

# LA DIFERENCIA FÁSICA: CLAVE FUNDAMENTAL DE LA ESTEREOPSIS

**JOSÉ ANTONIO AZNAR CASANOVA  
Y JOSEP FERRÉ I BORRULL**  
Universidad de Barcelona

## Resumen

En este trabajo se propone la utilización de unos nuevos estímulos, los *estereogramas de enrejado*, para experimentar y estudiar la estereopsis visual. Este tipo de estereogramas aporta dos ventajas adicionales respecto a los propuestos por Julesz (1960, 1964). Por un lado, nos han permitido describir los estímulos en terminos de ondas (luminancia, amplitud, frecuencia espacial, orientación y fase). Así, hemos puesto de manifiesto, mediante un experimento psicofísico, que la propiedad del estímulo que elicitla la estereopsis es la diferencia fásica entre los dos estereogramas del par. También, se describen las relaciones existentes entre las variables disparidad geométrica, desfase y similitud de estereogramas. Por otro lado, se muestran posibles aplicaciones del análisis de Fourier, en el dominio de la frecuencia espacial. Así, mostramos la contribución de los espectros de amplitud y fase en la estereopsis mediante imágenes sintetizadas a partir de estos espectros. Además, con base en este análisis, es posible obtener un mapa que haga explícita la profundidad, recurriendo a la operación de correlación entre las dos imágenes del par estereo.

**Palabras clave:** Estereopsis, estereogramas de enrejado, análisis de Fourier, espectro de fase,

## Abstract

In this paper the use of grating's stereos pairs is proposed in order to study the stereopsis. This type of stimulus brings two additional advantages to respect random-dot stereo pairs proposed for Julesz (1960, 1964). In the first place, allow to analyze the stimulus as a wave, in terms of mean luminance, amplitude, spatial frequency, orientation and phase. Thus, across a psychophysical experiment we have shown the property of stimulus to be able to produce the stereopsis: the phases' difference between two stereo pairs. Also, are described the relations between geometrical disparity, disphase and likeness of stereo pairs. In the second place, we have shown the possibles applications of the Fourier analysis into spatial frequency's dominion. Thus, we have shown the contribution to the stereopsis of amplitude's and phase's spectrum through images synthesized from this spectrums. Moreover, based on this type of analysis is possible to obtain a map which makes the depth explicit, using the correlation operation between two stereo images.

**Key words:** Stereopsis, gratings stereograms, Fourier analysis, spectral phase, visual psychophysic.

## Introducción

El renovado interés actual por la investigación de la visión de la profundidad ya no viene dado, exclusivamente, por la necesidad substantiva de comprender y explicar este fenómeno perceptivo, sino que también existe una creciente demanda en la investigación aplicada, desde la I.A. (Inteligencia Artificial), en áreas tales como la visión computacional y la robótica.

La estereopsis visual o visión estereoscópica, en cuanto consecuencia derivada de la binocularidad, posibilita la extracción de información relativa a la distancia entre las superficies y objetos o entre éstos y el observador. Por lo común, se admite desde los enfoques inferencial y constructivista que la información sobre la profundidad y distancia resulta obtenida a partir de las disparidades retinianas.

En relación con la estereopsis, señalaba Julesz (1964) que los investigadores de este campo se habían ocupado asiduamente, ya sea de cuantificar las relaciones entre disparidad binocular y profundidad percibida por sujetos humanos, ya sea de determinar los límites de la disparidad o de la escala métrica de estimación subjetiva de la profundidad del espacio y proponía que el problema relevante debía radicar en explicar cómo logra el S.V.H. (Sistema Visual Humano) emparejar correctamente los puntos apropiados de las dos imágenes que conforman el par estéreo. Aspecto que se abordaría, fundamentalmente desde las aproximaciones Computacional-representacional y psicofísica (véase Julesz, 1986, para una revisión). Atribuía la elección de los problemas antes enumerados a las limitaciones inherentes al tipo de estímulos (estereogramas) utilizados. Precisamente, criticaba el hecho de que se hubiesen usado representaciones similares a pares de diapositivas fotografiadas con cierta disparidad (desplazando horizontalmente la cámara fotográfica), tales como dibujos de escenas reales, fotografías, ya que, en la percepción de la profundidad mediante este tipo de estímulos, interactúan numerosas claves monoculares, de manera que resulta imposible controlarlas. En un primer intento de superar estas dificultades, se recurrió al uso de delgadas líneas verticales (Ogle, 1962) como único detalle que constituía los estereogramas. El propio Julesz (1960), propuso un original paradigma experimental, utilizando una nueva clase de estímulos: los estereogramas de puntos aleatorios generados por un ordenador (matrices binarias de dimensión 100 x 100). Los resultados obtenidos mediante este procedimiento de investigación evidenciaban que era posible la estereopsis, aún cuando se hallaban ausentes todas las claves que él denominó de familiaridad (tradicionalmente denominadas monoculares o pictóricas), excepto la disparidad binocular (a veces llamada disparidad retiniana, paralaje binocular o disparidad geométrica). La gran ventaja que se derivaba de la utilización de los estereogramas de puntos aleatorios en el estudio de la estereopsis era, naturalmente, que se podía prescindir de todas las claves, salvo la disparidad geométrica horizontal, y que la forma crítica sólo se manifiesta al observador cuando la contempla estereoscópicamente, es decir, cuando entra la imagen de cada uno de los dos estereogramas del par (el izquierdo y el derecho) por cada ojo, estando separados físicamente los dos campos visuales. La forma en relieve que se percibe no resulta captada si el sujeto observa los estereogramas monocularmente, e incluso al observarlos binocularmente, si no realiza una convergencia ocular forzada, lo que requiere entrenamiento.

A nuestro juicio, un avance de interés, en el intento de esclarecer la estereopsis, debería partir de la aplicación del análisis de Fourier, en primer lugar a la descripción de los estereogramas que elicitán el fenómeno estereoscópico y, una vez demostrada su utilidad y las ventajas que se evidencian mediante el uso de este modelo formal matemático, aplicarlo a la descripción del funcionamiento del S.V.H., en lo que a este fenómeno se refiere.

Como es sabido, es posible describir cualquier estímulo visual a partir de un conjunto limitado de importantes elementos, tales componentes elementales de las señales visuales (imágenes) son los denominados enrejados sinusoidales (sinusoidal grating). En efecto, el teorema de Fourier afirma, formalmente, que cualquier función de onda (y el perfil de luminancia de un enrejado y cualquier imagen lo son) puede descomponerse en infinitos componentes de onda sinusoidal. A partir de este teorema, se deduce que cualquier onda puede sintetizarse integrando entre sí sus componentes de onda sinusoidal, cuyas amplitudes y fases vienen dadas por la transformada de Fourier (TF) de dicha imagen (Bracewell, 1978). Por consiguiente, una imagen cualquiera puede expresarse como una determinada combinación lineal de enrejados sinusoidales.

Ciertamente, ya ha transcurrido un cuarto de siglo desde que Campbell y Robson (1968) aplicaran el análisis de Fourier, utilizando estímulos de enrejado, para obtener la FSC (Función de Sensibilidad al Contraste) del S.V.H., variando la frecuencia espacial de dichas señales. Las investigaciones realizadas en esta línea han llevado a ciertos autores (Campbell y Robson, 1968; Graham y Nachmias, 1973; Maffei y Fiorentini, 1973; Daugman, 1984; etc.) a sugerir que el S.V.H. practica una especie de análisis de Fourier sobre la entrada estimular.

Para mostrar la viabilidad del análisis de Fourier en la descripción de los estímulos susceptibles de producir la visión estereoscópica, hemos diseñado y elaborado lo que denominaremos *estereogramas de enrejado*, los cuales permiten un ilimitado número de variaciones en sus parámetros característicos (Luminancia media y amplitud, [o contraste], frecuencia espacial, orientación y fase) y acceder, mediante el análisis de Fourier, a la manipulación de tales parámetros.

Cualquier estímulo de enrejado puede ser descrito, analíticamente, mediante la siguiente función:

$$f(x,y) = L_0 + A \cdot \cos [ 2\pi \rho ( x \cdot \cos \alpha + y \cdot \sen \alpha ) + \varphi ]$$

Donde:

$L_0$  = luminancia media.

$A$  = amplitud de la onda.

$\rho$  = frecuencia espacial

$\alpha$  = orientación del enrejado.

$\varphi$  = fase del enrejado.

$(x,y)$  = coordenadas cartesianas del pixel en la imagen.

Determinados estereogramas pusieron de manifiesto que la disparidad retiniana era susceptible de generarse artificialmente. Para ello bastaba con producir cierta disparidad geométrica entre los dos estereogramas del par estéreo, lo que equivale a desplazar horizontalmente cierta porción de uno, o los dos estereogramas. Ello, a su vez, produce un desplazamiento de la fase de un estereograma con respecto al otro. Por consiguiente, el tipo de estereogramas aquí propuesto facilitará, en gran medida, la verificación de la hipótesis de que la única clave necesaria y suficiente para elicitar el fenómeno de la estereopsis visual sea la *diferencia fásica* o *desfase* entre los dos estereogramas del par. Así, en lo sucesivo, con el término *desfase* haremos referencia, aquí, a la diferencia de fases entre dos estímulos de enrejados.

$$\text{Desfase} = \varphi_n - \varphi_m ; \text{ Donde: } \varphi_n \in \text{Enrejado-2 y } \varphi_m \in \text{Enrejado-1.}$$

Por supuesto, es posible expresar el *desfase* en términos de *disparidad geométrica* (variable utilizada tradicionalmente) o número de pixeles que ha sido desplazada la porción central del estereograma, mediante una transformación lineal de dicho desfase.

Como primer objetivo y a fin de poner de manifiesto la relevancia en la estereopsis de la diferencia de fase entre las dos imágenes del par estereoscópico o *desfase*, obtendremos las correlaciones cruzadas entre cada par de estereogramas, haciendo variar dicho desfase en un rango que abarque, al menos dos ciclos de la variación de luminancia del enrejado. Para ello, se calcularán las TF de cada estereograma, se hallará el conjugado de la TF de uno de ellos y se efectuará el producto entre esta imagen conjugada y la TF del otro estereograma del par. Al obtener la  $TF^{-1}$  (Transformada inversa de Fourier) y visualizarla, deberíamos obtener un mapa de correlación entre estos dos elementos, o dicho en otros términos, del desplazamiento relativo entre las dos imágenes. Además, el máximo valor de pixel de la imagen de este mapa debe servir de indicador del parecido entre ambas (Domingo, 1993) [Para una ampliación de estas operaciones sobre la imagen véase González y Wintz, 1977; Bracewell, 1978]. El valor de este máximo, por consiguiente, expresará la *similitud* entre los dos elementos del par

estereoscópico. Y la tabla de valores de similitud, obtenida mediante la correlación de cada par de estereogramas de enrejado, constituirá el dominio de la función de similitud.

La representación gráfica de la similitud y el desfase, en función de la disparidad geométrica, puede poner de manifiesto la relación entre estas variables de naturaleza geométrica. Esperamos, así, mostrar las posibilidades que ofrece esta técnica de análisis del estímulo adaptada al fenómeno de la estereopsis visual y revelar el papel relevante que desempeña la diferencia fásica en ello.

Adicionalmente, llevaremos a cabo un experimento psicofísico de estimación de la profundidad subjetiva en el espacio, haciendo uso de los estereogramas de enrejado aquí propuestos. Naturalmente, el fin que perseguimos con ello es el de relacionar los parámetros geométricos (desfase y disparidad) con los valores de las estimaciones subjetivas de las personas.

El segundo objetivo, no por ello de menor interés, que nos proponemos es el de evidenciar el papel fundamental que juega el *espectro de fase* en la experiencia personal de la estereopsis. Más concretamente, analizar la importancia de la diferencia entre los espectros de fase de los dos estereogramas que conforman el par estereoscópico.

Las diferentes funciones que desempeñan los espectros de amplitud y de fase han sido señaladas por diferentes autores (Piotrowsky y Campbell, 1982; Caelli y Bevan, 1982; etc). Sierra-Vázquez (1992) subraya la importancia del espectro de fase, en cuanto que representa la estructura espacial fundamental de la imagen. Oppenheim y Lim (1981) mostraron que, la mayor parte de las veces, el espectro de fase contiene la información esencial de una imagen y que este espectro, por sí sólo, es suficiente para reconstruir completamente la imagen. Por esta razón, el espectro de fase ha sido aplicado con bastante éxito al reconocimiento de patrones. Horner y Gianino (1984), trataron de elaborar un filtro para el reconocimiento de imágenes que fuese más eficiente que los tradicionalmente empleados, utilizando como estímulos, únicamente, la información contenida en el espectro de fase de la TF de la imagen a identificar. No obstante, Juvells et al. (1990) advierten que, efectivamente, en imágenes tomadas del medio natural, las cuales muestran, en el espectro de amplitud, una forma similar en el dominio de la frecuencia espacial, el espectro de fase contiene la información más relevante sobre la estructura espacial de la imagen. Sin embargo, cuando se procesan computacionalmente imágenes digitalizadas, que presentan diferencias substanciales entre los espectros de amplitud, entonces, el espectro de fase no es el factor esencial. Si bien, admiten estos autores que encontrar imágenes que difieran tanto en sus distribuciones de amplitudes no es frecuente entre imágenes naturales, por ello, hubieron de recurrir para mostrar este efecto a imágenes artificiales con apariencia periódica, como por ejemplo los enrejados sinusoidales, a fin de combinar el espectro de fase de una imagen natural con el espectro de amplitud de un enrejado, o viceversa. Llegaron así a evidenciar que, sólo en estos casos, el espectro de amplitud contiene la información más relevante.

Sabido es que la TF de una imagen es una función bivariada sobre el conjunto de los números complejos  $[F(u,v)]$  que satisface:

$$f(x,y) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} F(u,v) e^{i2\pi(ux+vy)} \cdot du \cdot dv$$

Y se puede definir como:

$$F[f(x,y)] = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y) e^{-i2\pi(ux+vy)} \cdot dx \cdot dy$$

Donde  $u$  y  $v$  son las frecuencias espaciales asociadas a las dos direcciones  $(x,y)$  del dominio espacial, respectivamente.

Las imágenes sobre las que trabajamos se modelizan mediante funciones reales bidimensionales; mientras que su transformada (TF) es, en general, una función definida en el conjunto de los números complejos. Por consiguiente, es posible separar la parte real y la parte imaginaria:

$$\mathcal{F} [f(x,y)] = F(u,v) = \text{Re}[F(u,v)] + i.\text{Im}[F(u,v)]$$

Otra manera útil de expresar la TF es la denominada *módulo-argumental*, en la que:

- El *módulo* de la TF o *espectro de amplitud* se define por:  $|F(u,v)| = \sqrt{\text{Re}^2 + \text{Im}^2}$ .
- Y el *argumento* o *espectro de fase* se define por:  $\Phi(u,v) = \text{Arctg}(\text{Im}/\text{Re})$ .

Además, la TF de una imagen satisface la propiedad de ser simétrica respecto a los ejes  $u$  y  $v$ , por lo que también puede expresarse así:

$$\mathcal{F} [f(x,y)] = \frac{1}{2} \int_0^{+\infty} \int_{-\infty}^{+\infty} |F(u,v)| \cdot e^{i.\Phi(u,v)} \cdot e^{i2\pi(ux+vy)} \cdot du.dv$$

Ahora, mediante el análisis de Fourier se nos ofrece la posibilidad de analizar los estímulos de un modo nuevo y que facilita las manipulaciones que pueden efectuarse sobre las imágenes digitalizadas (procesamiento computacional de imágenes digitalizadas). Por ejemplo, podemos restar los espectros de fase de los dos estereogramas y obtener una nueva imagen combinando dos funciones, el espectro de amplitud de uno de ellos con el *espectro de la diferencia fásica* para, tras aplicar la  $TF^{-1}$ , observar los efectos producidos sobre la imagen. También, podemos sintetizar dos nuevas imágenes, una combinando el espectro de amplitud del estereograma izquierdo con el de fase del estereograma derecho y, otra a la inversa. Esta estrategia es habitualmente utilizada para analizar la importancia de ambos espectros (amplitud y fase) en cuanto al contenido de información relevante de la imagen. Por supuesto, además de estos filtrados basados en el espectro de fase, es posible realizar diversos filtrados basados en el espectro de amplitud (paso-bajo, paso-alto, paso-banda) en cada uno de los elementos del estereograma o alguna combinación de ellos.

Un aspecto importante, a nuestro juicio, en estas manipulaciones de los espectros de la imagen aparece al examinar la forma que adopta la operación de *correlación cruzada* en el dominio de Fourier. Ciertamente, si se expresan los espectros en forma módulo-argumental, obtendremos que el módulo de la operación de *correlación* entre dos imágenes es el producto de los módulos, en tanto que, el argumento de ésta es la diferencia de argumentos (o fases). De manera que, si llamamos a las transformadas (de Fourier) de  $f$  y  $g$ , respectivamente,  $F$  y  $G$ :  $\mathcal{F} [f] = F$  y  $\mathcal{F} [g] = G$ , la operación de correlación, como anteriormente describíamos, se expresaría así:

$$\mathcal{F} [f \star g] = F \cdot G^* \quad (1)$$

Donde  $G^*$  es el conjugado de  $G$ .

Expresando, ahora,  $F$  y  $G$  en forma módulo-argumento tenemos:

$$F = |F| \cdot \exp (+i \cdot \Phi_F)$$

$$G^* = |G| \cdot \exp (-i \cdot \Phi_G)$$

Adviértase el cambio de signo debido a la conjugación.

Si sustituimos en (1) estas expresiones y operamos, resulta:

$$\mathcal{F} [f \star g] = |F| \cdot |G| \cdot \exp (i (\Phi_F - \Phi_G))$$

En ella se observa que *la correlación cruzada de dos imágenes cualquiera es una operación en la cual queda registrada la información sobre la diferencia de fase y, por consiguiente, según nuestra hipótesis, debe contener la información acerca de la profundidad.* Si ésto es así, podremos afirmar que la imagen de la correlación viene a representar algo parecido a una especie de mapa de profundidades, expresadas en niveles de gris.

## Procedimiento de elaboración de estereogramas de enrejado

El procedimiento de elaboración del tipo de estímulos que aquí designamos con el término *estereogramas de enrejado* se hace explícita a través de la descripción de los estereogramas que nosotros utilizamos. Evidentemente, variando los mencionados parámetros que sirven para caracterizar a los estímulos de enrejado que lo componen, puede generarse una pluralidad cuasi-ilimitada de estereogramas de esta modalidad. En este trabajo, al objeto de controlar las variables intervinientes, cada par de estereogramas fue diseñado de acuerdo con las siguientes propiedades:

Las dimensiones de cada elemento del par estereográfico eran de 256x256 píxeles [aquí y en lo sucesivo, las dimensiones son expresadas en formato 'columna x fila']. Cada estereograma fue obtenido por el solapamiento de un enrejado vertical, que sirve de porción central, y cuyas dimensiones son de 80x256 píxeles, sobre otro enrejado horizontal, que sirve de fondo periférico, y cuyas dimensiones son de 256x256 píxeles. La porción central es llamada así porque oculta, al superponerse, una determinada porción del enrejado horizontal periférico, y se ubica en el centro de éste, dejando ver tan sólo sus porciones laterales (izquierda y derecha). Concretamente, esta porción central se localiza en las coordenadas (79,0) [esquina superior izquierda] y (159,255) [esquina inferior derecha], por lo que delimita un rectángulo central de 80x256 píxeles, quedando a izquierda y derecha de éste otros dos rectángulos de 88x256 píxeles [téngase en cuenta que las coordenadas de la imagen se hallan en el rango 0-255 píxeles, tanto en el eje de ordenadas como en el de abscisas]. Por consiguiente, podría decirse que cada estereograma consta de dos regiones, los *laterales*, que actúan de contexto periférico constante y la *central* cuyo interior (parte de un 'infinito' enrejado vertical) visible en cada caso (al concretarse en un par de estereogramas) es variable.

Las dos regiones laterales, ubicadas a cada lado de la región central, y constantes en los dos estereogramas, están constituidas por un enrejado de forma de onda cuadrada (en este trabajo), con los siguientes valores de los parámetros: una frecuencia espacial igual a  $\rho_1 = 6.4$  ciclos/img. [adoptaremos como unidad de medida el N° de ciclos/imagen, en lugar del número de ciclos/g.a.v. (grados de ángulo visual). Manteniendo fijo el tamaño de la imagen (256x256)], es decir, una longitud de onda ( $\lambda$ ) de 40 píxeles, una orientación  $\alpha_1 = 90^\circ$  (barras horizontales), el máximo contraste posible [según la fórmula de Michelson-1927:  $m = A / L_o = (L_{max} - L_{min}) / (L_{max} + L_{min})$ ]. o sea  $m=1$ , lo que en nuestro caso implica dos hechos. Primero, que la luminancia máxima será la mayor que permita el dispositivo de presentación y la luminancia mínima nula ( $L_{max} - L_{min} = L_{max}$ ), por lo que la luminancia media será  $L_o = (L_{max} / 2)$ . Segundo, que la amplitud será cuantitativamente igual a la luminancia media ( $L_o$ ). La fase será nula ( $\phi_1 = 0$  radianes) en la dirección ortogonal a la variación de luminancias y sentido de arriba hacia abajo.

La región central, que solapa parcialmente al enrejado horizontal, es siempre constante en el enrejado izquierdo y variable sólo en fase para el derecho. También se halla constituida por un enrejado de forma de onda cuadrada. Sin embargo, algunos de los parámetros que lo caracterizan difieren de los expresados para el enrejado que constituye las regiones laterales. En efecto, la frecuencia espacial será de  $\rho_2 = 12,8$  c/img. ( $\lambda = 20$  píxeles), la orientación de las barras es de  $\alpha_2 = 0^\circ$  (verticales). No obstante, tienen en común (esta región central con las

laterales) el contraste ( $m=1$ ) y, consiguientemente, la luminancia media ( $L_0$ ) y la amplitud ( $A$ ). La fase de la porción central del estereograma izquierdo, la cual se mantiene siempre constante, es nula ( $\varphi_2 = 0$  radianes) en la dirección perpendicular a la variación de luminancias y sentido de izquierda a derecha. Mientras que, la fase de la porción central del estereograma derecho es la variable independiente, la cual, por tratarse de un estímulo cuyo perfil de luminancia resulta ser una onda periódica puede variar en un rango definido ( $\varphi_3 = 0-2\pi$  radianes o entre  $0^\circ-360^\circ$  sexagesimales). Obviamente, el desfase entre los dos estereogramas (izquierdo y derecho) viene dado por la diferencia de fase entre los dos enrejados que ocupan la porción central del estereograma, a saber:

$$\text{DESFASE} = \varphi_3 - \varphi_2.$$

La Figura-1 muestra un ejemplo de par de estereogramas de enrejados, susceptible de elicitar la estereopsis. El estereograma izquierdo, que en nuestro caso es siempre constante, tiene la fase nula (0 radianes o  $0^\circ$  sexagesimales), es decir, no ha sido desplazada la porción central ningún pixel. El estereograma derecho, el que (en nuestro caso) varía, tiene aquí una fase de  $3\pi/2$  radianes ( $270^\circ$  sexag.), es decir, tres cuarto de ciclo (0,75c.) o, lo que es lo mismo, la porción central (enrejado) ha sido desplazado hacia la derecha 15 pixeles. Naturalmente, el desfase entre los dos estereogramas ( $\varphi_3 - \varphi_2$ ) es de  $3\pi/2$  radianes ( $270^\circ$ ).

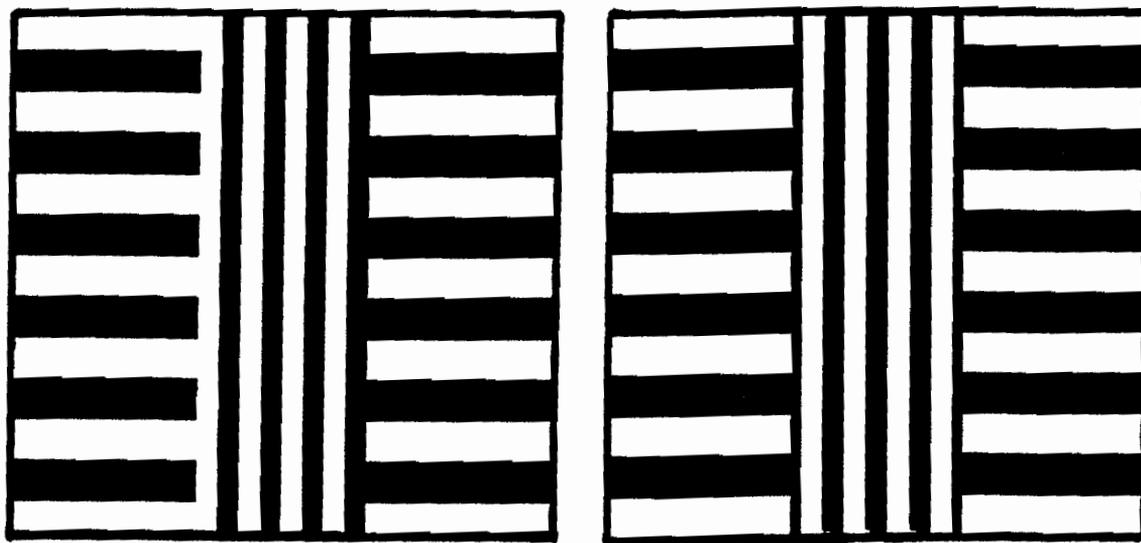
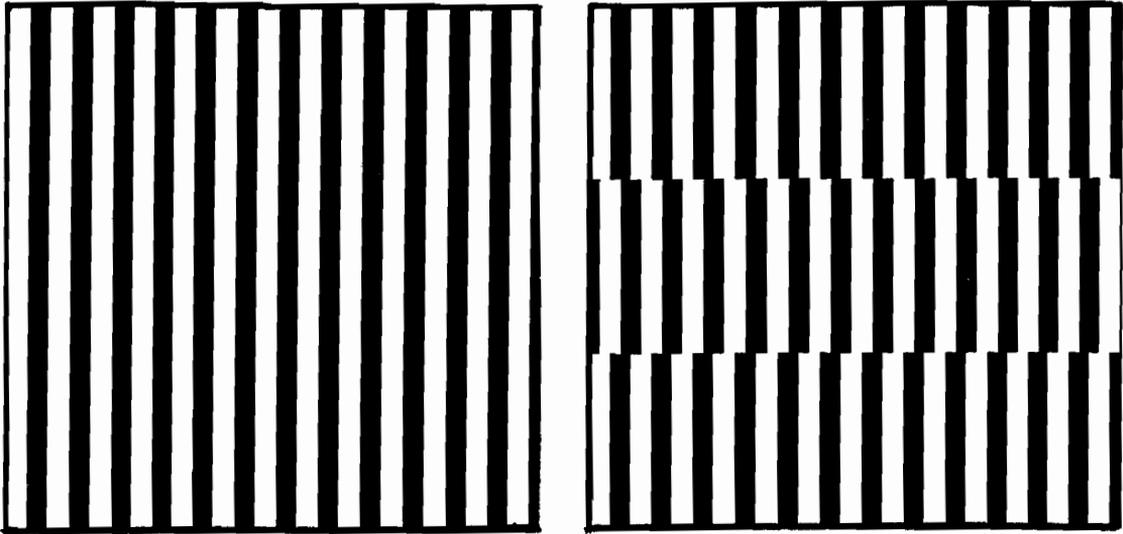


FIGURA 1.- Ejemplo de un par de estereogramas de enrejado que presentan un desfase de  $3\pi/2$  radianes o 15 pixeles de disparidad. A) Estereograma izquierdo ( $\varphi_2 = 0$  rad.). B) Estereograma derecho ( $\varphi_3 = 3\pi/2$  rad.).

Obsérvese que, a diferencia del procedimiento utilizado por Julesz (véase Julesz, 1986; para una revisión de trabajos sobre estereopsis), la porción central de cada estereograma, en nuestra propuesta, siempre se localiza en las mismas coordenadas espaciales. En otras palabras, el citado autor, al describir la elaboración de sus estereogramas de puntos aleatorios, desplazaba la porción central, de uno o de ambos estereogramas, cierto número de pixeles hacia la derecha o hacia la izquierda, lo que daba lugar a que dicha porción se experimentase como hundiéndose detrás de la porción periférica o flotando delante de ésta, respectivamente. En nuestro caso,

la porción central se comporta como una ventana, la cual muestra (en diferentes estereogramas derechos) un segmento de enrejado de  $\rho = 12,8$  c/img. (de subdimensión  $80 \times 256$  pixeles), que comienza, para cada estereograma derecho generado, en distinta fase ( $\varphi_3$  variable entre  $0$  y  $2\pi$  rad.) y que siempre comienza en la misma fase nula ( $\varphi_2 = 0$  radianes) en el estereograma izquierdo. Por tanto, la disparidad geométrica es obtenida de un modo diferente al empleado por Julesz (1964).

En la Figura-2, mostramos otra variante de par estéreo de enrejado, todavía con mayor grado de simplificación en su composición.



*FIGURA 2.- Otro ejemplo de par estereo de único enrejado, el cual presenta un desfase de  $3\pi/2$  radianes o 15 pixeles de disparidad ( $\rho = 12,8$  c/l)*

En efecto, la simple inspección visual de estos estereogramas nos informa de que éstos se hallan constituidos por un único enrejado, si bien en el estereograma dispar (el derecho), se ha variado la fase del tercio central horizontal, de modo que los parámetros del estereograma izquierdo son:  $m = 1$ ;  $\rho = 12,8$  c/i;  $\lambda = 20$  pix.,  $\alpha = 0^\circ$  y  $\varphi = 0$  rad.; mientras que los del estereograma derecho son los mismos, excepto la fase del citado tercio central ( $\varphi = 3\pi/2$  rad.;  $Dg = 15$  pix.).

### **Análisis de la relación entre variables geométricas del par de estereogramas de enrejado**

Dado que la fase del estereograma izquierdo es  $\varphi_2 = 0$  radianes y que la del derecho es variable [ $\varphi_3 = r$  radianes ( $\forall r \in \mathbb{R}$ ;  $0 \leq r \leq 2\pi$ )], haciendo variar la fase del estereograma derecho ( $\varphi_3$ ) entre  $0-4\pi$  radianes (40 pixeles), obtendremos los valores de la función de desfase entre ambos enrejados. Además, como en nuestro caso un ciclo cualquiera, del enrejado utilizado para formar la porción central, es de 20 pixeles, fácilmente podemos transformar la 'disparidad geométrica' ( $Dg$ ) o número de pixeles (del 'infinito' enrejado vertical de  $\rho = 12,8$  c/img.;  $\lambda = 20$

pix.) desplazados horizontalmente para insertar en dicha región central del estereograma derecho, a valor de la fase ( $\varphi_n$ ), expresado en radianes, mediante la siguiente fórmula:  $\varphi_n = (Dg/\lambda)2\pi$ . En nuestro caso:  $\varphi_n = (Dg/10)\pi$ .

Por otra parte, hallamos las correlaciones entre cada par de imágenes que componen un par de estereogramas y averiguamos el valor máximo de pixel del mapa de correlación resultante, el cual, como se dijo, expresa la similitud o parecido entre ambas imágenes.

Una vez obtenidos los valores de las funciones de *desfase*, *disparidad* y *similitud*, podemos representar la variación en desfase y similitud entre los estereogramas, en función de la 'disparidad geométrica'.

En el Gráfico-1 se muestra esta representación, el eje de ordenada izquierdo ( $Y_1$ ) contiene la escala de la variable *similitud* y el derecho ( $Y_2$ ) la escala de valores del *desfase* ( $\varphi_3 - \varphi_2$ ) entre el estereograma izquierdo y el derecho. La abscisa contiene la Dg en pixeles.

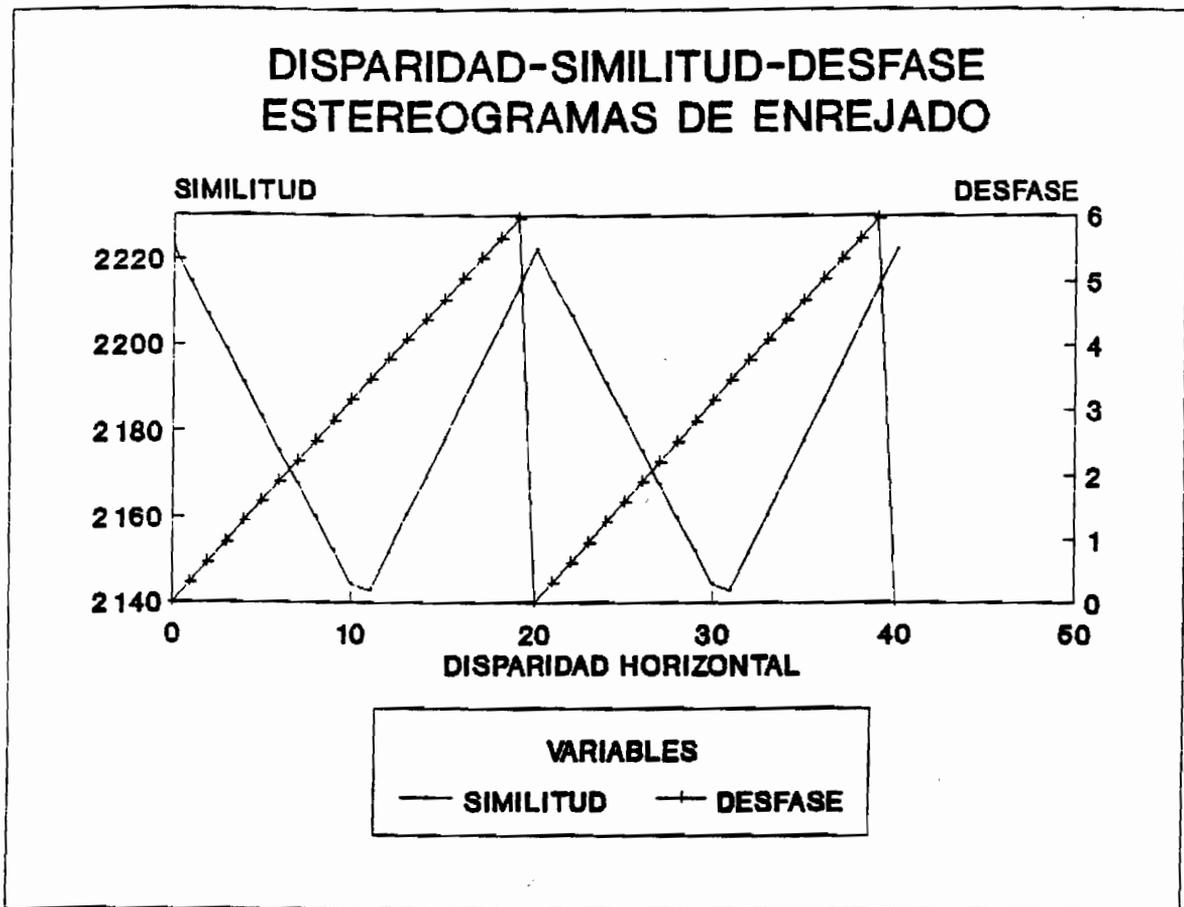


GRÁFICO 1.- Relación entre Desfase y Similitud, según la disparidad geométrica establecida en una serie de 40 estereogramas de enrejado

En este gráfico podemos observar como la periodicidad, con que aplicamos la disparidad geométrica a la serie constituida por 40 pares de estereogramas, que contiene dos ciclos completos, se refleja en la función de similitud. Esta función con forma de onda periódica muestra una longitud de onda de 20 píxeles, lo que, a su vez, coincide con la periodicidad del enrejado de la porción central (aquí utilizado). Por consiguiente, puede afirmarse que la función de similitud indica, como su nombre advierte, el parecido entre cada par de estereogramas considerado, teniendo en cuenta el período.

Con respecto al desfase, la función de similitud decrece hasta llegar a ser mínima para una fase de  $\pi$  radianes (disparidad geométrica= 10 píxeles) y luego vuelve a crecer hasta indicar el máximo parecido cuando comienza a repetirse un nuevo ciclo (20 píxeles), es decir, para un desfase de  $2\pi$  radianes.

En consecuencia, cada par de estereogramas queda caracterizado por el valor que presenta en estas tres variables, disparidad ( $D_g$ ), desfase ( $\varphi_3 - \varphi_2$ ), frecuencia espacial ( $\rho$ ) o, si se prefiere longitud de onda ( $\lambda$ ) y similitud (Máximo pixel). Es importante reparar en que la fase del estereograma derecho, variable que nosotros manipulamos, viene determinada por el cociente entre la  $D_g$  y  $1/\rho$  (o  $\lambda$ ); por tanto, cualquier variación en uno de estos tres parámetros afectará a la profundidad percibida. Así, los sujetos experimentarán el mismo grado de profundidad cuando, por ejemplo, presentemos un par estéreo de  $\lambda = 20$  pix., con una  $D_g = 10$  pix., que cuando presentemos un par estéreo de  $\lambda = 40$  pix., con una  $D_g = 20$  píxeles. No obstante, en sucesivos trabajos, conviene manipular otros parámetros tales como provocar diferencias en las frecuencias espaciales entre los dos enrejados componentes, diferencias de luminancia en las porciones centrales, diferencias de orientación de las barras, etc, de cada enrejado (central).

## Experimento Psicofísico

Puesto que la disparidad geométrica ( $D_g$ ) existente en cada par de estereogramas se halla en relación lineal con el desfase ( $\varphi_3 - \varphi_2$ ), tendremos que:

$$(\varphi_3 - \varphi_2) = \frac{D_g \cdot 2\pi}{\lambda}$$

Para tratar de verificar si, como cabe esperar, la *estimación de la profundidad subjetiva* se halla en relación directa con el citado *desfase*, realizaremos un experimento de emparejamiento de magnitudes, de acuerdo con la psicofísica directa de Stevens (1966).

## Método

### Sujetos

Un sólo sujeto, uno de los autores de este trabajo, realizó las estimaciones de la profundidad subjetiva percibidas en la serie de estereogramas.

### Material

Elaboramos un programa de ordenador, mediante el cual generamos 40 pares de estereogramas de enrejado, los cuales se diferenciaban unos de otros en el desfase que producía la disparidad geométrica existente en el estereograma derecho. El rango de dispa-

ridades de éste fluctuaba desde 0 píxeles ( $\varphi_3 = 0$  radianes) hasta 39 píxeles ( $\varphi_3 = 4\pi$  radianes), es decir, comprendía dos ciclos completos de la variación de luminancias del enrejado de la porción central. Los restantes parámetros eran los descritos anteriormente, en el apartado correspondiente al procedimiento de elaboración de estereogramas de enrejado.

## **Aparatos**

Un ordenador i486-33 MHz, RAM 4 Mbytes, con pantalla monocromática y tarjeta gráfica SVGA (1024x768 píxeles), el cual nos permitía tanto generar los estereogramas como actuar de presentador de estímulos.

## **Procedimiento**

Se utilizaron dos bloques de igual número de estereogramas. Cada uno de estos bloques constaba de 40 estereogramas, y un programa de ordenador los presentaba en orden aleatorio. Por consiguiente, la prueba constaba de 80 ensayos.

La secuencia experimental se iniciaba cuando la pantalla mostraba un par de estereogramas de enrejado. Se hacía visualizar al sujeto cada uno de los estereogramas del par mediante un panel separador, uno con cada ojo, ayudándose de unos prismas de  $12^\Delta$  y  $15^\Delta$ , para facilitar la convergencia ocular requerida y lograr el efecto estereoscópico. Mientras observaba, el sujeto ajustaba la abertura de un calibre hasta que éste igualaba la profundidad percibida (emparejamiento de modalidad cruzada). A continuación, introducía la medición registrada por el calibre a través del teclado del ordenador y comenzaba otro ensayo. El valor de cada estimación de profundidad resultaba del promedio de las estimaciones de cada par de ensayos homólogos (uno de cada bloque).

## **Resultados**

Como es lógico y evidente, dado que cada 20 píxeles de disparidad geométrica se iniciaba un nuevo ciclo de variación de luminancias de la porción central, las estimaciones de profundidad se ajustaban básicamente a tal periodicidad.

Ahora bien, en el transcurso de la prueba de estimación, el sujeto observó que en el rango comprendido entre 11-18 píxeles de disparidad geométrica existían dos posibles percepciones estereoscópicas de cada par de estereogramas de enrejado, las cuales vienen determinadas cognitivamente. Es decir, según la banda (blanca o negra) del enrejado de la porción central a la que adjudicase el papel de fondo o de figura, podía percibirse la porción central 'delante de' o 'detrás de' las porciones laterales. Así, cuando las bandas blancas del enrejado central actuaban de figura y las bandas negras del mismo a modo de fondo, dichas bandas blancas eran percibidas como 'detrás de' las porciones laterales pertenecientes al enrejado horizontal. Mientras que, cuando las bandas negras de la porción central, juntamente con las blancas adyacentes, actuaban de figura, éstas eran percibidas como 'delante de' las porciones laterales.

En el Gráfico-2, representamos la relación observada entre las tres variables. En el eje de abscisa colocamos la disparidad geométrica, en el eje de ordenadas derecho ( $Y_2$ ) la diferencia de fase entre los dos estereogramas del par o desfase y en el eje de ordenadas izquierdo la estimación de profundidad subjetiva.

Este gráfico nos muestra con claridad como, en nuestro caso, la estimación de profundidad se ajusta de modo casi perfecto al desfase, si la porción central es percibida como estando

'más cerca de' (Estimación-C), con respecto al observador, que las porciones laterales. Sin embargo, si dicha porción central es percibida como estando 'más lejos de' (Estimación-L), con respecto al mismo punto de referencia, que las dos porciones laterales, entonces la función de estimación se halla en relación inversa con el desfase.

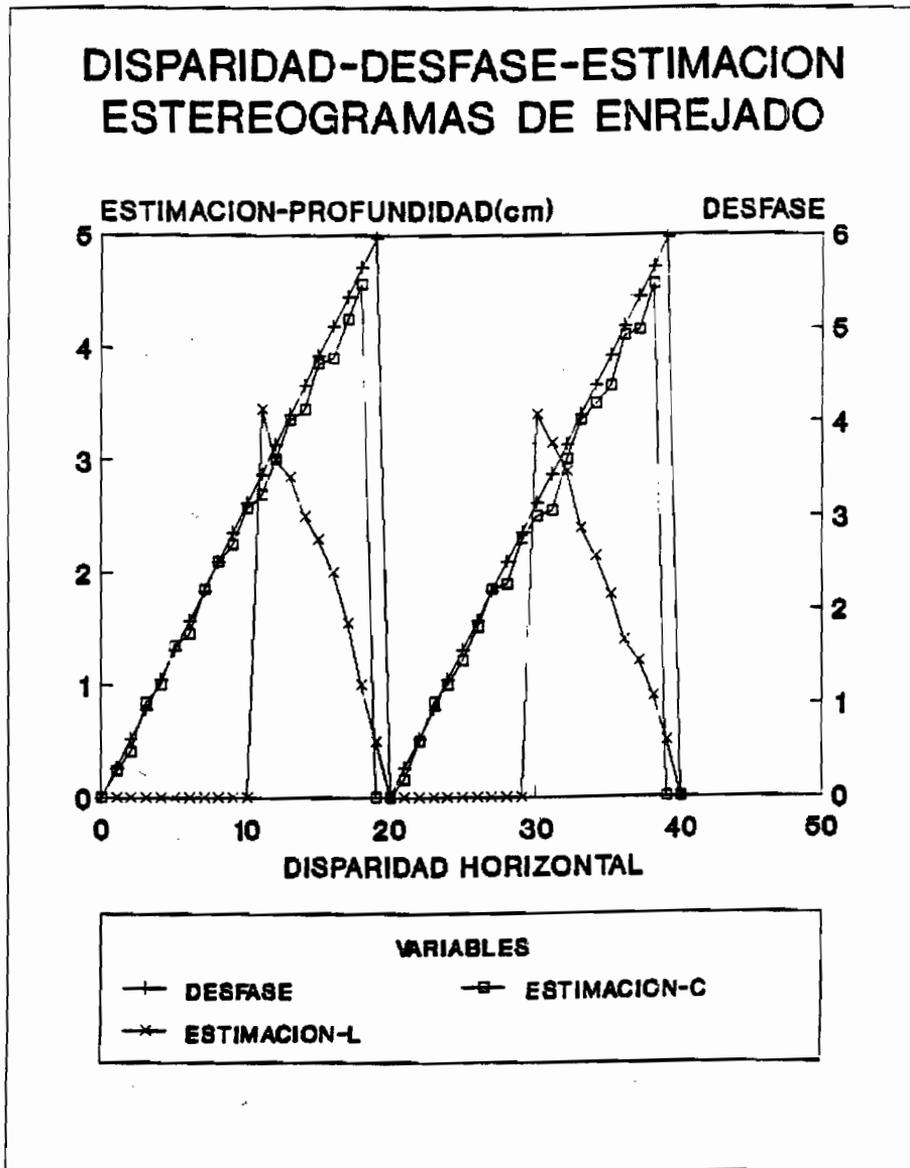


GRÁFICO 2.- Relación entre Desfase y Estimación de la profundidad del espacio (Estimación-C= cerca y Estimación-L= lejos), según la disparidad establecida en una serie de 40 estereogramas de enrejado.

## Influencia del espectro de fase en el procesamiento de imágenes digitalizadas

Para mostrar como contribuye la diferencia entre los espectros de fase de las TF de dos estereogramas, que forman par, utilizaremos una disparidad geométrica de 15 píxeles, es decir, un desfase de  $3\pi/2$  radianes ( $270^\circ$  sexag.).

Una vez obtenida la TF de cada uno de estos dos estereogramas, separamos los espectros de amplitud y los de fase. La diferencia entre los espectros de amplitud de estas dos imágenes, como era de esperar, es nula. Sin embargo, los espectros de fase son distintos, como puede observarse al visualizar la Figura-3.

En ella, como en todos los espectros de fase del tipo de estereogramas aquí utilizados, presentan la apariencia de un enrejado oblicuo con forma de onda 'en diente de sierra'. Sin embargo, al compararlos observaremos que, si bien el enrejado oblicuo (de las imágenes de los espectros de fase) es prácticamente igual en los dos espectros (y también para diferentes disparidades), difieren en la línea horizontal central del plano de Fourier, la cual contiene información sobre la periodicidad de la variación de luminancia del estereograma y, por tanto, del grado de asimetría que produce la disparidad establecida en cada estímulo.

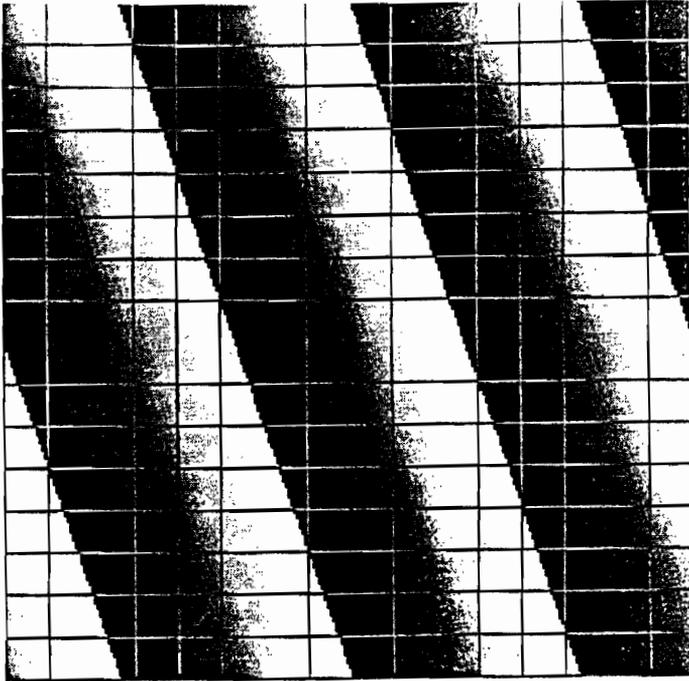
En general, cuando la  $D_g = 0$  píxeles, la línea horizontal central del espectro de fase y, consiguientemente, la porción central del estereograma (enrejado vertical) es estrictamente antisimétrica respecto a un eje vertical que pase por el centro geométrico del enrejado central. Es decir, si a cierta distancia de dicho eje, y a su izquierda, hay una franja blanca, a la misma distancia, pero hacia la derecha, habrá una franja negra. Cuando la  $D_g = 10$  píxeles, también existe antisimetría (tanto en la línea horizontal central del espectro de fase como en el dominio espacial del enrejado), pero la complementaria, ya que un desplazamiento horizontal de medio ciclo de este enrejado da lugar a que en la localización donde exista una franja blanca en él, habrá una franja negra en el de  $D_g = 0$  pix. y viceversa. Las disparidades con valores comprendidos en los intervalos 1-4 pix. y 6-9 pix. presentan dicha línea horizontal central del espectro de fase como reflejando los diversos grados de asimetría de la porción central (enrejado vertical). Mientras que las disparidades comprendidas en los intervalos 11-14 pix. y 16-19 pix. presentan asimetrías complementarias de las anteriores, según la correspondencia siguiente: (1,11), (2,12), (3,13), (4,14), (6,16), (7,17), (8,18) y (9,19). Véase, por ejemplo, en la Figura-4 los espectros de fase correspondientes a estereogramas con disparidades de 2 y 12 píxeles.

Sin embargo, la línea horizontal central del espectro de fase no aparece cuando las disparidades son de 5 o de 15 píxeles. La razón es obvia, sólo en estos casos ( $\varphi_3 = \pi/2$  y  $3\pi/2$  radianes, respectivamente) la porción central del enrejado vertical es totalmente simétrica (véase en Figura-3 el espectro de fase correspondiente a una  $D_g = 15$  pix.)

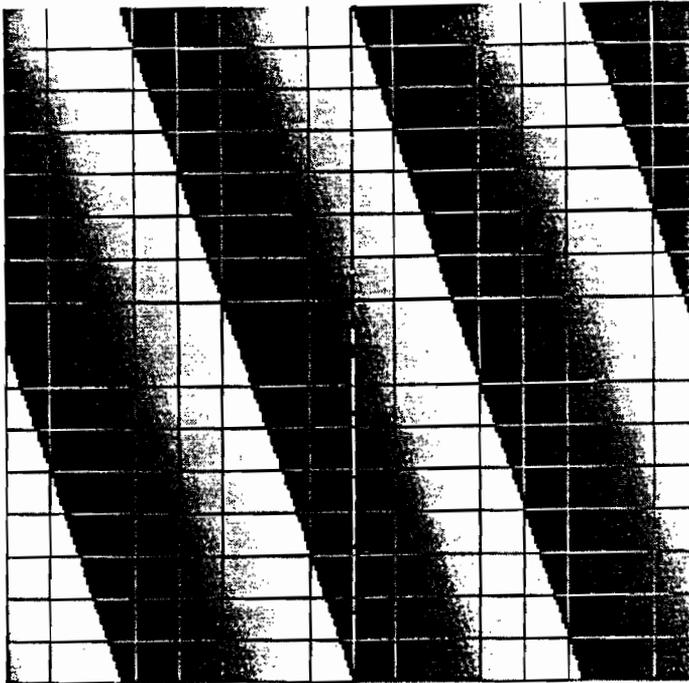
La Figura-5 muestra el espectro de amplitud de la correlación entre un par de estereogramas cuyo desfase es de  $3\pi/2$  radianes ( $270^\circ$  sexag. o  $D_g = 15$  píxeles), el cual explicita el desplazamiento relativo entre los dos estereogramas del par o, en nuestra opinión, describe la profundidad en niveles de gris.

Esta figura realza ciertos valores de gris en la porción central de la imagen, los cuales resultan más extremados (blanco o negro) que el gris promedio predominante en las porciones laterales. El hecho de que sean, precisamente, aquellos que son percibidos por los sujetos en profundidad sugiere que se trata de una especie de *mapa de profundidad*.

Finalmente, la Figura-6 contiene la imagen sintetizada a partir del espectro de amplitud del estereograma izquierdo ( $D_g = 0$  pix) y utilizando como espectro de fase la diferencia entre los espectros de fase de los dos estereogramas del par (Desfase =  $3\pi/2$  radianes o  $D_g = 0$  y 15 pix.).

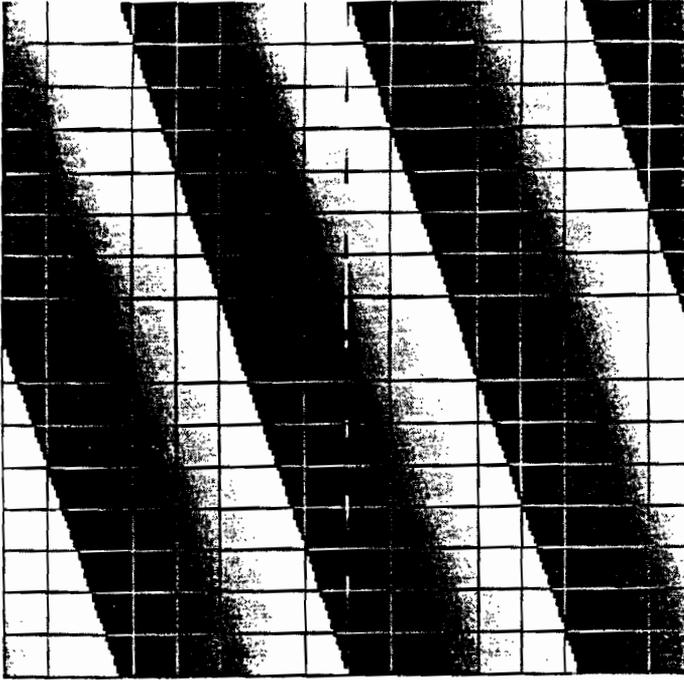


B

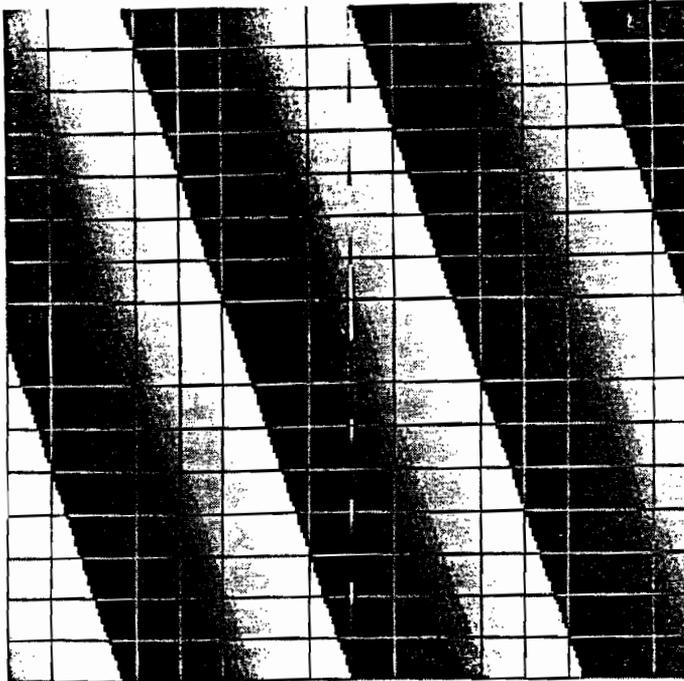


A

FIGURA 3.- Espectros de fase de dos estereogramas de enrejado. A) Estereograma con  $D_g = 0$  píxeles.  
B) Estereograma con  $D_g = 15$  píxeles.

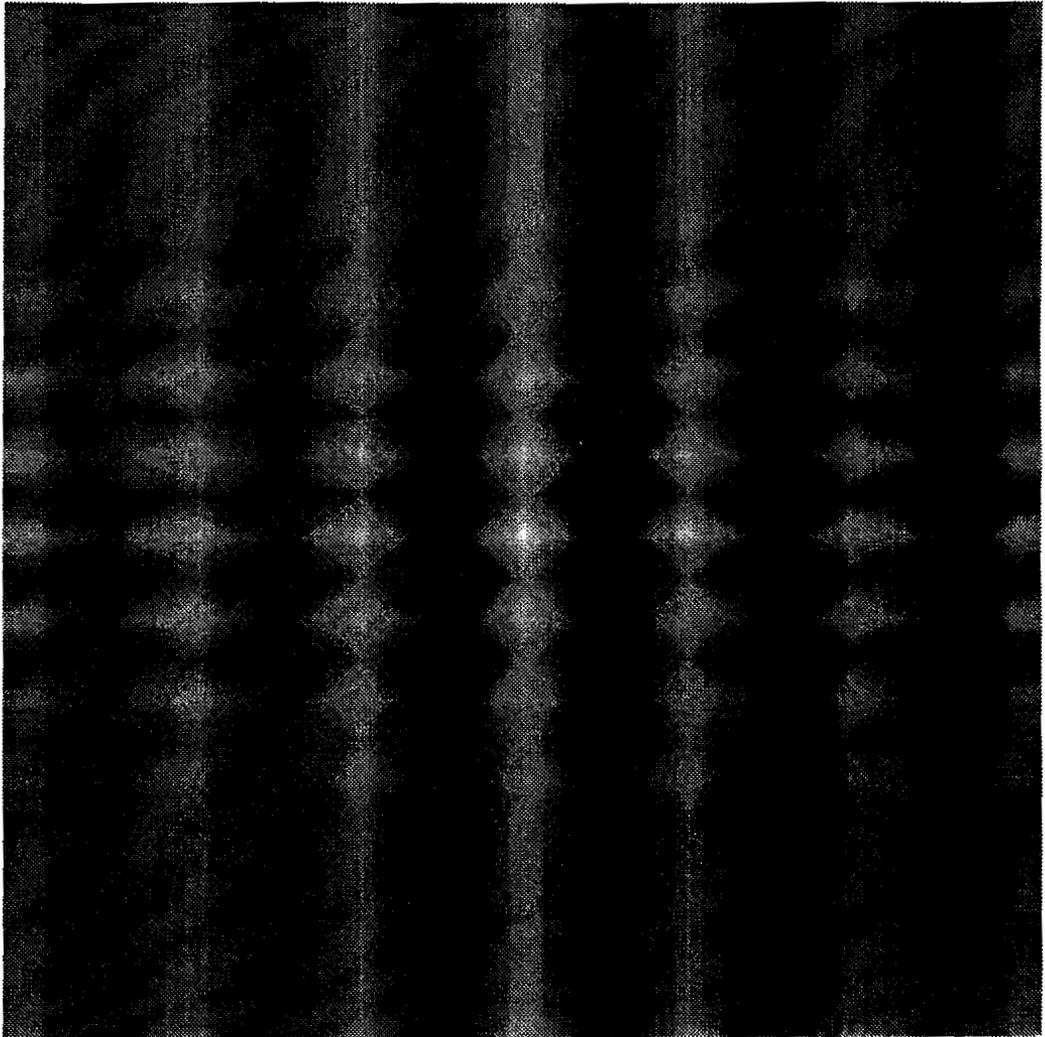


11

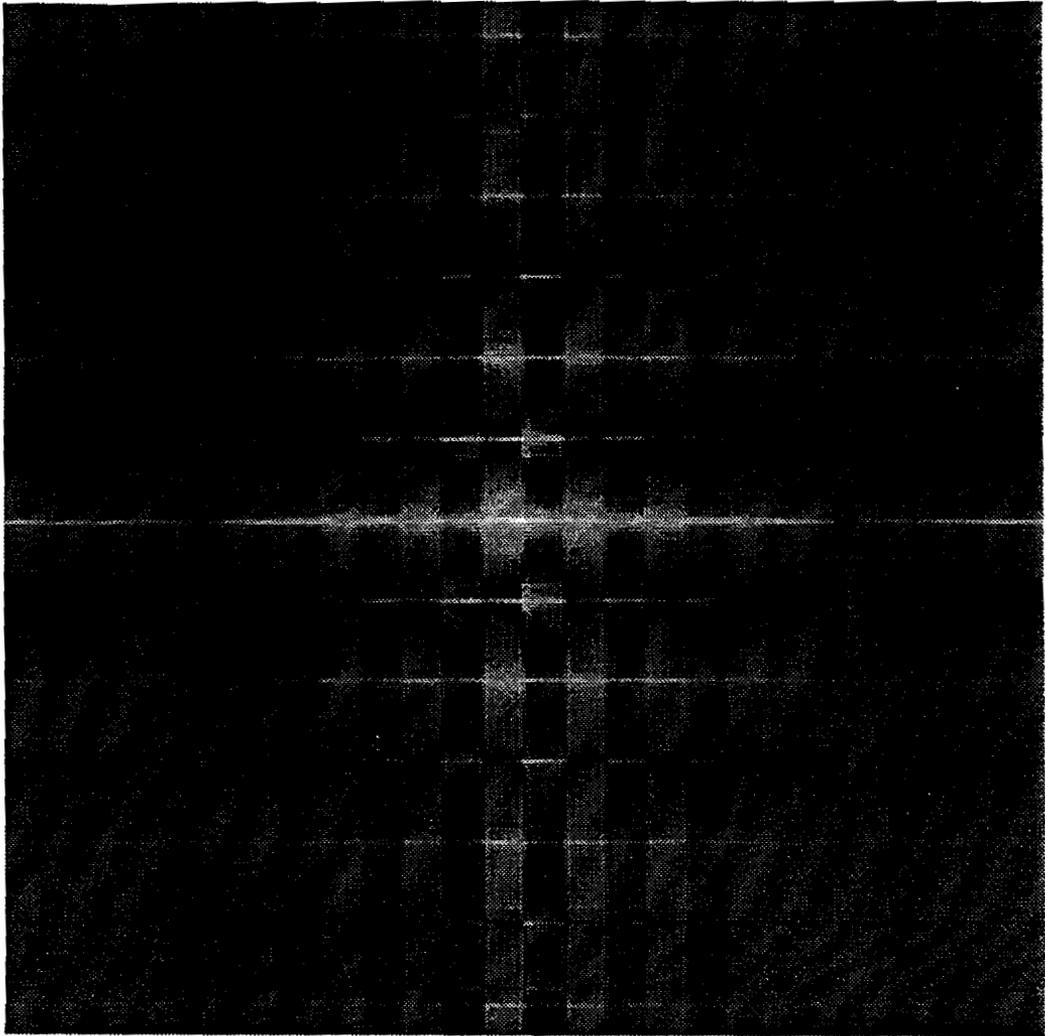


A

FIGURA 4.- Espectros de fase de dos estereogramas de enrejado. A) Estereograma con  $Dg= 2$  pixeles.  
B ) Estereograma con  $Dg= 12$  pixeles.



**FIGURA 5.-** Espectro de amplitud de la correlación entre dos estereogramas de enrejado que difieren ( $D_g$ ) 15 píxeles. 'Mapa de profundidad'.



**FIGURA 6.-** *Imagen sintetizada al combinar el espectro de amplitud del estereograma izquierdo ( $D_g = 0$  pix.) y utilizando como espectro de fase la diferencia entre los espectros de fase de un par estereo con  $Desfase = 3/2$  rad. o 15 pixeles.*

En ella se aprecia cierto realce de la porción central, frente a las laterales, resultando una imagen simétrica tanto horizontal como verticalmente.

## Discusión

En este trabajo se propone una metodología alternativa para la investigación de la esteropsis visual, recurriendo a los *estereogramas de enrejado*. Se ha descrito como pueden ser generados estos estímulos y sugerido las posibles variaciones que ofrece.

El análisis de las variables *disparidad geométrica*, *desfase* y *similitud*, muestra la existencia de una relación lineal entre las dos primeras y una relación periódica entre éstas y la similitud de la serie de 40 pares de estereogramas de enrejado estudiados.

El experimento psicofísico llevado a cabo apoya la idea de una relación casi perfecta entre la estimación de la profundidad espacial subjetiva y el desfase. Además, nos ha mostrado como pueden lograrse dos percepciones posibles (sucesivas, no simultáneas) de la porción central (delante y detrás de las porciones laterales), siendo el determinante de éstas, al parecer, una combinación de procesos cognitivos internos con factores estimulares dependientes de las diferencias de fase del par estereoscópico. Como consecuencia de este experimento, se pone de manifiesto que la profundidad percibida depende, fundamentalmente de la fase (más concretamente, de la diferencia fásica), la cual hemos visto que es directamente proporcional a la  $Dg$  e inversamente al período  $1/\rho$  (o si se prefiere a  $\lambda$ ), dado que la fase expresa la razón  $Dg/\lambda$ . Por ello, es posible experimentar la misma sensación de profundidad ante estereogramas que presenten igual cociente ( $Dg/\lambda$ ).

La aplicación del análisis de Fourier al tipo de estímulos que elicitan la esteropsis visual, nos ha permitido acercarnos con mayor detalle al estudio de la estructura espacial contenida en cada par de estereogramas. La TF de éstos ha puesto de manifiesto como la información relevante para el logro del fenómeno estereoscópico no radica en los espectros de amplitud, que son idénticos, sino en los espectros de fase. Estas diferencias en el espectro de fase han sido evidenciadas al comparar la línea horizontal central de los espectros de fase de los estereogramas, la cual contiene información acerca de las diferencias en el grado de simetría, antisimetría o asimetría de las porciones centrales del enrejado vertical. Además, por tratarse aquí de estímulos periódicos se ponen de manifiesto las regularidades cíclicas que la variación de fase en el dominio espacial produce en el espectro de fase del dominio de Fourier. Naturalmente, la diferencia entre los espectros de fase de un par de estereogramas de enrejado deberá indicar el grado de concordancia/discrepancia de las 'metrías' (simetría, asimetría, antisimetría) y, en consecuencia, el grado de desplazamiento relativo entre ambos estereogramas. Lo que, a su vez, en el *dominio cognitivo* debe ser experimentado como 'sensación' de profundidad.

La estrategia de síntesis de imágenes, a partir de los espectros de amplitud y fase de las dos imágenes del par de estereogramas (o alguna combinación de ellos) explicita la contribución específica de cada uno de estos espectros a la estructura espacial de la imagen resultante.

El hallazgo que mayor interés nos merece es el relativo a la obtención de mapas de profundidades mediante la correlación de pares de estereogramas. Así como la obtención de un indicador del parecido o similitud existente entre ambos elementos del par, el cual nos viene dado por el máximo valor de pixel de la correlación. Puesto que estos mapas de profundidad se hallan expresados en niveles de gris (valores de luminancia), convendrá adentrarse en el estudio cuantitativo de las distribuciones de tales valores de luminancia.

Como se ha dicho, anteriormente, también será conveniente investigar los efectos, que tienen lugar sobre el fenómeno de la estereopsis, el hecho de variar conjuntamente la diferencia fásica y, bien la frecuencia espacial, bien la orientación, e incluso el contraste, de cada estereograma del par.

## Agradecimientos

Expresamos nuestro agradecimiento al 'Computer Vision Group' de la Unitat de Processament d'Imatges i d'Inteligència Artificial' (Departament d'Informàtica. Universitat Autònoma de Barcelona) por habernos facilitado el entorno de 'Tratamiento de imágenes' ViLi (Vision Lisp) por ellos elaborado.

También agradecemos al profesor Sierra-Vázquez sus precisiones y sugerencias tras la lectura del primer manuscrito. En cualquier caso, los errores que pudieran persistir son responsabilidad nuestra.

## Referencias

- Bracewell, R.N. (1978). *The Fourier Transform and its Applications*. N.Y.:MacGraw-Hill.
- Caelli, T. y Bevan, P. (1982). Visual sensitivity to two-dimensional spatial phase. *Journal of the Optical Society of America*, 72, pp. 1375-1381.
- Campbell, F.W. y Robson, J.G. (1968). Applications of Fourier analysis to the ability of grating. *Journal of Physiology*, 197, pp. 551-566.
- Daugman, J.C. (1984). Spatial visual channels in the Fourier plane. *Vision Research*, 24, pp. 891-910.
- Domingo, A. (1993). *Tratamiento digital de imágenes*. Madrid: Anaya Multimedia.
- González, R.C. y Wintz, P. (1977). *Digital Image Processing*. MA: Addison-Wesley, Reading.
- Graham, N. y Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: A comparison of single channel and multichannel models. *Vision Research*, 11, pp. 251-259.
- Homer, J.L. y Gianino, P.D. (1984). Phase only matched filtering. *Applied Optical*, 23, pp. 812-816.
- Julesz, B. (1960). Binocular Depth Perception of computer-generated patterns. *Bell System Technical Journal*, 39, pp. 1125-1162.
- Julesz, B. (1964). Binocular Depth Perception without Familiarity Cues. *Science* 145, pp. 356-362.
- Julesz, B. (1986). Stereoscopic vision. *Vision Research*, 9, pp. 1601-1612.
- Juvels, S.; Vallmitjana, A.; Carnicer, A. y Campos, J. (1991). The role of amplitude and phase of the Fourier transform in the digital image processing. *American Journal of Physics*, 59 (8), pp. 744-748.
- Maffei L. y Fiorentini, A. (1973). The visual cortex as a spatial frequency analyser. *Vision Research*, 13, pp. 1255-1267.
- Ogle, K.N. (1962). The Visual Space Sense. *Science*, Vol 135, Nº 3506, pp. 763-771.
- Oppenheim, A.V. y LIM, J.S. (1981). The importance of phase in signals. *Proceeding of the IEEE*, 69, pp. 529-541.
- Piotrowsky, L.N. y Campbell, F.W. (1982). A demonstration of the visual importance and flexibility of spatial frequency amplitude and phase. *Perception*, 115, pp.337-346.
- Sierra-Vázquez, V. (1992). Procesamiento visual inicial: aspectos biológicos, psicofísicos y computacionales del análisis espacial de imágenes por el sistema visual humano. En J. Mayor y J.L. Pinillos (Eds.): *Atención y percepción*. Madrid: Alhambra Universidad. pp. 163-312.
- Stevens, S.S. (1966). Quantifying the sensory experience. En K. Feferabend y G. Maxwell (Eds.): *Mind, Matter and Method*. Minnesota: University of Minnesota Press.