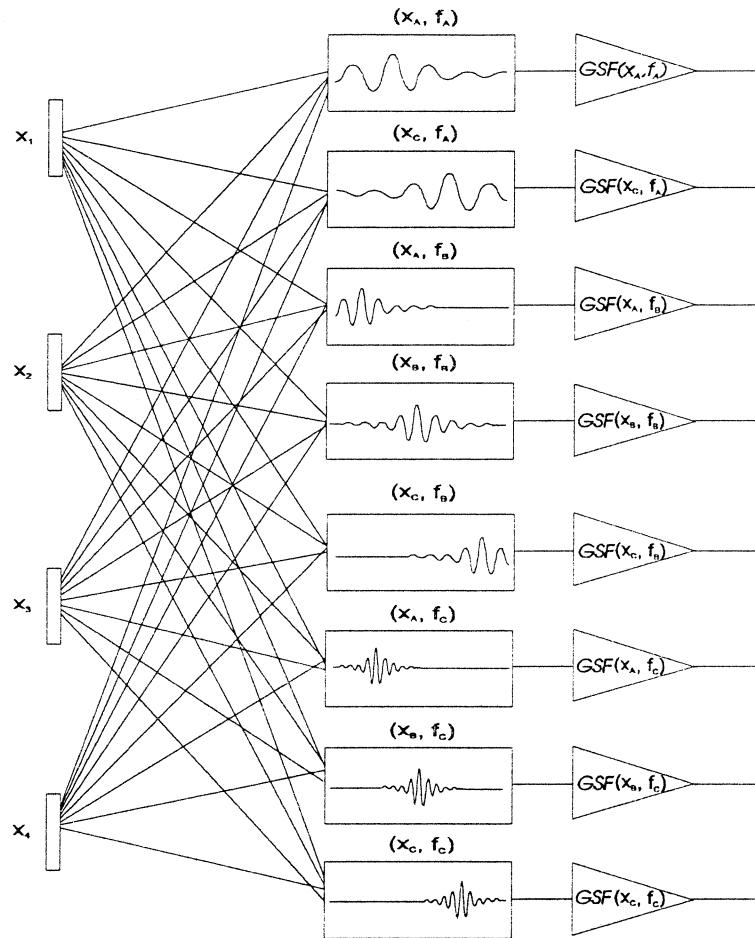


Fig. 19: Caracterización de los procesos T y R en el dominio de Gabor. La transformación T se produce mediante la actuación de un sistema de detectores sensibles a cada función base de la representación de Gabor. La eliminación de información R se modeliza mediante una atenuación lineal de la señal procedente de cada detector espacio-frecuencial.

a la señal procedente de los diferentes detectores de Gabor es lineal a nivel umbral, y dependiente de la frecuencia y posición espacial a la que esté sintonizado cada detector. De esta forma, el comportamiento global del sistema estará determinado por la función que dé cual es la importancia relativa concedida por el sistema a los diferentes coeficientes de la transformada de

Gabor. En resumen, tras aceptar que T es una transformada de Gabor, proponemos que R es un proceso de filtrado lineal en el dominio de Gabor correspondiente.

Definimos Función de Sensibilidad a estímulos de Gabor (GSF) a la función de pesos que indica cual es la atenuación relativa que el sistema impone sobre cada una de las componentes de la Transformada de Gabor (TG) de la imagen incidente. (fig. 19)

Lo que proponemos, pues, es que la codificación y distorsión efectuada por el sistema visual viene dada por:

$$f(p) = T[i(x)](p) = R [T [i(x)](p)](p) = GSF(x,f) * TG[i(x)](x,f) \quad (4.2)$$

4.2.3 Determinación experimental de la Función de Sensibilidad a estímulos de Gabor.

Atendiendo exclusivamente a la definición de la GSF , la función filtro en el dominio de Gabor de cualquier sistema lineal puede hallarse a partir del cociente de las amplitudes de entrada y de salida de las funciones base:

$$GSF(x,f) = A'(x,f)/A(x,f) \quad (4.3)$$

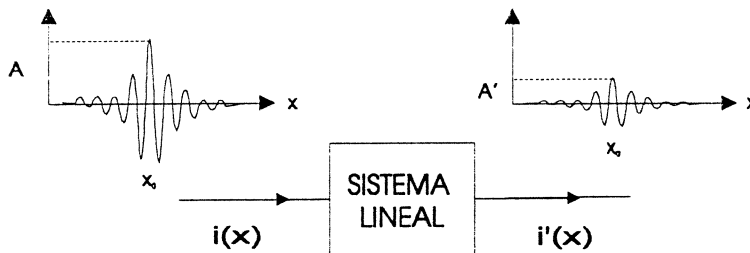


Fig. 20: Medida intuitiva de la GSF: medida de la relación entre las amplitudes de entrada y salida de las funciones base.

Esta medida trivial de los coeficientes de la GSF no es aplicable en sistemas en los que no se tiene acceso directo a la señal de salida como es el caso del sistema visual. Este problema en la determinación experimental de la función filtro se presenta también en el caso de plantear la modelización en el dominio de Fourier. En nuestro caso podemos adoptar una solución análoga

a la utilizada en el caso de la *CSF*[28]: si suponemos que la mínima amplitud (de la señal que sale del sistema de adquisición) necesaria para la detección es independiente de la frecuencia y la posición espacial, la amplitud mínima del estímulo necesaria para la detección tiene que ver con la atenuación introducida por el sistema sobre esa componente. Es decir, suponemos que para amplitudes umbrales se cumple:

$$GSF(\mathbf{x}, f) * A_u(\mathbf{x}, f) = \delta \quad (4.4)$$

donde $A_u(\mathbf{x}, f)$ es la mínima amplitud para que estímulo de Gabor $g(\mathbf{x}', \mathbf{x}, f)$ sea perceptible, y δ es el umbral mínimo de detección para las señales que salen del sistema de adquisición. En ese caso, la *GSF* en cada punto p del espacio conjunto será inversamente proporcional a la amplitud umbral de la función base correspondiente $g(\mathbf{x}, p)$:

$$GSF(p) \propto 1/A_u(p) \quad (4.5)$$

Según esto, la determinación experimental de la *GSF* involucra la medida de las amplitudes¹ mínimas necesarias para la detección de las funciones de Gabor de la base elegida.

4.3 RELACIONES MATEMÁTICAS ENTRE LA *CSF* Y LA *GSF*.

La *CSF* y la *GSF* son dos caracterizaciones del comportamiento de un mismo sistema en dos dominios de representación diferentes. Si el comportamiento global del sistema está análogamente caracterizado mediante la aplicación de cualquiera de esos filtros en sus dominios correspondientes, ambos filtros deben estar relacionados. En este apartado obtenemos cual debe ser esa relación para cualesquiera funciones

¹ Si trabajamos con funciones normalizadas en energía y con extensión espacial dependiente de la frecuencia, las amplitudes pueden ser diferentes en función de la frecuencia: para las funciones de alta frecuencia, y pequeña extensión espacial, la amplitud será grande, mientras que para las funciones de baja frecuencia, y gran extensión espacial, la amplitud será mucho menor. Esta dependencia de la amplitud con la frecuencia obliga a definir una amplitud normalizada para cada frecuencia, refiriendo las amplitudes de una determinada función a la amplitud máxima para esa frecuencia. Se entiende que las amplitudes a las que se hace referencia en las expresiones 4.4 y 4.5 están normalizadas en este sentido.

filtro R_1 y R_2 definidas en los dominios de representación inducidos por las transformadas T_1 y T_2 . Aplicando estas relaciones generales al caso que nos ocupa podremos obtener teóricamente, sin determinar los umbrales de detección de las funciones de Gabor, los coeficientes de la *GSF* a partir de los datos de las *CSFs* disponibles.

4.3.1 Relaciones entre dos funciones filtro R_1 y R_2 definidas en los dominios inducidos por las transformadas T_1 y T_2 .

Consideremos un sistema lineal S caracterizado, equivalentemente, por las funciones filtro $R_1(p)$ y $R_2(q)$ definidas en los dominios de parámetros $p \in P$ y $q \in Q$ inducidos por las transformaciones lineales T_1 y T_2 . El efecto de S sobre una señal $i(x)$ vendrá dado por cualquiera de estas expresiones:

$$i'_{R_1}(x) = T_1^{-1}[R_1(p) * T_1[i(x)](p)](x) \quad (4.6)$$

$$i'_{R_2}(x) = T_2^{-1}[R_2(q) * T_2[i(x)](q)](x) \quad (4.7)$$

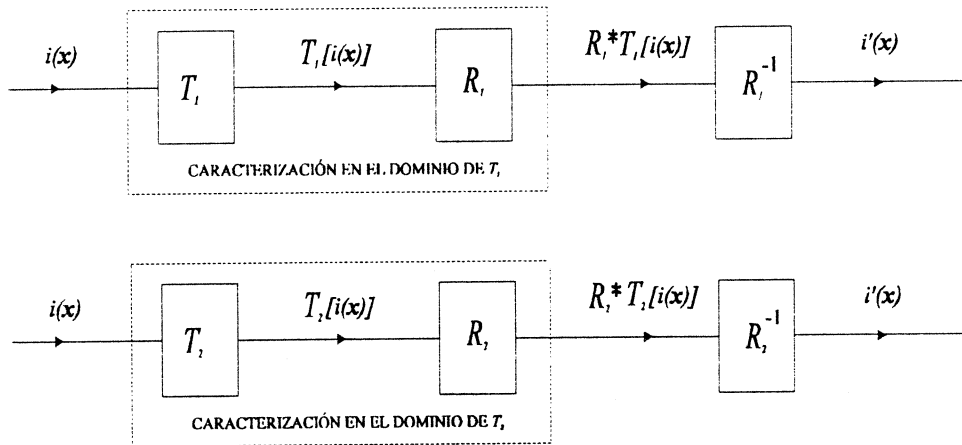


Fig. 21: Caracterización del proceso T en dominios diferentes.

Para que efectivamente se cumpla la equivalencia entre ambas caracterizaciones deberá cumplirse que la señal reconstruida a partir de las dos señales codificadas sea la misma, es decir:

$$T_1^{-1}[R_1(p) * T_1[i(x)](p)](x) = T_2^{-1}[R_2(q) * T_2[i(x)](q)](x) \quad (4.8)$$

La expresión 4.8 es la relación fundamental que liga las funciones filtro que caracterizan a un sistema en cualquier dominio de representación. A partir de ella podemos obtener la expresión explícita de $R_1(\mathbf{q})$ en función de $R_2(\mathbf{p})$.

Aplicando T_1 a ambos lados de 4.8, tenemos:

$$R_1(\mathbf{p}) * T_1[i(\mathbf{x})](\mathbf{p}) = T_1 [T_2^{-1}[R_2(\mathbf{q}) * T_2[i(\mathbf{x})](\mathbf{q})](\mathbf{x})](\mathbf{p}) \quad (4.9)$$

donde, para una imagen cualquiera $T_1[i(\mathbf{x})](\mathbf{p})$ tiene valores distintos de cero para un conjunto importante de puntos \mathbf{p} , de forma que no podemos despejar directamente ningún $R_1(\mathbf{p})$.

Sin embargo, si consideramos que $i(\mathbf{x})$ es una de las funciones base de la transformación T_1 no existirán interacciones complejas entre los coeficientes de la transformada y tendremos:

$$i(\mathbf{x}) = g(\mathbf{x}, \mathbf{p}') \Rightarrow T_1[g(\mathbf{x}, \mathbf{p}')](\mathbf{p}) = \delta(\mathbf{p} - \mathbf{p}') \quad (4.10)$$

con lo cual, la función definida en el dominio \mathbf{P} por ambos miembros de 4.9 solo tiene un valor significativo en \mathbf{p}' , y entonces podemos despejar $R_1(\mathbf{p}')$:

$$R_1(\mathbf{p}') = T_1 [T_2^{-1}[R_2(\mathbf{q}) * T_2[g(\mathbf{x}, \mathbf{p}')](\mathbf{q})](\mathbf{x})](\mathbf{p}') \quad (4.11)$$

Como vemos, mediante la expresión 4.11 podemos calcular el filtro R_1 , equivalente al filtro R_2 , en todos los puntos del dominio \mathbf{P} sin más que ir filtrando mediante R_2 las funciones base de T_1 e ir aplicando las transformadas correspondientes.

4.3.2 Relaciones entre la GSF y la CSF.

Como hemos dicho, la CSF y la GSF son dos caracterizaciones, como filtro lineal, del comportamiento de un mismo sistema en dos dominios de representación diferentes, con lo que la expresión 4.11 puede servir para establecer una relación entre ambas funciones.

En este caso, tenemos:

<i>DOMINIO CONJUNTO</i>	<i>DOMINIO FRECUENCIAL</i>
$T_1[\dots](\mathbf{p}) = TG[\dots](\mathbf{x}', \mathbf{f})$	$T_2[\dots](\mathbf{q}) = TF[\dots](\mathbf{f})$
$R_1(\mathbf{p}) = GSF(\mathbf{x}', \mathbf{f})$	$R_2(\mathbf{q}) = CSF(\mathbf{f})$

$$g(x,p) = g(x,x',f) \qquad g(x,q) = \exp(i2\pi fx)$$

Teniendo todo esto en cuenta pueden plantearse un par de ecuaciones análogas a 4.11:

$$GSF(x',f') = TG[TF^{-1}[CSF(f) * TF[g(x,x',f)]](x)](x',f') \quad (4.12)$$

$$CSF(f) = TF[TG^{-1}[GSF(x',f') * TG[\exp(i2\pi fx)](x)]](f) \quad (4.13)$$

4.3.3 Limitaciones de la relación entre la *GSF* y la *CSF*.

La principal restricción de las ecuaciones 4.12 y 4.13 procede del hecho de que relacionan objetos definidos en dominios de diferente dimensión, con lo cual, cierta información se pierde al efectuar el tránsito en cualquier sentido.

Obtención de la *GSF* a partir de la *CSF*

En efecto, el cálculo de la *GSF* equivalente a una determinada *CSF* mediante la ecuación 4.12 nos dará una *GSF* espacialmente invariante. Esto era de esperar ya que la formulación de Fourier no puede dar cuenta de las no homogeneidades de un sistema espacialmente variante, con lo cual, estas no pueden ser obtenidas al aplicar 4.12 porque, simplemente, no están presentes en la *CSF*.

En realidad las *CSFs* que se utilizan habitualmente solo tienen validez local: una determinada *CSF* solo es válida en la zona donde la respuesta impulsional del sistema variante se mantiene próxima a $TF^{-1}[CSF(f)](x)$. De esta manera, resulta que la *GSF* espacialmente invariante salida de una *CSF* mediante 4.12 tiene tan solo validez local, siendo exclusivamente aplicable a la zona en la que es válida la *CSF* inicial.

Esto no quiere decir que no podamos aprovechar las ventajas de la formulación espacio-frecuencial, en cuanto su capacidad para dar cuenta de las no homogeneidades de un sistema. La solución a este problema está en introducir una atenuación diferenciada en función de la posición para los coeficientes de un mismo significado frecuencial, y así obtener una respuesta impulsional espacialmente variante. Ya hemos dicho que el comportamiento espacialmente variante no puede extraerse de una única *CSF*, pero sí de varias *CSFs*, válidas en diferentes zonas del espacio, siguiendo el proceso descrito en el apartado 4.1.2 (fig 17). Las

no homogeneidades pueden incluirse en una única GSF calculando los valores de las $GSFs$ espacialmente invariantes equivalentes a las diferentes $CSFs$ locales, y combinando estos valores en las posiciones del dominio conjunto que les correspondan.

Obtención de la CSF a partir de la GSF

Si empleamos una GSF general (que incluya inhomogeneidades espaciales) para el cálculo de su CSF equivalente mediante la ecuación 4.13, estas inhomogeneidades serán promediadas al proyectar el dominio conjunto 4D (espacio 2D y frecuencia 2D) sobre el dominio frecuencial 2D. De esta forma se tendrá una CSF promedio que subestima la sensibilidad en la fovea y sobreestima la sensibilidad del sistema en las zonas extrafoveales.

Para evitar la eliminación de las inhomogeneidades por promediado, lo mejor es calcular $CSFs$ locales a partir de la respuesta impulsional del sistema para una cierta área:

$$CSF_{x_0} = TF[TG^{-1}[GSF(x', f') * TG[d(x-x_0)](x', f')](x)] \quad (4.14)$$

Comentarios finales

Resumiendo el conjunto de relaciones y matizaciones posteriores, podemos decir que con las expresiones 4.12 y 4.14 podemos calcular la GSF (espacialmente invariante) y la CSF (válida en la proximidad de un cierto x_0) a partir del conocimiento de la CSF y la GSF respectivamente.

Como vemos, las relaciones 4.12 y 4.14 no están limitadas por la elección de una base $\{g(x, x', f)\}$ concreta. Este hecho permite pues, una cierta libertad en la elección de las funciones base, porque las relaciones 4.12 y 4.14 indican que siempre vamos a poder hallar la caracterización adecuada de nuestro sistema independientemente de la base utilizada¹. El hecho de que la base concreta no sea un factor crítico en la modelización del sistema nos ayudará en el capítulo siguiente: al definir nuestra representación, no tendremos que preocuparnos por ajustarnos con precisión absoluta a un determinado tipo de base, sino que

¹ Siempre que sea completa.

bastará con reproducir a grandes rasgos las características de sintonizado en frecuencia y anchuras espaciales y de banda. Aunque no va a hacerse aquí, sería muy interesante contrastar el resultado de la *GSF* obtenida indirectamente a partir de la *CSF* mediante la expresión 4.12, con lo que se obtuviese en una medida experimental directa de la *GSF*. En la expresión 4.12 se supone implícitamente que la respuesta del sistema visual a un estímulo complejo (de Gabor) viene determinada totalmente por la forma en que el sistema responde a los estímulos simples (sinusoides) que lo constituyen, lo cual no tiene porque ser cierto.

5. DETERMINACIÓN DE LA *GSF* 4D A PARTIR DE LA *CSF* 2D DE LA ZONA FOVEAL

Después de describir con detalle cual es la caracterización propuesta, definir la función filtro en cualquier dominio de Gabor y analizar sus relaciones con la *CSF*, en este capítulo obtendremos la *GSF* 4D (correspondiente a la zona foveal y por lo tanto espacialmente invariante) de dos observadores concretos en un dominio de Gabor concreto.

Ese dominio de Gabor concreto es el inducido por una base separable inspirada en los datos fisiológicos sobre las características espectrales de los campos receptivos de los detectores corticales. La separabilidad de la base permitirá utilizar el algoritmo descrito en 2.2.3. Veremos como la elección de una base poco densa, incompleta, da lugar a transformaciones con pérdida de información y analizaremos el significado de los coeficientes obtenidos al aplicar el citado algoritmo con la base propuesta definitivamente.

No efectuamos una medida directa de la *GSF* debido a que el equipo experimental para la generación de los estímulos adecuados no está aun puesto a punto. Así pues, hemos obtenido las *GSFs* a partir de medidas de la *CSF* 2D, determinando sus valores através de la relación 4.12. El cómputo numérico de la expresión 4.12 implica tener valores de la *CSF* 2D en todo el dominio frecuencial. En este capítulo se describe una técnica experimental sencilla para obtener de forma bastante fiable la *CSF* 2D en cualquier punto del dominio de Fourier.

De los algoritmos y los datos sobre las *CSFs* 2D de dos observadores obtenemos las *GSF* 4D correspondientes. La bondad de los resultados se pone de manifiesto comprobando, mediante diferentes simulaciones, la equivalencia entre las caracterizaciones en ambos dominios.

5.1 REQUISITOS PARA LA DETERMINACIÓN DE LA *GSF* 4D.

El objetivo de este trabajo es, además de proponer la caracterización del sistema visual como un filtro lineal en un dominio de Gabor, determinar cuantitativamente los valores de dicho filtro.

Un requisito previo a la determinación de la *GSF* es decidir el modelo de canales de Gabor que va a utilizarse. Recuérdese que el dominio conjunto y el algoritmo de la *TG* dependen de la base elegida. Los valores concretos de la *GSF* dependerán también de dicho modelo porque la atenuación ejercida por un sistema sobre cada componente de la señal en dos dominios conjuntos diferentes será en general diferente. Así pues, unos determinados valores de la *GSF* de un sistema están ligados a una base concreta (y a un algoritmo de cálculo de la transformada), con lo cual, al plantearse la medida o dar los valores del filtro hay que especificar cual es la base respecto de la que se está haciendo la transformada.

Una vez elegida la base, lo ideal sería determinar de forma experimental los valores del filtro en los puntos del dominio elegido. Esto implica, como vimos, la medida de las amplitudes umbrales de detección de las funciones de la base. Esa medida requiere la generación de los estímulos de Gabor con cualquier frecuencia, orientación y posición espacial y su presentación en un monitor calibrado. Lo tedioso de las medidas experimentales¹ y el hecho de que no se disponía de todo el equipo experimental para la presentación de los estímulos hizo aconsejable que en primera instancia se abordase tan solo el cálculo de la *GSF* a partir de las medidas de la *CSF* según la expresión 4.12.

En esta situación, los elementos necesarios para el cálculo de *GSFs* son tres:

- Modelo de canales de Gabor: conjunto de funciones base sobre el que realizar la descomposición de la señal.
- Algoritmo para el cálculo de los coeficientes de la transformada sobre la base elegida.

¹ Para conseguir una caracterización que nos permitiese la utilización numérica de la función de transferencia en principio habría que muestrear más o menos exhaustivamente dos dominios bidimensionales.

- Datos experimentales de los filtros de Fourier, y no solo curvas más o menos cualitativas, sino un muestreo exhaustivo del dominio frecuencial.

5.2 SELECCIÓN DE LA BASE: MODELO DE CANALES DE GABOR.

El problema de la definición de una base de funciones de Gabor consiste en fijar un conjunto de valores de los parámetros $p \in P$ tales que el conjunto de funciones de Gabor $\{g(x,p)\}$ sea una base de base completa de R^2 e incluya las propiedades básicas de los canales de detección del sistema visual.

Para que la base sea completa algún detector debe dar señal con independencia de la posición espacial o la composición frecuencial del estímulo, por eso el sistema de detectores a plantear debe recubrir completamente tanto el dominio frecuencial como el espacial.

Como las funciones de Gabor están espacialmente localizadas, será necesaria la presencia de varias funciones, con una determinada selectividad frecuencial, para recubrir todo el dominio espacial.

Por tanto, la elección de una base implica:

- La elección de un sistema de detectores cuyas bandas de paso induzcan un cierto recubrimiento del dominio frecuencial.
- La elección de la posición y el grado de solapamiento de las funciones campo receptivo en el dominio espacial para cada una de las bandas frecuenciales consideradas.

5.2.1 Recubrimiento del dominio frecuencial.

Los elementos de nuestra base deben reproducir las características fundamentales de los mecanismos frecuenciales de detección[9,23,41]:

- Existencia de cuatro o cinco picos frecuenciales de sintonización (cuatro o cinco mecanismos) separados en octavas (de forma que la frecuencias de sintonizado sigan esta relación: $f_{i+1} = 2f_i$).

-.Anchura de banda del orden de una octava.

-.Anchura en orientación constante (independiente de la frecuencia de sintonizado), y situada en torno a 30° .

Ya han sido planteados varios modelos de recubrimiento del dominio frecuencial que reproducen estas propiedades[6,16,46-47,67]:

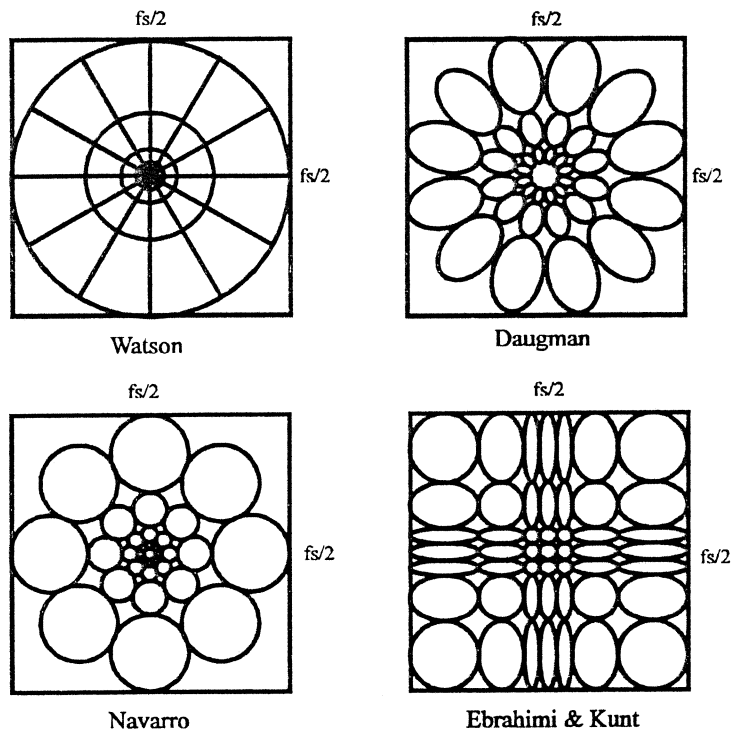


Fig. 22: Modelos de recubrimiento del dominio frecuencial con los espectros de los mecanismos de detección.

Algunos de ellos incluyen características adicionales presentes en el sistema visual como la no separabilidad polar[16,67], mientras que otros[6,46,47] se permiten simplificaciones como la separabilidad polar o la cartesiana para facilitar la formulación del problema.

Más allá de su mayor o menor adecuación a los datos fisiológicos, todos estos cumplen las propiedades esenciales de los canales visuales, ajustándose a la distribución energética y de varianza de las imágenes naturales: la densidad de detectores

es mayor en el centro del dominio frecuencial (zona de bajas frecuencias), y la separación entre mecanismos y su anchura crecen con la frecuencia.

Nosotros hemos elegido un recubrimiento mediante funciones separables (análogo al de Ebrahimi & Kunt[6]) porque así es posible la aplicación del algoritmo matricial descrito en 2.2.3.

En nuestro caso, el dominio de ha discretizado con una frecuencia de muestreo $f_s = 60$ cl/deg, de forma que la frecuencia máxima del dominio frecuencial discreto sea aproximadamente la frecuencia de corte del sistema visual.

En el recubrimiento separable elegido, se han empleado siete frecuencias centrales en cada dirección 1D (horizontal y vertical): $-3/8 f_s$, $-1/6 f_s$, $-1/18 f_s$, 0 , $1/18 f_s$, $1/6 f_s$, $3/8 f_s$, es decir 49 bandas frecuenciales 2D en total. Las anchuras de las bandas 1D son: $\Delta f = f_s/36$ para las de baja frecuencia ($|f| \leq 1/18 f_s$), $\Delta f = f_s/12$ para las de frecuencia $|f| = 1/6 f_s$, y $\Delta f = f_s/4$ para las de más alta frecuencia.

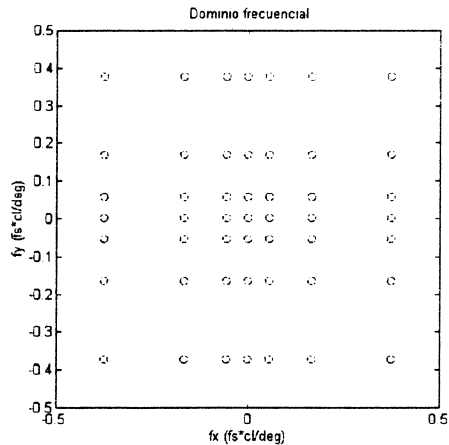


Fig. 23: Frecuencias centrales de las funciones base utilizadas. (Frecuencia de muestreo $f_s = 60 \text{cl/deg}$.)

5.2.2 Recubrimiento del dominio espacial.

Como ya hemos dicho, para conseguir una base completa es necesario que, para cada banda frecuencial, existan diferentes

detectores situados en diferentes posiciones espaciales de forma que el dominio espacial quede totalmente cubierto.

Esto puede conseguirse distribuyendo las funciones campo receptivo en una retícula cartesiana en el dominio de posiciones. El problema es establecer cual debe ser dicha retícula.

Las anchuras espaciales de las funciones de cada banda estarán determinadas por la inversa de su anchura frecuencial: pequeña extensión para las funciones de alta frecuencia y gran extensión para las de baja, pero en principio no existe ninguna relación entre el tipo de función y la retícula de posiciones espaciales a utilizar.

Para generar y controlar fácilmente los recubrimientos del dominio espacial imponemos la condición de solapamiento¹ constante: en cada dirección se situarán, dentro del dominio considerado, el número de funciones necesario para que el solapamiento entre funciones vecinas sea un determinado porcentaje de la norma de la función².

Manteniendo constante el solapamiento entre funciones vecinas se consigue que el número de funciones por unidad de área crezca con la frecuencia. Para alcanzar un determinado nivel de solapamiento las funciones de alta frecuencia deben situarse muy próximas (debido a su escasa extensión espacial) mientras que cuando baja la frecuencia (y crece la extensión espacial), debe haber más distancia entre las funciones para que se siga manteniendo la tasa de solapamiento impuesta. De esta manera, con un solo parámetro se controla el número de funciones diferentes que hay que utilizar para cada frecuencia y su colocación en el dominio espacial.

Dado un determinado recubrimiento frecuencial, lo único que nos falta por hacer es fijar el nivel de solapamiento que debe existir entre funciones base vecinas y de la misma banda. Evidentemente el valor del solapamiento no puede ser cualquiera: estamos limitados por la exigencia de que la base sea completa. Si por ejemplo imponemos un solapamiento demasiado pequeño las retículas resultantes serán muy poco

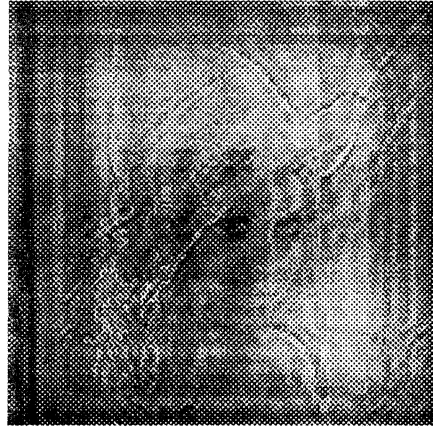
¹ El solapamiento entre las funciones se calcula mediante el producto escalar entre ellas.

² Las funciones base se toman de norma (energía) unidad.

tupidas, con lo que habrá tan pocas funciones base que no podrán dar cuenta de toda la información contenida en la señal. Por otra parte, no conviene utilizar un solapamiento demasiado grande para no generar una base muy redundante que implicaría un número excesivo de funciones base.



(24.a)



(24.b)



(24.c)



(24.d)

Fig. 24: Efecto de la variación del solapamiento espacial entre las funciones base. Las figuras 24.b, 24.c y 24.d son reconstrucciones de la original (24.a) a partir de los coeficientes de las transformadas de Gabor con recubrimiento del dominio frecuencial análogo al de la figura 23, pero diferentes solapamientos espaciales: 70% (24.b), 75% (24.c) y 85%(24.d).

El solapamiento óptimo dependerá en general del recubrimiento planteado en el dominio frecuencial. Es posible que un recubrimiento del dominio frecuencial con un gran solapamiento reduzca el solapamiento espacial necesario para hacer completa la base propuesta.

En nuestro caso, el solapamiento utilizado definitivamente fue el mínimo necesario para tener una reconstrucción perfecta: un 85%. En la figura 24 se observan los resultados de señales reconstruidas a partir de las transformadas respecto de bases con el recubrimiento frecuencial planteado y diferentes solapamientos en el espacio.

Trabajando con imágenes de $128*128$ (dominio de 2 grados de lado), el solapamiento del 85% impone que el dominio espacial se recubra uniformemente en cada dirección con: 7 funciones para las bandas 1D con frecuencias $f \leq 1/18 f_s$, 19 funciones para las bandas 1D con $f = 1/6 f_s$ y 51 funciones para las bandas 1D con $f=3/8 f_s$. En total se usan 161 funciones 1D en cada dirección del dominio frecuencial, es decir $161*161$ funciones separables 2D.

Como se ya se dijo al analizar las posibles modelizaciones del procesado de bajo nivel, lo más seguro es que los procesos de reducción de información tengan ya lugar al cambiar de dominio (mediante una transformación incompleta). La figura 24 constituye un ejemplo de como la variación de un único parámetro de la base controla el grado de información perdido en la transformación. Como se puede intuir, aunque introducir la posibilidad de una base incompleta es mucho más realista, la complejidad aumenta porque en ese caso la definición de la base resulta muy crítica, pudiéndose conseguir similares distorsiones mediante bases con parámetros diferentes.

5.3 TRANSFORMADA DE GABOR RESPECTO DE NUESTRA BASE.

Como se ha dicho antes, la elección de la base ha estado determinada por la voluntad de utilizar el algoritmo matricial expuesto en 2.2.3.

Recordemos que en ese algoritmo la transformada viene dada por una matriz de coeficientes F cuyas dimensiones coinciden con el número de funciones base 1D utilizadas en la dirección y

y la x respectivamente (en nuestro caso 161 en cada eje). En aquel momento cuando no se especificaba ninguna base en particular, se dedujo que cada elemento F_{ik} controla la contribución de la función 2D $g_{kl}(x)$ que se obtiene del producto exterior de la columna l de la matriz G_1 (que contiene las funciones base en la dirección y) y la fila k de la matriz G_2 (que contiene las funciones base en la dirección x). Por tanto, como se dijo entonces, el significado espacio-frecuencial de los coeficientes de la transformada F depende de la ordenación concreta de las funciones base en las matrices G_i .

En nuestro caso, las funciones base 1D se han ordenado como sigue: la matriz G_1 (128*161) está formada por 7 submatrices de columnas. Cada una de dichas submatrices corresponde a una de las frecuencias centrales consideradas en la dirección y (es decir la frecuencia central de todas las funciones base 1D que están contenidas en una determinada submatriz es la misma). Estas submatrices frecuenciales están ordenadas en función de la frecuencia: la de menor frecuencia ($f = -3/8 f_s$) ocupa las primeras columnas de G_1 y la de mayor frecuencia ($f = 3/8 f_s$) las últimas. Dentro de cada submatriz, cada columna contiene los valores de una función base 1D centrada en una posición y_0 diferente. Las diferentes funciones de la misma frecuencia se han ordenado en función de su posición central: las centradas en los valores más bajos del dominio espacial ocupan las primeras columnas de la submatriz, y conforme se avanza en las columnas de la submatriz, las funciones consideradas ocupan posiciones más avanzadas.

En la matriz G_2 (161*128) el criterio de ordenación ha sido similar solo que en este caso funciones base se sitúan en las filas de G_2 . Así pues, la matriz G_2 está formada por 7 submatrices de filas. Cada una de dichas submatrices corresponde a una de las frecuencias centrales consideradas en la dirección x (es decir la frecuencia central de todas las funciones base 1D que están contenidas en una determinada submatriz es la misma). Estas submatrices frecuenciales están ordenadas en función de la frecuencia: la de menor frecuencia ($f = -3/8 f_s$) ocupa las primeras filas de G_2 y la de mayor frecuencia ($f = 3/8 f_s$) las últimas. Dentro de cada submatriz, cada fila contiene los valores de una función base 1D centrada en una posición x_0 diferente. Las diferentes funciones de la misma frecuencia se han ordenado en función de su posición central: las centradas en los valores más

bajos del dominio espacial ocupan las primeras columnas de la submatriz, y conforme se avanza en las columnas de la submatriz, las funciones consideradas ocupan posiciones más avanzadas.

En la matriz G_2 (161*128) el criterio de ordenación ha sido similar solo que en este caso funciones base se sitúan en las filas de G_2 . Así pues, la matriz G_2 está formada por 7 submatrices de filas. Cada una de dichas submatrices corresponde a una de las frecuencias centrales consideradas en la dirección x (es decir la frecuencia central de todas las funciones base 1D que están contenidas en una determinada submatriz es la misma). Estas submatrices frecuenciales están ordenadas en función de la frecuencia: la de menor frecuencia ($f = -3/8 f_s$) ocupa las primeras filas de G_2 y la de mayor frecuencia ($f = 3/8 f_s$) las últimas. Dentro de cada submatriz, cada fila contiene los valores de una función base 1D centrada en una posición x_0 diferente. Las diferentes funciones de la misma frecuencia se han ordenado en función de su posición central: las centradas en los valores más bajos del dominio espacial ocupan las primeras columnas de la submatriz, y conforme se avanza en las columnas de la submatriz, las funciones consideradas ocupan posiciones más avanzadas.

La figura 25 muestra un clarificador (?) esquema de la situación de las funciones base 1D en las matrices G_1 y G_2 .

Ahora que ya conocemos el criterio de ordenación de las funciones base 1D en las matrices G_i ya podemos analizar el significado espacio-frecuencial de los coeficientes de F (y el de los coeficientes de la GSF que actuarán sobre los elementos de la transformada de Gabor). Recuérdese que el significado de cada elemento F_{lk} está dado por la posición y la frecuencia de la función 2D que surge del producto de la columna l de G_1 y la fila k de G_2 . Véanse unos ejemplos del significado de cada coeficiente (a que función da lugar) en la figura 26.

En resumen: en la matriz F (de 161*161) coeficientes, se tendrán 49 submatrices de diferente significado frecuencial 2D. El significado frecuencial de cada submatriz está recogido en la figura 27.a. Los elementos de cada submatriz controlan la contribución de las funciones de esa frecuencia en diferentes posiciones espaciales. Las posiciones espaciales concretas dependen de cada submatriz ya que como hemos visto, para cada

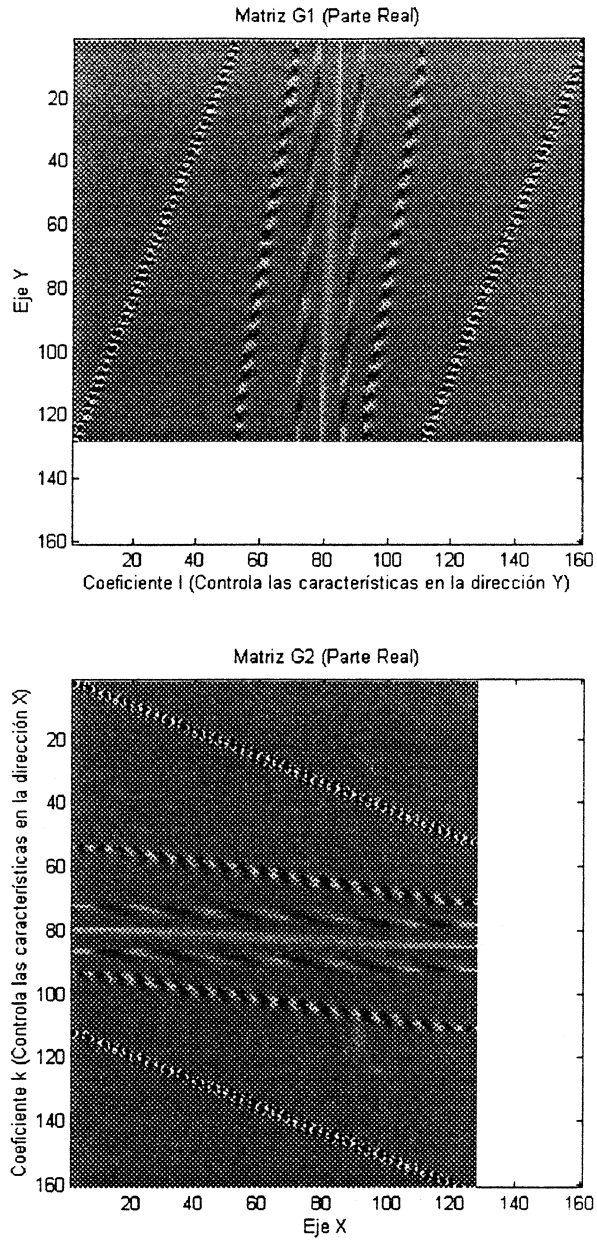
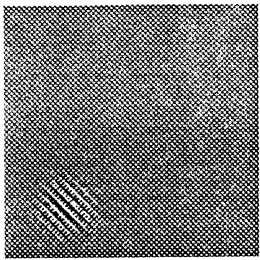
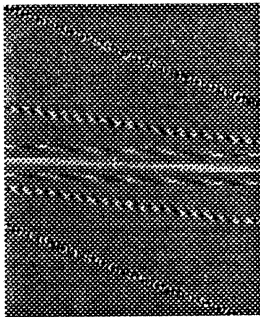


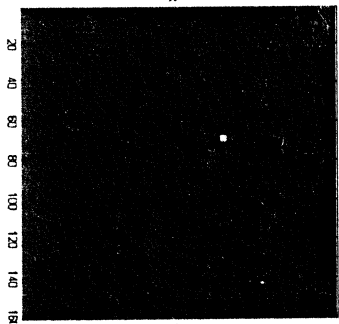
Fig. 25: Parte real de los elementos de las matrices G_1 y G_2 que participan en la TG y contienen los elementos de la base. Obsévese como existen representantes de cada banda frecuencial repartidos por todo el espacio.



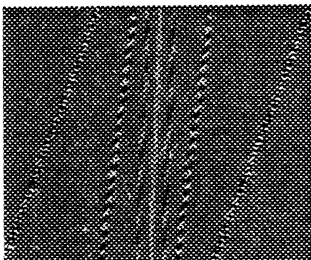
=



*



*

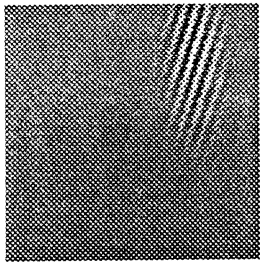


$R_{67 56}(x)$

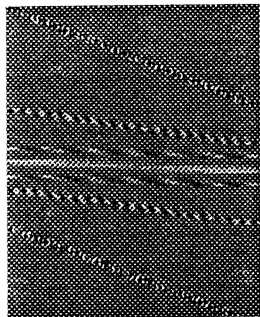
G_1

$F_{56 67} = I$

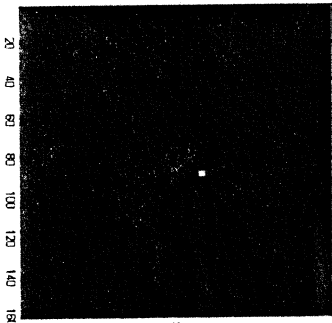
G_2



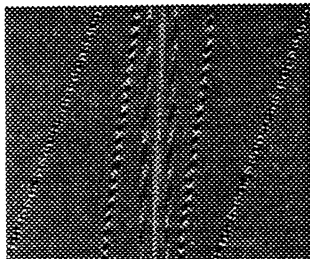
=



*



*



$R_{86 66}(x)$

G_1

$F_{66 86} = I$

G_2

Fig. 26: Significado de los coeficientes de la transformada (matriz F): Cada elemento de F selecciona una columna de G_1 y una fila de G_2 para generar la función base que le corresponde. En estos ejemplos vemos como los coeficientes $F_{56 67}$ y $F_{66 86}$ (el resto de la matriz es cero) dan lugar a las funciones $R_{67 56}(x)$ y $R_{86 66}(x)$ respectivamente.

$$\begin{pmatrix}
 (-3/8, -3/8)fs/2 & (-3/8, -1/6)fs/2 & (-3/8, -1/18)fs/2 & (-3/8, 0)fs/2 & (-3/8, 1/18)fs/2 & (-3/8, 1/6)fs/2 & (-3/8, 3/8)fs/2 \\
 (-1/6, -3/8)fs/2 & (-1/6, -1/6)fs/2 & (-1/6, -1/18)fs/2 & (-1/6, 0)fs/2 & (-1/6, 1/18)fs/2 & (-1/6, 1/6)fs/2 & (-1/6, 3/8)fs/2 \\
 (-1/18, -3/8)fs/2 & (-1/18, -1/6)fs/2 & (-1/18, -1/18)fs/2 & (-1/18, 0)fs/2 & (-1/18, 1/18)fs/2 & (-1/18, 1/6)fs/2 & (-1/18, 3/8)fs/2 \\
 (0, -3/8)fs/2 & (0, -1/6)fs/2 & (0, -1/18)fs/2 & (0, 0)fs/2 & (0, 1/18)fs/2 & (0, 1/6)fs/2 & (0, 3/8)fs/2 \\
 (1/18, -3/8)fs/2 & (1/18, -1/6)fs/2 & (1/18, -1/18)fs/2 & (1/18, 0)fs/2 & (1/18, 1/18)fs/2 & (1/18, 1/6)fs/2 & (1/18, 3/8)fs/2 \\
 (1/6, -3/8)fs/2 & (1/6, -1/6)fs/2 & (1/6, -1/18)fs/2 & (1/6, 0)fs/2 & (1/6, 1/18)fs/2 & (1/6, 1/6)fs/2 & (1/6, 3/8)fs/2 \\
 (3/8, -3/8)fs/2 & (3/8,)fs/2 & (3/8, -1/18)fs/2 & (3/8, 0)fs/2 & (3/8, 1/18)fs/2 & (3/8, 1/6)fs/2 & (3/8, 3/8)fs/2
 \end{pmatrix}$$

Fig. 27.a: Significado frecuencial (frecuencia central de las funciones base) de las submatrices de nuestro dominio de Gabor

Fig. 27.b: Significado espacial (posición central de las funciones base) de los elementos que constituyen cada submatriz de nuestro dominio de Gabor.

$$\begin{pmatrix}
 x < 0 & x > 0 \\
 y < 0 & y < 0 \\
 \\
 x < 0 & x > 0 \\
 y > 0 & y > 0
 \end{pmatrix}$$

banda frecuencial se tendrá un muestreo espacial diferente. Como regla general puede diremos que en cada submatriz de F pueden establecerse cuatro cuadrantes que se corresponden con los cuadrantes del dominio espacial y cuyos elementos siguen la misma ordenación que en aquel. Véase la figura 27.b

5.4 DETERMINACIÓN EXPERIMENTAL DE LOS FILTROS EN EL DOMINIO DE FOURIER.

5.4.1 Consideraciones generales.

Después de tener bien determinada la base y el algoritmo de cálculo de la transformada solo nos faltan los valores de los filtros en el dominio de Fourier para poder presentar algún ejemplo concreto de *GSF*.

Nuestro objetivo es plantear una caracterización que sea cuantitativamente operativa. Por tanto, el resultado debe ser una determinada función de transferencia que nos permita hallar *numéricamente* la salida del sistema ante *cualquier entrada bidimensional*:

$$i'(x) = TG^{-1}[GSF(x', f') * TG[i(x)](x', f')](x) \quad (5.1)$$

El cómputo numérico de la expresión 4.12 para el cálculo de los $161 * 161$ puntos de la *GSF* implica tener valores de la *CSF* en *todo* el dominio frecuencial para filtrar las funciones base $g_{\mu}(x)$. Una ventaja interesante del tratamiento 2D de la *CSF* es que las caracterizaciones obtenidas, tanto la de Fourier como la *GSF* resultante, incluirán el comportamiento no isótropo del sistema en orientación. Además, lo más importante, es considerar la bidimensionalidad intrínseca de una caracterización del sistema visual en el dominio de Fourier: si hemos de filtrar numéricamente señales 2D, lo lógico, en principio, es intentar determinar el filtro en todo el dominio en el que está definido. Para ello no basta la medida usual de la *CSF* mediante un conjunto reducido de redes sinusoidales de orientación fija[70], sino que será necesario determinar los contrastes umbrales de sinusoides con diferentes orientaciones.

Tanto énfasis para algo que parece tan obvio viene a cuento de que en cierta literatura[71,72] no es práctica habitual intentar implementar numéricamente los modelos que se proponen, con

lo cual, las argumentaciones resultan demasiado vagas. Uno se da realmente cuenta de cual debe ser la práctica experimental, o cuales son los límites de la argumentación abstracta, cuando intenta implementar numéricamente un modelo con los datos experimentales de que se dispone y ve que aun le faltan demasiados datos.

Así pues, la necesidad de datos de la *CSF* en todo el dominio frecuencial impuesta por la voluntad de implementar numéricamente el modelo, nos llevó a desarrollar una técnica experimental sencilla para determinar de forma rápida la *CSF 2D*[73].

Como es conocido, la sensibilidad del sistema visual a cada componente frecuencial $f=(f_x, f_y)$ en el dominio de Fourier viene definida por la inversa del contraste minimamente perceptible de un estímulo sinusoidal de esa frecuencia. Por ello, para determinar la *CSF* en el dominio frecuencial 2D hay, en principio, que determinar el contraste umbral de un conjunto de sinusoides cuyas frecuencias muestreen más o menos exhaustivamente dicho dominio. Como supimos después, ya se habían efectuado medidas de la *CSF 2D* aplicando directamente esa idea: median los contrastes umbrales de redes sinusoidales situadas según una retícula cartesiana en el dominio de Fourier[41,42].

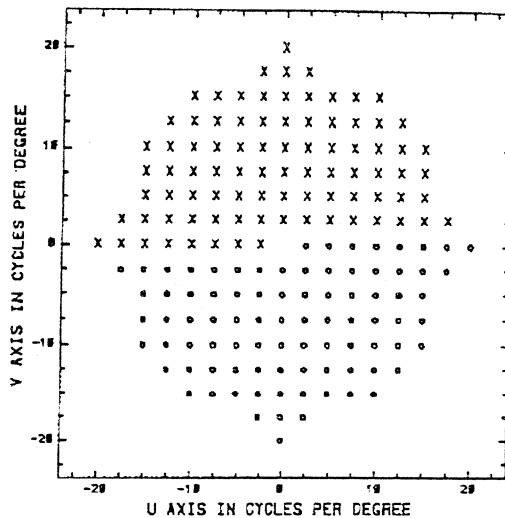


Fig. 28: Retícula de muestreo del dominio frecuencial para la medida de la *CSF* bidimensional (Harvey et al. [42])

Sin embargo, este método no es óptimo porque no explota el hecho de que la dependencia angular de la sensibilidad al contraste es mucho más suave que la dependencia radial (frecuencial). La idea básica de nuestro método experimental es que si la superficie a determinar no presenta una fuerte dependencia angular, su forma quedará bien definida a partir de unos pocos cortes 1D en diferentes direcciones angulares, y una posterior interpolación lineal para los ángulos intermedios:

$$CSF(|f|, \theta) = CSF(|f|, \theta_1) + \frac{CSF(|f|, \theta_2) - CSF(|f|, \theta_1)}{(\theta_2 - \theta_1)} (\theta - \theta_1)$$

(5.2)

donde θ es un ángulo intermedio entre dos direcciones θ_1 y θ_2 donde se han determinado sendos cortes de la *CSF*. Evidentemente, si medimos los contrastes umbrales con un conjunto de sinusoides de una sola orientación, como es lo habitual, no obtendremos información acerca de las asimetrías angulares del sistema (estaremos realizando un estudio exclusivamente 1D). Como es lógico, cuanto mayor sea el número de cortes determinados, tanto mejor definida quedará la superficie de la *CSF*, pero, para que el método sea práctico, hay que alcanzar una solución de compromiso entre la fiabilidad de los resultados y la comodidad del método.

Para estimar el número de perfiles radiales de la *CSF* que son necesarios para dar cuenta de las asimetrías angulares del sistema visual, determinamos para los dos observadores (JM emélope: con comportamiento angular uniforme, y RZ astigmático -3dp a 0°: con un comportamiento angular fuertemente asimétrico), tres perfiles angulares de la *CSF* mediante sinusoides orientadas a 0, 45, y 90 grados.

5.4.2 Práctica experimental.

Los diferentes puntos de la *CSF* de cada observador se obtuvieron calculando el contraste umbral de redes cuadradas de frecuencia y orientación variable generadas en un monitor DM2 Joyce Display Oscilloscope con fósforo blanco WA y una resolución de 1024*1024 a 8 bits/pix. Los patrones estuvieron controlados por un PC a través de un generador de redes VSG2 basado en un DSP TMS320C25.

Una vez fijada la orientación, el programa de control, presenta aleatoriamente redes de distinta frecuencia. El contraste umbral se determinó mediante el método de elección forzosa de dos alternativas. Se utilizaron redes de frecuencias 0.5, 1.8, 3.7, 7.3, 14.6 y 29.2 ciclos/grado, con una luminancia media del orden de 300cd/m². No se utilizaron pupilas artificiales ni mecanismos de control de los movimientos oculares. La excentricidad del test era nula (visión foveal). La distancia de observación se fijó en 1m. En esas condiciones el test presentado subtendía 8°.

5.4.3 Filtros en el dominio de Fourier, respuestas impulsionales y simulaciones.

Con objeto de incrementar el número de puntos disponibles para aplicar la expresión (5.2), ajustamos los valores de los cortes de la *CSF* mediante una exponencial, para la caída en la zona de altas frecuencias, y un polinomio del orden adecuado en la zona de las bajas frecuencias. De esta manera, podemos disponer de valores de la *CSF* para toda $|f|$ en los planos $\theta = 0^\circ, 45^\circ$ y 90° . La figura 29 muestra, ya en el dominio frecuencial 2D, los resultados experimentales correspondientes a los cortes radiales de la *CSF* de los dos observadores así como las curvas ajustadas.

A partir de dichos datos vamos a obtener las funciones filtro (mediante 5.2) y las respuestas impulsionales (mediante la TF^{-1} de las funciones filtro), utilizando uno, dos, o los tres cortes de la *CSF*. Con ello conseguimos dos cosas:

- Poner de manifiesto las ventajas del tratamiento bidimensional al ver como la utilización de más de un corte de la *CSF* proporciona información valiosa acerca de las asimetrías del sistema en orientación.
- Estimar el esfuerzo experimental mínimo (número mínimo de cortes a determinar), para obtener una caracterización que recoja adecuadamente dichas asimetrías.

Las figuras 30 y 31 muestran el resultado de aplicar la expresión (5.2) considerando respectivamente: el corte de orientación 0° , el de 0° y el de 90° , y los tres. La información aportada por cada nuevo corte introducido y las diferencias entre ambos observadores pueden analizarse con mayor claridad a partir de

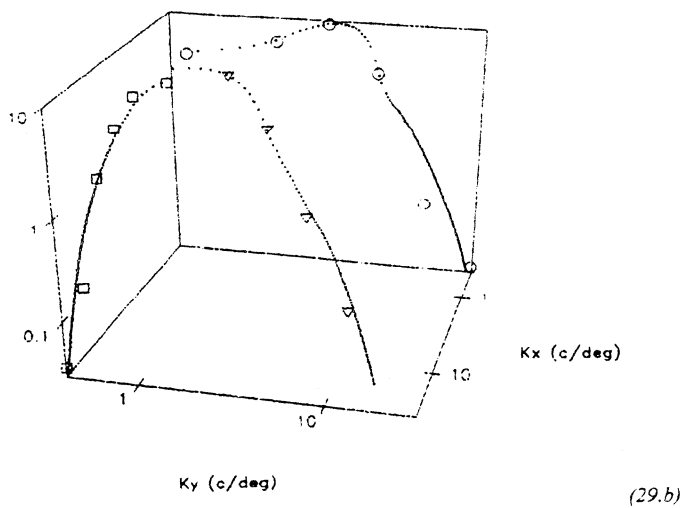
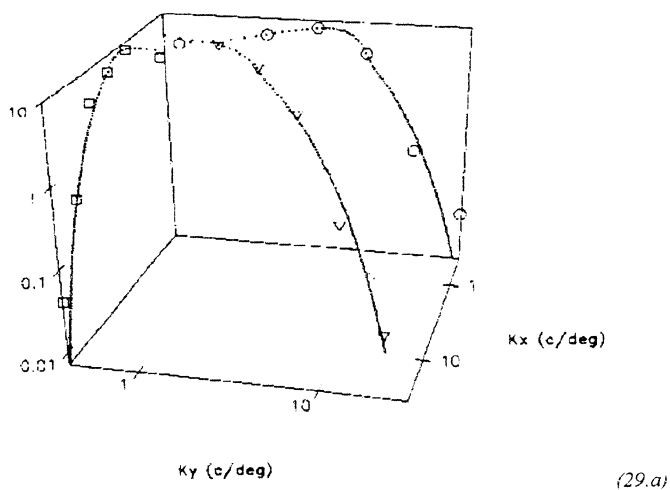
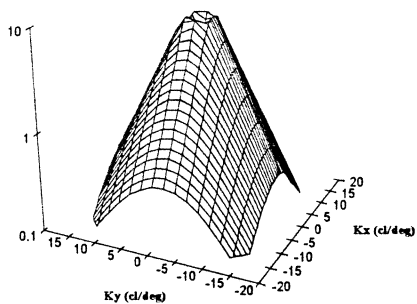
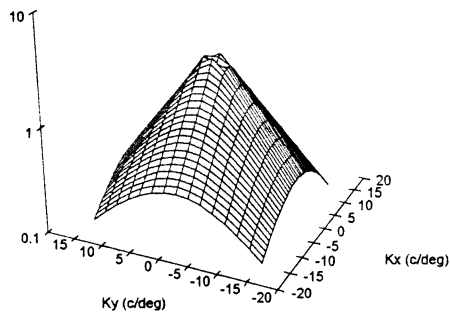


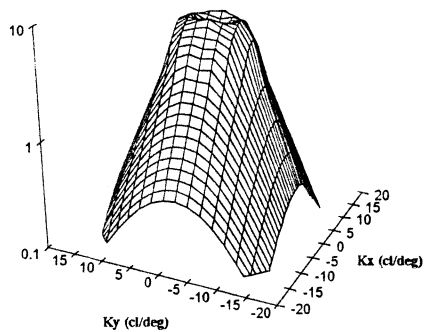
Fig. 29: Cortes radiales experimentales de las CSFs 2D de los observadores JM (Emétrope) -29.a- y RZ (astigmático) -29.b-. Ambos ejes (los de frecuencias y el de amplitud) están en escala logarítmica como es usual en las representaciones 1D. En adelante representaremos los ejes frecuenciales en escala lineal debido a que será necesario incluir el origen.



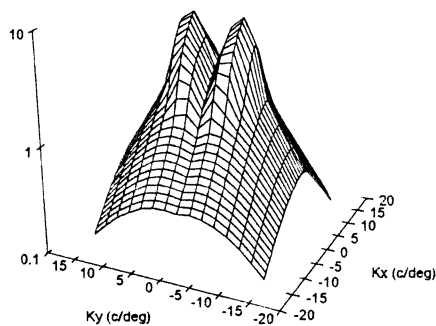
(30.a)



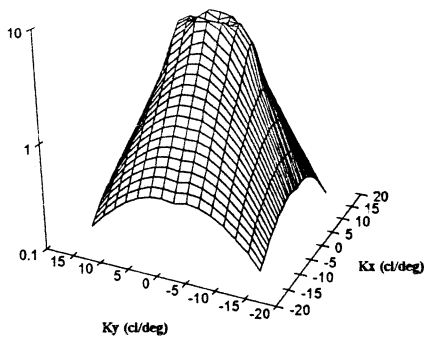
(31.a)



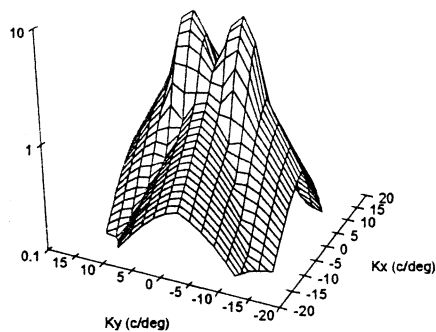
(30.b)



(31.b)

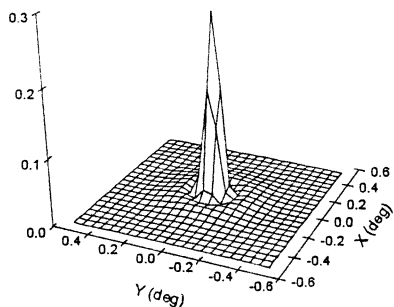


(30.c)

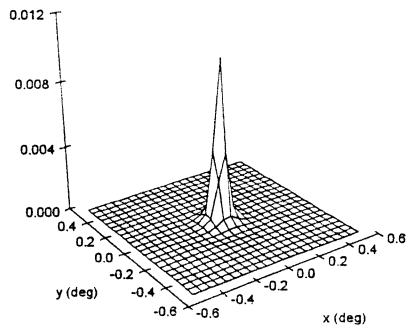


(31.c)

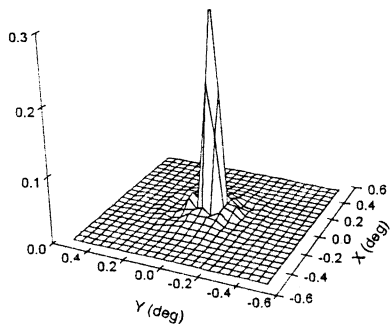
Fig. 30: CSFs 2D de JM obtenidas a partir de los cortes de 0° (30.a), 0° y 90° (30.b) y $0^\circ, 45^\circ$ y 90° (30.c).
 Fig. 31: CSFs 2D de RZ obtenidas a partir de los cortes de 0° (31.a), 0° y 90° (31.b) y $0^\circ, 45^\circ$ y 90° (31.c).



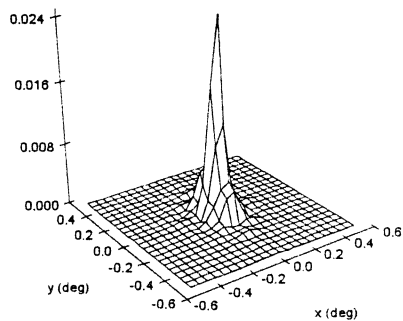
(32.a).



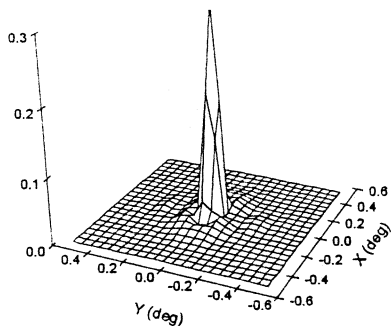
(33.a)



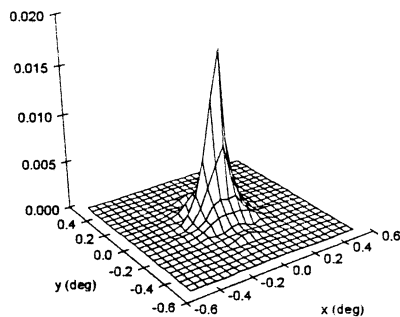
(32.b).



(33.b).



(32.c).



(33.c).

Fig. 32: Respuestas impulsionales del observador JM obtenidas a partir de las funciones de transferencia representadas en la fig. 30.

Fig. 33: Respuestas impulsionales del observador RZ obtenidas a partir de las funciones de transferencia representadas en la fig. 31. Por motivos de claridad en la impresión se han decimado las señales representadas, de forma que son menos evidentes las diferencias entre las PSFs del observador emétrepe y el astigmático. Las PSFs definitivas (figuras 32.c y 33.c) se muestran también en las figuras 38 y 39.

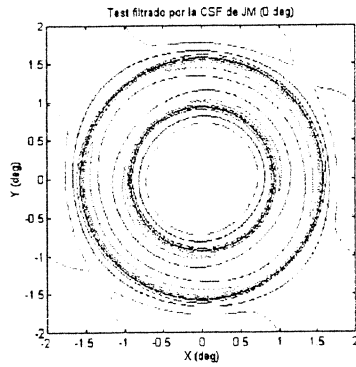
las respuestas impulsionales y las simulaciones de la distorsión sufrida por un test complejo.

La escala del dominio espacial obtenido mediante la TF^{-1} está dada por el tamaño del dominio frecuencial y el número de puntos utilizado. Como en las DFTs el dominio frecuencial se extiende entre $-f_s/2$ y $f_s/2$ [74] (siendo $f_s = N/X_{\max}$ la frecuencia de muestreo), y la *CSF* se determinó en un dominio discreto de $N=256$ puntos entre -30cl/grad y 30cl/grad , la frecuencia de muestreo será 60cl/grad , el ángulo subtendido por el lado del dominio espacial será $X_{\max} = N/f_s = 4.250$ grados. Como se ve, no extendemos los resultados obtenidos del cálculo más allá de lo determinado por el experimento (test de 8°).

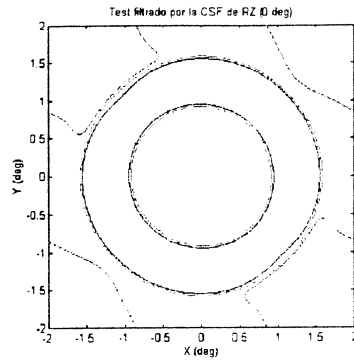
Las figuras 32 y 33 muestran las respuestas impulsionales (en la zona foveal) de los observadores JM y RZ obtenidas a partir de los filtros de 30 y 31. Las figuras 34 y 35 muestran el resultado de filtrar un test circular con dichos filtros.

Los resultados para el observador emélope tanto en la función filtro (fig. 30) como en la respuesta impulsional (fig. 32) o en la simulación (fig. 34), revelan las limitaciones de la caracterización 1D (mediante un único corte de la *CSF*): en las figuras 30.a, 32.a y 34.a se tiene un comportamiento independiente de la dirección que no responde totalmente a la realidad. La introducción de un nuevo perfil en la caracterización corrige los resultados previos de forma cualitativamente importante: mediante ese nuevo perfil se rompe la simetría de revolución poniéndose de manifiesto un suave ensanchamiento de la respuesta impulsional en la dirección x (fig. 32.b), mayor ancho de banda en la dirección y (fig. 30.b), y una respuesta más nítida al resolver los bordes horizontales (fig. 34.b). Con la introducción de un tercer corte en la caracterización, se tienen algunas modificaciones de los lóbulos laterales de la respuesta impulsional, pero estas correcciones no aportan información cualitativa adicional. Podemos concluir que en el caso de observadores emétopes, la medida de la *CSF* en dos direcciones perpendiculares basta para poner de manifiesto las pequeñas asimetrías que puedan estar presentes en el sistema.

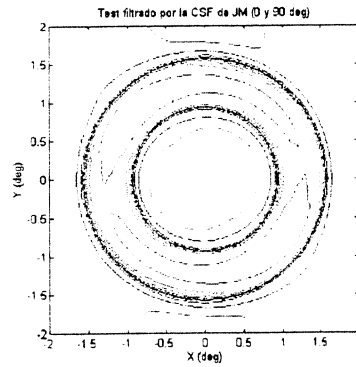
Para el observador astigmático la insuficiencia de la caracterización mediante un solo corte de la *CSF* es patente. Es evidente que para un sistema con fuerte dependencia angular como es el astigmático, no es válida la caracterización dada por



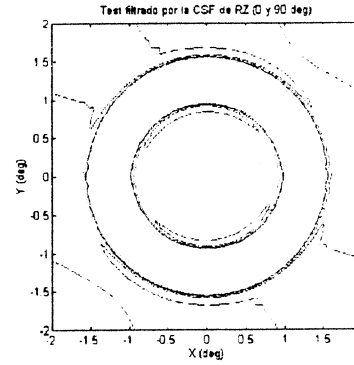
(34.a)



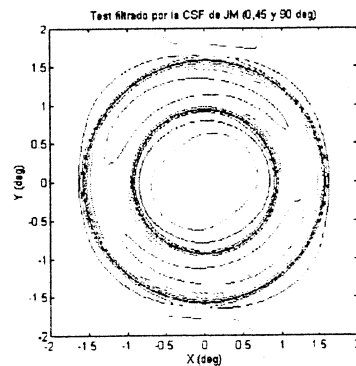
(35.a)



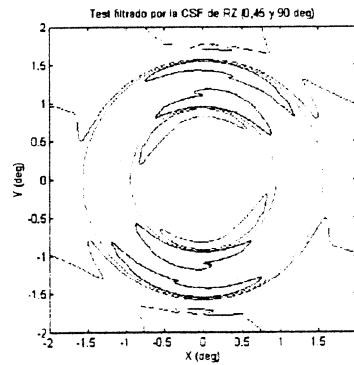
(34.b)



(35.b)



(34.c)



(35.c)

Fig. 34: Simulación de la imagen de un test circular percibida por el observador JM asumiendo cada uno de los filtros de la figura 30.

Fig. 35: Simulación de la imagen de un test circular percibida por el observador RZ asumiendo cada uno de los filtros de la figura 31.

las figuras 31.a o 33.a ni el comportamiento reflejado en 35.a. Si introducimos las medidas en la dirección perpendicular (90° , la dirección de máxima calidad para el observador RZ), la asimetría de la *CSF* y la respuesta impulsional se ponen de manifiesto de forma espectacular (figs. 31.b y 33.b), dando lugar a un comportamiento en el que se observa una dirección en la que la percepción es muy deficiente (fig. 35.b). El uso de un tercer perfil radial modifica los resultados previos de forma importante. Al contrario que en el caso emélope este nuevo corte sí introduce información adicional: se acentúa el efecto de la anisotropía (figs. 31.c y 33.c), confirmando la existencia de una dirección preferente muy definida, de forma que el rango angular de nitidez del test percibido se reduce (fig 35.c). En el caso de RZ, la introducción de nuevos cortes en las direcciones de 22.5° y 67.5° no añade nueva información. Estos resultados sugieren que la transición entre las direcciones de mejor y peor visibilidad en un observador astigmático es en general no lineal. En conclusión, para obtener una buena caracterización para este tipo de observadores es necesario introducir al menos una tercera medida en una orientación intermedia entre las direcciones de máxima sensibilidad y su perpendicular.

5.4.4 Comentarios acerca del método. *CSFs* definitivas.

Como resumen de este apartado dedicado a explicar la determinación experimental de los filtros en el dominio de Fourier podemos concluir que:

- Es necesario poner énfasis en la bidimensionalidad de la caracterización de Fourier para explotar toda la información que pueden ofrecer los filtros en ese dominio.
- Los resultados muestran que para obtener una caracterización adecuada de la variación de la *CSF* con la frecuencia y la orientación basta, en el caso de observadores emétopes, con medir la sensibilidad en un par de direcciones perpendiculares y luego interpolar mediante 5.2.
- En el caso de observadores astigmáticos (sistemas muy asimétricos), es necesaria la determinación de una sección adicional de la *CSF* entre las direcciones de máxima (o mínima) sensibilidad y su perpendicular.

Los filtros definitivos (los que se utilizarán en el cálculo de las *GSFs*) son los que incluyen todas las medidas experimentales: el de la figura 30.c para JM y el de la figura 31.c para RZ. Como, por motivos de coste computacional, la base y el algoritmo descritos en 5.2 y 5.3 están definidos para operar con dominios espaciales de 128*128 puntos, se aplicará un proceso de decimación standard[74] a las funciones de transferencia (ya que se calcularon en un dominio de 256*256 puntos).

Estos filtros de Fourier representan el comportamiento de ambos observadores en la región foveal, donde sí es válida la suposición de homogeneidad del sistema.

5.5 CÁLCULO DE LAS *GSFs* EQUIVALENTES A LAS *CSFs* FOVEALES.

Una vez elegida la base, concretado el algoritmo y determinadas experimentalmente unas funciones filtro particulares en el dominio de Fourier, ya podemos obtener las *GSFs* equivalentes a dichas *CSFs*.

5.5.1 Resultados.

El resultado de aplicar la ecuación 4.12 filtrando mediante la *CSF* de JM cada una de las funciones $g_{\mathbf{w}}(\mathbf{x})$, se muestra en la figura 36.

El esquema de la figura 27 da el significado espacio-frecuencial de cada coeficiente. Recordando que las submatrices internas corresponden a las bajas frecuencias y las externas a las altas frecuencias, podemos ver que los pesos están organizados de forma que el filtro resultante es un pasa-banda: se tiene un máximo de sensibilidad para las funciones con un a frecuencia central $|f|=1.67c/l/\text{grad}$, y menor sensibilidad para las funciones de frecuencia inferior o superior. La resolución frecuencial en el dominio de Gabor es menor que en el dominio de Fourier, debido a que las funciones de Gabor son de banda ancha. En las funciones de una cierta banda se mezclan los efectos de las asimetrías de la *CSF* en esa banda. Como veremos en 5.5.2, esto no implica que se haya perdido la información sobre las asimetrías del sistema en orientación, sino que, dichas asimetrías

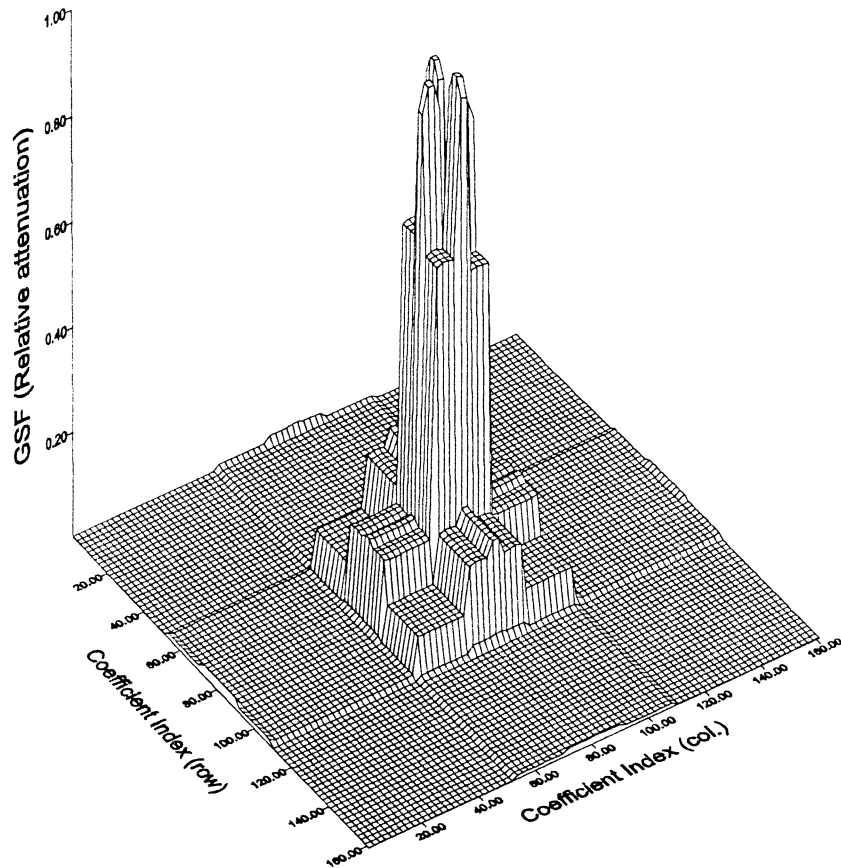


Fig. 36: GSF del observador JM calculada en todos los puntos del dominio de Gabor a partir del filtro de la figura 30.c.

de la función filtro en el dominio de Fourier se perciben con menor claridad en este nuevo dominio.

Dentro de cada submatriz, el valor de los coeficientes es aproximadamente constante. Esto quiere decir que la atenuación sufrida por las funciones de cada frecuencia espacial es independiente su localización en el espacio. Este es el comportamiento que se esperaba al obtener el filtro de Gabor a partir de un filtro de Fourier: como dijimos en 4.3.3, si partimos de una caracterización espacialmente invariante como la de

Fourier, la transformación 4.12 que calcula la caracterización equivalente, nos proporcionará un filtro espacialmente invariante.

Existen pequeñas discrepancias respecto de la invariancia espacial de la atenuación, y dichas diferencias se hacen más acusadas en los coeficientes cuyas funciones tienen baja frecuencia (están muy extendidas en el espacio) y están situadas hacia los extremos del dominio espacial. Esta desviación sistemática puede estar causada por la finitud del dominio discreto considerado de forma que se produzca una especie de *spectral leakage*[74]. El *spectral leakage* en la DFT usual es el efecto de distorsión del espectro que se produce al tratar con dominios discretos de longitud (o en el caso 2D área) finita: para las señales que se ven recortadas por el dominio finito, el efecto de dicha finitud es análogo al efecto de eventanar la señal en un dominio infinito mediante una ventana $w(x)$ cuadrada:

$$g(x_{finito}) = w(x_{infinito}) g(x_{infinito}) \quad (5.3)$$

al obtener el espectro mediante la DTF, el espectro de la función quedará distorsionado al convolucionarse con el espectro de la ventana:

$$G(f_{finito}) = W(f_{infinito}) \otimes G(f_{infinito}) \quad (5.4)$$

Este efecto será tanto más importante cuanto mayor parte de la señal caiga fuera del dominio finito, de ahí que la distorsión afecte fundamentalmente a las funciones base situadas en la parte exterior del dominio espacial.

Estas desviaciones están generadas por el método de cálculo empleado, no están relacionadas con las características reales del sistema. El valor real de la atenuación para las funciones en las posiciones extremas del dominio espacial será el mismo que el obtenido para el resto de las funciones que sí están totalmente contenidas en el dominio considerado.

De acuerdo con los resultados obtenidos, al calcular el filtro en el dominio de Gabor equivalente a un filtro dado en el dominio de Fourier, no es necesario calcular explícitamente todos los coeficientes de la *GSF*, tan solo es imprescindible calcular el coeficiente central de cada submatriz frecuencial (el coeficiente correspondiente a una función base totalmente contenida en el dominio finito). Como el filtro de Fourier es espacialmente invariante, podemos extrapolar la atenuación obtenida para la

posición central al resto de los elementos de cada submatriz frecuencial. De esta forma, se ahorra tiempo de cálculo y se evitan los efectos de bordes sobre las funciones con posiciones más exteriores.

Los resultados de las *GSFs* de ambos observadores utilizando este procedimiento se muestran en la figura 37.

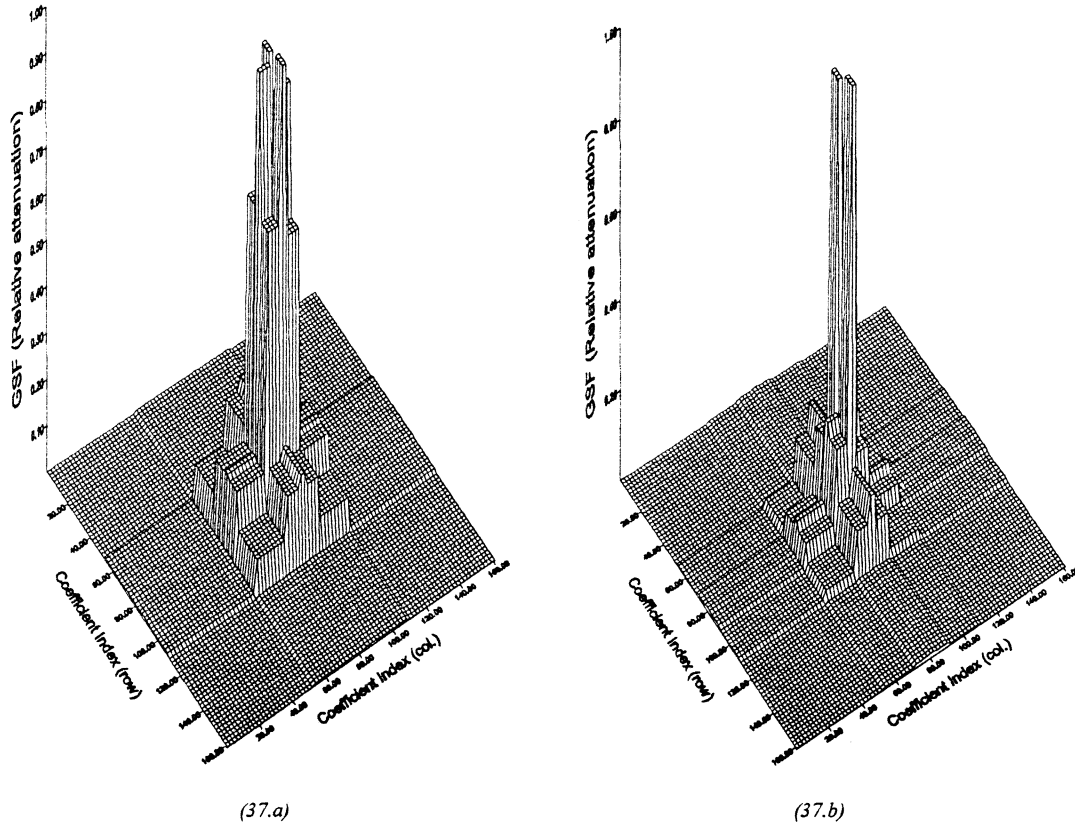


Fig. 37: *GSFs* de los observadores JM (a) y RZ (b) calculadas a partir de un solo valor por submatriz del dominio de Gabor.

5.5.2 Equivalencia entre las *GSFs* calculadas y los filtros de Fourier iniciales.

Para demostrar la equivalencia entre las *GSFs* calculadas y las *CSFs* a partir de las que se calcularon, simularemos la respuesta impulsional del sistema y el filtrado de imágenes complejas mediante los filtros en el dominio de Gabor, comparando dichas

simulaciones con las análogas calculadas a partir de los filtros de Fourier experimentales.

La respuesta impulsional de un sistema caracterizado mediante una *GSF* puede calcularse mediante la expresión 5.1 tomando como señal incidente una función impulso unidad: $i(x) = \delta(x - x_0)$. En el caso de la caracterización de Fourier la respuesta impulsional viene dada directamente por la TF^{-1} de la función filtro.

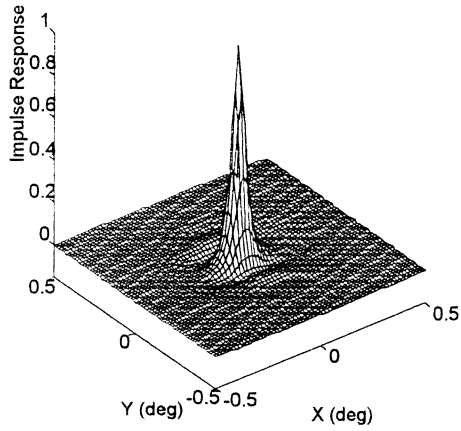
En las figuras 38 y 39 podemos ver las *PSFs* de ambos observadores obtenidas de las *GSFs* de la figura 38 y de las *CSFs* de las figuras 30.c y 31.c.

Como podemos ver, tanto la anchura de la respuesta impulsional de ambos observadores como la deformación en la dirección x de la *PSF* del observador astigmático son bien reproducidas por la caracterización de Gabor. Así pues, aunque la resolución frecuencial es menor en el caso de la caracterización en el dominio conjunto y la presencia de las asimetrías es menos evidente, vemos que la transformación al nuevo dominio conserva la información respecto de las asimetrías en orientación. Las diferencias en los lóbulos secundarios de las respuestas impulsionales generadas por las *GSFs* darán lugar a ruido en la parte externa de las imágenes filtradas.

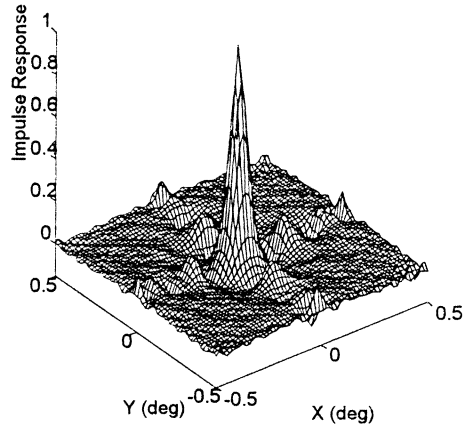
En el caso del procesado de imágenes complejas también se demuestra el acuerdo entre ambas caracterizaciones. En la figura 40 se muestra el procesado de una imagen natural por parte de los filtros correspondientes a JM. La figura 41 muestra el resultado de la actuación de los filtros de RZ sobre un test circular para mostrar los efectos de las no uniformidades en orientación del observador astigmático.

La degradación general introducida por la *GSF* en el caso de JM es equivalente a la introducida por la *CSF*: la distorsión cuadrática media es del orden del 12%. En el caso de RZ vemos que el filtro de Gabor reproduce muy bien el comportamiento astigmático: los bordes horizontales aparecen mucho más nitidamente resueltos que los verticales como era de esperar con la respuesta impulsional obtenida. En este caso, la diferencia cuadrática media entre las dos imágenes 41.b y 41.c es del 18%.

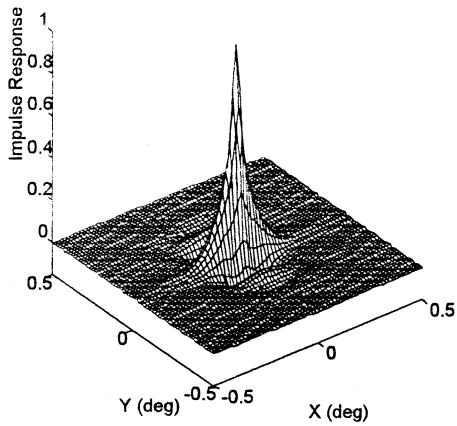
Podemos concluir este capítulo afirmando que con la base propuesta y el algoritmo utilizado es posible obtener una caracterización del comportamiento del sistema visual a nivel



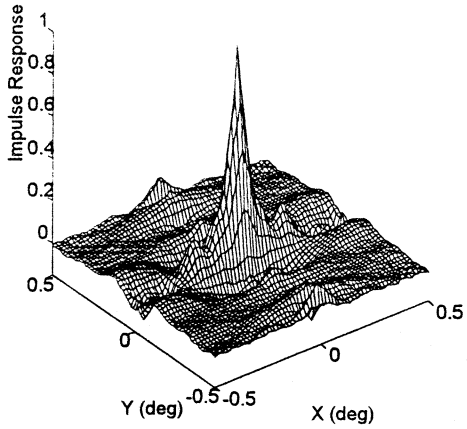
(38.a)



(38.b)



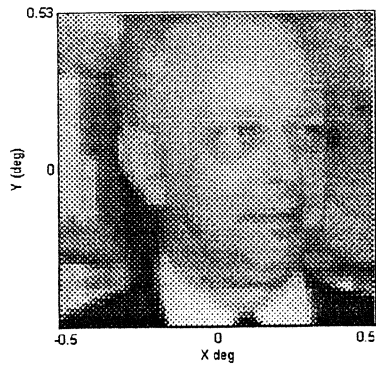
(39.a)



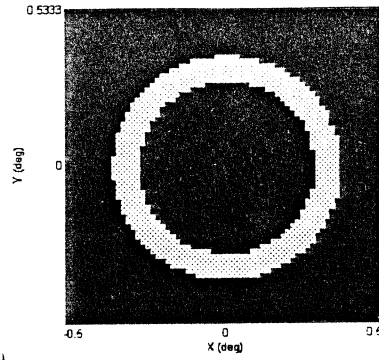
(39.b)

Fig. 38: (a) PSF del observador JM obtenida a partir del filtro de Fourier de la figura 30.c. (b) PSF análoga obtenida filtrando una delta de Dirac mediante la GSF de la figura 37.a

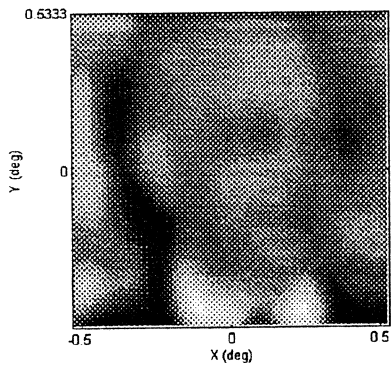
Fig. 39: (a) PSF del observador RZ obtenida a partir del filtro de Fourier de la figura 31.c. (b) PSF análoga obtenida filtrando una delta de Dirac mediante la GSF de la figura 37.b



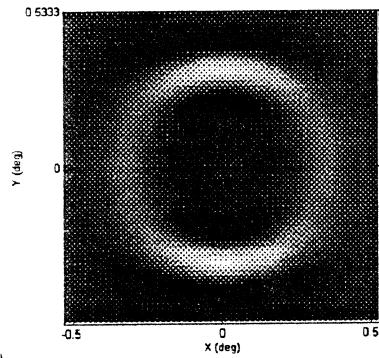
(40.a)



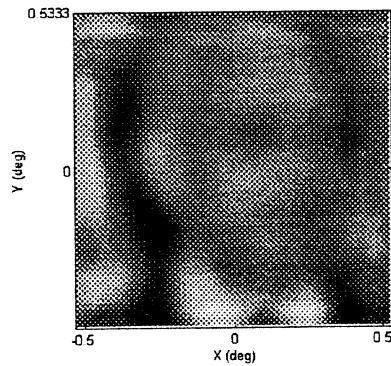
(41.a)



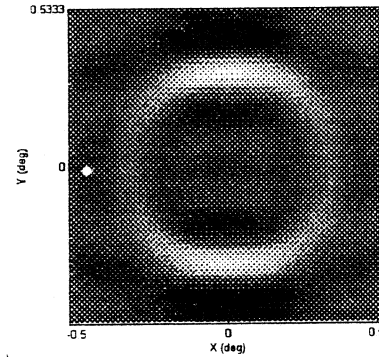
(40.b)



(41.b)



(40.c)



(41.c)

Fig. 40: Simulación del procesado de imágenes complejas (observador JM): (40.a) imagen original. (40.b) imagen tratada con el filtro de Fourier de la figura 30.c. (40.c) imagen tratada con el filtro de Gabor equivalente (fig. 37.a)

Fig. 41: Simulación del procesado de imágenes complejas (observador RZ): (41.a) imagen original. (41.b) imagen tratada con el filtro de Fourier de la figura 31.c. (41.c) imagen tratada con el filtro de Gabor equivalente (fig. 37.b)

umbral, y para la zona foveal, al menos tan potente como la caracterización de Fourier. La demostración de esto son las funciones de la figura 37 y su equivalencia con los filtros de Fourier (figs. 30.c y 31.c) confirmada mediante la reproducción de la respuesta impulsional y el comportamiento general ante imágenes complejas.

En el capítulo siguiente exploraremos con un ejemplo concreto las ventajas de esta nueva formulación, viendo como, a diferencia de la caracterización de Fourier, puede incluir comportamientos espacialmente variantes.

6. GSF DE UN SISTEMA ESPACIALMENTE VARIANTE

Hasta ahora hemos planteado una nueva caracterización de las operaciones desarrolladas en las primeras etapas del proceso visual (mediante una transformada de Gabor y la aplicación posterior de una *GSF*) y hemos relacionado dicha caracterización con una caracterización previa (la que utiliza la *TF* y la aplicación de la *CSF*), demostrando que la nueva caracterización es capaz de incluir tanta información sobre el sistema como la caracterización de Fourier.

Sin embargo, hasta el momento, no hemos explotado ninguna de las ventajas teóricas que supuestamente tiene el procesado espacio-frecuencial (apdo. 3.2.1). Evidentemente, si la única ventaja fuese que el esquema propuesto es más acorde con los datos fisiológicos correspondientes a los campos receptivos de las células corticales, no valdrían la pena las complicaciones matemáticas (y eventualmente experimentales) que son necesarias para el establecimiento de esta caracterización. Por eso, en este último capítulo, aplicaremos a un caso concreto el proceso descrito en 4.1.2 para la caracterización de sistemas espacialmente variantes.

En el ejemplo que proponemos, el sistema puede parecer un tanto artificial, pero basta para demostrar que, refinando quizá la base empleada, las ideas expuestas en 3.2.1 pueden llegar a implementarse numéricamente de forma práctica.

6.1 SÍNTESIS DE UN SISTEMA ESPACIALMENTE VARIANTE EN EL DOMINIO DE GABOR.

En esta sección demostraremos que es posible introducir, de manera controlada, inhomogeneidades espaciales en una *GSF*. De esa forma, el sistema caracterizado mediante dicha función tendrá un comportamiento espacialmente variante.

Según la misma técnica que emplearemos aquí, podrían incluirse los datos existentes sobre la respuesta impulsional del sistema visual a diferentes excentricidades[75,76] para establecer un modelo realista que permitiese incluir los efectos de la visión en campos amplios. La inclusión de zonas extrafoveales requeriría trabajar con dominios espaciales más amplios e introducir, para cada banda de la base considerada, nuevas funciones en posiciones del espacio más alejadas, incrementándose el tamaño de la matriz de coeficientes de la *TG*.

Para evitar tener que utilizar un mayor número de coeficientes, y debido a que tan solo pretendemos mostrar el potencial de la caracterización, nuestro ejemplo sintético presentará un comportamiento espacialmente variante en la zona foveal, aunque dicho comportamiento no sea muy realista¹.

En nuestro caso, vamos a generar un observador simulado a partir de los datos de los observadores reales de manera que el observador sintético presente un comportamiento emétrope o astigmático en diferentes zonas del espacio. Con objeto de que en las simulaciones queden perfectamente claros los dos comportamientos, planteamos un caso simple: comportamiento astigmático (anisótropo) en el semiplano espacial superior y comportamiento emétrope (isótropo) en el semiplano inferior.

Para obtener el sistema no homogéneo tenemos que romper la uniformidad existente en los coeficientes de las diferentes submatrices frecuenciales de los filtros de la figura 37, es decir, hay que introducir una atenuación espacialmente variante para las funciones base de la misma frecuencia. El proceso seguido se muestra en la figura 42 (análoga al esquema de la figura 17, pero fijando ya un dominio de Gabor y una relación entre los filtros en el dominio de Fourier y los de Gabor)

¹ Habitualmente se suele suponer que en la región foveal el sistema es homogéneo.

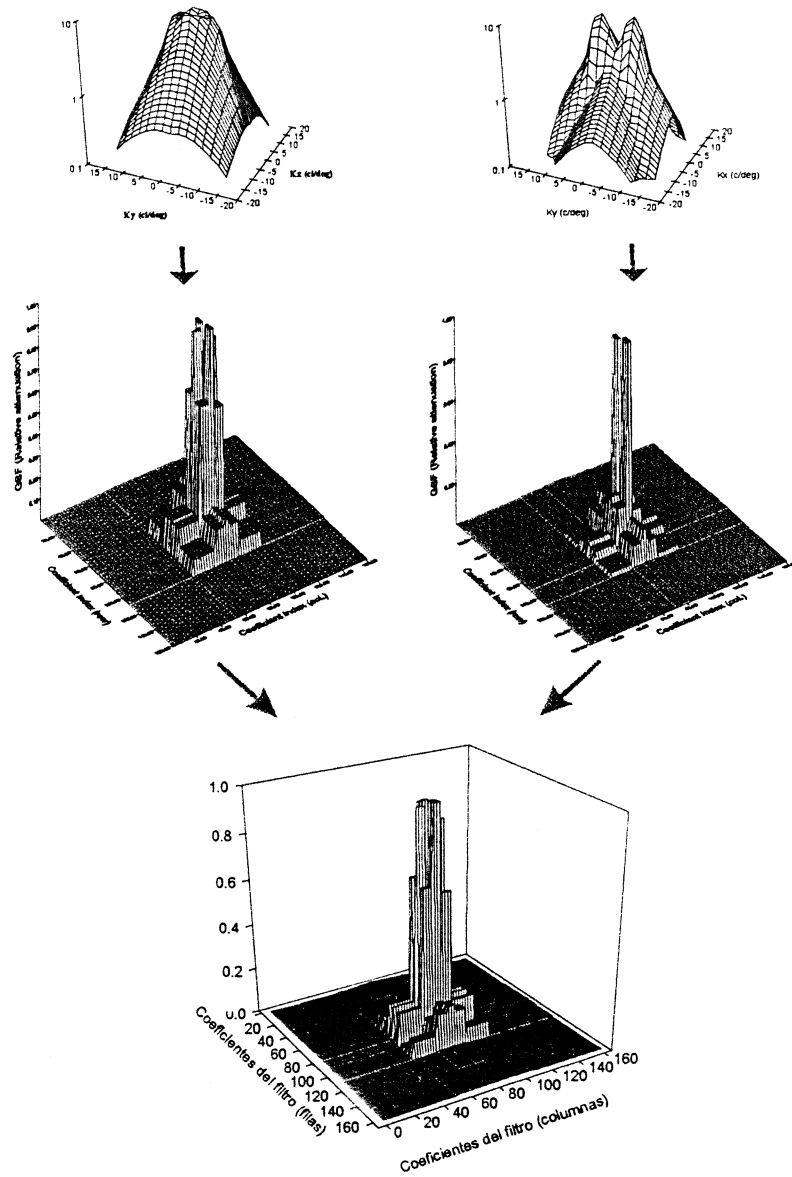


Fig. 42: Síntesis del sistema espacialmente variante propuesto en este apartado: se combinan en las posiciones correspondientes a los semiplanos superior e inferior, los valores de los filtros de Gabor equivalentes a los filtros de Fourier requeridos.

La función filtro del sistema sintético resultará de considerar los valores de la *GSF* de RZ en los puntos correspondientes al semiplano superior (la primera mitad de las filas de cada submatriz frecuencial¹), y la *GSF* de JM en los puntos del semiplano inferior (las últimas filas de cada submatriz frecuencial).

Como ya anunciamos al hablar en general de esta manera de conseguir sistemas espacialmente variantes, la combinación directa de los coeficientes no da buenos resultados cuando se trabaja con bases no ortogonales. En el apartado siguiente comprobaremos ese mal comportamiento e indicaremos una posible solución al problema.

6.2 COMPORTAMIENTO DEL SISTEMA SINTÉTICO.

Como la diferencia fundamental entre los dos comportamientos incluidos en el sistema inhomogéneo es su sensibilidad a orientaciones, utilizaremos nuevamente el test circular de la figura 41.a para ver con claridad como el sistema distorsiona angularmente el test en los dos semiplanos considerados.

Si aplicamos el filtro resultante del proceso mostrado en la figura 42 sobre la transformada de Gabor del test circular, se obtiene la imagen de la figura 43.

Al utilizar un grado de solapamiento entre funciones base tan elevado (vease apdo. 5.2.2), tenemos una base fuertemente no ortogonal, con lo que el valor de cada coeficiente depende de los valores de los coeficientes de su entorno de manera que no es posible imponer grados de atenuación muy diferentes sobre funciones muy próximas. Al trasladar al dominio de Gabor los filtros de JM y RZ y combinarlos directamente, aparece ruido en la zona de transición entre los filtros² Esto es debido a que las

¹ Según el significado espacial de los coeficientes de cada submatriz mostrado en el esquema de la figura 27.b, los coeficientes del semiplano superior están dados por la primera mitad de filas de cada submatriz frecuencial.

² Esta zona puede parecer grande en el ejemplo de la fig. 43 (de hecho toda la imagen es ruidosa), pero hay que pensar que las funciones de baja frecuencia implicadas en la transición tienen una gran extensión espacial.

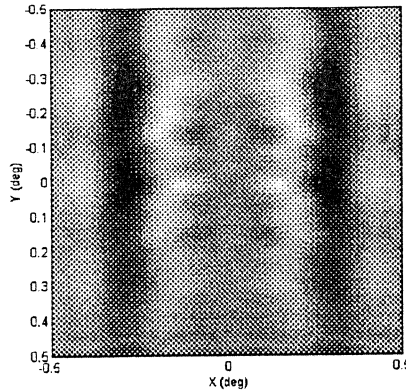


Fig. 43: Test circular filtrado por la GSF sintetizada según el proceso descrito en la fig.42 (combinando directamente los valores de los dos filtros sin preocuparnos de las interacciones entre coeficientes.)

oscilaciones de las funciones menos atenuadas no son canceladas por sus vecinas más atenuadas.

Este problema puede resolverse suavizando las diferencias de atenuación bruscas entre funciones próximas. En este caso, hemos efectuado dicho suavizado promediando localmente los coeficientes del filtro de la fig. 42 (convolucionando la matriz de coeficientes con una gaussiana estrecha). El filtro resultante del suavizado se muestra en la figura 44.

El comportamiento de este filtro sobre el test circular comparado con la actuación de los filtros espacialmente invariantes de JM y RZ se muestra en la figura 45.

Como se ve, con el suavizado de los coeficientes se ha eliminado el ruido y se ha conseguido el efecto deseado: el sistema sintetizado tiene una diferente sensibilidad angular en los diferentes semiplanos. En el semiplano superior la definición de los bordes del test está limitada a un pequeño rango angular como corresponde al comportamiento del filtro de RZ. Por otra parte, en el semiplano inferior el rango angular nítido es mucho mayor como corresponde al comportamiento isótropo de JM. La transición entre los dos comportamientos es suave debido al promediado efectuado sobre los coeficientes de la combinación directa de los filtros originales.

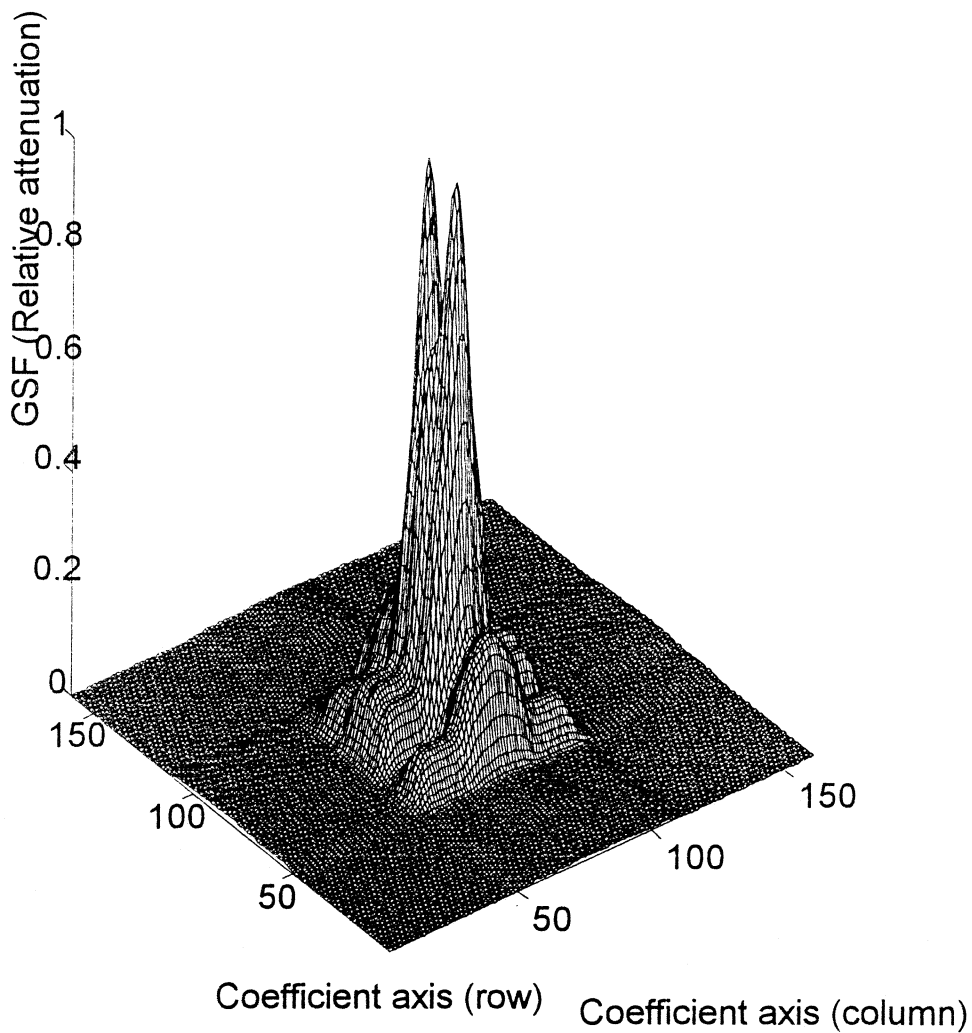


Fig. 44: GSF espacialmente variante con las transiciones espaciales suavizadas para evitar el ruido debido a la interacción entre coeficientes muy próximos con atenuaciones muy diferentes.

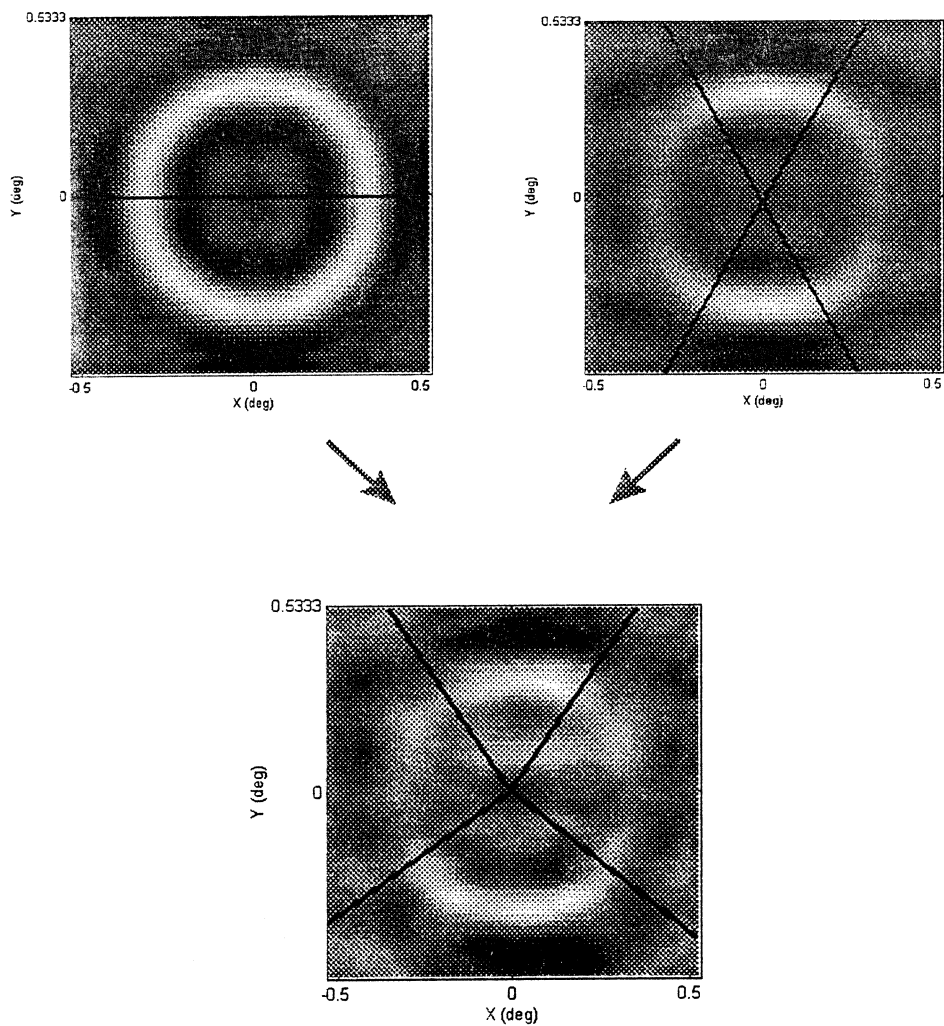


Fig. 45: Test circular procesado por el filtro espacialmente variante de la figura 44. En la parte superior de la figura hemos representado el resultado de aplicar los filtros espacialmente invariantes de partida para que se aprecie como el resultado obtenido es una combinación de los resultados previos en diferentes posiciones espaciales.

El efecto mostrado en la figura 43 y la solución del promediado implican que, con la base planteada, debemos restringirnos a caracterizar sistemas cuya variación espacial no sea muy brusca. Esto no es un grave problema en el caso del sistema visual, porque las variaciones del comportamiento en función de la excentricidad recogidas en la literatura[75,76] son suaves.

Este ejemplo sencillo demuestra la capacidad de la formulación conjunta en general, y en particular del método expuesto en 4.1.2 para caracterizar y sintetizar sistemas con comportamiento espacialmente variante. A falta de un estudio teórico detallado, la solución a las dificultades impuestas por la no ortogonalidad es una cuestión de *experiencia*. Sobre el esquema general planteado es posible efectuar diferentes variaciones para optimizar los resultados:

- Pueden probarse otros tipos de promediado sobre los coeficientes del filtro en el dominio de Gabor para encontrar cual es la máxima variación espacial con un nivel de ruido tolerable. Pueden plantearse promediados diferentes para las diferentes submatrices frecuenciales, porque es de suponer que el algoritmo permitirá variaciones espaciales más rápidas sobre la atenuación de las funciones de alta frecuencia (poca extensión espacial) que sobre las de baja frecuencia (gran extensión espacial).
- Una posibilidad, no explorada, para reducir la no ortogonalidad de la base (a parte de modificar directamente el grado de solapamiento) puede ser controlar el valor a partir del cual consideramos significativa o no, para la representación, la contribución de un cierto elemento de la base. Recordemos que a través del límite inferior permitido a los valores singulares correspondientes a los diferentes elementos de la base podían eliminarse redundancias que dieran divergencias en la transformada (apdo. 2.2.3 y referencia [18]).
- Por último, puede plantearse la utilización de un conjunto de funciones base más apropiado: otro recubrimiento del dominio frecuencial u otro tipo de función separable de soporte compacto. Como hemos visto, todo el formalismo desarrollado (algoritmo de la transformada y relación entre los filtros) no exige la utilización de una forma funcional

concreta para $g(x)$ ni un recubrimiento concreto de los dominio espacial y frecuencial.

Todas estas cuestiones indican que, aunque se ha dado un primer paso en el sentido de conseguir una caracterización computacionalmente implementable del sistema visual en un dominio conjunto, la cuestión sigue abierta y aun hay que refinar el modelo para hacerlo más manejable.

7. CONCLUSIONES.

Asumiendo la caracterización del sistema visual como un codificador de bajo nivel cuya misión es el cómputo de la transformación \mathbf{T} composición de la transformación T de cambio de dominio y la transformación R de reducción de información, en este trabajo hemos propuesto una modelización del proceso R asumiendo razonadamente que la transformación T es una proyección completa de la señal topográfica sobre un dominio conjunto espacio-frecuencial.

Las aportaciones concretas de nuestro trabajo han sido:

- Haber propuesto que el comportamiento del sistema visual puede caracterizarse, a nivel umbral, mediante un filtro lineal definido en un dominio de Gabor (la *GSF*).
- Haber encontrado la relación de equivalencia entre dos funciones filtro R_1 y R_2 definidas en los dominios inducidos por las transformaciones lineales T_1 y T_2 .
- Utilizando las relaciones propuestas se ha calculado, en un dominio de Gabor concreto, las *GSFs* equivalentes a las *CSFs* de dos observadores reales. Se ha comprobado la equivalencia entre ambas caracterizaciones reproduciendo con los filtros de Gabor, las respuestas impulsionales y el comportamiento general del sistema ante imágenes complejas.
- Se ha analizado, en el caso concreto de los filtros en el dominio de Fourier y filtros en un dominio espacio-frecuencial, cuales son las limitaciones de las relaciones de

equivalencia propuestas, proponiendo métodos para superar dichas limitaciones.

- Se ha aplicado uno de esos métodos para superar la homogeneidad impuesta a la GSF por el carácter homogéneo de la CSF , de forma que hemos sintetizado un sistema espacialmente variante.
- Lateralmente, para obtener los datos experimentales necesarios para el cálculo de los filtros en el dominio de Gabor, se ha propuesto un método experimental sencillo para la determinación de la CSF en todo su dominio bidimensional. Los resultados señalan que para el caso de observadores isótropos se obtiene una buena caracterización midiendo la CSF en dos direcciones perpendiculares del dominio frecuencial, mientras que para los observadores que presenten un comportamiento anisótropo es necesario determinar la CSF al menos en tres direcciones.

8. REFERENCIAS.

1. C.E. SHANNON. A mathematical theory of communication. Bell Syst. Tech. J. **27**, 373-423 (1948)
2. R.C. GONZALEZ & P. WINTZ. Digital image processing. Chap. 6: Image encoding. Addison-Wesley, Reading, Massachusetts. (1987)
3. D. KERSTEN. Predictability and redundancy of natural images. J. Opt. Soc. Am. A **4**, 12, 2395-2400 (1987)
4. J. MALO, A.M. PONS & J.M. ARTIGAS. Bit allocation algorithm for codebook design in vector quantization fully based on human visual system non linearities for suprathreshold contrasts. Elect. Lett. (En prensa)
5. J. MALO & A.M. PONS. Bit allocation by the human visual system inferred from contrast incremental thresholds of sinusoidal gratings. ECVF 95'. (Tübingen 21-25 Aug. 1995)
6. T. EBRAHIMI & M. KUNT. Image compression by Gabor expansion. Optical Engineering. **30**, 7, 873-880 (1991)
7. D. GABOR. Theory of communication. JIEE **93**, 429-457. (1946)
8. H. NYQUIST. Certain factors affecting telegraph speed. Bell System Technical Journal. Vol. 7. p 535. (1928)
9. J.G. DAUGMAN. Uncertainty relation for resolution in space, spatial frequency and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. A **2**, 7, 1160-1169 (1985)
10. J.D. GASKILL. Linear systems, Fourier transforms and optics. Chap. 9: Two-dimensional convolution and Fourier transformation. John Wiley & Sons. New York (1978)

11. A. PAPOULIS. Systems and transforms with applications in optics. Chap. 6: Band-limited signals and uncertainty principle. McGraw Hill (1968)
12. D.G. STORK & H.R. WILSON. Do Gabor functions provide appropriate descriptions of visual cortical receptive fields?. J. Opt. Soc. Am. A 7, 8, 1368-1373 (1990)
13. C.K. CHUI. An introduction to Wavelets. Academic Press. London (1992)
14. D.A. POLLEN & S.F. RONNER. Phase relationships between adjacent simple cells in the cat. Science. 212, 1409-1411 (1981).
15. K. FUKUNAGA. Statistical pattern recognition. Chap 9: Feature extraction and linear mapping for signal representation. Chap 10: Feature extraction and linear mapping for classification. Acad. Press. Boston (1990).
16. J.G. DAUGMAN. Complete discrete 2-D Gabor Transform by neural networks for image analysis and compression. IEEE Trans. Acous., Speech, and Signal process. 36, 7, 1169-1179 (1988)
17. W. H. PRESS, B.P. FLANNERY, S.A. TEUKOLSKY & W.T. VETTERLING. Numerical Recipes in C: the art of scientific computing. Cambridge University Press. Chap.14: Modeling of data (1992)
18. W. H. PRESS, B.P. FLANNERY, S.A. TEUKOLSKY & W.T. VETTERLING. Numerical Recipes in C: the art of scientific computing. Cambridge University Press. Chap. 2: Solution of linear algebraic equations (1992)
19. A.B. WATSON. Detection and recognition of spatial forms. In Physical and biological processing of images. 100-114. Braddick & Sleigh eds. Springer. Berlin (1983)
20. W.S. GEISLER & D.B. HAMILTON. Sampling-theory analysis of spatial vision. JOSA A, 3, 1, 62-70 (1986)
21. J.G. DAUGMAN. Two-dimensional spectral analysis of cortical receptive field profiles. Vision Res. 20, 847-856 (1980)
22. R.W. RODIECK. Quantitative analysis of cat retinal ganglion cell response to visual stimuli. Vision Res. 5, 583-601 (1965)
23. R.L. DE VALOIS, D.G. ALBRECHT & L.G. THORELL. Spatial frequency selectivity of cells in macaque visual cortex. Vision Res. 22, 545-559 (1982)

IEEE Trans. Biom Eng.
Vol. 36 pp. 107-109
(1989)

24. D.A. POLLEN & S.F. RONNER. Visual cortical neurons as localized spatial frequency filter. *IEEE Trans. Syst., Man, Cybern.*, **13**, 5, 907-916, (1983).
25. M.A. WEBSTER & R.L. DE VALOIS. Relationship between spatial frequency and orientation tuning of striate-cortex cells. *J. Opt. Soc. Am. A* **2**, 1124-1132 (1985)
26. F.W. CAMPBELL & J.G. ROBSON. Application of the Fourier analysis to the visibility of gratings. *J. Physiol.* **213**, 157-174 (1968)
27. F.W. CAMPBELL & D. GREEN. Optical and retinal factors affecting visual resolution. *J. Physiol.* **181**, 576-593 (1965)
28. B.E.A. SALEH. Optical information processing and the human visual system. Chapter 10 of *Applications of optical Fourier transforms*. Ed. H. Stark. Academic Press. New York (1982)
29. J.M. ARTIGAS, P. CAPILLA, A. FELIPE & J. PUJOL. Óptica fisiológica: psicofísica de la de la visión. Cap. 13: La función de sensibilidad al contraste. Ed. McGraw Hill. (En prensa)
30. K. DeVALOIS, R.L. DeVALOIS & E.W. YUND. Responses of striate cortex cells to grating and checker-board patterns. *J. Physiol.* **291**, 483-505. (1979)
31. D. FIELD. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* **4**, 12, 2379-2394 (1987)
32. D.C. KNILL & D. FIELD. Human discrimination of fractal images. *JOSA A*, **7**, 6, 1113-1123 (1990)
33. F. MULLER. Distribution shape of two-dimensional DCT coefficients of natural images. *Elect. Lett.* **29**, 22, 1935-1936 (1993)
34. A. COHEN & J. FROMENT. Image compression and multiscale approximation. In *wavelets and Applications*. p 181-206. Y. Meyer ed. Springer. Berlin (1992)
35. M. ANTONINI, M. BARLAUD & P. MATHIEU. Digital image compression using vector quantization and the wavelet transform. In *Wavelets and Applications*. p 160-174. Y. Meyer ed. Springer. Berlin (1992)
36. H.S. WU, R.A. KING & R.I. KITNEY. Improving the performance of the quadtree-based image approximation via the generalized DCT. *Elect. Lett.* **29**, 10, 887-888 (1993)

37. C. GONZALO, J.M. GIL & L.R. BERRIEL-VALDÓS. Definition of a new joint representation for space variant filtering: Discrete Wigner-Walsh Transform. *Actas del VI Simposium Nacional de Reconocimiento de Formas y Análisis de Imágenes*. p 336-343. (1995)
38. J. MALO & A.M. PONS. Linear Filtering in a 4D Gabor domain: a compact formulation for a space variant image processing. *Actas del VI Simposium Nacional de Reconocimiento de Formas y Análisis de Imágenes*. p 344-349. (1995)
39. J. MALO, A. FELIPE, A.M. PONS & J.M. ARTIGAS. Numerical characterization of the human visual system by a weighting function in the Gabor domain: the Gabor stimuli Sensitivity Function (GSF) and its relation to the CSF. (Enviado al JOSA A).
40. H. MOSTAFAVI & D.J. SAKRISON. Structure and properties of a single channel in the human visual system. *Vision Res.* **16**, 957-968 (1976)
41. J.G. DAUGMAN. Spatial visual channels in the Fourier plane. *Vision Res.* **24**, 891-910 (1984)
42. L.O. HARVEY & V.V. DOAN. Visual masking at different polar angles in the two-dimensional Fourier plane. *JOSA A*, **7**, 1, 116-127 (1990)
43. J.G. DAUGMAN. Six formal properties of two-dimensional anisotropic visual filters: structural principles and frequency/orientation selectivity. *IEEE Trans. Syst., Man, and Cybern.*, **13**, 5, 882-887 (1983)
44. S. MARCELJA. Mathematical description of the response of simple cortical cells. *J. Opt. Soc. Am.* **70**, 11, 1297-1300 (1980)
45. D. MARR. *Vision*. San Francisco, CA: W.H. Freeman (1982)
46. A.B. WATSON. The cortex transform: rapid computation of simulated neural images. *Comput. Vision Graph. and Image Process.* **39**, 311-327. (1987)
47. A.B. WATSON. Efficiency of a model human image code. *JOSA A* **4**, 12, 2401-2417 (1987)
48. C.W. HELSTROM. An expansion of a signal in Gaussian elementary signals. *IEEE Trans. Inf. Theory* **12**, 81-82, (1966)
49. M.J. BASTIAANS. Gabor's signal expansion and degrees of freedom of a signal. *Optica Acta* **29**, 9 1223-1229 (1982)
50. A.J.E.M. JANSSEN. Bargmann transform, Zak transform and coherent states. *J. Math. Phys.* **23**, 5, 720-731 (1982)

51. I. DAUBECHIES. Orthonormal bases of compactly supported wavelets. Commun. on pure and applied math. **41**, 909-996 (1988)
52. S.G.MALLAT. Multiresolution approximation and wavelet orthonormal bases of $L^2(\mathbb{R})$. Trans. Am. Math. Soc. **315**, 69-87 (1989)
53. S.G. MALLAT. A theory for multiresolution signal decomposition: the wavelet representation. IEEE Trans. Pattern Anal. and Machine Intell. **11**, 7, 674-693 (1989).
54. G. CRISTÓBAL & R. NAVARRO. Space and frequency variant image enhancement based on a Gabor representation. Patt. Rec. Lett. **15**, 272-277 (1994)
55. M. PORAT & Y.Y. ZEEVI. The generalized Gabor scheme of image representation in biological and machine vision. IEEE Trans. Pattern Anal. and Machine Intellig. **10**, 4, 452-467 (1988)
56. D.J. FIELD, A. HAYES & R.F. HESS. Contour integration by the human visual system: evidence for a local "association field". Vision Res. **33**, 2, 173-193 (1993)
57. L. GAUDART, J. CREBASSA & J.P. PETRAKIAN. Wavelet transform in human visual channels. Appl. Opt. **32**, 22, 4119- 4127 (1993)
58. A.C. BOVIK, M. CLARK & W.S. GEISLER. Multichannel Texture analysis using localized spatial textures. IEEE Trans. on Pattern Anal., and Machine Intell. **12**, 1, 55-73 (1990)
59. M.S. LANDY & J.R. BERGEN. Texture segregation and orientation gradient. Vision Res. **31**, 4, 679-691 (1991)
60. R. NAVARRO, O. NESTARES, J. PORTILLA & A. TABERNERO. Texture analysis and synthesis based on biological vision. Actas del VI Simposium Nacional de Reconocimiento de Formas y Análisis de Imágenes. p 663-664. (1995)
61. U. POLAT & D. SAGI. The architecture of perceptual spatial interactions. Vision Res. **34**, 1, 73-78 (1994)
62. R.F. HESS & D.R. BADCOCK. Metric for separation discrimination by the human visual system. JOSA A, **12**, 1, 3-16 (1995)
63. Special Issue on Wavelet Analysis. IEEE Transactions on Information Theory. Vol. **38**, N° 2 March 1992
64. H.H. SZU. Special Issue on Wavelet Applications. Proc. of the SPIE Vol. 2242 (1994)

65. M.R. PORTNOFF. Time frequency representation of digital signals and systems based on Short Time Fourier analysis. *IEEE Trans. Acoust. Speech, Signal Proc.* **28**, 55-69 (1980)
66. R. WILSON & H.E. KNUTSOSON. Uncertainty and inference in the visual system. *IEEE Trans. Syst., Man, Cybern.* **18**, 305-312 (1988)
67. R. NAVARRO & A. TABERNERO. Gaussian wavelet transform: two alternative fast implementations for images. *Multidimensional Sys. Signal Proc.* **2**, 421-436 (1991)
68. W.K. PRATT. Digital image processing. Chap. 7: Superposition and convolution. John Wiley & Sons. New York (1991).
69. L.A. OLZAK & J.P. THOMAS. Handbook of perception and human performance. Vol. 1: Sensory processes and perception. Chap. 7: Seeing spatial patterns. John Wiley (1986)
70. M. BUADES. Funciones de Sensibilidad al Contraste en presencia de moteado coherente. Tesis doctoral. Dpt. d'Òptica Univertsitat de València. (1994)
71. C.R. INGLING & E. MARTINEZ-URIEGAS. The relationship between spectral sensitivity and spatial sensitivity for the primate r-g X-Channel. *Vision Res.* **23**, 12, 1495-1500 (1983)
72. J.L. BERDEGAY, A. VILA, J. QUEVEDO, J. GUTIERREZ y P. ALBERTOS. Curso de introducción a las tecnologías fuzzy y sus aplicaciones. (Seminario impartido durante las Jornadas sobre Tecnologías Asesoramiento y Formación para las PYMES. Universidad Politécnica de Valencia, 3-6 Mayo 1995)
73. J. MALO, A. FELIPE, M.J. LUQUE & J.M. ARTIGAS. On the intrinsic two-dimensionality of the CSF and its measurement. *Journal of Optics.* **25**, 3, 93-103 (1994)
74. J.G. PROAKIS & D.G. MANOLAKIS. Digital signal processing. Principles, algorithms and applications. Chap. 6, p. 450. Macmillan Publ. Co. (1992)
75. D.H. KELLY. Receptive-field-like functions inferred from large-area psychophysical measurements *Vision Res.* **25**, 12, 1895-1900. (1985)
76. R. NAVARRO, P. ARTAL & D. WILLIAMS. Modulation transfer of the human eye as a function of retinal eccentricity *J. Opt. Soc. Am. A* **10**, 2, 201-212. (1993)