

Universidad Nacional de La Plata

Facultad de Ciencias Exactas

Departamento de Física



**Disminución de ruido y corrección del
desenfoque en imágenes de objetos 3D**

Tesis Doctoral

Patricia Laura Torroba

Agosto de 2004

Universidad Nacional de La Plata

Facultad de Ciencias Exactas

Departamento de Física



**Disminución de ruido y corrección del
desenfoque en imágenes de objetos 3D**

Tesis para optar al grado de Doctora en Física presentada por :

Patricia Laura Torroba

Director: Dr. Héctor Rabal

Codirector: Dr. Mario Garavaglia

Lugar de realización: Centro de Investigaciones Ópticas, CIOp

Agosto de 2004

*A mi hijo
y a mis sobrinos*

Índice

| | |
|--|----------|
| Índice..... | i |
| Agradecimientos..... | v |
| 1. Introducción general..... | 1 |
| 1.1 Objetivos..... | 1 |
| 1.2 Esquema general..... | 3 |
| 1.3 Bibliografía..... | 6 |
| 2 Disminución de ruido en imágenes aplicando “La media de orden fraccionario”..... | 7 |
| 2.1 Introducción..... | 7 |
| 2.2 La media de orden h..... | 9 |
| 2.3 La media de orden h fraccionario usando ventanas de 9 y 8 pixels y ruido sal y pimienta..... | 13 |
| 2.4 Criterios de calidad..... | 16 |
| 2.4.1 Criterios de calidad objetivos..... | 16 |
| 2.4.2 Criterio de calidad subjetivo..... | 17 |
| 2.5 Evaluación del efecto de la media de orden h fraccionario en la disminución de ruido sal y pimienta..... | 18 |
| 2.6 La media de orden h fraccionario usando ventanas de 9 y 8 pixels y ruido gaussiano..... | 21 |
| 2.6.1 Análisis de los bordes..... | 21 |
| 2.7 Evaluación de la media de orden h fraccionario en preservar bordes..... | 23 |
| 2.7.1 Promedio local..... | 24 |
| 2.7.2 Varianza local..... | 25 |
| 2.7.3 Raíz cuadrada del error cuadrático medio..... | 26 |

| | | |
|----------|---|-----------|
| | 2.7.4 Respuestas a líneas angostas..... | 27 |
| | 2.8 Evaluación psicofísica..... | 27 |
| | 2.9 Bibliografía..... | 28 |
| 3 | Respuesta en frecuencia de un sistema óptico desenfocado..... | 31 |
| | 3.1 Introducción..... | 31 |
| | 3.2 Análisis de Fourier en la teoría de formación de imágenes incoherentes..... | 32 |
| | 3.3 Función de Transferencia Optica de un sistema libre de aberraciones..... | 44 |
| | 3.4 Función de Transferencia Optica de un sistema desenfocado..... | 38 |
| | 3.5 Evaluación de la respuesta D(s)..... | 44 |
| | 3.6 Bibliografía..... | 48 |
| 4 | Corrección del desenfoque usando el criterio de máxima visibilidad..... | 51 |
| | 4.1 Introducción..... | 51 |
| | 4.2 Profundidad de enfoque y profundidad de campo..... | 53 |
| | 4.3 Criterios conocidos asociados al desenfoque..... | 59 |
| | 4.3.1 Contenido Estructural..... | 59 |
| | 4.3.2 Entropía..... | 60 |
| | 4.3.3 Entropía corregida..... | 61 |
| | 4.3.4 Criterio de extremo de intensidad..... | 61 |
| | 4.3.5 Varianza..... | 64 |
| | 4.4 Criterios propuestos asociados al desenfoque..... | 65 |
| | 4.4.1 Gradiente..... | 65 |
| | 4.4.2 Desvío medio aritmético local..... | 65 |
| | 4.4.3 Visibilidad..... | 66 |
| | 4.5 Experiencia..... | 67 |
| | 4.6 Análisis de los resultados..... | 67 |
| | 4.6.1 Entropía..... | 68 |
| | 4.6.2 Entropía corregida..... | 68 |
| | 4.6.3 Criterio de extremo de intensidad..... | 69 |
| | 4.6.4 Varianza..... | 70 |
| | 4.6.5 Gradiente..... | 70 |
| | 4.6.6 Desvío medio aritmético..... | 71 |

| | | |
|----------|---|------------|
| 4.6.7 | Visibilidad..... | 71 |
| 4.7 | Corrección del desenfoque usando el criterio de la visibilidad..... | 73 |
| 4.8 | Estudio teórico de la visibilidad en función del desenfoque para fuentes sinusoidales..... | 80 |
| 4.9 | Bibliografía..... | 84 |
| 5 | Corrección del desenfoque usando iluminación estructurada..... | 89 |
| 5.1 | Introducción..... | 89 |
| 5.2 | Esquema experimental..... | 91 |
| 5.3 | Experiencias utilizando objetos 2D..... | 92 |
| 5.4 | Estudio teórico de la visibilidad en función del desenfoque para un objeto sinusoidal doblemente desenfocado..... | 95 |
| 5.5 | Corrección del desenfoque usando imágenes de objetos 3D..... | 99 |
| 5.6 | Bibliografía..... | 103 |
| 6 | El modelo visual de Cornsweet y Yellott..... | 105 |
| 6.1 | Introducción | 105 |
| 6.2 | Modelo visual de Marr y Hidreth..... | 106 |
| 6.3 | Modelos visuales y operadores lineales..... | 109 |
| 6.3.1 | Motivación psicofísica..... | 110 |
| 6.3.2 | Ruido fotónico y operadores de volumen constante..... | 112 |
| 6.4 | Modelo visual de Cornsweet y Yellott..... | 115 |
| 6.4.1 | Formalismo del modelo..... | 116 |
| 6.5 | Comparación entre las respuestas del sistema visual y el modelo IDS..... | 120 |
| 6.5.1 | Bandas de Mach..... | 122 |
| 6.5.2 | Función de Sensibilidad al Contraste..... | 125 |
| 6.5.3 | Ley de Ricco..... | 130 |
| 6.5.4 | Ley de Weber..... | 131 |
| 6.5.5 | Adaptación..... | 132 |
| 6.6 | Bibliografía..... | 133 |
| 7 | El modelo de volumen aproximadamente constante..... | 137 |
| 7.1 | Introducción..... | 137 |
| 7.2 | Formalismo del modelo de volumen aproximadamente constante..... | 138 |

| | | |
|----------|---|------------|
| 7.3 | Las bandas de Mach y el modelo ACVO..... | 141 |
| 7.4 | Respuestas del ACVO a barras verticales..... | 145 |
| 7.5 | La Función de Sensibilidad al Contraste y el ACVO..... | 146 |
| 7.6 | La ley de Weber y el ACVO..... | 149 |
| 7.7 | La señal. Aplicación de las bandas de Mach..... | 151 |
| 7.8 | El valor de la señal para perfiles sinusoidales teóricos..... | 152 |
| 7.9 | Experiencia..... | 154 |
| 7.10 | Bibliografía..... | 157 |
| 8 | Corrección del desenfoque usando la matriz de Co-ocurrencia y la Función Estructura..... | 159 |
| 8.1 | Introducción..... | 159 |
| 8.2 | La matriz de Co-ocurrencia y su Momento de Inercia..... | 160 |
| 8.3 | El Momento de Inercia de la COM de perfiles de intensidades teóricos desenfocados..... | 163 |
| 8.4 | La Función Estructura..... | 165 |
| 8.5 | Experiencia..... | 166 |
| | 8.5.1 Objetos 2D..... | 167 |
| | 8.5.2 Objetos 3D..... | 168 |
| 8.9 | Bibliografía..... | 171 |
| 9 | Conclusiones y perspectivas..... | 173 |
| | Bibliografía..... | 183 |
| | Lista de publicaciones..... | 187 |

Agradecimientos

En primer lugar deseo expresar mi agradecimiento al Dr. Héctor Rabal por la dirección de esta Tesis. Su orientación y apoyo permanentes fueron fundamentales para el desarrollo de mi trabajo. También quiero agradecerle al Dr. Mario Garavaglia por sus enseñanzas y su colaboración en todos los aspectos.

Al Dr. Mario Gallardo, Director del Centro de Investigaciones Ópticas (CIOp), le agradezco la oportunidad de realizar mi trabajo de Tesis en la Institución.

Quiero agradecer muy especialmente a Lili Cap por las horas de trabajo compartido, durante las cuales surgió una gran amistad.

Asimismo, mi agradecimiento a todo el personal del Ciop: Investigadores, Profesionales, Becarios, Técnicos, Administrativos y de Servicios, con quienes compartí gratos momentos alrededor de una taza de café.

Por último, quiero agradecer al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) por otorgarme una Beca Doctoral y a la Facultad de Ingeniería de la Universidad Nacional de La Plata por la adjudicación de una Semidedicación a la Investigación con lugar de trabajo en el Centro de Investigaciones Ópticas.

La Plata agosto de 2004

Capítulo 1

Introducción General

1.1 Objetivos

En los sistemas formadores de imágenes hay muchas causas de la degradación de la misma. Los tipos de degradación que afectan sólo el nivel de gris del pixel, sin introducir un borrono espacial son denominados degradaciones puntuales. Otros tipos, los cuales introducen borrono, son llamados degradaciones espaciales. Existen además otras clases, relacionadas con efectos temporales y cromáticos.

Estas degradaciones aparecen en una gran variedad de aplicaciones tales como en el reconocimiento astronómico, donde las imágenes obtenidas están deterioradas por turbulencia atmosférica, en imágenes tomadas con microscopios electrónicos, debido al tipo de muestras utilizadas , aberraciones del sistema óptico y el movimiento entre la cámara y el objeto. Otro ejemplo son las imágenes radiográficas médicas que presentan

baja resolución y contraste, debido a los sistemas formadores de imágenes que utilizan rayos-x.

En este trabajo se propone estudiar dos características ligadas a la formación de imágenes, el ruido como ejemplo de degradaciones puntuales y el desenfoque en escenas de objetos de 3D como ejemplo de degradaciones espaciales y cómo corregirlos.

Inicialmente investigamos el problema de disminuir el ruido en imágenes aplicando una serie de filtros no lineales. Esto involucra tener que formular criterios de calidad que den alguna estimación óptima del resultado deseado. Además, cuando las imágenes son observadas por humanos, el conocimiento del sistema visual puede ser muy útil para determinar la calidad subjetiva de una imagen o la fidelidad de la misma con respecto a una escena original. Entonces, primero analizamos criterios objetivos relacionados con medidas de magnitudes físicas y segundo criterios subjetivos que tienen en cuenta el sistema visual humano.

Por otro lado, el sistema visual parece no darse cuenta ni del ruido ni del desenfoque. Entonces, nos interesó analizar cómo resuelve la naturaleza a través del sistema visual ambos defectos. En este contexto estudiamos el modelo visual de Marr [1.1] y el modelo de Cornsweet y Yellott [1.2].

Encontramos que el modelo de Cornsweet contiene un tratamiento del ruido y su resolución. Luego, nos pareció interesante analizar si este modelo se podía aplicar para corregir el desenfoque, ya que es poco o nada lo que se conoce sobre el mecanismo de control del sistema visual humano para el ajuste del enfoque visual.

Para estudiar el desenfoque y cómo corregirlo fue necesario describir matemáticamente como actúa físicamente este fenómeno inevitable. A continuación, ensayamos varios criterios para corregir dicho defecto. Este análisis nos condujo a introducir franjas proyectadas sobre la imagen y este hecho al doble desenfoque. De ésta manera encontramos que si se tenía en cuenta el efecto del tamaño finito de los

fotorreceptores en el modelo de Cornsweet, algunas experiencias como las Bandas de Mach [1.3], se podían explicar mejor.

Cuando esta modificación se tuvo en cuenta, las Bandas de Mach resultaron más realistas pero se perdía la ley de Weber [1.4]. Propusimos entonces, un cambio para recuperar esta ley fundamental. Mediante esta última innovación se pudo asignar una medida numérica a la respuesta de este modelo modificado y las Bandas de Mach pudieron ser utilizadas como indicadores del desenfoque. De este modo, el mismo modelo servía a la vez para adaptarse al ruido y describir el desenfoque.

Por último, encontramos otra herramienta de cálculo, la función estructura, como criterio para corregir el desenfoque. Este parámetro produjo resultados de similar calidad con los anteriores con un costo computacional muy inferior.

1.2 Esquema general

Esta Tesis consta de nueve capítulos cuyos lineamientos generales se presentan a continuación.

El capítulo 2 trata el problema de eliminar ruido en imágenes utilizando operaciones de filtrado no lineales y una lineal que se obtienen a partir de una expresión matemática llamada “Media matemática generalizada de orden h ”. Inicialmente se define esta expresión y se analizan algunos casos particulares para distintos valores del parámetro h . Luego, para tener una idea del efecto que ellos producen sobre las diferentes escalas de los detalles de la imagen, se los aplica a sinusoides de diferentes frecuencias que contienen ruido sal y pimienta. A continuación, se evalúa el comportamiento de estos filtros con respecto a sus capacidades para eliminar ruido utilizando distintos criterios de calidad objetivos y otro subjetivo. También se estudia la habilidad que tienen estos filtros en disminuir ruido y preservar bordes cuando se los aplica sobre imágenes que contienen ruido gaussiano. En la última parte de este capítulo se realiza una evaluación visual de la

calidad de las imágenes filtradas y se comparan los resultados de aplicar un criterio objetivo con el subjetivo.

En los siguientes capítulos empezamos a estudiar el efecto del desenfoque sobre imágenes planas y sobre imágenes de objetos 3D iluminados incoherentemente. Analizamos por un lado, como se desenfoca, que efectos se producen y por otro como modificarlos.

En el capítulo 3 comenzamos con el estudio teórico del efecto del desenfoque en imágenes iluminadas incoherentemente. Para hacer un estudio teórico de imágenes que han sido degradadas por desenfoque es esencial conocer el análisis de Fourier en la teoría de formación de imágenes. En este capítulo se desarrolla la parte necesaria de este análisis, para obtener la respuesta en frecuencia de un sistema óptico desenfocado iluminado incoherentemente. Esta respuesta debe ser especificada para describir completamente el sistema formador de imágenes y posteriormente obtener la distribución de intensidad en el plano imagen de objetos que han sufrido desenfoque. En la última parte se evalúa la función respuesta en el caso particular de una pupila rectangular.

El objetivo central del capítulo 4 es estudiar teórica y experimentalmente criterios que determinen la posición del plano mejor enfocado dentro de un conjunto de imágenes planas con desenfoque progresivo. Para un mejor entendimiento del tema, primero se definen algunas magnitudes que están frecuentemente asociadas al desenfoque tales como: profundidad de campo, f-número, apertura angular y se analizan algunos criterios conocidos: el contenido estructural, la entropía, la entropía corregida, el criterio de extremo de la intensidad y la varianza. A continuación, se ensayan criterios nuevos propuestos por nosotros, el gradiente, el desvío medio aritmético local y la visibilidad. El contenido estructural y el criterio de extremo de la intensidad, se investigan desde el punto de vista teórico aplicando el análisis de Fourier desarrollado en el capítulo tres. Luego, se realizan una serie de experiencias en las que se aplican los criterios estudiados y se comparan sus comportamientos en función del desenfoque. A partir de los resultados, se elige como criterio alternativo original el de máxima visibilidad y se aplica en imágenes de objetos 3D

para construir una imagen en la cual el desenfoque está parcialmente corregido.

En el capítulo 5 se propone introducir iluminación estructurada sobre la escena para mejorar la calidad de la imagen resultante en donde el desenfoque está parcialmente corregido. Para ello, se realizan una serie de experiencias en las que proyectamos franjas sobre los objetos bajo estudio para incorporar la iluminación estructurada, produciéndose en el arreglo experimental un efecto de doble desenfoque. A continuación, para realzar aun más las componentes de alta frecuencia le aplicamos a dichas imágenes el operador derivada. Luego, desarrollamos un formalismo teórico que analiza el comportamiento de la visibilidad en función del desenfoque para una fuente sinusoidal doblemente desenfocada. Por último, trabajamos con imágenes de objetos 3-D con franjas proyectadas derivadas y sin derivar y corregimos el desenfoque aplicando el criterio de máxima visibilidad teniendo en cuenta las imágenes obtenidas con iluminación común.

Cuando se estudia un problema de imágenes suele ser muy útil fijarse cómo lo solucionó la naturaleza a través del sistema visual. Entonces, investigamos el sistema visual humano para saber cómo hace éste para solucionar los defectos causados por el ruido y el desenfoque. En el capítulo 6 se desarrollan una serie de conceptos previos relacionados con modelos visuales para que el lector se sitúe y pueda entender lo que sigue. Principalmente desarrollamos el modelo visual inventado por Cornsweet y Yellott conocido como “Suma espacial dependiente de la intensidad” (IDS). Este modelo propone una manera en que interactúan las células en la retina. Este mecanismo es simple y predice muchas características de la visión humana.

En el capítulo 7 aplicamos el modelo IDS, teniendo en cuenta en él una modificación, de manera tal de obtener una medida que indique las zonas mejor enfocadas dentro de un conjunto de imágenes con desenfoque progresivo. Comenzamos con el desarrollo del formalismo del modelo modificado, llamado por nosotros Modelo de Volumen Aproximadamente Constante (ACVO). A continuación, lo aplicamos a imágenes de entrada unidimensionales : rampa de luminancia, barras verticales y a un conjunto sinusoides de diferentes frecuencias y estudiamos las repuestas del modelo ACVO a dichas

entradas. Luego, definimos una medida que está asociada a la magnitud del realce de bordes introducida por el modelo y la aplicamos como método para corregir el desenfoque. En este capítulo, también hacemos un análisis teórico que avala los resultados experimentales.

En el capítulo 8 presentamos otras dos aproximaciones que hemos desarrollado de la corrección del desenfoque en un conjunto de imágenes de escenas 3D, usando el momento de Inercia de la matriz de Co-ocurrencia y la Función Estructura. Estos criterios se aplican a simulaciones por computadora, a imágenes de objetos planos y a objetos 3D para construir una imagen en donde el desenfoque está parcialmente corregido.

Finalmente en el Capítulo 9 se sintetizan los resultados obtenidos y paralelamente se plantean algunos desarrollos que darían continuidad a la presente investigación.

1.3 Bibliografía

- [1.1] Marr, D. Y. and Hildreth, E.: Theory of edge detection. Proc. R. Soc. London Ser. **B 207**, 187-217, (1980).
- [1.2] T. Cornsweet and J. Yellott: Intensity Dependent spatial summation. J. Opt. Soc. Am, **2**, 1769-1786, (1985).
- [1.3] F. Ratliff. Contour and contrast. Scien. Am. **226**, 6 , 91-101, (1972).
- [1.4] S. Hecht: The visual discrimination of intensity and the Weber-Fechner law. J. Gen. Physiol. **7**, 235-265, (1924).

Capítulo 2

Disminución de ruido en imágenes aplicando “La media de orden fraccionario”

2.1 Introducción

El problema de eliminar ruido en imágenes ha dado origen a operaciones de filtrado, lineales y no lineales que presentan diferente comportamiento en preservar bordes, reducir ruido y en pérdida de resolución. Por medio de muchos estudios experimentales se ha demostrado que aunque los filtros lineales poseen buena capacidad para disminuir el ruido [2.1-2.2], difunden los bordes y atenúan las líneas delgadas presentes en la imagen original debido a que realizan la operación lineal del promediado. Por otro lado, el filtrado no lineal es un método conocido, muy bueno para eliminar ruido y preservar bordes. Esta clase de filtro se ha hecho popular y ha alcanzado algunos resultados interesantes en muchas aplicaciones en el procesamiento de imágenes.

El filtro mediana, utilizado por primera vez por Tukey para suavizado de datos, es un conocido filtro no lineal que ha demostrado ser muy eficiente en reducir ruido impulsivo, en preservar bordes y es computacionalmente muy rápido [2.3]. Por eso, ya está incluido en muchos softwares comerciales.

Debido a que el filtro mediana tiene como único parámetro libre el tamaño de la ventana, no es muy fácil controlar el grado de suavizado que este produce. Por otro lado, la media aritmética es también útil en reducir ruido pero tiene como consecuencia pérdida de resolución y suavizado de bordes [2.4]. Ambos efectos pueden ser combinados en otros filtros y un estudio comparativo de ellos se puede encontrar en la Referencia [2.5].

En general, los filtros muestran una eficiencia óptima solamente para una determinada clase de ruido de una imagen específica. Para aplicaciones futuras de dichos filtros en el procesamiento de imágenes, es conveniente rever y analizar todos estos filtros y evaluar objetivamente su comportamiento en cuanto a su capacidad de disminuir el ruido como en sus habilidades en preservar bordes.

La media aritmética, la mediana y la moda son tres estadísticas características que pueden ser expresadas como casos particulares de una misma expresión matemática llamada: “Media matemática generalizada de orden h ” [2.6]. El orden h es usualmente un número entero, generalmente pequeño por ejemplo, $h=0$, $h=1$, $h=2$. La posibilidad de variarlo de manera continua no ha sido, que nosotros sepamos, estudiada y si bien el desarrollo matemático contenido en la demostración teórica es complejo, podría esperarse que valores reales de h en algunos intervalos tuvieran un comportamiento intermedio al de sus extremos. El único estudio teórico que se ha encontrado en la bibliografía es el de la mediana [2.3].

Uno de los propósitos de esta Tesis es disminuir el ruido en las imágenes y para ello, proponemos en este capítulo que el orden h puede ser usado como un parámetro adicional para obtener una familia de diferentes filtros no lineales (con la excepción de la media aritmética, la cual es lineal) cuyo efecto esperamos que se encuentre, para valores de h

adecuadamente elegidos, entre los obtenidos por la mediana y la media aritmética; y entre la moda y la mediana [2.7]. Ensayaremos entonces de manera numérica las propiedades de estos filtros relacionadas con la eliminación de ruido utilizando ventanas de 9 y 8 pixels. Primero, aplicamos esta familia de filtros en imágenes que contienen ruido sal y pimienta y a continuación sobre imágenes con ruido gaussiano. Para evaluar el efecto de los mismos calculamos varios criterios de calidad objetivos relacionados con medidas de magnitudes físicas y proponemos un criterio original subjetivo asociado con un el sistema visual humano. Además, analizamos la capacidad que tienen estos filtros en preservar bordes y líneas angostas usando estadísticas locales. Finalmente, realizamos una evaluación visual preliminar de las imágenes filtradas.

2.2 La media de orden h .

Vamos a comenzar definiendo la expresión matemática de la media generalizada de orden h .

Sea $\{a_i\}$ un conjunto de N números positivos, en nuestro caso los valores de intensidad de una imagen que se hallan dentro de una dada ventana y p_i pesos positivos tal que su suma es 1. Entonces, se puede demostrar que el valor x_h (para un conjunto particular de valores de h) que minimiza la expresión:

$$M_h(x) = \sqrt[h]{\sum_{i=1}^{i=N} p_i |x - a_i|^h} \quad (2.1)$$

es la moda, la mediana, la media aritmética, o la media de los valores extremos. Cuando todos los p_i son iguales ($p_i = 1/N$) y $h = 0$, x_0 es la moda; para $h = 1$, x_1 es la mediana; para $h = 2$, x_2 es la media aritmética y para $h \rightarrow \infty$, es la media de los valores extremos de los valores del conjunto. Estos casos particulares pueden ser rigurosamente demostrados [2.6].

Con el objeto de entender mejor como actúa el procedimiento empleado para hallar la media de orden h y extenderlo a valores reales, proponemos efectuar una transformación lineal a la ecuación (2.1). Para ello consideramos:

Primero, sea el conjunto ordenado a_i

$$\begin{aligned} \{a_i\} &= \{a_0, a_1, \dots, a_n\} \quad \text{tal que :} \\ a_0 &\leq a_1 \leq a_2 \leq \dots \leq a_n \end{aligned} \quad (2.2)$$

Luego, si se aplica la transformación lineal:

$$z_i = \frac{a_i - a_0}{a_n - a_0} \quad (2.3)$$

a la ecuación (2.1), se obtiene una expresión análoga dada por:

$$\tilde{M}_h(z) = \sqrt[h]{\sum_{i=1}^{i=N} |z - z_i|^h} \quad (2.4)$$

donde, el conjunto de valores z_i está en el intervalo $[0,1]$ y todos los p_i fueron elegidos iguales a $\frac{1}{N}$. Esta última ecuación es minimizada por un valor z_i , similarmente factoreado y desplazado con respecto a la ubicación del valor de x_h que minimiza el resultado en la ecuación (2.1). Por otra parte, si $h \geq 0$ el valor que minimiza $\tilde{M}_h(z)$ es también el mismo que lo hace con la expresión subradical y por lo tanto, la raíz no necesita ser calculada. Luego, para $h \geq 0$ los valores de z que minimizan la ecuación (2.4) están también en el intervalo $[0,1]$ de modo que, si por simplicidad se define la variable $\varepsilon_i = |z - z_i|$, ésta va a estar en el mismo intervalo. La Figura 2.1 muestra el gráfico de ε^h en función de ε para distintos valores del parámetro h .

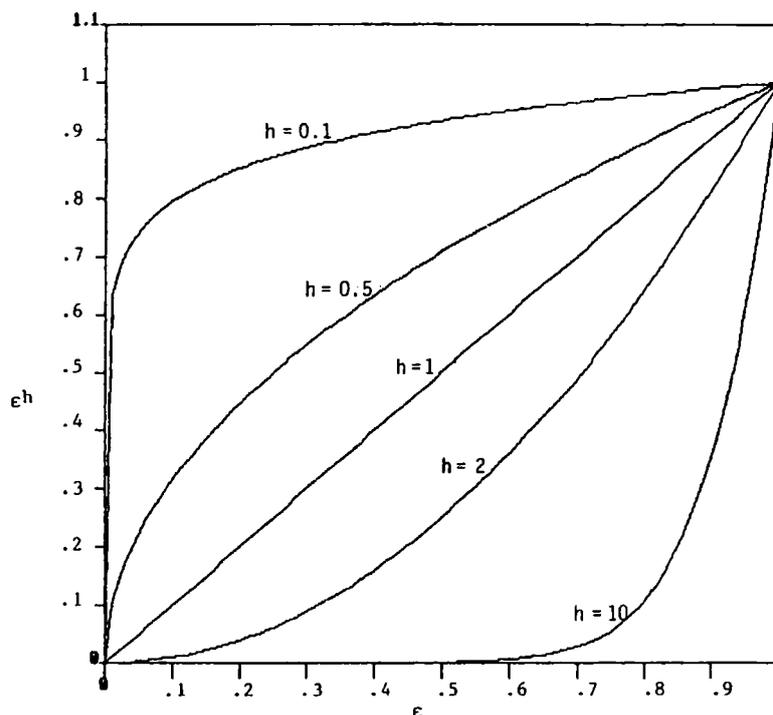


Figura 2.1 ε^h en función de ε para distintos valores del parámetro h

Es inmediato que para $h \rightarrow 0$ el valor de ε que menos contribuye a la expresión (2.4) es $\varepsilon = 0$. Si algunos valores a_i se repiten, el más frecuente será aquel que contribuya con el mayor número de ceros a la ecuación (2.4). Por lo tanto para $h \rightarrow 0$, el valor a_i más frecuente (la moda) minimiza la ecuación (2.4).

Por otro lado, para $h \rightarrow \infty$, los únicos valores de ε que no pueden ser despreciados son aquellos que tienen un valor cercano a 1. Pero, cada valor de z contribuirá con dos de tales términos a la expresión (2.4), uno debido a z^h y otro de la forma $|z-1|^h$. La suma de ambos términos se muestra en la figura 2.2. Es evidente que $z = \frac{1}{2}$ minimiza la suma de ambas contribuciones y este valor se transforma utilizando la ecuación (2.3) en $a_i = \frac{1}{2}(a_n + a_0)$.

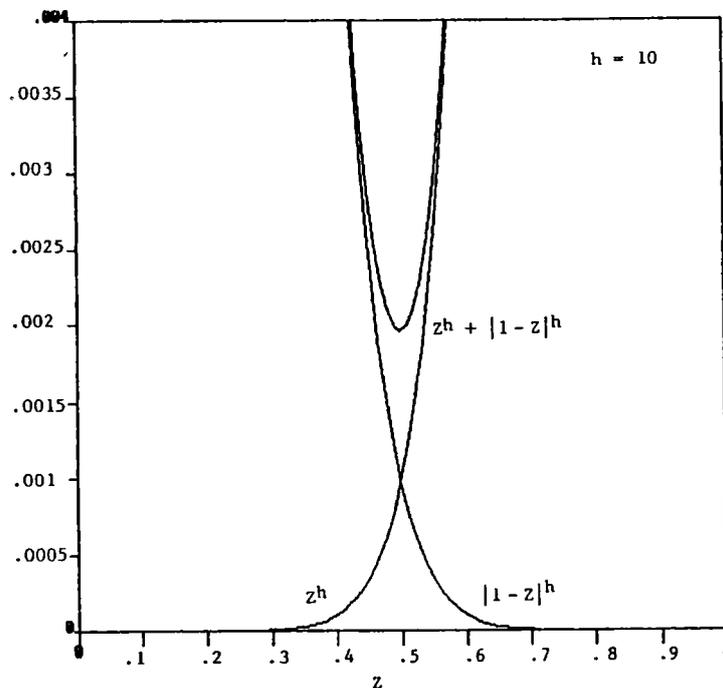


Figura 2.2 $z^h + |1-z|^h$ para altos valores de h .

En el caso $h = 1$, se considerará por el momento sólo la contribución de a_0 y de a_n en la ecuación (2.1) o en su análoga (2.4). En la Figura 2.3 se muestra la contribución de $|z|$ y de $|z-1|$ separadamente y el hecho de que su suma es 1. Esto significa que no interesa cual sea el valor de z que se elija para que su contribución a (2.1) o a (2.4) sea 1. Por lo tanto, no tendremos en cuenta los valores de a_0 y de a_n ya que no cambian el valor de x o de z que minimiza la ecuación (2.1) o la (2.4). Se puede iterar con el mismo resultado este razonamiento con los nuevos extremos a_1 y a_{n-1} y continuar hasta que quede sólo un elemento si n es impar o dos si es par. El valor que minimiza (2.1) y (2.4) es la mediana del conjunto original, lo cual es obvio en el primer caso (n impar) y se puede demostrar en el segundo (n par). Entre estos valores extremos de h existe un continuo de órdenes fraccionarios y enteros, donde la importancia del elemento central en el conjunto elegido, se corre hacia los extremos a medida que el orden h se incrementa.

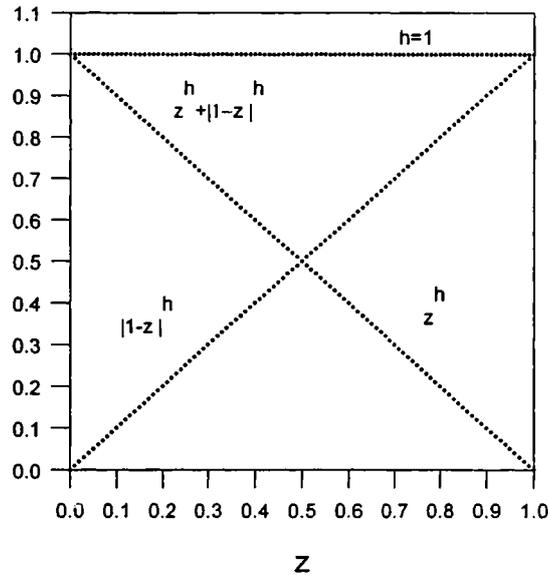


Figura 2.3. $z^h + |1-z|^h$ para $h=1$.

Entonces, estudiamos el efecto de los filtros de la media de orden fraccionario en la eliminación de ruido, para los valores de h comprendidos en el intervalo $[0,1]$, entre la moda y la mediana; y en el intervalo $[1,2]$ entre la mediana y la media aritmética, correspondientes a las curvas que están en la figura 2.1. Estos filtros son casos particulares de la familia de los llamados filtros del tipo M [2.5]

2.3 La media de orden h fraccionario usando ventanas de 9 y 8 pixels y ruido sal y pimienta

Generamos, como imágenes de ensayo, sinusoides con ruido a las cuales les aplicamos los filtros con distintos valores de h y calculamos algunos criterios de calidad. Comenzaremos, estudiando el comportamiento de estos filtros usando ruido sal y pimienta. Luego, generaremos imágenes similares a las de la referencia [2.3] con ruido gaussiano y seguiremos un procedimiento cuyas líneas generales se parecen a las de dicho trabajo aunque con variantes propias (el uso de la ventana de 8 pixels, el criterio subjetivo) para que los resultados puedan ser comparados.

Como los filtros son no lineales, no es posible definir una función de transferencia. Sin embargo, se puede tener una idea del efecto que ellos producen sobre las diferentes escalas de los detalles de la imagen si se los aplica a sinusoides de diferentes frecuencias. Por lo tanto, se generaron imágenes de 128×128 pixels, que contienen 256 niveles de gris y que consisten en sinusoides de diferentes frecuencias. A continuación se les agregó ruido sal y pimienta de la siguiente manera:

- a) Se utilizó para ello un generador de números aleatorios para obtener una distribución uniforme de la posición del ruido.
- b) Se reemplazó en los lugares aleatoriamente elegidos el valor del pixel por un valor preestablecido más alto (sal) o uno más bajo (pimienta), para un dado porcentaje de ruido. Luego, sobre estas imágenes se calculó la media de orden h , utilizando ventanas de 3×3 pixels, y se reemplazó el valor del pixel central por el valor asignado. Se trabajó sobre sinusoides de frecuencias: 0, 0.015, 0.03, 0.06, 0.08, 1.2, 0.15 ciclos/pixel. En la Figura 2.4 se muestra una ilustración del efecto que tienen estos filtros para diferentes valores de h .

Además, calculamos otro conjunto de filtros también propuesto por nosotros, utilizando la media de orden h , aplicado a ventanas de 3×3 pixels sin incluir en el cálculo el valor del pixel central. Nos referiremos a estos filtros como *ventanas de 8 pixels* y procesan la imagen de la siguiente forma:

- a) Calculan la media de orden h usando una ventana de 8 pixels.
- b) Comparan el valor calculado en a) con el valor del pixel central de la ventana
- c) Si la diferencia obtenida en b) es mayor que un dado porcentaje, entonces reemplazan el pixel sin procesar por el calculado. En caso contrario, el pixel central queda sin modificarse.

Esta operación está basada en la suposición de que si el valor del pixel central es muy diferente del obtenido con la media de orden h , es probable que dicho valor central se deba a la presencia de ruido. El umbral fue en la mayoría de los casos del 100%, esto es, sólo fueron reemplazados aquellos pixels cuyos valores diferían con respecto al valor calculado de la media de orden h en más de dicho umbral. La figura 2.5 muestra el efecto de estos filtros sobre sinusoides de distinta frecuencia para diferentes valores de h .

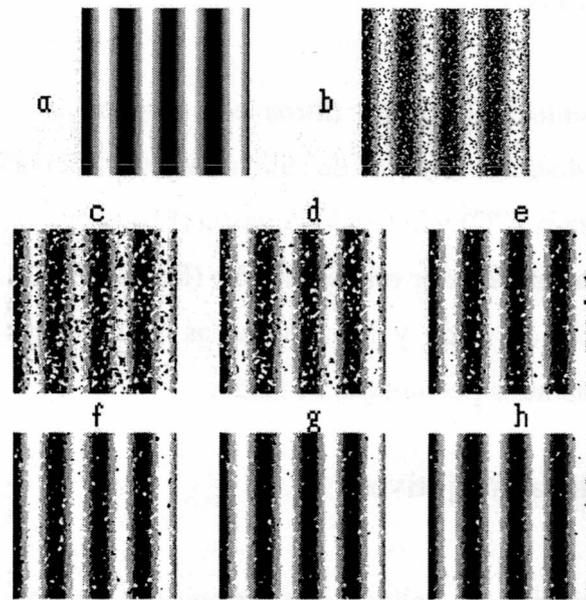


Figura 2.4. Ventana de 9 pixels (a) Imagen original, $f = 0.03$ ciclos/pixel, (b) Resultado de agregar ruido sal y pimienta a (a), porcentaje de sal = 20 %, nivel de sal = 250, porcentaje de pimienta = 15%, nivel de pimienta = 10. (c), (d), (e), (f), (g), y (h) resultados de aplicar estos filtros a la imagen (b) para: (c) $h=0$, (d) $h=0.05$, (e) $h=0.2$, (f) $h=0.3$ (g) $h=0.6$ y (h) $h=1$.

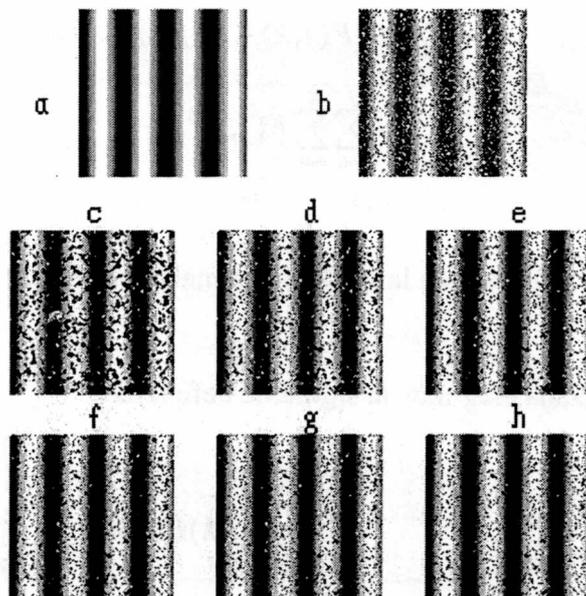


Figura 2.5. Ventana de 8 pixels (a) Imagen original, $f = 0.03$ ciclos/pixel, (b) Resultado de agregar ruido sal y pimienta a (a), porcentaje de sal = 20 %, nivel de sal = 250, porcentaje de pimienta = 15%, nivel de pimienta = 10. (c), (d), (e), (f), (g), y (h) resultados de aplicar estos filtros a la imagen (b) para los mismos valores de h que en la figura 4. El valor umbral es 100%.

2.4 Criterios de calidad

Para evaluar el comportamiento de estos filtros con respecto a sus capacidades para eliminar ruido, utilizamos distintos criterios de calidad objetivos: el error cuadrático medio (ECM), la correlación cruzada (CC) y la ley de potencia (PL) [2.8] y proponemos un nuevo criterio subjetivo que lo llamamos error relativo medio (ERM). Para analizar los resultados usamos las imágenes libres de ruido y las procesadas. Los cálculos se repitieron para diferentes valores de h y distintos porcentajes de ruido.

2.4.1 Criterios de calidad objetivos

Nos referiremos a los criterios de calidad que tratamos a continuación como criterios objetivos, debido a que estos están relacionados con medidas de magnitudes físicas de la imagen (distancia o energía para el ECM y ángulo para la CC).

a) El error cuadrático medio se define como:

$$ECM = \frac{\sum_{j=1}^J \sum_{k=1}^K [F(j,k) - \hat{F}(j,k)]^2}{\sum_{j=1}^J \sum_{k=1}^K F(j,k)^2} \quad (2.5)$$

donde $F(j,k)$ es el valor del pixel en la imagen original (libre de ruido) y $\hat{F}(j,k)$ es el de la imagen filtrada.

b) Para la correlación cruzada elegimos la siguiente definición:

$$CC = \frac{\sum_{j=1}^J \sum_{k=1}^K F(j,k) \hat{F}(j,k)}{\left(\sum_{j=1}^J \sum_{k=1}^K (F(j,k))^2 \sum_{j=1}^J \sum_{k=1}^K (\hat{F}(j,k))^2 \right)^{1/2}} \quad (2.6)$$

La CC fue normalizada a su máximo valor (1), mientras que el valor obtenido en la definición dada en la referencia [2.8] no está acotado.

c) La Ley de Potencia se define como:

$$LP = \frac{\sum_{j=1}^J \sum_{k=1}^K [F(j,k)^\nu - \hat{F}(j,k)^\nu]^2}{\sum_{j=1}^J \sum_{k=1}^K F(j,k)^{2\nu}} \quad (2.7)$$

donde ν es un valor entre 0 y 1.

2.4.2 Criterio de calidad subjetivo

Cuando las imágenes son observadas por humanos, el conocimiento del sistema visual (VS) si bien sujeto a un grado muy importante de variabilidad individual, puede ser muy útil para determinar la calidad subjetiva de una imagen o la fidelidad de la misma con respecto a una escena original. Por lo tanto proponemos un nuevo criterio de calidad subjetivo que tiene en cuenta las características promedio del sistema visual humano. Este criterio, llamado *error relativo medio (ERM)*, está basado en la premisa de que un error en un nivel de gris no es percibido de la misma manera para los niveles de intensidad altos como lo es para los bajos. Entonces, basado en la analogía con la ley de Weber-Fechner [2.9] (ver capítulos 6 y 7), el error relativo medio (ERM) podría ser un criterio más adecuado para la evaluación subjetiva de la calidad de la imagen. Lo definimos como:

$$ERM = \sum_{j=1}^J \sum_{k=1}^K \frac{|F(j,k) - \hat{F}(j,k)|}{F(j,k)} \quad (2.8)$$

donde las barras indican el valor absoluto y cada término es calculado siempre que $F(j,k) \neq 0$. La ley de Weber-Fechner [2.9] establece que la relación: $\Delta I/I$ es constante dentro de amplios límites de intensidad (u otros estímulos); donde I es la intensidad para la cual el ojo está adaptado y ΔI es el mínimo incremento de esa intensidad requerido para que la variación sea perceptible.

2.5 Evaluación del efecto de la media de orden h fraccionario en la eliminación de ruido sal y pimienta.

El primer resultado que encontramos es que todos los criterios objetivos evaluados, interpretados en su propio contexto, dan aproximadamente la misma información. Elegimos uno de ellos (el ECM) para describir los resultados de aplicar estos filtros a las sinusoides. La Figura 2.6 y la Figura 2.7 muestran el ECM para diferentes frecuencias y valores de h correspondientes a ventanas de 9 y 8 pixels respectivamente.

Segundo, los mejores resultados con respecto a la reducción de ruido para el ECM se obtienen para ventanas de 9 pixels y $h = 1.3$ para las frecuencias más altas y se corren hasta $h = 1.2$ para las más bajas. Para estos valores de h la mejora con respecto a $h=1$ (mediana) es aproximadamente del 20%.

Tercero, las frecuencias espaciales bajas se corrigen mejor que las altas, esto es una consecuencia del tamaño finito de la ventana.

Cuarto, para valores bajos de h la ventana de 8 pixels no es tan buena como la de 9 para la reducción de ruido.

Obtuvimos que usando ventana de 9 pixels el ECM es sistemáticamente menor que con ventanas de 8 pixels. (t test [2.10] 10 grados de libertad, $P < 0.001$). Para el criterio de la correlación cruzada se obtuvo la relación inversa que significa lo mismo.

2.5 Evaluación del efecto de la media de orden h fraccionario en la eliminación de ruido sal y 19 pimienta

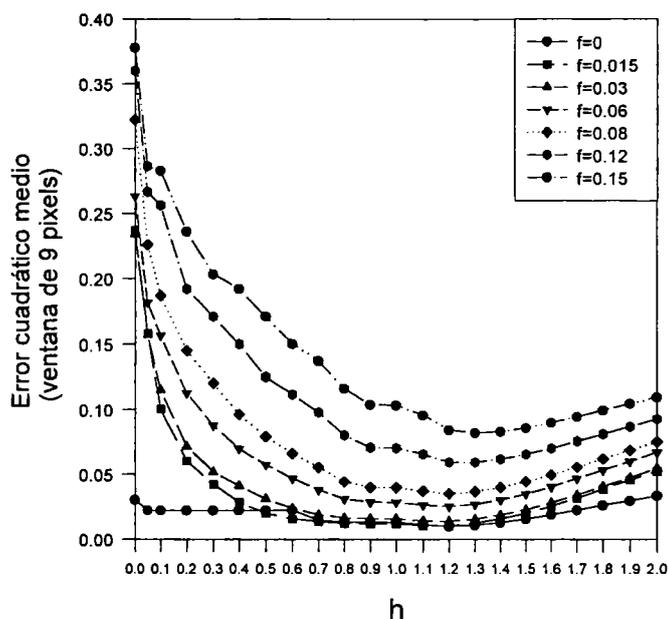


Figura 2.6. Error cuadrático medio para ventanas de 9 pixels, distintas frecuencias, f (expresadas en ciclos por pixel) y h en el intervalo $[0,2]$.

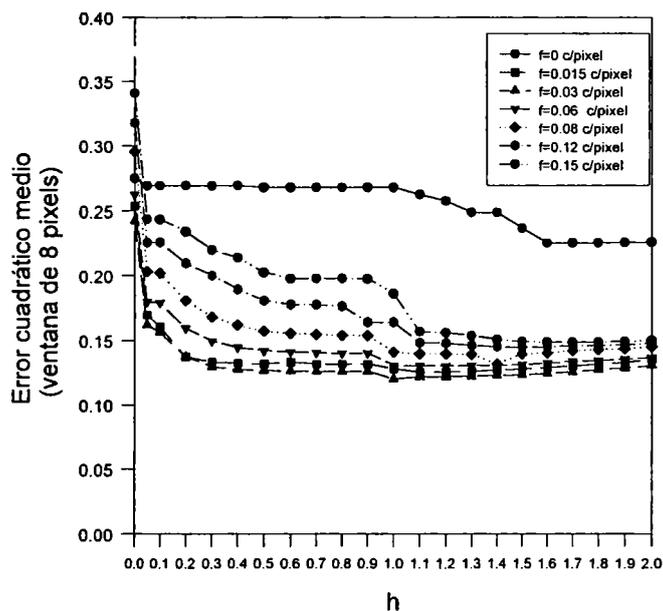


Figura 2.7. Error cuadrático medio para ventanas de 8 pixels, distintas frecuencias, f (expresadas en ciclos por pixel) y h en el intervalo $[0,2]$.

Quinto, encontramos que las ventanas de 8 pixels son mejores que las de 9 pixels para la reducción de ruido con respecto al ERM, como se puede observar en las figuras 2.8 y 2.9.

Por último, el valor mínimo del ERM para ambos tipos de ventanas está en $h = 1$, esto coincide con el buen comportamiento del conocido filtro mediana.

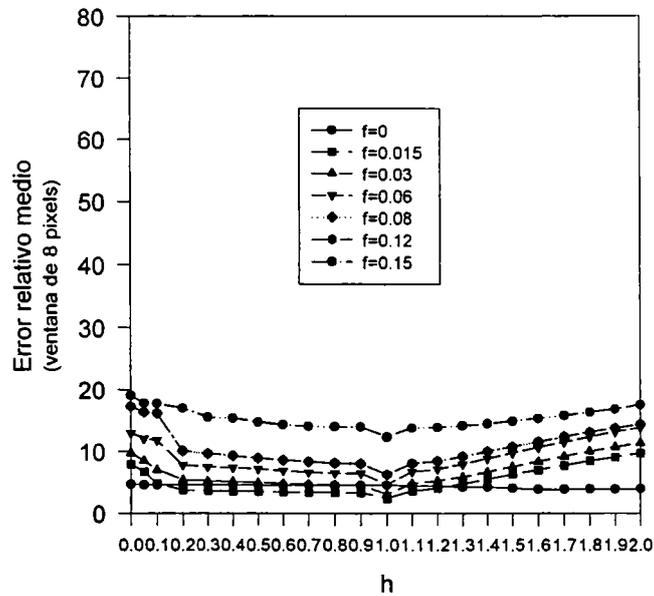


Figura 2.8. Error relativo medio para ventanas de 8 pixels, distintas frecuencias, f (expresadas en ciclos por pixel) y h en el intervalo $[0,2]$.

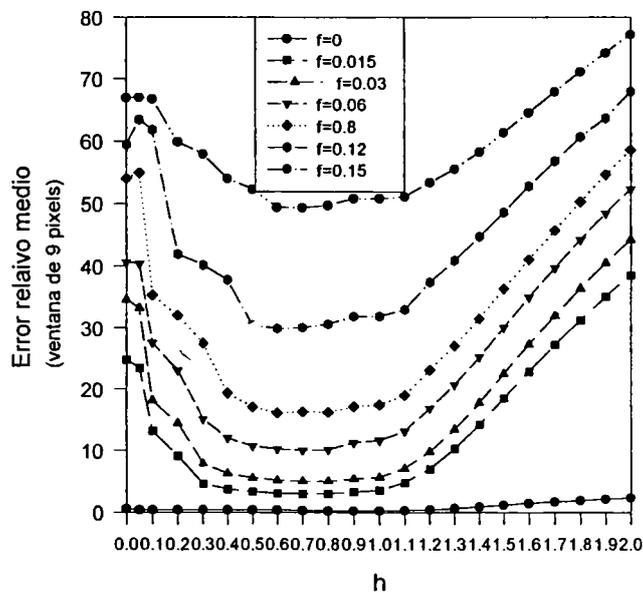


Figura 2.9. Error relativo medio para ventanas de 9 pixels, distintas frecuencias, f (expresadas en ciclos por pixel) y h en el intervalo $[0,2]$.

2.6 La media de orden h fraccionario usando ventanas de 9 y 8 pixels y ruido gaussiano.

En esta sección vamos a analizar el efecto de estos filtros en relación con la disminución de ruido y con preservar bordes, cuando se aplican en imágenes que contienen ruido gaussiano. En primer lugar generamos la imagen I de la referencia [2.5] y luego le agregamos ruido aditivo gaussiano de manera que los resultados mostrados en dicha referencia puedan ser comparados con los de la media de orden h fraccionario. Este ruido lo obtuvimos a partir de la suma de tres números aleatorios con distribución uniforme. Para generar dicho ruido con un determinado valor medio m y varianza σ^2 usamos la siguiente expresión [2.11]

$$x = 2\sigma(y_1 + y_2 + y_3) + m - 3\sigma \quad (2.9)$$

donde los y_i son números aleatorios con distribución uniforme en el intervalo $[0,1]$ y x es la variable aleatoria con distribución aproximadamente gaussiana. A este valor de x lo multiplicamos por 127 y se lo sumamos a las imágenes. La posición del ruido fue generada de la misma manera que antes en diferentes porcentajes. El resultado de la adición fue acotado a 255 y a 0. En nuestro caso elegimos $m = 0$ y $\sigma^2 = 19$.

Un ejemplo del efecto que producen estos filtros con ventanas de 9 pixels para dos porcentajes de ruido diferentes se muestra en la Figura 2.10.

2.6.1 Análisis de los bordes

Ahora estudiamos la habilidad que tienen estos filtros para preservar bordes recurriendo a la Figura 2.10. En ella elegimos una fila que contiene una función escalón de 14 pixels. Luego, corremos una ventana unidimensional sobre ella y calculamos el promedio local, la varianza local y el error cuadrático medio entre la imagen filtrada y la original libre de ruido.

Comenzamos definiendo el promedio local como:

$$M_1 = \frac{1}{w} \sum_{j=j_0}^{j=j_0+w} F(j, k_1) \quad (2.10)$$

donde, k_1 es el número de fila y w es el ancho de la ventana. Ahora bien, si comparamos el promedio local de la imagen de entrada con el de la salida obtenemos información de cuan bien el correspondiente filtro preserva los bordes. Nosotros usamos en los cálculos $w = 3$.

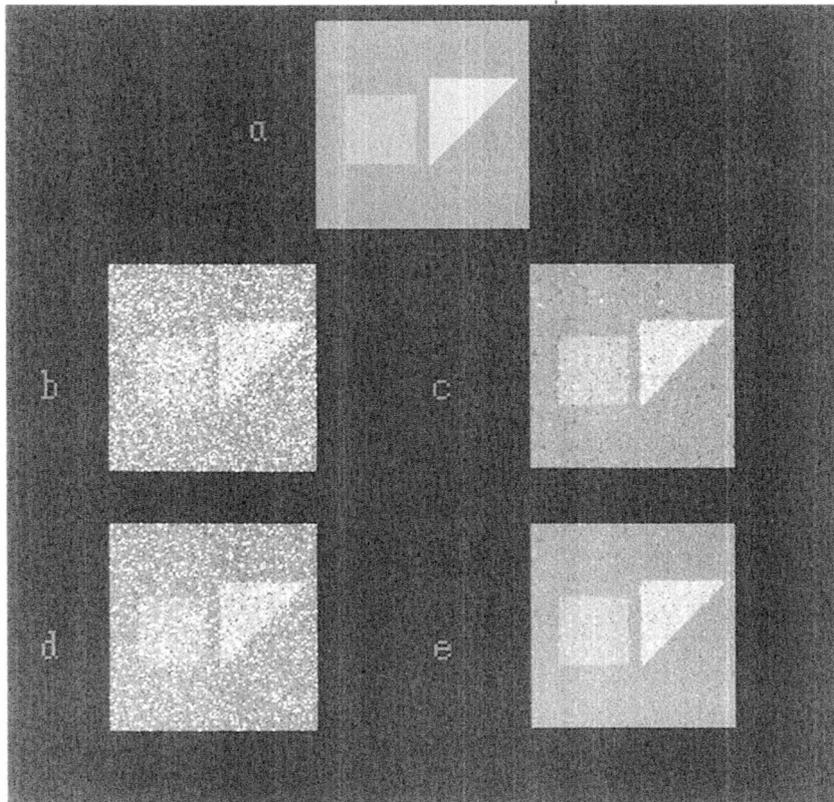


Figura 2.10 (a) Imagen original de la ref. [2.3]. (b) Resultado de agregar a (a) 55% de ruido Gaussiano. (c) Imagen filtrada con ventana de 9 pixels y $h = 0.5$. (d) Resultado de agregar 55% de ruido Gaussiano a (a) y (e) Imagen filtrada con ventana de 9 pixels y $h = 0.8$.

Continuamos definiendo la varianza local como:

$$V_1 = \sum_{j=j_0}^{j=j_0+w} [F(j, k_1) - M_1]^2 \quad (2.11)$$

donde M_1 está definido en la ecuación (2.10). A partir de la comparación entre la varianza local de la salida y la de la imagen original libre de ruido se obtiene la difusión que produce cada filtro en el borde o en los puntos de discontinuidad.

Por último, representamos al valor local del error cuadrático medio en cada subregión por RCM y lo definimos como:

$$RCM_1 = \left\{ \sum_{j=j_0}^{j=j_0+w} [\hat{F}(j, k_1) - F(j, k_1)]^2 \right\}^{1/2} \quad (2.12)$$

donde $\hat{F}(j, k_1)$ representa el pixel filtrado y $F(j, k_1)$ el valor en la imagen libre de ruido. Este cálculo da información sobre la habilidad que tiene el filtro tanto en preservar el borde como en eliminar ruido. Esta operación indica la desviación de la imagen filtrada con respecto a la original

2.7 Evaluación del efecto de la media de orden h fraccionario en preservar bordes

Generamos la imagen I de la Referencia [2.5], en la medida que pudimos de acuerdo con la información disponible, de manera tal que nuestros resultados pudieran ser comparados directamente con los obtenidos con otros filtros no lineales. Utilizamos una ventana unidimensional de 3 pixels aplicada sobre la función escalón anteriormente señalada para estudiar la respuesta en los bordes y calculamos el promedio local, la varianza local y el error cuadrático medio de acuerdo a las definiciones mencionadas.

2.7.1 Promedio local

Aplicamos los filtros de orden h a la imagen I con distintos porcentajes de ruido aditivo Gaussiano. Usamos ventanas de 8 y 9 pixels.

Comenzamos agregando 10% de ruido aditivo y continuamos incrementándolo en pasos de 5% hasta llegar al 55%. Usando ventanas de 9 pixels y hasta este último porcentaje observamos un excelente comportamiento de estos filtros para todos los valores de h , no encontramos diferencias entre la imagen libre de ruido y la filtrada (ver Figura 2.11). Con 60% de ocupación y $h = 0.7$ aparece una distorsión en los bordes como se puede ver en la Figura 2.12.

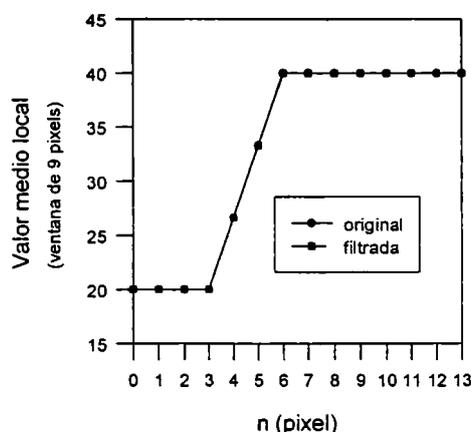


Figura 2.11. Promedio local de la respuesta filtrada de un escalón. Resultado de agregar 55% de ruido Gaussiano y usar un filtro con ventana de 9 pixels. n indica la posición del píxel central de la ventana.

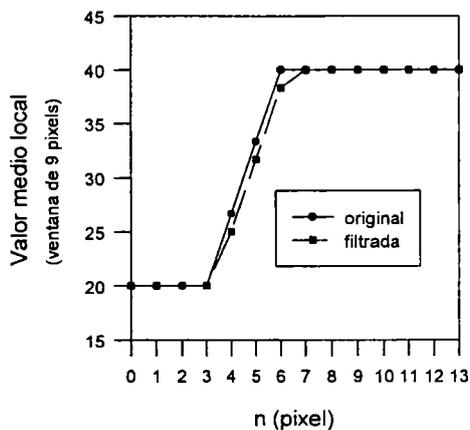


Figura 2.12. Promedio local de la respuesta filtrada de un escalón. Ventana de 9 pixels, $h = 0.7$ y 60% de ruido Gaussiano.

A medida que el porcentaje de ruido aumenta, la distorsión es cada vez más severa. Usando ventanas de 8 pixels la distorsión comienza a aparecer con porcentajes de ruido menores, tan bajos como 20% y $h = 1$ (Figura 2.13).

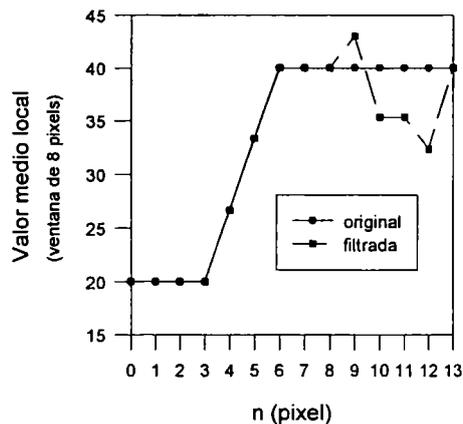


Figura 2.13. Promedio local de la respuesta filtrada de un escalón. Ventanas de 8 pixels y 20% de ruido Gaussiano.

2.7.2 Varianza local

El ancho de la varianza local muestra el grado de difusión del borde como resultado del efecto del filtro. Nuevamente, encontramos excelentes resultados usando ventanas de 9 pixels y ruido Gaussiano hasta 55% de ocupación (Figura 2.14). La varianza local filtrada tiene en el borde el mismo ancho y la misma amplitud que la original.

A medida que se incrementa el porcentaje de ruido por encima del 60%, comienza nuevamente la distorsión (Figura 2.15), con 65% hay sobre un lado del borde una importante distorsión y con 85% hallamos distorsión en los dos bordes. Usando ventanas de 8 pixels una difusión importante aparece para 20% de ruido y valores altos de h .

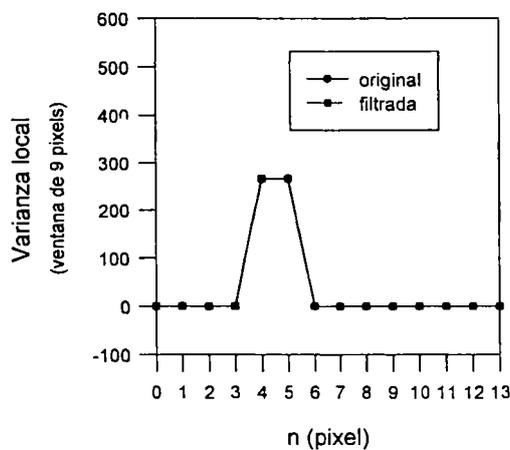


Figura 2.14. Varianza local de la respuesta filtrada de un escalón. Resultado de agregar 55% de ruido Gaussiano, $h = 0.7$ y usar ventanas de 9 pixels

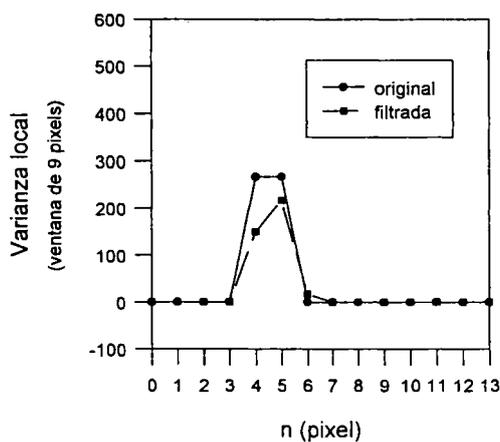


Figura 2.15. Varianza local de la respuesta filtrada de un escalón. Ventanas de 9 pixels, $h = 0.7$ y 60% de ruido Gaussiano.

2.7.3 Raíz cuadrada del error cuadrático medio

Aplicando la raíz cuadrada del error cuadrático medio, encontramos las mismas conclusiones que con la varianza local.

2.7.4 Repuestas a líneas angostas

Para estudiar la resolución que tienen estos filtros sobre una línea angosta, usamos la imagen II de la referencia [2.5]. Todos los filtros, tanto los de 8 como los de 9 pixels no lograron resolver una línea de 1 pixel de ancho.

2.8 Evaluación psicofísica

A continuación realizamos una evaluación visual preliminar de las imágenes filtradas. Comenzamos mostrando una imagen en un monitor de TV a veintitrés observadores. Ésta estaba formada por tres sinusoides libres de ruido ubicadas en la zona central y por sus correspondientes imágenes filtradas con ventanas de 8 y 9 pixels en sus laterales (Figura 2.16).

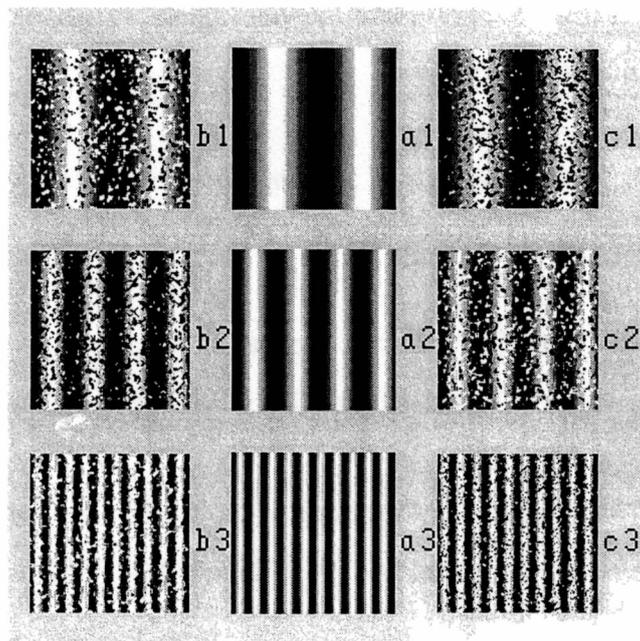


Figura 2.16. Imagen usada en la evaluación psicofísica. (a_1) , (a_2) y (a_3) son las tres imágenes originales de las sinusoides libres de ruido. (c_1) , (b_2) y (c_3) son las filtradas con ventanas de 8 pixels: (c_1) ERM=7865, ECM=0.253; (b_2) ERM=9666, ECM=0.242 y (c_3) ERM=10.004, ECM=0.18. (b_1) (c_2) y (b_3) son las imágenes filtradas con ventanas de 9 pixels. (b_1) MRE=24.661, ECM=0.236; (c_2) ERM=34.532, ECM=0.234 y (b_3) ERM=32.007, ECM=0.144. Las imágenes obtenidas con ventanas de 9 pixels tiene mayor ERM y menor ECM que sus correspondientes con ventanas de 8 pixels.

Las imágenes fueron elegidas de tal manera que una tuviera un ECM mayor que la otra y que lo contrario valiera para el ERM. Luego, se les pidió a los observadores que eligieran la imagen que les resultara más parecida a su original. Las respuestas indicaron que los observadores clasificaron consistentemente como las más similares a aquellas con menor ERM y no a las de menor ECM, cuando la diferencia entre aquellos errores no era pequeña. Este hecho estaría señalando que el ERM se relaciona mejor con la calidad de la imagen percibida que el ECM. Además, las ventanas de 8 pixels mostraron menor ERM que las de 9. Por lo tanto, esta evidencia parecería sugerir que cuando el aspecto visual es el motivo del filtrado, esto es, cuando no está orientado hacia medidas metrológicas, las ventanas de 8 pixels son las más adecuadas para la reducción de ruido.

2.9 Bibliografía.

- [2.1] S.T.Huang Ed, “*Two-Dimensional Digital Signal Processing I*” *Linear Filters*, Springer-Verlag, Berlín (1981).
- [2.2] S.T.Huang Ed, “*Two-Dimensional Digital Signal Processing II*” *Transforms and Median Filters*, Springer-Verlag, Berlín (1981).
- [2.3] B. I. Justusson, “ Median filtering: statistical properties ” in *Two Dimensional Digital Processing II*, T. S. Huang, Ed., 161-196, Springer Verlag, Berlín (1981).
- [2.4] B. R. Frieden, “ Computacional methods of probability and statistics ” Chapter 3, in *The Computer in Optical Research*, 164-166, Springer Verlag, Berlín (1980).
- [2.5] Y. S. Fong, C. A. Pomalaza-Ráez, and X: H. Wang. “Comparision study of nonlinear filters in image processing application”, *Opt. Eng.*, Vol 28, N° 7, pp 749-760, (1989).

- [2.6] J. Barral Souto, “El modo y otras medias: casos particulares de una misma expresión matemática”, Cuadernos de Trabajo N°3, Universidad Nacional de Buenos Aires (1938).
- [2.7] P. Torroba, N. Cap, H. Rabal, W. Furlan. “Fractional order mean in image processing”, *Opt. Eng.*, vol 33, N° 2, 528-534, (1994).
- [2.8] W. K. Pratt, “Sample image quality measure”, “Chapter 7” in *Digital Image Processing*, pp 162-185, New York, Willey (1978).
- [2.9] S. Hecht, “The visual discrimination of intensity and the Weber –Fechner law”, *J. Gen. Physiol.* 7, 241-268 (1924).
- [2.10] B. R. Frieden “The student t- test on the mean”, Chapter 12, in *Probability, Statistical and Data Testing*, 319-384, Springer Verlag, Berlín 1983.
- [2.11] B. R. Frieden, “Fourier methods in probability”, Chapter 4 in *Probability, Statistical and Data Testing*, 93-95, Springer Verlag, Berlín (1983):

Capítulo 3

Respuesta en frecuencia de un sistema óptico desenfocado

3.1 Introducción

Vamos a estudiar el efecto del desenfoco sobre imágenes planas y sobre imágenes de objetos 3D iluminadas incoherentemente. En este capítulo comenzamos con un desarrollo teórico del desenfoco. Analizaremos por un lado, como se desenfoca, que efectos se producen y por otro como modificarlos. Veremos como disminuir el desenfoco o como exagerarlo o hacerlo más sensible.

Para hacer un estudio teórico de imágenes que han sido degradadas por desenfoco es esencial conocer el análisis de Fourier en la teoría de formación de imágenes. En este capítulo se desarrolla la parte necesaria de este análisis, para poder obtener la respuesta en frecuencia de un sistema óptico desenfocado iluminado incoherentemente. Esta respuesta, debe ser especificada para describir completamente el sistema formador de imágenes y

posteriormente obtener la distribución de intensidad en el plano imagen de objetos que han sufrido desenfoco. Además se evalúa la función respuesta en el caso particular de una la pupila rectangular.

3.2 Análisis de Fourier en la teoría de formación de imágenes incoherentes

Se encuentra experimentalmente que muchos fenómenos físicos comparten la propiedad básica de que su respuesta a varios estímulos que actúan simultáneamente es idénticamente igual a la suma de las respuestas que cada estímulo produce individualmente. Estos fenómenos son denominados lineales y la propiedad que ellos comparten es la linealidad. Además, la linealidad de la ecuación de onda que describe la propagación de la luz conduce a considerar a las operaciones ópticas de formación de imágenes en transformaciones de distribuciones de luz "objeto" en distribuciones de luz "imagen".

Cuando la iluminación sobre el objeto es incoherente, las respuestas impulsivas en el plano imagen varían en forma estadísticamente independiente. Por lo tanto, deben ser sumadas en potencia o en base a intensidad. Puesto que la intensidad de cualquier respuesta impulsiva es proporcional a la intensidad del punto objeto que le dio origen, se sigue que un sistema formador de imágenes incoherente es un mapeo lineal en intensidad y que la respuesta impulsiva del mapeo de intensidad es proporcional al módulo al cuadrado de la respuesta impulsiva coherente [3.1].

Un sistema lineal formador de imágenes se dice que es espacialmente invariante si la imagen de una fuente puntual objeto cambia sólo en su posición y no cambia en su forma funcional a medida que la fuente puntual cambia de posición en el campo objeto. En la práctica, los sistemas formadores de imágenes son sólo parcialmente invariantes o también llamados isoplanáticos sobre el campo objeto, pero es posible dividirlo a este campo en pequeñas regiones dentro de las cuales es aproximadamente invariante.

Para describir completamente el sistema formador de imágenes, debe ser

especificada la respuesta impulsiva para cada una de las regiones isoplanáticas; pero si la porción particular de interés del campo objeto es suficientemente pequeña, es suficiente considerar sólo la zona sobre el eje del sistema óptico y en los alrededores.

Si además de considerar sistemas lineales e invariantes, la iluminación es espacialmente incoherente, la distribución de intensidad en el plano imagen se obtiene integrando la intensidad de las imágenes de difracción asociadas con cada punto del objeto. Si $I_g(x_0, y_0)$ es la intensidad geométrica ideal en (x_0, y_0) , $|h(x_i, y_i)|^2$ es la intensidad de la imagen de difracción de la fuente puntual, conocida como *función de punto extendido* (PSF), h la respuesta impulsiva coherente; la intensidad de la imagen en (x_i, y_i) , se obtiene de:

$$I_i(x_i, y_i) = \iint |h(x_i - x_0, y_i - y_0)|^2 I_g(x_0, y_0) dx_0 dy_0 \quad (3.1)$$

Esta ecuación expresa que un sistema formador de imágenes está descrito por la imagen de la fuente puntual localizada a través del campo objeto. Esta integral, por ser una convolución, es conveniente expresarla como

$$I_i = |h|^2 * I_g \quad (3.2)$$

La simplicidad de los sistemas invariantes se pone de manifiesto cuando se aplica la transformada de Fourier a la ecuación (3.1). Transformando ambos lados de la ecuación (3.2) y aplicando el teorema de la convolución, el espectro $I_i(f_x, f_y)$ e $I_g(f_x, f_y)$ del sistema de salida y entrada están relacionados por

$$I_i(f_x, f_y) = H(f_x, f_y) I_g(f_x, f_y) \quad (3.3)$$

donde

$$I_i(f_x, f_y) = \iint I_i(x_i, y_i) \exp[-2\pi i(f_x x_i + f_y y_i)] dx_i dy_i$$

e

$$I_g(f_x, f_y) = \iint I_g(x_i, y_i) \exp[-2\pi i(f_x x_i + f_y y_i)] dx_i dy_i$$

son las transformadas de Fourier de las intensidades imagen y objeto respectivamente y \mathbf{H} es la transformada de Fourier de la respuesta impulsiva incoherente dada por:

$$\mathbf{H}(f_x, f_y) = \iint |h(x_i, y_i)|^2 \exp[-2\pi i(f_x x_i + f_y y_i)] dx_i dy_i \quad (3.4)$$

La función \mathbf{H} indica el efecto del sistema en el dominio de las frecuencias. Para obtener la intensidad de la imagen en (x_i, y_i) (ecuación (3.1)) se ha reemplazado la convolución por otra operación más simple que consiste en realizar la transformada de Fourier multiplicar y aplicar la transformada inversa a I_i , es decir:

$$F^{-1}(I_i(f_x, f_y)) = F^{-1}(\mathbf{H}(f_x, f_y) I_g(f_x, f_y)) = I_i(x_i, y_i)$$

que es en definitiva lo que se necesita calcular.

Para hacer un estudio teórico de imágenes que han sido degradadas por desenfoque es necesario entonces conocer \mathbf{H} , la respuesta en frecuencia de un sistema óptico desenfocado.

3.3 Función de Transferencia Óptica de un sistema libre de aberraciones.

La calidad visual de una imagen depende del contraste o de las intensidades relativas de la porción portadora de información de la imagen y de la porción presente en el fondo, por este motivo es conveniente reemplazar I_g , I_i y \mathbf{H} por su versión normalizada.

Un campo con intensidad constante en el plano objeto tiene en su espectro de Fourier I_g frecuencia cero y da origen a una imagen con intensidad constante. El espectro de frecuencias normalizado de I_g e I_i va a estar definido por

$$G_g(f_x, f_y) = \frac{I_g(f_x, f_y)}{I_g(0,0)}$$

$$G_i(f_x, f_y) = \frac{I_i(f_x, f_y)}{I_i(0,0)}$$

En forma similar la función de transferencia normalizada H va a estar dada por

$$D(f_x, f_y) = \frac{H(f_x, f_y)}{H(0,0)} \quad (3.5)$$

Esta función D es conocida como función de transferencia óptica (FTO). Por lo tanto D queda expresada como:

$$D(f_x, f_y) = \frac{F\{|h|^2\}}{F\{|h|^2\}_{f_x=0, f_y=0}}$$

Puesto que la función de transferencia coherente y la función de transferencia incoherente involucran a la respuesta impulsiva coherente h entonces, debe existir alguna relación entre ambas. La función de transferencia coherente está dada por [3.1]

$$C(f_x, f_y) = F\{h\} \quad (3.6)$$

donde $F\{h\}$ indica la transformada de Fourier de h . Haciendo uso de esta última ecuación y aplicando el teorema de autocorrelación se obtiene [3.2]

$$\mathbf{D}(f_x, f_y) = \frac{\iint \mathbf{C}(\xi', \eta') \mathbf{C}^*(\xi' - f_x, \eta' - f_y) d\xi' d\eta'}{\iint |\mathbf{C}(\xi', \eta')|^2 d\xi' d\eta'}$$

Si se hace el siguiente cambio de coordenadas

$$\xi = \xi' - \frac{f_x}{2} \quad \eta = \eta' - \frac{f_y}{2}$$

resulta la expresión simétrica

$$\mathbf{D}(f_x, f_y) = \frac{\iint \mathbf{C}\left(\xi + \frac{f_x}{2}, \eta + \frac{f_y}{2}\right) \mathbf{C}^*\left(\xi - \frac{f_x}{2}, \eta - \frac{f_y}{2}\right) d\xi d\eta}{\iint |\mathbf{C}(\xi, \eta)|^2 d\xi d\eta} \quad (3.7)$$

Esta ecuación es válida para sistemas incoherentes con y sin aberraciones. Las propiedades más importantes que se derivan de esta ecuación son las siguientes :

$$1. -\mathbf{D}(0,0) = 1$$

$$2. -\mathbf{D}(-f_x, -f_y) = \mathbf{D}^*(f_x, f_y)$$

$$3. -|\mathbf{D}(f_x, f_y)| \leq |\mathbf{D}(0,0)|$$

Para un sistema coherente libre de aberraciones se tiene que

$$\mathbf{C}(f_x, f_y) = P(f_x, f_y) \quad (3.8)$$

en donde P es transmitancia de la pupila.

Para sistemas ópticos incoherentes se obtiene usando la ecuación (3.8)

$$\mathbf{D}(f_x, f_y) = \frac{\iint P(\xi + \frac{f_x}{2}, \eta + \frac{f_y}{2}) P(\xi - \frac{f_x}{2}, \eta - \frac{f_y}{2}) d\xi d\eta}{\iint P(\xi, \eta) d\xi d\eta} \quad (3.9)$$

en el denominador se reemplazó P^2 por P debido a que se considera que la transparencia de la pupila toma el valor uno o cero. Esta última integral tiene la siguiente interpretación geométrica:

El numerador representa el área de superposición de dos funciones pupila desplazadas, una centrada en $(\frac{f_x}{2}, \frac{f_y}{2})$ y la otra en $(-\frac{f_x}{2}, -\frac{f_y}{2})$. El denominador normaliza dicha área por el área total de la pupila.

$$\mathbf{D}(f_x, f_y) = \frac{\text{área de superposición}}{\text{área total}} \quad (3.10)$$

Para calcular la FTO de un sistema limitado por difracción, se pueden realizar directamente los pasos indicados en esta interpretación, como se muestra en la Figura 3.1.

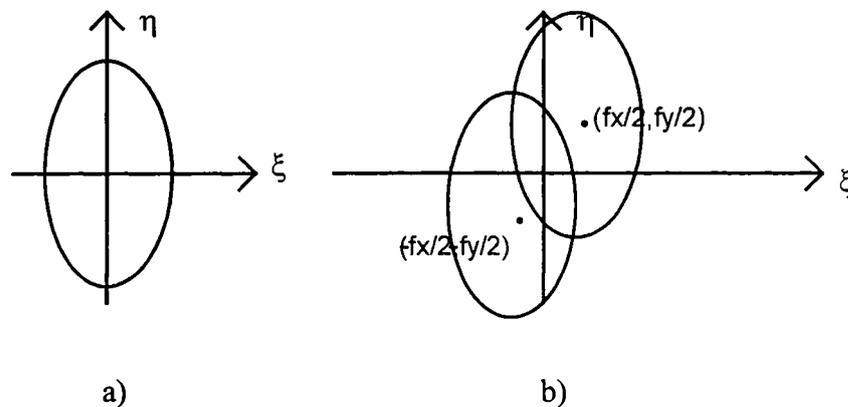


Figura 3.1. Interpretación geométrica de la Función de Transferencia Optica (FTO) de un sistema limitado por difracción. (a) La función pupila. El área total es el denominador de la FTO. (b) Dos pupilas desplazadas. El área de superposición es el numerador de la FTO.

3.4 Función de Transferencia Óptica de un sistema desenfocado

Cuando un sistema formador de imágenes está limitado sólo por difracción, la respuesta impulsiva coherente consiste en el patrón de difracción de Fraunhofer de la apertura de salida centrada alrededor del punto imagen ideal. Este hecho sugiere una manera de introducir las aberraciones. Cuando existe un error de fase en el frente de ondas, se puede pensar que la pupila de salida está iluminada por una onda esférica ideal y que existe dentro de la apertura una placa que produce un corrimiento de fase de tal manera que deforma el frente de onda que sale de la apertura. Si el error de fase en el punto (x,y) en la pupila de salida está representado por $k.W(x,y)$, donde $k = \frac{2.\pi}{\lambda}$, λ la longitud de onda y W es la distancia óptica entre el frente de onda de referencia ideal y el emergente, entonces la transmitancia compleja \mathbf{P} va a estar dada por

$$\mathbf{P}(x,y) = P(x,y).exp(i.k.W(x,y)) \quad (3.11)$$

en donde \mathbf{P} es la función pupila generalizada. La respuesta impulsiva de un sistema formador de imágenes coherente es el patrón de difracción de la abertura con transmitancia \mathbf{P} . La respuesta impulsiva de un sistema incoherente que presenta aberraciones es nuevamente el módulo al cuadrado del resultado coherente.

La función de transferencia óptica expresada en función de \mathbf{P} va a estar dada por:

$$D(f_x, f_y) = \frac{\iint \mathbf{P}\left(\xi + \frac{f_x}{2}, \eta + \frac{f_y}{2}\right) \cdot \mathbf{P}^*\left(\xi - \frac{f_x}{2}, \eta - \frac{f_y}{2}\right) d\xi d\eta}{\iint \mathbf{P}(\xi, \eta) \cdot \mathbf{P}^*(\xi, \eta) d\xi d\eta} \quad (3.12)$$

Si se considera que la transparencia de la pupila es uniforme y toma el valor uno dentro de la abertura de la lente, la función \mathbf{P} tiene la forma:

$$P(x,y) = \begin{cases} \exp(ikW(x,y)) & \text{dentro de la abertura de la lente} \\ 0 & \text{afuera} \end{cases}$$

Es la integral dada por la ecuación (3.12) la que debe ser evaluada para encontrar la respuesta de un sistema óptico que presenta aberraciones o desenfoco. Si se consideran dos círculos unidad en el plano (x,y), centrados en los puntos $\left(\pm \frac{\alpha}{2}, \pm \frac{\beta}{2}\right)$ como muestra la Figura 3.2, el integrando en la ecuación (3.12) se anula afuera de la región común a estos círculos porque $P(x,y)=P^*(x,y)=0$ para $x^2 + y^2 > .$

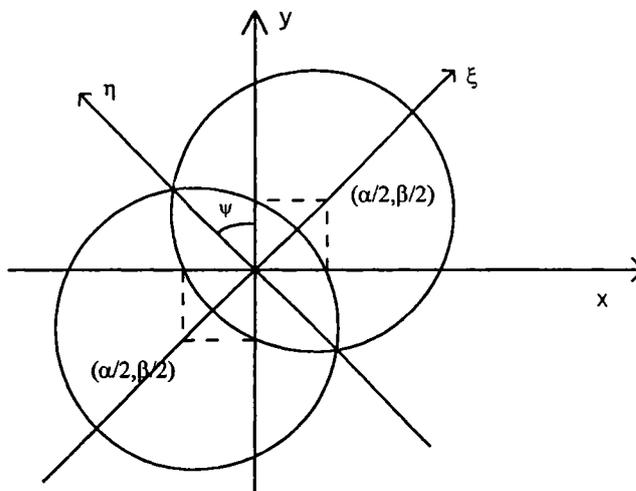


Figura 3.2. Región de integración para el par de frecuencias (alpha,beta).

Si los ejes coordenados (x,y) se reemplazan por (xi,eta), de manera que el eje xi pase por los puntos $\left(\pm \frac{\alpha}{2}, \pm \frac{\beta}{2}\right)$, estando el origen en el punto medio de su unión, la ecuación (3-12) se convierte en:

$$D(f_x,0) = \frac{\iint P\left(\xi + \frac{f_x}{2}, \eta\right) \cdot P^*\left(\xi - \frac{f_x}{2}, \eta\right) d\xi d\eta}{\iint P(\xi, \eta) \cdot P^*(\xi, \eta) d\xi d\eta} \tag{3.13}$$

donde $f_x = \sqrt{\alpha^2 + \beta^2}$ y $\mathbf{P}(\xi, \eta)$ se obtiene de $\mathbf{P}(\alpha, \beta)$ por medio de las sustituciones:

$$\xi = x \cos \psi + y \sin \psi$$

$$\eta = y \cos \psi - x \sin \psi$$

$$\text{siendo } \psi = \tan^{-1} \left(\frac{\beta}{\alpha} \right)$$

por lo tanto sin pérdida de generalidad se pueden considerar sólo objetos cuya forma sea de estructuras lineales; es decir objetos cuya intensidad sea constante a lo largo de un eje y varíe respecto del eje perpendicular. Las consideraciones anteriores muestran que la FTO para el par de frecuencias (α, β) , es la misma que para una sola frecuencia $(f_x, 0)$, siempre que la pupila esté rotada un ángulo ψ . Alternativamente, se puede interpretar el resultado identificando a la respuesta del sistema para el par de frecuencias (α, β) con la respuesta para la frecuencia (f_x) , de una estructura lineal cuya dirección esté inclinada un ángulo ψ con la sección meridional de la pupila. Este resultado, el cual se aplica a cualquier forma de la función pupila, no solamente simplifica el álgebra en la evaluación de la función respuesta en cualquier caso, sino que implica el importante corolario: toda la información relacionada con las propiedades de la formación de imágenes de cualquier sistema se puede obtener en la práctica usando estructuras lineales. Más aún, el resultado nos permite pensar en términos del caso más fácilmente visualizable de un objeto unidimensional.

La expresión (3.13) se convierte en

$$D(s, 0) = \frac{\iint \mathbf{P}(\xi + \frac{s}{2}, \eta) \cdot \mathbf{P}^*(\xi - \frac{s}{2}, \eta) d\xi d\eta}{\iint \mathbf{P}(\xi, \eta) \cdot \mathbf{P}^*(\xi, \eta) d\xi d\eta} \quad (3.14)$$

en donde s es la frecuencia espacial reducida relacionada con la frecuencia espacial f_x en el plano imagen por

$$s = \left(\frac{\lambda}{n \sin u} \right) f_x \quad (3.15)$$

en donde λ es la longitud de onda y u es la abertura angular.

En nuestro caso estudiamos como aberración el error de enfoque. La condición de foco correcto está dada por la ley de las lentes:

$$\frac{1}{d_0} + \frac{1}{d_i} - \frac{1}{F} = 0 \quad (3.16)$$

en donde d_0 es la distancia entre el objeto y la lente, d_i , entre la lente y la imagen y F la distancia focal. Cuando el plano imagen está fuera de foco vale la relación más general:

$$\frac{1}{d_0} + \frac{1}{d_i} - \frac{1}{F} = e \quad (3.17)$$

en donde e es el error de enfoque. Haciendo uso de las ecuaciones (3.11) y (3.12) y comparando con la respuesta impulsiva de un sistema formador de imágenes bajo luz coherente dada por [3.3]:

$$h(x_i, y_i; x_0, y_0) \cong \iint P(x, y) \exp \left[i \frac{k}{2} \left(\frac{1}{d_0} + \frac{1}{d_i} - \frac{1}{f} \right) (x^2 + y^2) \right] \\ \exp \left\{ -ik \left[\left(\frac{x_0}{d_0} + \frac{x_i}{d_i} \right) x + \left(\frac{y_0}{d_0} + \frac{y_i}{d_i} \right) y \right] \right\} dx dy$$

$W(x, y)$ va a estar dado por:

$$(x, y) = \frac{e(x^2 + y^2)}{2} \quad (3.18)$$

Para estudiar el problema de la formación de imágenes libre de otras aberraciones en la presencia de desenfoco, la función pupila toma la forma

$$\mathbf{P}(x, y) = \begin{cases} \exp\{ikw(x^2 + y^2)\} & \text{dentro de la pupila} \\ 0 & \text{fuera} \end{cases} \quad (3.19)$$

en donde el coeficiente $w = \frac{e}{2}$ es una medida del defecto de enfoque; es la longitud del camino óptico entre el frente de onda emergente y la esfera de referencia centrada sobre el eje axial en el punto I como se puede ver en la figura 3.3. Los signos de w y z se toman positivos si el frente de onda está enfrente de la esfera de referencia. Este es el caso mostrado en la Figura 3.3.

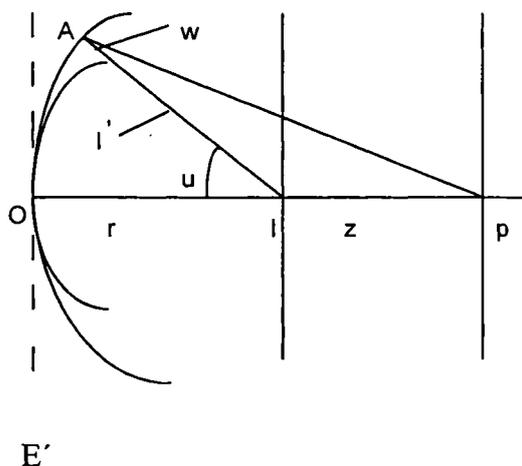


Figura 3.3. Sistema óptico con defecto de enfoque. E' denota la pupila de salida, I' el rayo extremo en el espacio imagen. La esfera de referencia S' está centrada sobre el eje en el punto fuera de foco P'. El frente de onda S' converge hacia el punto enfocado I. La cantidad desenfocada está especificada con la distancia fuera de foco z, o la distancia óptica w entre S y S'.

Es conveniente expresar la función de transferencia óptica en términos de la longitud de camino óptico w y de la frecuencia normalizada s porque esta función puede ser

calculada sin el conocimiento de los parámetros ópticos del sistema. Por otro lado una vez conocida la FTO se debe establecer una relación entre w y z de modo que se pueda calcular la función de transferencia óptica para un sistema específico.

Del triángulo AIP se sigue que

$$(AP)^2 = (AI)^2 + (IP)^2 - 2(AI)(IP)\cos(\pi - u)$$

esta ecuación se puede escribir como

$$(r + z)^2 = (r + w)^2 + z^2 - 2(r + w)z\cos(\pi - u)$$

donde la distancia imagen $OI = r$. Resolviendo para w

$$w = -r - z\cos u + (r^2 + 2zr + z^2\cos^2 u)^{1/2}$$

para pequeños ángulos y desenfocos la aproximación

$$z^2\cos^2 u = z^2$$

puede ser reemplazada en la ecuación anterior obteniendo

$$w = z(1 - \cos u)$$

y utilizando la aproximación trigonométrica para pequeños ángulos

$$\sin \frac{u}{2} = \frac{1}{2} \sin u$$

se encuentra que:

$$w = \frac{1}{2} z \operatorname{sen}^2 u \quad (3.20)$$

la cual da la relación entre w y z que es la distancia entre el plano de enfoque y el de fuera de foco.

3.5 Evaluación de la función respuesta $D(s)$

Si se sustituye la función pupila dada por la ecuación (3.19) en el integrando de la ecuación (3.14) se obtiene:

$$\exp(ikw) \left\{ \left[\left(\xi + \frac{s}{2} \right)^2 + \eta^2 \right] - \left[\left(\xi - \frac{s}{2} \right)^2 + \eta^2 \right] \right\} = \exp(ia\xi) \quad (3.21)$$

en donde $a = 2.k.w|s| = 4.\pi.w|s|/\lambda$. La integral de la expresión (3.21) dada por el numerador de ecuación (3.14), es cero fuera de la región común a las dos pupilas centradas en $\left(\pm \frac{s}{2}, 0 \right)$.

Si se considera una apertura rectangular de semiancho 1 y semilongitud y_0 , se debe resolver:

$$D(s) = \frac{\iint_A \exp(ia\xi) d\xi d\eta}{\iint_{A_t} d\xi d\eta}$$

en donde A es el área común de las dos pupilas y A_t es el área total de una sola. Los límites de ξ son $\pm \left(1 - \frac{s}{2} \right)$ y los de η son $\pm y_0$. Por lo tanto la FTO está dada por:

$$D(s) = \frac{\text{sen}\left\{a\left(1 - \frac{s}{2}\right)\right\}}{a} \quad (|s| \leq 2)$$

$$= 0 \quad (|s| > 2)$$
(3.22)

Se calculó esta función para distintos valores del defecto de foco correspondientes a:

$$w = n \left(\frac{\lambda}{\pi} \right)$$

en donde n varía entre 1 y 50. El valor más grande de $n=50$ denota un defecto de foco de $w = 15.9\lambda$. Para un sistema de apertura numérica $\text{sen } u = 0.10$, $n'=1$ y $\lambda = 0.5\mu$ (ecuación 3-20) le corresponde un defecto de enfoque de $z = \pm 1.6\text{mm}$.

Una característica de estas curvas, como se puede observar en las Figuras 3.6, 3.7 y 3.8 es el rápido deterioro de la respuesta de la lente para altas frecuencias aún para pequeños valores de n .

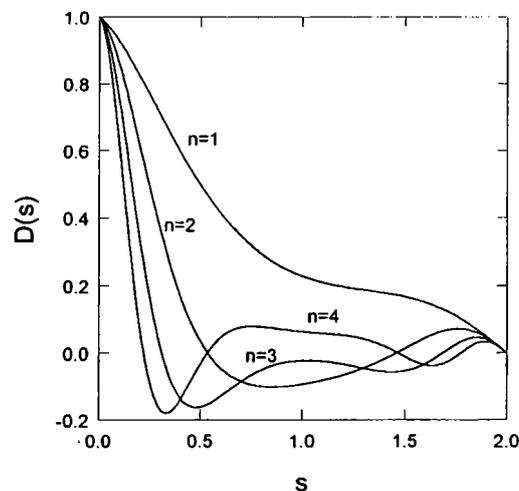


Figura 3.6. Función respuesta (FTO) de una abertura rectangular en función de la frecuencia s , para distintos valores de defecto de enfoque $w = n \left(\frac{\lambda}{\pi} \right)$. En este caso n varía entre 1 y 4.

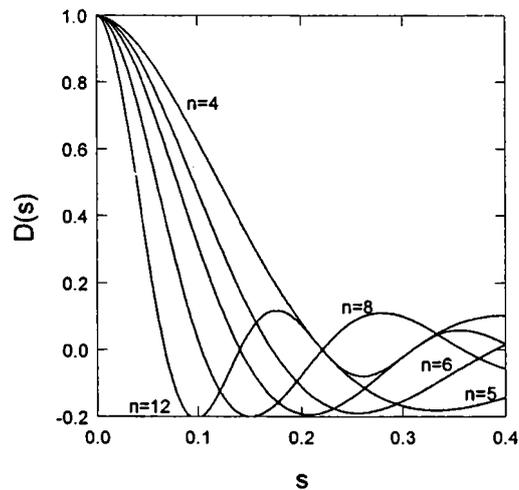


Figura 3.7. Función respuesta (FTO) de una apertura rectangular en función de la frecuencia s . En este caso n varía entre 4 y 12.

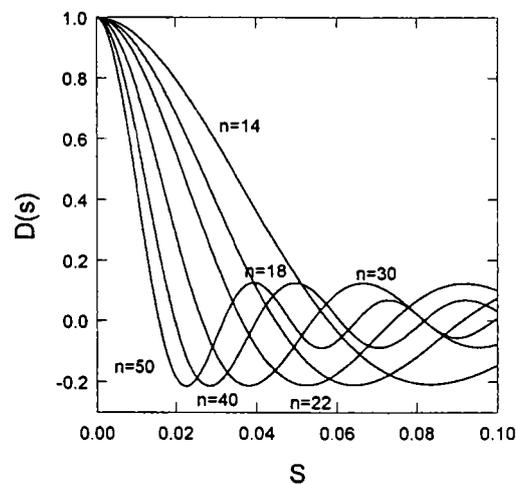


Figura 3.8. Función respuesta (FTO) de una apertura rectangular en función de la frecuencia s . En este caso n varía entre 14 y 50.

Para valores de n mayores que 2 se produce un cambio de signo en la FTO que indica una inversión de contraste. La región en la que la OTF tiene signo negativo se llama *resolución espúrea*. Se concluye de los gráficos que cuando el sistema óptico está fuera de foco se produce una atenuación del contraste y el número de veces que se invierte aumenta para las altas frecuencias.

Se estudió también el comportamiento de la FTO en función del desenfoco z para distintos valores de la apertura numérica y para una frecuencia s fija. En la Figura 3.9 y en

la 3.10 se puede ver dicha función en donde se usó como parámetro la apertura angular. A medida que esta aumenta (desde (a) hasta (c)), el deterioro de la respuesta de la lente se produce para apartamientos menores de la posición de foco correcto. Es decir que la FTO se hace más sensible al desenfoque cuando aumenta la apertura numérica. Esta sensibilidad se hace aún mayor cuando aumenta la frecuencia s que se está analizando. Este resultado se puede observar si se compara la Figura 3.9 en la cual se utilizó $s=0.01$ y la Figura 3.10 con $s=0.1$.

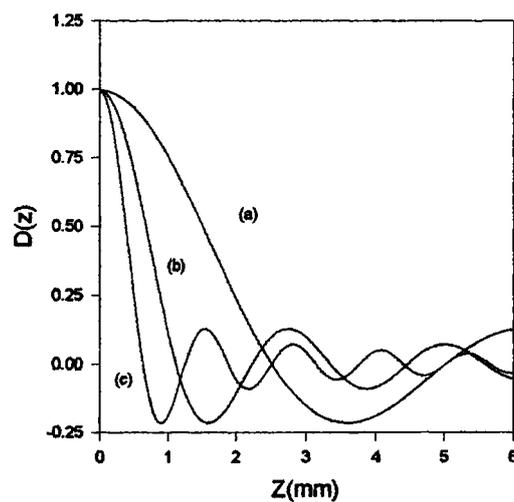


Figura 3.9. Función respuesta (FTO) de una apertura rectangular en función de la distancia de enfoque, para una frecuencia fija $s=0.01$. Se usó como parámetro la apertura angular (a) seno $u=0.1$, (b) seno $u=0.15$ y (c) seno $u=0.2$.

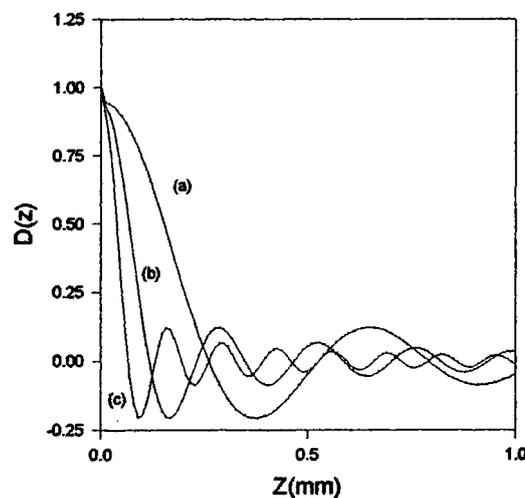


Figura 3.10. Función respuesta (FTO) de una apertura rectangular en función de la distancia de enfoque, para una frecuencia fija $s=0.1$. Se usó como parámetro la apertura angular (a) seno $u=0.1$, (b) seno $u=0.15$ y (c) seno $u=0.2$.

La función respuesta anteriormente calculada para una abertura rectangular es cualitativamente similar a la correspondiente usando una abertura circular [3.4]. Por este motivo no se han calculado resultados numéricos en este trabajo.

Se puede observar que cuando $w \rightarrow 0$ la ecuación (3.22) se reduce a:

$$D(s) = \begin{cases} 1 - \frac{s}{2} & (|s| \leq 2) \\ = 0 & (|s| > 2) \end{cases}$$

que corresponde a la FTO enfocada de sistemas limitados por difracción con pupila cuadrada [3.2].

Finalmente se puede ver en la referencia [3.5] que la aproximación geométrica no describe bien los sistemas limitados por difracción para las altas frecuencias y para cualquier desenfoco relevante ($w=50\lambda$). Sin embargo, cuando $w \geq 2\lambda$, la forma de las dos funciones de transferencia es similar para bajas frecuencias y la FTO geométrica es satisfactoriamente parecida a la exacta FTO para las frecuencias que se encuentran dentro del primer lóbulo lateral positivo. Para altas frecuencias, la FTO geométrica está desplazada con respecto a la exacta FTO, y este desplazamiento es función de la frecuencia y del defecto de enfoque.

3.6 Bibliografía

- [3.1] Joseph W. Goodman.: Introduction to Fourier Optics, Cap. 6, McGraw-Hill, New York, (1968).
- [3.2] Joseph W. Goodman.: Introduction to Fourier Optics, Cap. 2, McGraw-Hill, New York, (1968).
- [3.3] Joseph W. Goodman.: Introduction to Fourier Optics, Cap. 5, McGraw-Hill, New York, (1968).

- [3.4] H. H. Hopkins, "The Frequency Response of a Defocused Optical System", Proc. R. Soc. London Ser. A **231** pp 91-103 (1955).
- [3-5] A. Stokseth, "Properties of a Defocused Optical System", J. Opt. Soc. Am. **59** pp 1314-1321, (1969).

Capítulo 4

Corrección del desenfoque usando el criterio de máxima visibilidad

4.1 Introducción

En las primeras etapas de la fotografía, enfocar con una cámara era una tarea laboriosa que consistía en mover la lente hasta que la imagen en el visor estuviera en foco. En nuestros días la tecnología ha hecho que el enfoque se realice simplemente apretando un botón. Para poder obtener imágenes perfectamente enfocadas, las cámaras de enfoque automático emiten un haz de luz infrarrojo y a veces ultrasonidos sobre los objetos. A través de la diferencia de tiempo entre los haces transmitidos y reflejados, se calcula la distancia entre el objeto y la cámara. Esta información es entonces usada para controlar un motor y ajustar el enfoque de la lente de la cámara en la posición correcta. Si bien, esta técnica es excelente y simple, puede no funcionar bien en microscopios, telescopios, cámaras oftalmológicas o endoscópicas.

La aplicación de técnicas de procesamiento de imágenes en microscopía requiere trabajar con imágenes perfectamente enfocadas. Cuando se examinan secciones de distintas muestras mediante inspección visual, el enfoque manual del microscopio es frecuentemente muy impreciso y subjetivo para poder asegurar reproducibilidad. Esto es especialmente severo cuando se tienen que analizar un gran número de objetos en el campo de visión, por ejemplo en aplicaciones médicas, biología, industrias de semiconductores, etc y puede afectar los resultados de los programas que usan estas imágenes para mejorar análisis. La automatización del control de enfoque asegura la reproducibilidad de la experiencia.

En microscopía, uno de los inconvenientes más comunes que presentan las imágenes tridimensionales es su degradación debida al desenfoque. Cuando se observa una muestra que tiene un cierto espesor, en un microscopio convencional, la estructura fina está en ciertos objetos planos que están perfectamente enfocados, pero no se ven en los fuertemente desenfocados. Por lo tanto la observación de estructuras 3-D en un objeto con cierta profundidad, en una sola imagen es generalmente dificultosa. Durante los últimos años ha habido considerables esfuerzos para obtener una imagen de un objeto tridimensional de manera tal, que esté totalmente en foco.

En este capítulo se estudian algunos criterios conocidos y otros nuevos, propuestos por nosotros, que determinan cual es la imagen plana mejor enfocada dentro de un conjunto de imágenes planas con desenfoque progresivo. Estos criterios se investigan desde el punto de vista teórico, aplicando el análisis de Fourier desarrollado en el capítulo anterior y desde el punto de vista experimental. A partir de los resultados, se elige el criterio de máxima visibilidad y se aplica en imágenes de objetos 3D para construir una imagen en la cual el desenfoque está parcialmente corregido [4.1].

Antes de empezar a tratar algunos trabajos realizados relacionados con este tema, se definen algunas magnitudes de uso frecuente asociadas al desenfoque.

4.2 Profundidad de enfoque y profundidad de campo

Para un dado sistema óptico, existe un borrono en la imagen de un objeto (debido al desenfoque) que si es suficientemente pequeño no afecta el buen funcionamiento del sistema. La profundidad de enfoque es la distancia que puede ser corrida longitudinalmente la imagen con respecto a un plano de referencia y que introduzca un borrono aceptable. La profundidad de campo es la distancia que se puede correr el objeto hasta que se produzca un borrono aceptable y depende de la apertura numérica del sistema formador de imágenes. El tamaño admitido del borrono puede ser especificado como el diámetro lineal del punto desenfocado (Figura 4.1) o como el borrono angular, esto es el ángulo subtendido por el punto borrono desde la lente. Entonces, la magnitud lineal B y la angular β están relacionadas con la distancia D por:

$$\beta = \frac{B}{D} = \frac{B'}{D'}$$

para un sistema en el aire y donde los símbolos primados se refieren a las cantidades del lado de la imagen.

En la Figura 4.2 se puede ver que la profundidad de campo δ para un sistema de apertura A se obtiene de la relación:

$$\frac{\delta}{\beta(D \pm \delta)} = \frac{D}{A}$$

Esta expresión puede ser resuelta para δ , dando:

$$\delta = \frac{D^2 \beta}{(A \pm D\beta)}$$

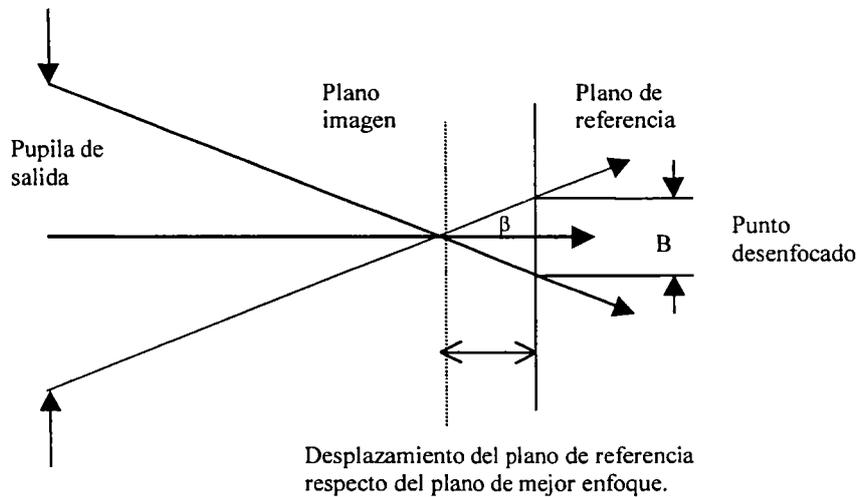


Figura 4.1: Cuando un sistema óptico se desenfoca, la imagen de un punto se empieza a borronear. El tamaño de este borronero está determinado por la apertura y el corrimiento del foco.

De esta relación surge que la profundidad de campo no es simétrica respecto del enfoque correcto. Cuando δ es pequeño comparado con la distancia D , la expresión anterior se reduce a:

$$\delta = \frac{D^2 \beta}{A}$$

Para el lado de la imagen, esta relación es:

$$\delta' = \frac{D'^2 \beta'}{A} = \frac{F^2 \beta}{A} = F\beta (F \text{ número})$$

donde la segunda y la tercera forma del lado derecho se aplican cuando la imagen está en el punto focal del sistema y F es la longitud focal. El “ F número” está dado por la relación entre la longitud focal F del sistema y el diámetro de la pupila de entrada [4.2]. Otra manera de expresar esta relación es utilizando la apertura numérica N.A., la cual se define como el seno de la semi-apertura angular en el espacio objeto multiplicada por el índice de refracción del espacio objeto. Ambas cantidades están relacionadas por:

$$F \text{ número} = \frac{1}{2 N.A.}$$

La profundidad de enfoque en un telescopio puede ser muy grande y en un microscopio es muy pequeña y disminuye cuando la apertura numérica N.A. del objetivo aumenta.

Durante los últimos años ha habido esfuerzos considerables para obtener una imagen con una extendida profundidad de campo, de manera tal de obtener toda la imagen tridimensional en foco.

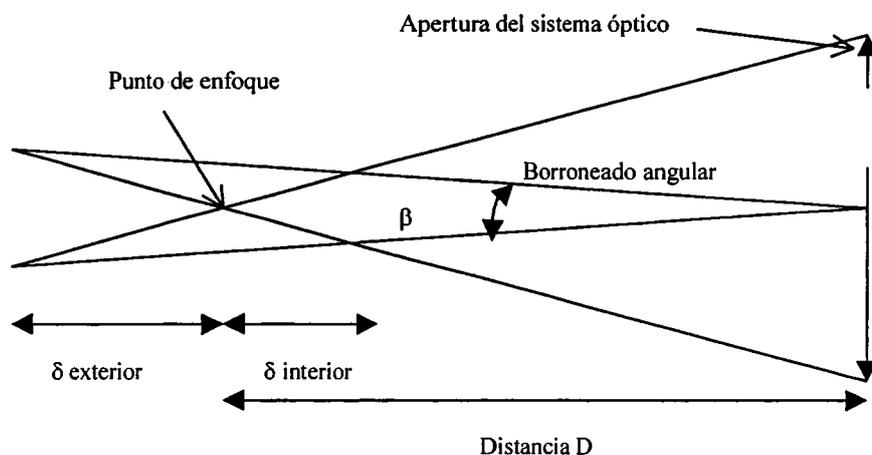


Figura 4.2: Relaciones empleadas para determinar la profundidad longitudinal de campo en términos de un borroado angular tolerable.

W.Welford [4.3] realizó un estudio teórico sobre las posibilidades de usar una apertura anular para aumentar la profundidad focal de un objetivo fotográfico. Si la profundidad de campo se obtiene usando este tipo de aperturas, la definición de imágenes de objetos puntuales aislados mejora considerablemente con respecto a las aperturas convencionales, en las zonas cercanas al eje. Este diseño es desventajoso cuando ciertas consideraciones experimentales exigen el uso de lentes que abarquen amplios ángulos con respecto al eje.

Poon y Motamendi [4.4] proponen y demuestran la posibilidad de extender la

profundidad de campo de un sistema incoherente formador de imágenes a través de una técnica opto-digital. Una apertura anular se coloca enfrente de una cámara de TV para incrementar la profundidad de campo del sistema. La pérdida del contraste de la imagen se compensa con el uso de filtros inversos en una computadora digital.

King y Berry [4.5] presentan una propuesta a la microscopía tridimensional que efectivamente aumenta la profundidad de campo de un microscopio compuesto, entre uno a dos órdenes de magnitud.

Haüslér [4.6] propone un método para incrementar la profundidad de campo que consta de dos pasos. En el primero, genera una imagen incoherente modificada de un objeto tridimensional la cual, aunque degradada, tiene la misma degradación en todos los planos objetos. En el segundo paso, se filtra esta imagen modificada en un procesador de imágenes coherente para obtener una imagen sin degradación con la profundidad de campo aumentada. Este método tiene algunas limitaciones originadas por la presencia de ruido.

M. Mino and Y. Okano [4.7] mejoran las propiedades de las imágenes de un sistema óptico desenfocado, usando aperturas cuyas amplitudes de transmitancia varían gradualmente desde el centro de la pupila hacia el borde. Este estudio está basado en la Función de Transferencia Óptica. Una de las aperturas estudiadas tiene una amplitud de transmitancia que decrece gradualmente desde el centro hacia el borde y en la otra decrece desde el borde hacia el centro. Los resultados de los cálculos muestran que cuando el sistema está desenfocado, en zonas cercanas al foco, la primera clase de las aperturas anteriormente mencionadas produce una imagen desenfocada mejorada con relación a la segunda clase.

Erhardt et al [4.8] desarrollan un procedimiento de reconstrucción basado en la teoría de sistemas lineales para imágenes 3-D en microscopía. Utiliza la función de transferencia óptica dada por Stokseth [3.5] y su antitransformada para la reconstrucción de la imagen. Este procedimiento permite mejoras significativas en la resolución espacial en las imágenes planas perpendiculares al eje óptico. Con este método se hacen visibles estructuras finas

que estaban superpuestas debido a la proyección desenfocada. Por otro lado, este procedimiento tiene el efecto adicional de incrementar la relación señal ruido en un factor de 3.

Los microscopios confocales producen muy buenas imágenes de objetos 3D escaneando todos los puntos en diferentes profundidades y eligiendo la profundidad que maximiza la luz dispersada, este fenómeno sucede en la posición de mejor foco para ese pixel. Generalmente la única información para la determinación del foco es la imagen. Por lo tanto, los mismos datos de la imagen deben ser usados para determinar el foco óptimo por medio del análisis computarizado de la misma. El objetivo del uso de algoritmos de enfoque automático es obtener imágenes mejores y reducir el esfuerzo del observador en enfocar las muestras en forma continua. La necesidad es más evidente cuando se realiza sobre áreas extensas, con más de una zona para ser inspeccionada y con diferentes niveles de profundidad.

Itoh et al [4.9] demuestran un método para extender la profundidad de campo de un microscopio óptico usando un mecanismo de control simple y técnicas de procesamiento digital de imágenes. El método permite obtener proyecciones de objetos 3-D y medir superficies 3-D procesando una sucesión de imágenes digitalizadas tomadas con un microscopio. Estadísticas locales de imágenes de objetos 3-D proveen la información sobre el objeto. También sugieren una modificación del sistema óptico para eliminar efectos no deseados causados por la iluminación cuando se miden superficies 3-D.

Sugimoto et al. [4-10] proponen dos métodos digitales capaces de componer imágenes, (tomadas con un microscopio) con una profundidad de foco incrementada. Uno, usa pseudocoloreado en la serie de imágenes planas y el otro considera la varianza local de cada una de las imágenes planas de la serie. El método de pseudocoloreado es útil cuando el número de imágenes es bajo. Para que el método sea beneficioso, en el caso de usar la varianza local, es necesario elegir el tamaño óptimo de la ventana de acuerdo a las características del objeto y de la resolución espacial requerida. En los resultados experimentales se puede apreciar la eficacia de los métodos. El sistema de procesamiento

de imágenes desarrollado por los autores juega un rol fundamental debido a su alta velocidad digital.

Pieper et al. [4.11] proponen aumentar la profundidad de campo procesando una sucesión de imágenes planas. Debido a la limitada profundidad de campo, cada imagen tiene una zona bien enfocada y otras fuera de foco. Los autores presentan tres algoritmos para seleccionar los segmentos en foco y forman una imagen mejorada. En el primero de ellos, suponen como primera aproximación que la intensidad toma un valor máximo o mínimo en cuando está en foco. El segundo, está basado en el hecho de que la intensidad promedio tomada sobre el conjunto de imágenes tiende a tomar valores que corresponden a zonas fuera de foco. Esto es cierto si la profundidad de campo es más angosta que la zona en que fueron escaneadas las imágenes. En el tercero, definen un operador diferencial no direccional. En las imágenes resultantes se observa un círculo de confusión en las zonas fuera del eje causada por el ligero corrimiento lateral producido por el desenfoque.

Joewono Widjaja et al. [4.12] proponen un nuevo método para mejorar el enfoque automático haciendo uso del análisis de wavelet. La razón de esta aplicación es que la transformada de "wavelet" es útil para realzar los bordes característicos de una imagen. Los autores toman las imágenes con una CCD conectada a una computadora y calculan las autocorrelaciones unidimensionales con los métodos convencionales de las wavelet. Finalmente, grafican las funciones de autocorrelación, comparan el ancho de sus picos y eligen a la imagen que tiene el menor ancho como a la mejor enfocada..

Suganda Jutamulia et al. [4.13] describen un mecanismo de enfoque automático basado en el concepto de autocorrelación. El método puede ser implementado ópticamente con detectores de anillos en el dominio de Fourier o electrónicamente con una cámara CCD en el dominio de la imagen. Además, usan un algoritmo que calcula tanto la transformada de Fourier como su dispersión. La imagen mejor enfocada es aquella que tiene el menor ancho del pico de correlación.

Roca et al. [4.14] presentan un método nuevo de enfoque automático dentro de un

proyecto general para automatizar la transmisión de un microscopio para inspeccionar tejidos citológicos. El criterio de enfoque está basado en el cálculo de variaciones de niveles de gris en dos imágenes detectadas por medio del histograma de contraste dinámico. Este criterio establece cual de las dos imágenes está mas enfocada. El algoritmo de búsqueda está basado en la suposición de simetría y unimodalidad del borroneado.

Comenzaremos con el análisis de algunos algoritmos conocidos que permiten determinar la posición del foco correcto.

4.3 Criterios conocidos asociados al desenfoque

A continuación vamos a estudiar algunos criterios conocidos y otros nuevos propuestos por nosotros, que determinan el plano mejor enfocado dentro de un conjunto de imágenes planas con desenfoque progresivo. Para respaldar estos criterios teóricamente es necesario conocer la teoría de sistemas lineales y el análisis de Fourier, lo cual provee la fundamentación sobre la cual está construida la teoría de formación de imágenes. La parte básica necesaria de este análisis relacionada con el desenfoque fue desarrollada en el capítulo 3 y se aplicará en el siguiente desarrollo.

C. Dähane y F. Lanz [4.15] describen algunos criterios para encontrar la posición óptima del foco en microscopía óptica; estos son el Contenido Estructural de la imagen, la Entropía y la Entropía corregida.

4.3.1 Contenido Estructural

El contenido estructural se define como:

$$S \approx \iint_A I_i(x_i, y_i)^2 dx_i dy_i$$

donde $I_i(x_i, y_i)$ es la intensidad en los puntos imagen (x_i, y_i) que ocupan el campo A. de

un objeto extendido, con intensidades $I_g(x_0, y_0)$. Si se aplica la transformada de Fourier a la ecuación anterior y se hace uso del teorema de Parseval [4.16] se obtiene que S es proporcional a

$$\iint |OTF(f_x, f_y)|^2 |I_g(f_x, f_y)|^2 df_x df_y$$

en donde $OTF(f_x, f_y)$ es la función de transferencia óptica incoherente e $I_g(f_x, f_y)$ es la transformada de Fourier de la intensidad del objeto $I_g(x_0, y_0)$. Las variables f_x y f_y son las frecuencias espaciales en las direcciones x e y respectivamente. Puesto que los dos factores son independientes uno de otro, y se ha tomado el módulo al cuadrado, la integral resultante presenta un máximo cuando cada integral se maximiza. Para un dado espectro de un objeto I_g la FTO toma su máximo valor en el rango espectral cuando el sistema óptico está perfectamente enfocado como se puede observar en las Figuras 3.6, 3.7 y 3.8 en el capítulo anterior. Cuando el sistema se desenfoca la integral de la FTO decrece. Esta función de transferencia es menor o igual que uno y lo mismo se verifica para el módulo al cuadrado de la FTO y como consecuencia el contenido estructural alcanza un máximo en el foco óptico.

4.3.2 Entropía

La entropía está dada por:

$$H = -\sum P_n \log(P_n)$$

donde (P_1, P_2, \dots, P_N) es el conjunto de probabilidades con $\sum_i P_i = 1$. La cantidad P_N se elige como la frecuencia relativa de ocurrencia de los niveles de gris en la imagen. Por lo tanto la entropía es un criterio que caracteriza a una sola imagen.

La cantidad H tiene interesantes propiedades, una de las cuales es ser una medida

razonable para la determinación del foco óptimo. Las propiedades de H son:

1. $H = 0$ si y solo si todos los P_N son cero menos uno que tiene el valor unidad.
2. Para un dado n , H es máximo e igual a $\log(N)$, cuando todos los P_N son iguales.
3. Esta propiedad muestra el comportamiento de la entropía entre los límites permitidos dados por 1.y 2. Cualquier cambio hacia la ecualización de las frecuencias (P_1, P_2, \dots, P_N) , produce un incremento de H .
4. Una transformación lineal de la intensidad $I' = aI + b$ da un término aditivo en la entropía $H = H + \log|a|$. Esto significa que H es invariante frente a cambios b del origen de la escala de intensidades. El término en a determina cambios de contraste en la imagen.
5. Transformaciones lineales del espacio de coordenadas, tales como traslaciones y rotaciones de la imagen dejan a H invariante.

De estas propiedades se sigue que la entropía es una medida sensible a la forma de la distribución en frecuencia de los niveles de gris en la imagen.

4.3.3 Entropía corregida.

Esta cantidad se define como:

$$H_{cor} = H - \log\left(\frac{\Delta I}{\Delta I_{cor}}\right)$$

Es sensible a los cambios de la intensidad de los contornos de la imagen. ΔI y ΔI_{cor} son las diferencias entre el máximo y el mínimo de intensidad de la imagen desenfocada y enfocada respectivamente.

4.3.4 Criterio de extremo de la intensidad

Haüsler et al [4.17] demuestra que la intensidad $I_i(x_i, y_i)$ en la posición (x_i, y_i) como función del desenfoque z siempre tiene un extremo en el mejor foco [4.18]. Basado en este

hecho establece un criterio simple para ser aplicado a un enfoque automático. Este autor propone medir la variación de intensidad en función del desenfoque para unos pocos pixels, en una serie de imágenes con diferente desenfoque. Para encontrar el mejor foco sugiere primero medir $I_n = I(x_{in}, y_{in}, z)$ en varias posiciones (x_{in}, y_{in}) . Entonces, el extremo que es común a todas las $I_n(z)$ determina el mejor foco. Supone iluminación incoherente, que no hay aberraciones y que el objeto se encuentra en un plano perpendicular al eje óptico.

El objeto $I_g(x_0, y_0)$ y la intensidad de la imagen están relacionadas por

$$I_i(x_i, y_i, z) = \iint |h(x_0, y_0, z)|^2 I_g(x_0 - x_i, y_0 - y_i) dx_0 dy_0 \quad (4.1)$$

donde $|h(x_0, y_0, z)|^2$ es la función de punto extendido característica del desenfoque). Para encontrar los valores que son extremos de $I_i(x_i, y_i)$ con respecto a z se debe resolver la siguiente ecuación

$$\frac{\partial I_i(x_i, y_i, z)}{\partial z} = \iint I_g(x_0 - x_i, y_0 - y_i) \frac{\partial |h(x_0, y_0, z)|^2}{\partial z} dx_0 dy_0 \quad (4.2)$$

Esta ecuación puede ser satisfecha por dos clases de soluciones: En la primera clase z_0 satisface:

$$\left. \frac{\partial |h(x_i, y_i, z)|^2}{\partial z} \right|_{z_0} = 0 \quad \text{para todo}(x_i, y_i) \quad (4.3)$$

Estas soluciones son independientes del objeto.

En la segunda clase z_e depende del objeto y de la posición (x_i, y_i) y como se busca un criterio que sea independiente del objeto no es necesario continuar con este tipo de

solución.

Para que la ecuación (4.3) sea satisfecha es necesario y suficiente que se cumpla:

$$\left. \frac{\partial D(v, z)}{\partial z} \right|_{z_0} = 0 \quad \text{para todo } v$$

donde $D(v, z)$ es la OTF (la transformada de Fourier de $|h(x_i, y_i, z)|^2$) y v es la frecuencia espacial normalizada. Esta función está relacionada con la función pupila por medio de la ecuación (3-14), donde P está dada por

$$P(x, y) = \begin{cases} P(x, y) \exp\{ikw(x^2 + y^2)\} & \text{dentro de la pupila} \\ 0 & \text{fuera} \end{cases}$$

y si w se expresa en función de z a través de la ecuación (3.20), la derivada de la función de transferencia incoherente $D(v, z)$ está relacionada con la función pupila P por

$$\frac{\partial D(v, z)}{\partial z} \propto \int P\left(v_1 - \frac{v}{2}\right) P^*\left(v_1 + \frac{v}{2}\right) v \cdot v_1 \cdot \exp\left(-\frac{2\pi}{\lambda} \cdot z \cdot \text{sen}^2 u \cdot v \cdot v_1\right) dv_1$$

donde u es la apertura angular. Para una función pupila hermítica los dos primeros factores $P \cdot P^*$ representan una función impar con respecto a v_1 . Consecuentemente, la integral se anula para toda v en $z = 0$ y siempre existe un extremo de la intensidad de la imagen en el mejor foco.

Una pregunta que puede surgir es ¿Existen otros valores extremos para $z_0 \neq 0$? Si se separa la exponencial dentro de la integral, en la parte real y la imaginaria se ve por razones de simetría que la integral no puede ser idénticamente nula para toda v y $z_0 \neq 0$. Por lo tanto la ecuación (4.3) tiene soluciones independientes de (x_i, y_i) sólo para $z = 0$;

solamente el extremo que es común a cada pixel del plano de mejor foco. Este criterio muestra la siguiente característica: depende de los pixels individuales y no del entorno.

4.3.5 Varianza

Satoshi A. Sugimoto et al. [4.10] proponen el uso de la varianza local como método digital para determinar la posición del enfoque correcto.

La función de transferencia óptica (FTO) de un sistema formador de imágenes desenfocado muestra características de un filtro espacial pasabajos. Cuando el desenfoque se incrementa las componentes de alta frecuencia espacial se atenúan. En el dominio espacial este fenómeno se reconoce por variaciones en la apariencia de la imagen es decir, la estructura fina de las imágenes en foco aparece en forma destacada y sucede lo contrario en las imágenes que están fuera de foco (véase por ejemplo la figura 5.9). La varianza puede ser usada como una medida para caracterizar la distribución espacial tanto en el plano de enfoque como en los desenfocados, en lugar de la FTO en el dominio de las frecuencias.

La varianza se define como:

$$\sigma^2 = \frac{1}{n} \left\{ \sum_{x=0}^{w_x} \sum_{y=0}^{w_y} (I(x,y) - \bar{I})^2 \right\}$$

donde $n = w_x \cdot w_y$, es el número total de pixels de la imagen; w_x y w_y son las dimensiones de la imagen e \bar{I} es el promedio de la intensidad dado por:

$$\bar{I} = \frac{1}{w_x \times w_y} \sum_{x=0}^{w_x} \sum_{y=0}^{w_y} I(x,y)$$

e $I(x,y)$ es la intensidad en la posición de pixel (x,y)

4.4 Criterios propuestos asociados al desenfoque.

Los parámetros físicos nuevos que se ensayaron para determinar la posición del mejor foco fueron los siguientes:

4.4.1 Gradiente

En nuestro caso el gradiente $A(x,y)$ se calculó en la siguiente forma:

$$A(x, y) = \left((\Delta x)^2 + (\Delta y)^2 \right)^{1/2}$$

Donde:

$$\Delta x = I(x+1, y) - I(x, y)$$

$$\Delta y = I(x, y+1) - I(x, y)$$

Este operador se aplicó en forma global, es decir a la imagen como un todo, con la idea de encontrar un valor numérico en la imagen asociado a los bordes (bordes son pixels adyacentes con distintos niveles de gris). Los resultados parciales obtenidos para cada posición de pixel se suman para conseguir un dato único para cada imagen.

4.4.2 Desvío medio aritmético local

El desvío medio aritmético local puede ser usado como otra medida que caracteriza la distribución espacial de una imagen. El desvío medio aritmético local se define como:

$$d.m.a._k(m, n) = \sum_{p=-w_x}^{w_x} \sum_{q=-w_y}^{w_y} \left| I(m+p, n+q) - a_{k(m, n)} \right|$$

donde a es la mediana. La mediana de tamaño n de una secuencia $\{x_i, i \in Z\}$ se define

para n impar como:

$$Y_i = \text{mediana } x_i = (x_{i-\nu}, \dots, x_i, \dots, x_{i+\nu}) \quad i \in Z$$

donde $\nu = \frac{(n-1)}{2}$ y Z denota el conjunto de todos los números enteros.

$2w_x + 1$ y $2w_y + 1$ son las dimensiones x e y de la ventana usada, k es el número de imagen.

En las experiencias se usó n impar.

4.4.3 Visibilidad

La visibilidad definida por Michelson [4.19] para franjas de interferencia se aplica en una manera levemente diferente. Se centra una ventana en cada posición de pixel, se miden los valores de intensidad para cada pixel que se encuentra en la ventana y la visibilidad local se define como:

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad (4.4)$$

I_{max} e I_{min} son los valores máximos y mínimos alcanzados por la intensidad.

Propiedades:

1. De esta definición surge que V pertenece al intervalo $[0,1]$.
2. Vale 1 si y sólo si $I_{min} = 0$.
3. Vale 0 cuando la región considerada es uniforme y distinta de cero.
4. No está definida cuando la región es uniforme e igual a cero. Estos casos raramente se encuentran en la práctica y la definimos igual a 1.

Se propone a la visibilidad como parámetro, debido a que la intensidad es un extremo en el mejor foco [4.17]. Algunos pixels alcanzarán su máximo valor mientras que otros

tomarán el mínimo, incrementando de esta manera el valor de la visibilidad local.

4.5 Experiencia

Las experiencias que se realizaron consistieron en tomar una serie de 50 imágenes de un objeto plano, con desenfoque progresivo en ambos sentidos, con una cámara de 50 mm. de distancia focal. Estas imágenes fueron digitalizadas en 256 niveles de gris y almacenadas en la memoria de una computadora. Las imágenes fueron obtenidas desplazando la cámara de manera de mantener constantes las condiciones de iluminación sobre el objeto y la geometría del sistema formador de imágenes. El objetivo de la cámara se fijó en su menor número F (2.8) con el fin de minimizar la profundidad de campo. La imagen número 25 fue considerada por el observador como la imagen que se encontraba en el foco correcto. El resto de las imágenes se obtuvieron incrementando la distancia a la cámara en pasos de 2 mm. En la figura 4.3 se muestran algunas de ellas.

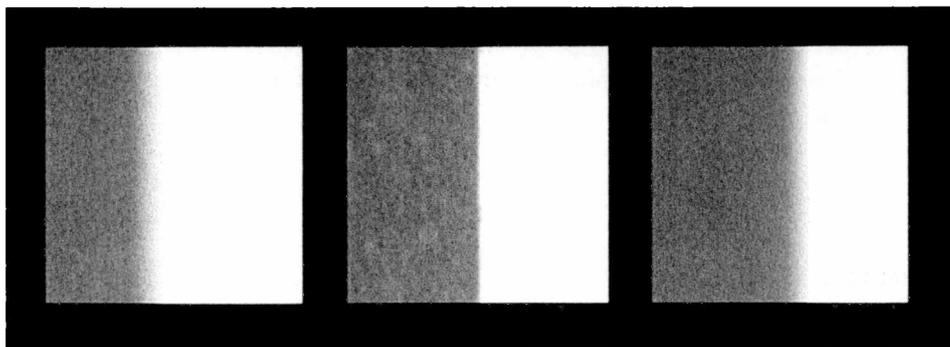


Figura 4.3. Imágenes de un borde desenfocado utilizadas en los cálculos de los parámetros físicos asociados al desenfoque.

4.6 Análisis de los resultados

Para estudiar comparativamente el comportamiento de los parámetros anteriormente mencionados, se aplicaron sobre el conjunto de imágenes planas con desenfoque progresivo. Los resultados obtenidos para cada criterio se graficaron en función del número de imagen, que es proporcional a la distancia z de desenfoque y fueron:

4.6.1 Entropía

En el algoritmo que calcula la entropía se generó un histograma para obtener la probabilidad de ocurrencia de cada nivel de gris, para luego ser utilizada en su definición. La entropía se muestra en la figura 4.4. En ella se observa un mínimo en la imagen número 22.

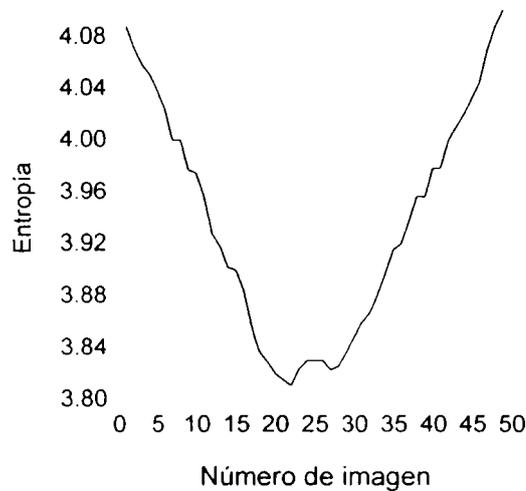


Figura 4.4. La entropía en función del desenfoque (número de imagen).

4.6.2 Entropía corregida

El algoritmo que se implementó es similar al anterior al cual se le agregó la búsqueda del máximo y mínimo de la intensidad en cada imagen y se eligió como imagen enfocada, aquella determinada por el criterio de la entropía. La entropía corregida se observa en la Figura 4.5. En este caso el mínimo se encuentra en la imagen 25, que coincide con el mejor foco según el criterio del observador.

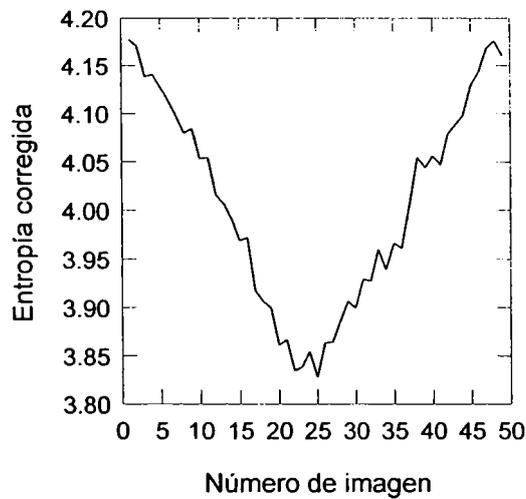


Figura 4.5. Entropía corregida en función del número de imagen.

4.6.3 Criterio de extremo de intensidad

El algoritmo implementado busca para un dado valor de pixel el extremo de la intensidad en función del desenfoco. La Figura 4.6 muestra la intensidad en función del desenfoco para dos posiciones diferentes de pixels. Se puede ver que el extremo común a dichas intensidades se encuentra aproximadamente en la imagen número 26.

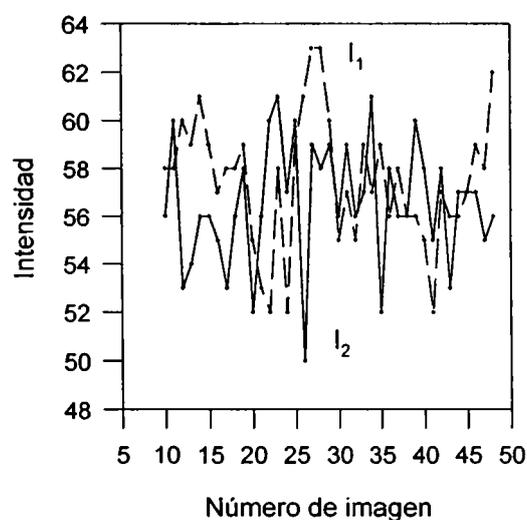


Figura 4.6. Intensidad en función del desenfoco. I_1 e I_2 corresponden a dos posiciones diferentes de pixels en las imágenes

4.6.4 Varianza

El algoritmo calcula el valor medio de la intensidad para toda la imagen, se lo resta a la intensidad de cada pixel, eleva al cuadrado y suma para toda la imagen. La varianza mostrada en la figura 4.7, tiene también un máximo en la imagen 25.

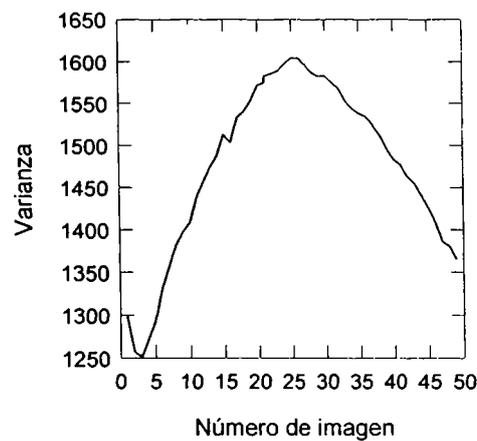


Figura 4.7. Varianza en función del desenfoque.

4.6.5 Gradiente

El gradiente se observa en la Figura 4.8. Este resultado no da ninguna información sobre el mejor foco. También se lo aplicó en forma local utilizando ventanas de 3x3 pixels. En este caso se suman los resultados parciales que se obtienen para cada ventana. Una consecuencia no deseable que surge al aplicar este parámetro es realzar el ruido.

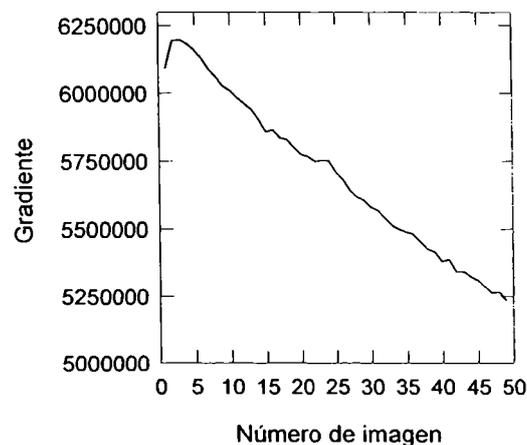


Figura 4.8. Gradiente en función del desenfoque.

4.6.6 Desvío medio aritmético

El algoritmo calcula la mediana para cada ventana, se lo resta al valor de intensidad del pixel central, toma el módulo y suma los resultados obtenidos en cada ventana para todos los pixels de la imagen. Se usaron ventanas de 3x3 pixels. El desvío medio aritmético local se muestra en la Figura 4.9. Un máximo agudo se puede ver en la posición correspondiente a la imagen 24.

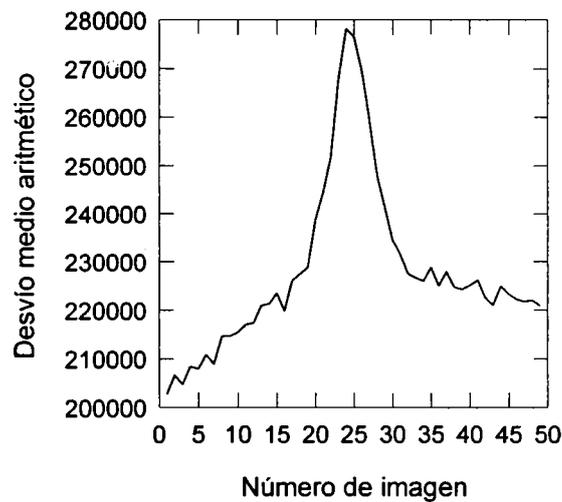
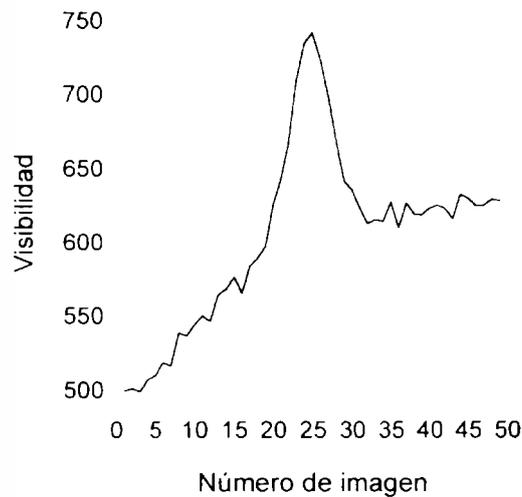


Figura 4.9. Desvío medio aritmético en función del desenfoque.

4.6.7 Visibilidad

El algoritmo que se utilizó busca en una dada ventana los valores máximos y mínimos de la intensidad, calcula la visibilidad y suma estas cantidades para toda posición de pixel de la imagen. Se usaron ventanas de 3x3 pixels. La visibilidad (graficada en la Figura 4.10), presenta un máximo en la posición que corresponde a la del foco óptico (imagen 25). Se puede observar que este máximo es más agudo que el de la varianza.



.Figura 4.10. Visibilidad en función del desenfoque.

Encontramos que los resultados eran sensibles a la presencia de ruido electrónico. Para disminuirlo se repitió la experiencia con otro objeto y se les aplicó a las imágenes un filtro promediador en el tiempo y un filtro mediana, antes de grabarlas. Este filtro realiza el promedio de 32 imágenes en las mismas condiciones experimentales luego se les aplicó el filtro mediana y por último se almacenaron en la computadora. Una de ellas se muestra en la Figura 4.11a. Se calculó la visibilidad de la misma manera que antes y el resultado se muestra como imagen en la Figura 4.12. Se observa en la imagen de mejor foco (número 25) un máximo mucho más destacado que cuando no se aplica el promediador y el filtro mediana.

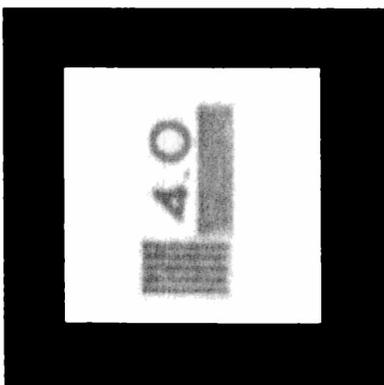


Figura 4.11 a). Imagen de un objeto plano usada en la experiencia.

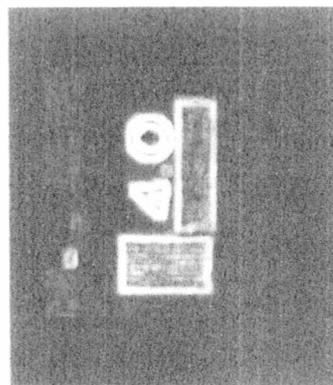


Figura 4.11 b). Distribución de la Visibilidad asociada a la figura 4.11 a).

Para dar una idea de como está distribuida la visibilidad en la imagen se construyó la Figura 4.11 b) asignando a cada pixel el valor escaleado de la visibilidad. Se puede observar que las áreas uniformes en la imagen original aparecen oscuras en la visibilidad, mientras que los bordes y los detalles dan lugar a regiones claras.

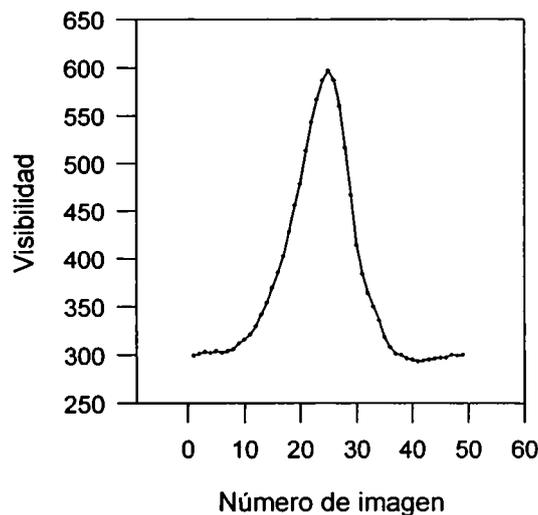


Figura 4.12. Visibilidad en función del desenfoque

Estas experiencias se repitieron con distintos objetos y se obtuvieron resultados consistentes para todos los parámetros. A partir del análisis de los resultados surgió la visibilidad como el mejor parámetro entre todos los propuestos.

4.7 Corrección del desenfoque usando el criterio de la visibilidad.

Anteriormente, encontramos que la visibilidad en función del desenfoque presenta un máximo en la posición del foco óptimo. Por lo tanto, sugerimos usar este criterio para seleccionar la intensidad de la imagen mejor enfocada para cada posición de pixel y construir con ella una nueva imagen en la cual el desenfoque esté parcialmente corregido.

A continuación, realizamos la siguiente experiencia: tomamos una serie de 20 imágenes con desenfoque progresivo, en este caso el objeto estaba constituido por dos

planos en diferentes posiciones del eje z, conteniendo cada plano una letra. Debido a que en la experiencia anteriormente realizada con objetos planos se encontró que los resultados eran sensibles a la presencia de ruido electrónico, se les aplicó a las imágenes un filtro promediador y un filtro mediana, antes de guardarlas en el frame de la computadora. Dentro de este conjunto, las imágenes a procesar fueron elegidas de tal manera que en ninguno de ellas se podían distinguir ambas letras simultáneamente. La Figura 4.13 muestra las cinco imágenes seleccionadas.

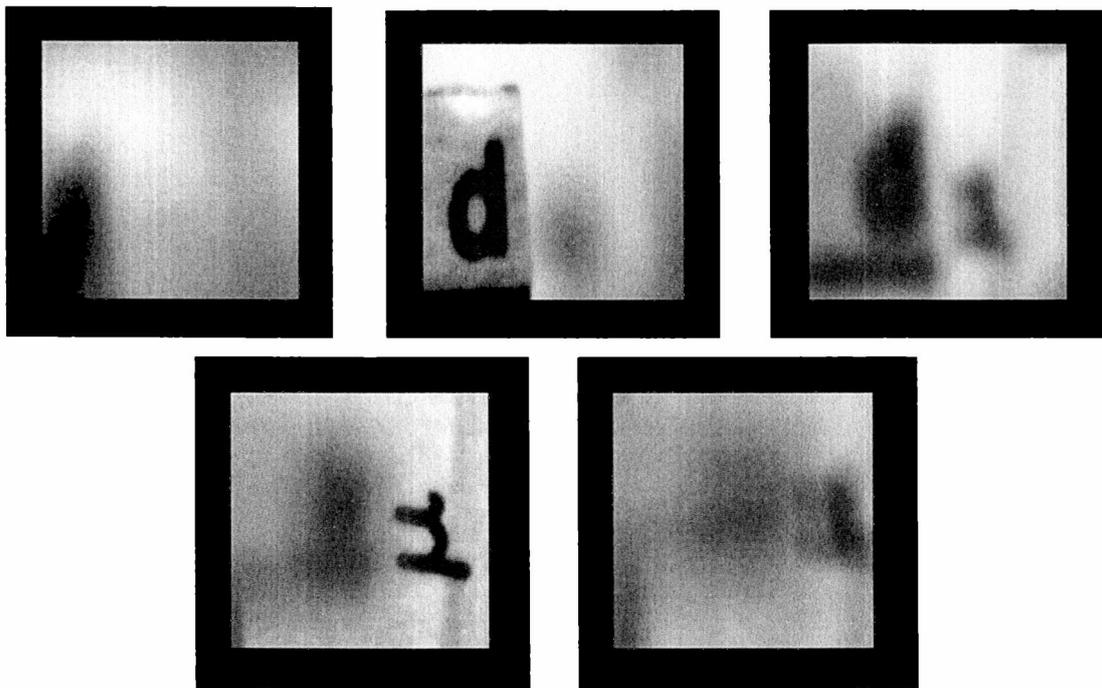


Figura 4.13. Las cinco imágenes seleccionadas de los dos objetos planos.

Estas imágenes fueron examinadas de la siguiente manera:

- a) Se centró una ventana en una dada posición, (i, j) y se calculó la visibilidad para todas las imágenes, en esa posición de pixel, en la dirección del desenfoque.
- b) Se buscó el valor máximo de la visibilidad entre todo el conjunto de imágenes, para esta misma posición de pixel. La imagen encontrada indicaba que en ella estaba el mejor foco.
- c) Se eligió el valor de intensidad del pixel central de esa ventana como el valor correspondiente a la nueva imagen corregida.
- d) Se repitió el proceso para todas las posiciones de pixel de toda la imagen.

La imagen resultante corregida se muestra en la Figura 4.14. En este ejemplo la ventana utilizada fue de 21x21 pixels. En este simple caso ambas letras pueden ser reconocidas y distinguidas claramente. Una zona confusa aparece en la región correspondiente a los bordes del plano frontal.

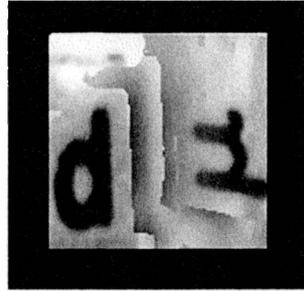


Figura 4.14. Imagen procesada.

La Figura 4.15 muestra las imágenes de visibilidad correspondientes a las de la Figura 4.13. Las imágenes más desenfocadas muestran baja visibilidad en todas partes. Ambos planos presentan valores altos de visibilidad en las áreas de las letras, especialmente en los bordes. En la imagen procesada la visibilidad más alta se encuentra en las áreas mejor enfocadas en ambos planos como se puede observar en la Figura 4.16.

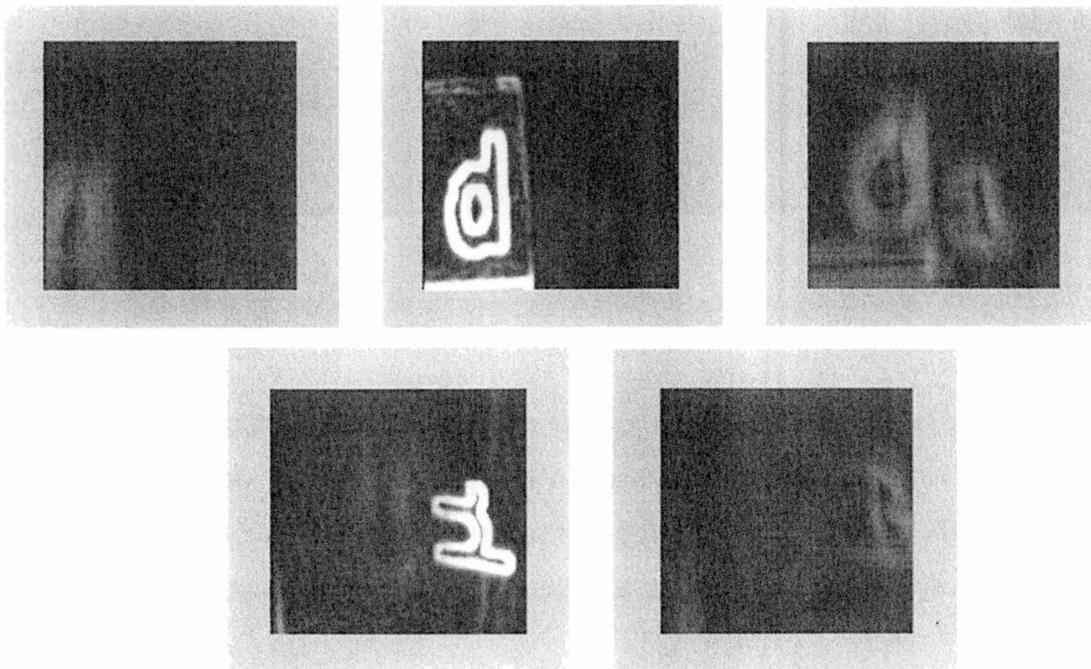


Figura 4.15. Imágenes de la distribución de la visibilidad asociadas a la figura 4.13

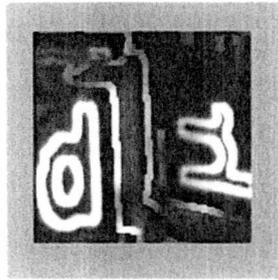


Figura 4.16. Imagen de la distribución de la visibilidad correspondiente a la figura 4.14.

La figura 4.17 muestra el efecto de cambiar el tamaño de la ventana. En todos los casos se utilizaron 15 imágenes. Es evidente que existe un compromiso entre el tamaño de la ventana y la resolución longitudinal. Ventanas más grandes producen buenas imágenes de áreas extendidas planas. En ventanas pequeñas se espera que el procesamiento realice mejor resolución en profundidad, debido a que ellas incluyen un menor número de píxeles desenfocados cuando la profundidad varía constantemente.

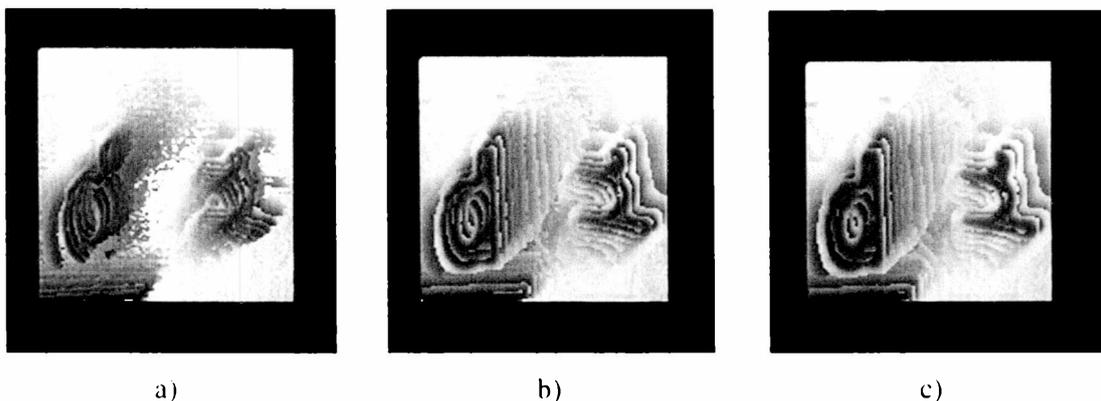


Figura 4.17. Imágenes resultantes empleando diferentes tamaños de ventana. a) ventana de 3x3 píxeles. b) ventana de 7x7 píxeles. c) ventana de 9x9 píxeles.

La siguiente experiencia consistió en aplicar el procedimiento a un objeto 3D en el que la profundidad varía de manera continua. En la figura 4.18 se pueden observar las imágenes de una regla inclinada con desenfoque progresivo usadas en esta experiencia. La imagen resultante corregida se muestra en la Figura 4.19, donde se usaron ventanas de 21x21 píxeles.



Figura 4.18. Imágenes de una regla inclinada con desenfoco progresivo usadas en la experiencia.

La Figura 4.20 muestra el resultado anterior al cual se le aplicó un estiramiento de histograma

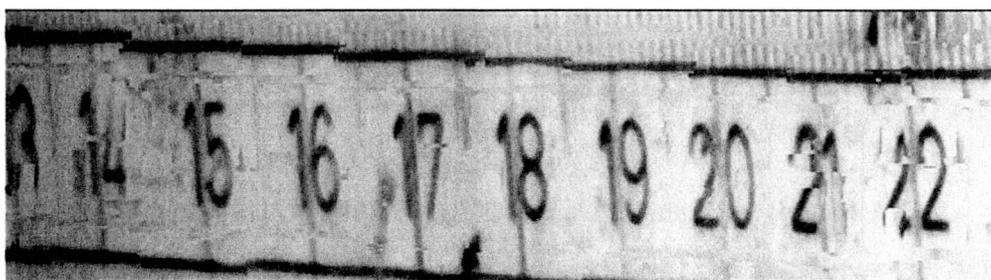


Figura 4.19. Imagen procesada resultante.

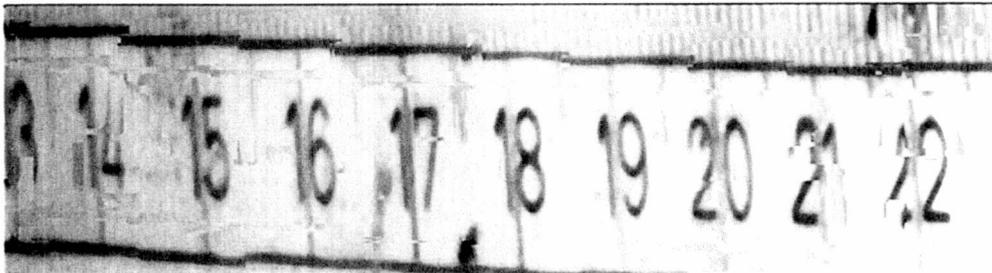


Figura 4.20. Resultado de haber aplícado a la imagen anterior un estiramiento del histograma.

Los números y las divisiones de la regla en la imagen resultante se aprecian mejor con respecto a cualquiera de las imágenes que fueron tomadas. Los resultados son mejores en la parte central de la imagen procesada y la calidad disminuye a medida que la distancia desde el centro aumenta. Tiende a aparecer una sombra en los números. Esto es debido a la proyección central producida por el sistema óptico desenfocado y disminuye un poco cuando aumenta el tamaño de la ventana. En la Figura 4.21 se muestra el resultado cuando el tamaño de la ventana es de 9x9 píxeles y se puede observar el efecto mencionado.

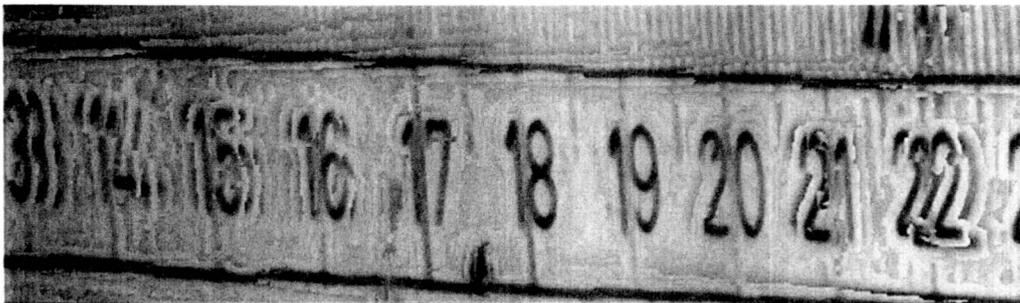


Figura 4.21. Resultado obtenido usando ventanas de 9x9 píxeles.

En la aproximación geométrica la función de punto extendido desenfocada es una versión demagnificada de la pupila del sistema formador de imágenes [4.20]. Si un objeto puntual está fuera del eje óptico, la función de punto extendido viaja en dirección radial a medida que el desenfoque progresa. En la Figura 4.22 se esquematiza esta situación. Cuando el objeto presenta grandes áreas uniformes, comparadas con el tamaño de la ventana, la visibilidad es baja. Si un detalle de alto contraste está cerca de la zona

uniforme, su versión desenfocada invadirá la posición de las áreas uniformes debido al mencionado desplazamiento de la función de punto extendido. Entonces el algoritmo elige el valor más alto de la visibilidad debido a la versión desenfocada del detalle en lugar de la baja visibilidad de la zona uniforme. En algunas aplicaciones este efecto puede ser reducido usando un sistema telecéntrico [4.21]. Este sistema óptico tiene un diafragma ubicado en el plano focal como se esquematiza en la figura 4.23 a) Se puede observar que el rayo principal, en la parte izquierda de la lente, es paralelo al eje. Si el sistema es usado para proyectar la imagen de una escala, se puede ver que un pequeño desplazamiento del enfoque de la escala no cambia la altura de la escala en la que pasa el rayo principal. Si se compara este sistema con otro que tenga la apertura en la lente, como el mostrado en la figura 4.23 b), un pequeño desenfoque causará un error proporcional en la altura del rayo.

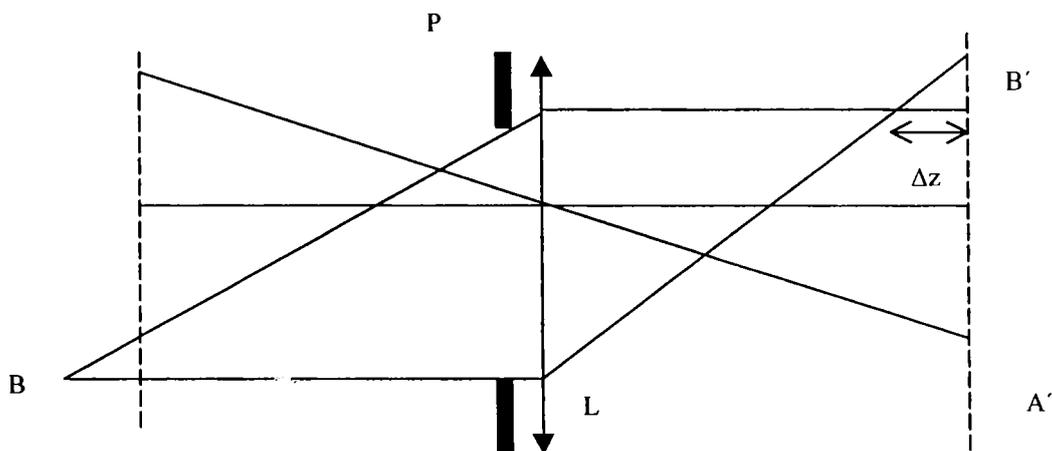


Figura 4.22. En este esquema, usando la aproximación geométrica, se muestra el camino en la dirección radial que recorre la función de punto extendido desenfocada.

Se concluye entonces que el método necesita "guías" para elegir el mejor foco. Este hecho también sucede en la mayoría de los sistemas de enfoque, así como en el sistema visual humano. El uso de iluminación estructurada de alto contraste podría ser una herramienta para mejorar este método de corrección del desenfoque como veremos mas adelante.

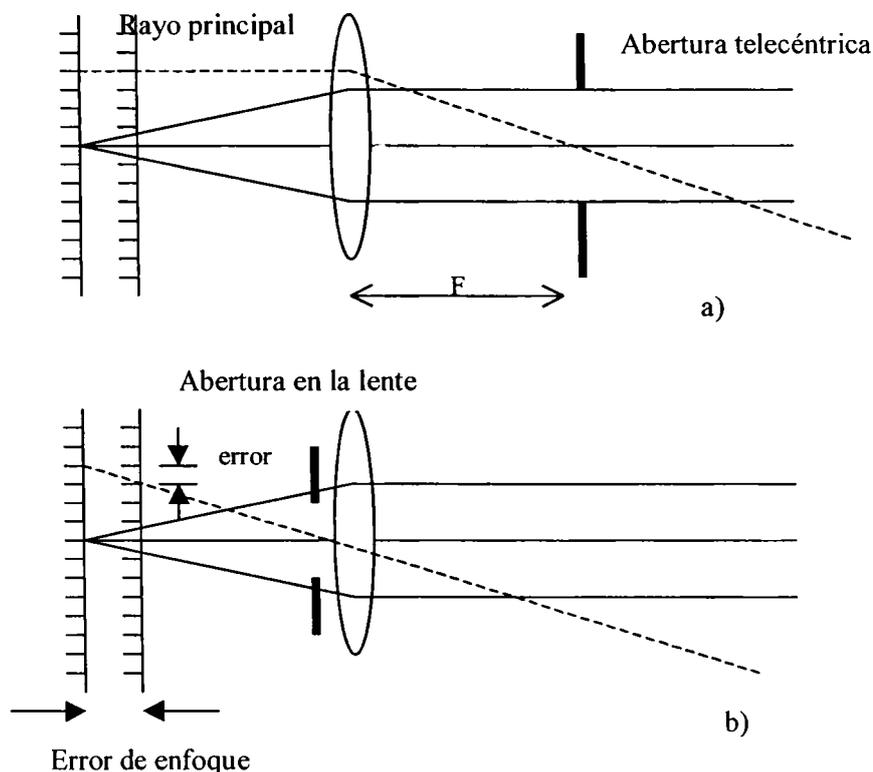


Figura 4.23. La apertura telecéntrica está localizada en el punto focal de del sistema de proyección, de modo que el rayo principal es paralelo al eje del objeto. Cuando el objeto está ligeramente fuera de foco no hay error en el tamaño de la imagen proyectada como lo hay en el sistema con la apertura en la lente, como el esquematizado en la parte inferior de la figura.

Por otro lado si el rango dinámico de los valores de intensidad es angosto, la probabilidad de encontrar dos o más máximos de visibilidad en la misma posición (i,j) , pero en diferentes z , no es despreciable. Para disminuir esta probabilidad habría que optimizar el uso del rango dinámico.

4.8 Estudio teórico de la visibilidad en función del desenfoque para fuentes sinusoidales.

Como la visibilidad, originalmente propuesta por Michelson, fue definida para franjas de interferencia proponemos hacer un estudio teórico de la misma en función del desenfoque, considerando como objeto una fuente sinusoidal unidimensional dada por

$$I_g(x) = \frac{1}{2}(1 - \cos(4\pi f_1 x)) \quad (4.5)$$

donde, f_1 es la frecuencia de la red.

Para obtener la visibilidad, definida por la ecuación (4.4), es necesario conocer la distribución de intensidad a la salida del sistema óptico. En este caso consideramos que la iluminación es incoherente. Por lo tanto, si I_g es la intensidad geométrica ideal en el plano objeto, dada por la ecuación (4.5) y $|h(x_i, y_i)|^2$ es la función de punto extendido la intensidad en el plano imagen va a estar dada por la ecuación (3.1). La relación expresada en esta ecuación se escribe en el dominio de Fourier a través de la expresión (3.3), donde la transformada de Fourier de I_g está dada por:

$$I_g(f_x, f_y) = \frac{1}{2}\delta(f_x, f_y) + \frac{1}{4}[\delta(f_x - 2f_1, f_y) + \delta(f_x + 2f_1, f_y)]$$

y la transformada de la función de punto extendido por:

$$H(f_x, f_y) = D(f_x, f_y)H(0,0)$$

donde $D(f_x, f_y)$ está dada por el módulo de la ecuación (3.22). Si suponemos que la apertura es rectangular de lado l , su área está dada por

$$H(0,0) = l^2$$

Si se expresa a $D(f_x, f_y)$ en función de la frecuencia de corte f_0 toma la forma:

$$D(f_x) = \left(1 - \frac{f_x}{2f_0}\right) \frac{\left| \operatorname{sen} \left(\frac{8\pi\omega}{\lambda} \frac{f_x}{2f_0} \left(1 - \frac{f_x}{2f_0}\right) \right) \right|}{\frac{8\pi\omega}{\lambda} \frac{f_x}{2f_0} \left(1 - \frac{f_x}{2f_0}\right)} \quad f_x \leq 2f_0 \quad (4.6)$$

donde f_0 está definida como

$$f_0 = \frac{\operatorname{sen} u}{\lambda}$$

y está relacionada con la frecuencia espacial reducida s por

$$s = \frac{f_x}{f_0}$$

Para obtener la distribución de intensidad de la imagen se debe aplicar la transformada inversa de Fourier a I_1

$$F^{-1}(I_1) = F^{-1}(D(f_x) \mathbf{H}(0,0) \mathbf{I}_g(f_x, f_y)) = I_i(x)$$

Luego de un poco de álgebra resulta:

$$I_i(x) = \frac{l^2}{2} \left\{ 1 - \left[\left(1 - \frac{f_1}{f_0}\right) \frac{\left(\operatorname{sen} \left(\frac{8\pi\omega}{\lambda} \frac{f_1}{f_0} \left(1 - \frac{f_1}{f_0}\right) \right) \right)}{\frac{8\pi\omega}{\lambda} \frac{f_1}{f_0} \left(1 - \frac{f_1}{f_0}\right)} \right] \cos 4\pi f_1 x \right\} \quad (4.7)$$

Ahora bien, para estudiar el comportamiento de la visibilidad en función del desenfoque, esta intensidad debe ser expresada en función de z en lugar de ω que es el máximo error en el camino óptico. Como se considera una apertura rectangular de lado l , ω está dado por [4.22]

$$\omega = \frac{e \cdot l^2}{8} \quad (4.8)$$

donde el defecto de enfoque e está dado por:

$$e = \frac{1}{d_o} + \frac{1}{d_i} - \frac{1}{F} \quad (4.9)$$

F es la distancia focal y d_o y d_i son las distancias objeto e imagen respectivamente. Cuando $e = 0$ y $d_o = d_c$, donde c significa la distancia objeto correcta (el mejor foco), entonces la ecuación (4.9) se convierte en

$$\frac{1}{F} = \frac{1}{d_o} + \frac{1}{d_i}$$

En general

$$d_o = d_c + z \quad (4.10)$$

donde z es la distancia desenfocada. Ahora, usando las ecuaciones (4.8) a (4.10) y eligiendo $d_i = 2F$, la distribución de intensidad puede ser expresada en función de z

$$I_i(x) = \frac{l^2}{2} \left\{ 1 - \left[\left(1 - \frac{f_1}{f_0} \right) \operatorname{sinc} \left(\left(\frac{\left(\frac{-z}{(2F+z)2F} \right) l^2}{\lambda} \right) \left(\frac{f_1}{f_0} \right) \left(1 - \frac{f_1}{f_0} \right) \right) \right] \cos 4\pi f_1 x \right\} \quad (4-11)$$

Los valores máximos y mínimos de esta intensidad para un dado plano de enfoque z , se obtienen cuando $\cos 4\pi f_1 x = \pm 1$. Bajo estas condiciones resulta:

$$V = \left(1 - \frac{f_1}{f_0}\right) \text{sinc} \left(\left(\left(\frac{z}{(2F+z)2F} \right)^2 \frac{f_1}{\lambda} \right) \left(1 - \frac{f_1}{f_0}\right) \right) \quad (4-12)$$

Por lo tanto, el máximo valor de la visibilidad va a coincidir con el máximo valor de la función seno cardinal. En la Figura 4.24 se ha graficado la visibilidad en función del desenfoque z y puede observarse que el máximo se encuentra en el origen, que corresponde a $z=0$. Los cálculos se hicieron con los siguientes valores: $\frac{f_1}{f_0} = 0.7$, $\lambda = 0.6\mu$ y $F = 20\text{cm}$.

Se concluye entonces que la visibilidad es máxima en la posición del foco correcto, que es lo que se quería demostrar.

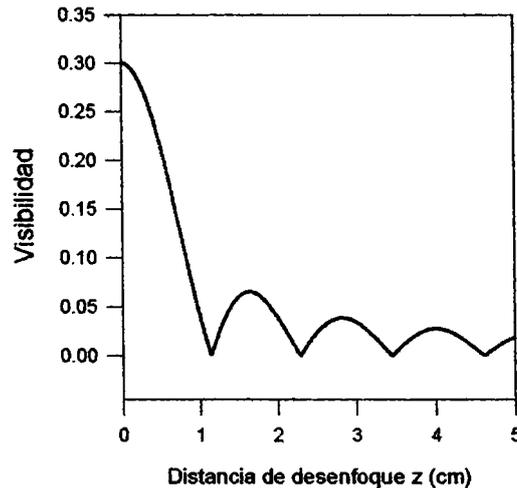


Figura 4.24. Visibilidad en función del desenfoque z (cm)

4.9 Bibliografía

- [4.1] Torroba P., Cap. N and Rabal H. :“Defocus correction using visibility criterion”, *Journal of Modern Optics*, **41**, (1994), 111-117.

- [4.2] Max Born and Emil Wolf: *Principles of Optics*, Sexta Edición, Pergamon Press, p. 187, (1980).
- [4.3] W. T. Welford, "Use of Annular Apertures to Increase Focal Depth", *J. Opt. Soc. Am.* vol **30** N° 8, 749-753, (1960).
- [4.4] Ting-Chung Poon and Masoud Motamendi, "Optical/digital incoherent image processing for extended depth of field", *Appl. Optics*, vol **26**, N° 21, 4612-4615, (1987).
- [4.5] M. C. King and D. H. Berry, "A Depth Scanning Microscope", *Appl. Optics*, vol **10**, N° 1, 208-210, (1971).
- [4.6] G. Häusler, "A method to increase the depth of focus by two step image processing", *Opt. Commun.*, vol **6**, N° 1, 38-42, (1972).
- [4.7] M. Mino and Y. Okano, "Improvement in the OTF of a defocused optical system through the use of shaded apertures", *Appl. Optics*, vol **10** N° 10, 2219-2225,(1971).
- [4.8] A. Erhardt, G. Zinser, D. Komitowski and J. Bille, "Reconstructing 3-D lighth microscopic images by digital image processing", *Appl. Optics*, vol **24** , N° 2, 194-200, (1985).
- [4.9] Kazuyoshi Itoh, Atsushi Hayashi and Yoshiki Ichioka, "Digitized optical microscopy with extended depth of field", *Appl. Optics*, vol **28**, N° 15, 3487-3493, (1989).
- [4.10] Satoshi A. Sugimoto and Yoshiki Ichioka, "Digital composition of images with increased depth of focus considering depth information", *Appl. Optics*, vol **24**, N° 14, 2076-2080, (1985).

- [4.11] R. J. Pieper and A. Korpel, "Image processing for extended depth of field", *Appl. Optics*, vol **22**, N° 10, 1449-1453, (1985).
- [4.12] Joewono Widjaja and Suganda Jutamulia, "Use of wavelet analysis for improving autofocusing capability", *Opt. Commun.*, vol **151**, 12-14, (1998).
- [4.13] Suganda Jutamulia, Toshimitsu Asakura, R. D. Bahuguna, and Panfilo C. De Guzmán, "Autofocusing based on power-spectra analysis", *Appl. Optics*, vol **33**, N° 26, 6210-6212, (1994).
- [4.14] Xavier Roca, Xavier Binefa, and Jordi Vitria, "New autofocusing algorithm for cytological tissue in microscopy environment", *Opt. Eng.* Vol **37**, N° 2, 635-641, (1998).
- [4.15] C. Dähne and F. Lanz, "A microscope focusing technique for a digital image analysis system", *Optik*, vol **55**, N° 4, 437-448, (1980).
- [4.16] J. W. Goodman, *Introduction to Fourier Optics*, MacGraw-Hill Book Company Cap 2, p. 10 (1968)
- [4.17] Ger. Häusler and Eva Körner, "Simple focusing criterion", *Appl. Optics*, vol **23**, N° 15, 2468-2469, (1984).
- [4.18] R. Torroba B. Ruiz and H. Rabal, "An entropy approach to light propagation", *Journal of Modern Optics*, vol 39, No 19, 1939-46 (1992).
- [4.19] Max Born and E. Wolf, *Principles of Optics*, sixth edition, New York, Pergamon Press, p. 267, (1980).
- [4.20] A. Stokseth, "Properties of a Defocused Optical System", *J. Opt. Soc. Am.* **59** pp 1314-1321, (1969).

[4.21] Marx Born and Emil Wolf, *Principles of Optics*, Pergamon Press Sexta Edición (1983) New York p. 187.

[4.22] J. W. Goodman, *Introduction to Fourier Optics*, MacGraw-Hill Book Company Cap 6, p. 124, (1968)

Capítulo 5

Corrección del desenfoque usando iluminación estructurada

5.1 Introducción

En el capítulo anterior estudiamos la visibilidad local en función del desenfoque y hallamos que ésta presenta un máximo en la posición correcta de enfoque como consecuencia de que la intensidad es un extremo en dicha posición. Luego, la maximización de la visibilidad fue utilizada como una indicación de la región mejor enfocada para posteriormente, construir una imagen resultante en donde el desenfoque está parcialmente corregido [4.1].

Los resultados obtenidos son muy buenos en la parte central de la imagen procesada y su calidad decrece cuando se incrementa la distancia al centro. Este defecto se debe a que la función de punto extendido viaja en dirección radial a medida que el desenfoque progresa. En los casos considerados el desenfoque es severo y podemos utilizar para dicha

función la aproximación geométrica [5.1, 5.2] que es una versión demagnificada de la función pupila. Cuando el objeto considerado tiene áreas uniformes grandes (comparadas con el tamaño de la ventana) la visibilidad es baja. Por lo tanto, si un detalle de alto contraste está cerca de una región uniforme su versión desenfocada invadirá la posición de las áreas uniformes debido al viaje radial que realiza la función de punto extendido. El algoritmo elegirá como máxima a la visibilidad que corresponde a la versión desenfocada del detalle en lugar de la baja visibilidad debida a la uniformidad del área. Entonces, para evitar las regiones enfocadas de bajo contraste el método necesita tener en todas partes detalles de alto contraste para “ayudar” al algoritmo a elegir el foco correcto.

El sistema visual humano también comete serios errores de enfoque cuando dichos detalles están ausentes los cuales, serían los responsables de la miopía del campo vacío [5.3] y de la ilusión de la luna [5.4]. Una manera de introducir estos detalles es proyectar iluminación estructurada sobre la escena.

En este capítulo se propone introducir iluminación estructurada sobre la escena para mejorar la calidad de la imagen resultante en donde el desenfoque está parcialmente corregido [5.5]. Para ello, realizamos una serie de experiencias utilizando un arreglo experimental parecido al propuesto por Häusler [5.6] para medidas fotográficas. Con este esquema experimental proyectamos franjas sobre los objetos bajo estudio para incorporar la iluminación estructurada en la escena. Primero, estudiamos el comportamiento de la visibilidad en función del desenfoque en objetos planos con franjas proyectadas. A continuación, para realzar aun más las componentes de alta frecuencia le aplicamos a dichas imágenes el operador derivada y comparamos los valores de la visibilidad de ambos grupos con los valores obtenidos bajo iluminación común. Posteriormente, desarrollamos un formalismo teórico que avala los resultados experimentales. Por último, trabajamos con imágenes 3-D con franjas proyectadas derivadas y sin derivar y corregimos el desenfoque aplicando el criterio de máxima visibilidad teniendo en cuenta las imágenes obtenidas con iluminación común.

5.2 Esquema experimental

El arreglo experimental utilizado para proyectar un patrón fotográfico de franjas o speckle sobre un objeto bajo iluminación incoherente se muestra en la Figura 5.1. La proyección de las franjas sobre el objeto se realiza a través de un separador de haz desde la misma dirección de observación, haciendo coincidir el eje de la cámara con el sistema de proyección. El proyector y la cámara están enfocados en el mismo plano de modo que el contraste de las franjas proyectadas es máximo en la intersección de dicho plano con el del objeto. Con este arreglo experimental el efecto del desenfoque es doble, debido a que se produce tanto en el sistema de proyección como en el sistema formador de imágenes de la cámara CCD. Seleccionamos el f /número de ambos dispositivos en sus valores mínimos para producir un rápido desenfoque y consecuentemente un rápido decrecimiento en el contraste de la estructura desenfocada.

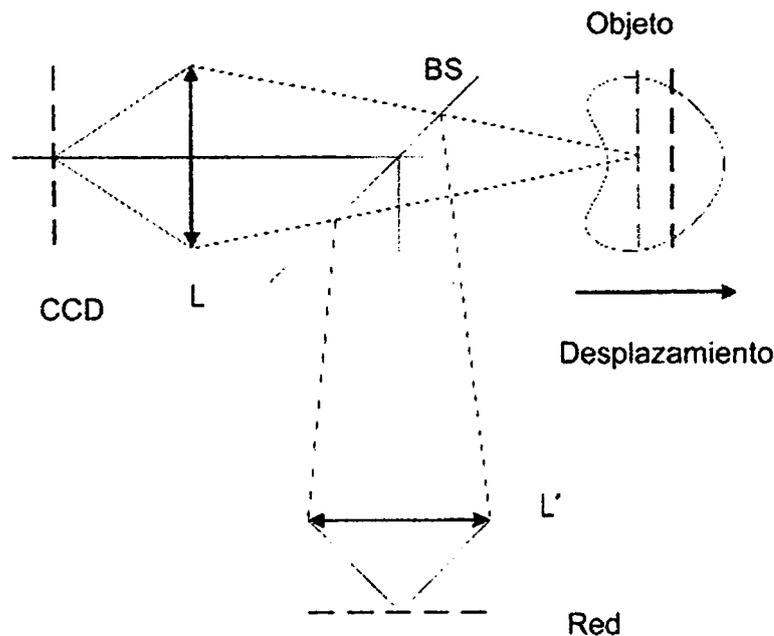


Figura 5.1. Esquema experimental para proyectar franjas sobre un objeto, utilizando iluminación incoherente.

La ventaja práctica más importante de este sistema confocal es que no existen puntos ocultos sobre la superficie de observación. Dichos puntos surgen cuando la observación y

la iluminación se realizan desde diferentes direcciones.

5.3 Experiencias utilizando objetos 2D

Realizamos una serie de experiencias para estudiar el comportamiento de la visibilidad usando un objeto plano (Figura 5.2.a)) perpendicular al eje óptico. Primero, registramos con una cámara CCD un conjunto de 20 imágenes de 128x128 pixels con desenfoque progresivo en ambas direcciones, las digitalizamos en 256 niveles de gris y las almacenamos en la memoria de una computadora. La imagen número 10 fue juzgada por el observador como la mejor enfocada y las restantes fueron obtenidas incrementando la distancia a la cámara en pasos de 3mm. En este primer caso no proyectamos franjas sobre el objeto.

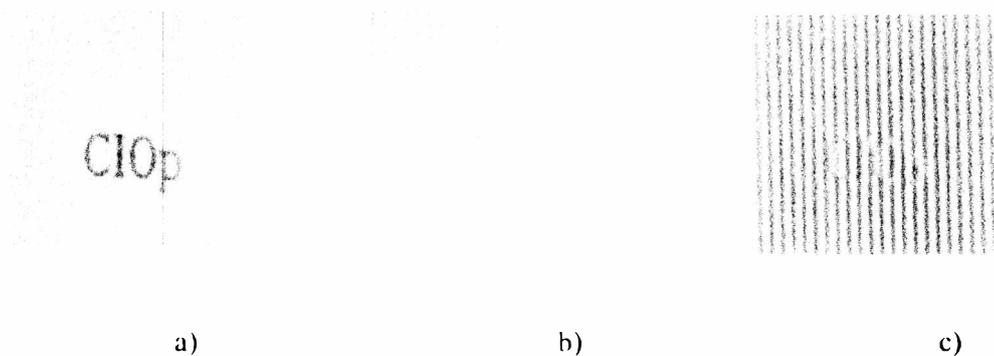


Figura 5.2. Una de las imágenes de un objeto plano utilizado en la experiencia. a) Imagen tomada con iluminación común. b) Imagen tomada con franjas proyectadas. c) Resultado de aplicar el operador derivada a la imagen b).

A continuación, proyectamos franjas de alto contraste a través de un beam splitter como se muestra en la Figura 5.1, para introducir los detalles de alta frecuencia. El f /número de la CCD y del dispositivo de proyección se eligieron en su menor valor (2.8) para minimizar la profundidad de campo y el objetivo de la lente utilizada fue de 50 mm. En estas condiciones se tomó un nuevo conjunto de 20 imágenes en las mismas posiciones que sus correspondientes imágenes registradas con iluminación común. Las imágenes fueron obtenidas desplazando el objeto para mantener la cámara y la iluminación en el mismo lugar. La Figura 5-2.b) muestra una de ellas.

Debido a que el sistema de proyección no era de muy buena calidad, las franjas proyectadas tenían bajo contraste. Para incrementar entonces el contenido relativo de frecuencias altas le aplicamos a las imágenes la derivada, mediante una convolución de las mismas con un núcleo derivativo en la dirección perpendicular a las franjas proyectadas. El operador digital sustrae los valores de los pixels vecinos en una dirección. Como el resultado de dicha operación puede ser negativo y de esta manera no representar un valor físico de la intensidad, el operador requiere de la definición de una constante aditiva que la llamamos ϕ , para poder ser representado como imagen. La Figura 5-2.-c) muestra una de estas imágenes, en este caso se usó $\phi=100$.

Calculamos la visibilidad para cada pixel usando ventanas de 11x11 pixels (de manera que estuviera contenido un período completo de las franjas) y luego las sumamos en cada una de las imágenes. La Figura 5.3 muestra la visibilidad resultante para cada imagen en función del desenfoque para los siguientes casos: a) desenfoque simple sin proyección de franjas, b) desenfoque doble con proyección de franjas y c) desenfoque doble con proyección de franjas derivadas.

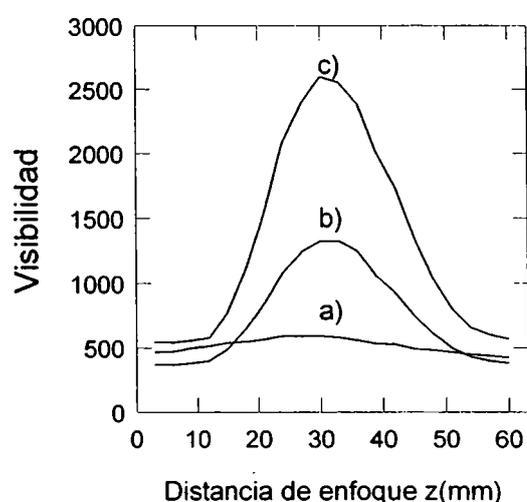


Figura 5-3. Visibilidad resultante para cada imagen en función de la distancia de enfoque. a) desenfoque simple. b) doble desenfoque con proyección de franjas. c) doble desenfoque con proyección de franjas derivadas.

El gráfico muestra un máximo cercano a la imagen que fue juzgada por el

experimentador como la mejor enfocada ($z=33$ cm). Se puede observar que los valores de la visibilidad, en promedio, son más grandes en las imágenes a las cuales le aplicamos el operador derivada a las franjas proyectadas. La variación de la visibilidad con la distancia de enfoque es más abrupta en este caso, lo que indica que el procedimiento es sensible a las frecuencias altas.

Anteriormente, mencionamos que cuando aplicamos el operador derivada a las imágenes existe una indeterminación en la constante ϕ . La Figura 5.4 muestra este comportamiento para diferentes valores de dicha constante.

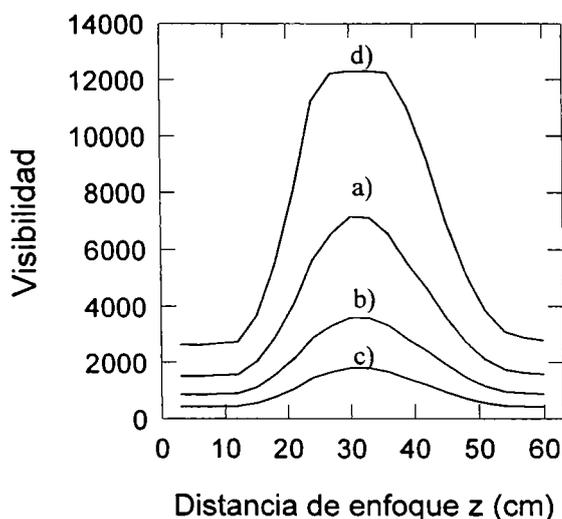


Figura 5.4. Visibilidad resultante para cada imagen en función del desenfoque para diferentes valores de la constante ϕ . a) $\phi=I_{\min}$. b) $\phi=2I_{\min}$. c) $\phi=4I_{\min}$. d) $\phi=0.5I_{\min}$.

En dicha figura se puede observar que cuando los valores de la constante son altos la visibilidad tiende a ser insensible al desenfoque. Para valores pequeños de ϕ , la visibilidad tiende a tomar valores más altos pero permanece constante en un amplio entorno de la imagen mejor enfocada. El valor más adecuado de ϕ es dependiente de la escena.

5.4 Estudio teórico de la visibilidad en función del desenfoque de un objeto sinusoidal doblemente desenfocado.

En la sección precedente estudiamos desde el punto de vista experimental, la visibilidad en función del desenfoque en el caso de un objeto doblemente desenfocado que contiene franjas proyectadas. A continuación haremos un desarrollo teórico que avala los resultados experimentales cuando el objeto bajo estudio sufre un doble desenfoque.

Realizamos una similitud teórica del arreglo experimental y consideramos que:

- a) La entrada al sistema óptico está dada por una fuente sinusoidal unidimensional dada por la Ecuación (4.5) que da origen a una distribución de intensidad a la salida dada por la Ecuación (4.10).
- b) Dicha salida es tomada como un nuevo objeto de entrada y hallamos la intensidad a la salida simulando un doble desenfoque.

La expresión de la intensidad de la senoide doblemente desenfocada se obtiene siguiendo los pasos realizados en el capítulo anterior, Sección 4.8. Luego de un poco de álgebra la intensidad a la salida tiene la siguiente dependencia:

$$I_2(x) = \frac{l^4}{2} \left\{ 1 - \left[\left(1 - \frac{f_1}{f_0} \right)^2 \operatorname{sinc}^2 \left(\left(\frac{\left(\frac{-z}{(2F+z)2F} \right) l^2}{\lambda} \right) \left(\frac{f_1}{f_0} \right) \left(1 - \frac{f_1}{f_0} \right) \right) \right] \cos 4\pi f_1 x \right\} \quad (5.1)$$

Para incrementar el contenido relativo de las altas frecuencias aplicamos el operador derivada a la Ecuación (5.1). Debido a que el resultado queremos representarlo como una imagen y la intensidad es una magnitud que no puede ser negativa, le sumamos una constante ϕ y se obtiene:

$$I_3(x) = -2\pi f_1 l^4 \left(1 - \frac{f_1}{f_0}\right)^2 \operatorname{sinc}^2 \left(\left(\frac{\left(\frac{-z}{(2F+z)2F} \right) l^2}{\lambda} \right) \left(\frac{f_1}{f_0} \right) \left(1 - \frac{f_1}{f_0} \right) \right) + \phi \quad (5.2)$$

Una vez halladas las distribuciones de intensidades I_2 e I_3 estamos en condiciones de hacer el estudio de la visibilidad. Haciendo uso de su definición, Ecuación (4.4), encontramos que la visibilidad en función del desenfoque z , para la intensidad I_2 está dada por:

$$V_2(z) = \left(1 - \frac{f_1}{f_0}\right)^2 \operatorname{sinc}^2 \left(\left(\frac{\left(\frac{-z}{(2F+z)2F} \right) l^2}{\lambda} \right) \left(\frac{f_1}{f_0} \right) \left(1 - \frac{f_1}{f_0} \right) \right) \quad (5.3)$$

El comportamiento de la visibilidad en función de z para el caso de doble desenfoque con franjas derivadas está dado por:

$$V_3(z) = \left(\frac{2\pi f_1 l^4 \left(1 - \frac{f_1}{f_0}\right)^2}{\phi} \right) \operatorname{sinc}^2 \left(\left(\frac{\left(\frac{-z}{(2F+z)2F} \right) l^2}{\lambda} \right) \left(\frac{f_1}{f_0} \right) \left(1 - \frac{f_1}{f_0} \right) \right) \quad (5.4)$$

En la Figura 5.5 se grafica la visibilidad en función del desenfoque. La curva a) corresponde al caso de un desenfoque simple tratado en el capítulo 4 y se ha incluido para comparar los resultados. En la curva b) se puede observar que V_2 , la visibilidad del objeto con doble desenfoque, toma su máximo valor en $z=0$, la cual corresponde a la posición de enfoque correcto y decrece a medida que aumenta z . En la curva c) V_3 , la visibilidad del objeto con doble desenfoque con franjas derivadas, toma un valor máximo que es mayor que en la curva a) y que en la b).. El valor de la constante ϕ fue elegido de manera tal que la intensidad I_3 tome su mínimo valor que es igual a cero. Llamamos a este valor óptimo y está dado por:

$$\phi_{opt} = 2\pi f_1 l^4 \left(1 - \frac{f_1}{f_0}\right)^2 \quad (5.5)$$

Los resultados obtenidos en la Figura 5.5 son consistentes con los hallados experimentalmente y mostrados en la Figura 5.3

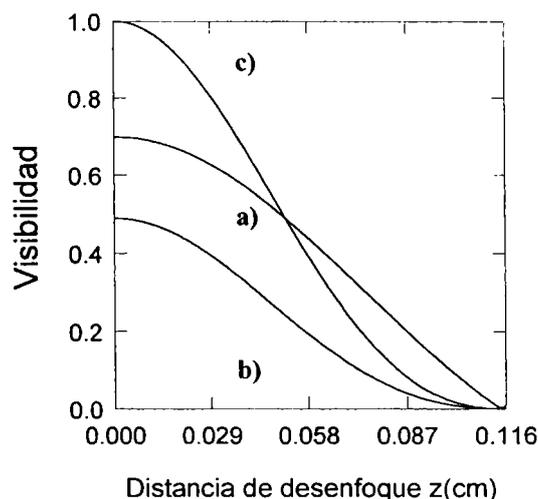


Figura 5.5. Gráfico de la visibilidad en función del desenfoque z (cm). a) Desenfoque simple. b) Doble desenfoque. c) Doble desenfoque con franjas derivadas.

La Figura 5.6 muestra la visibilidad en función del desenfoque de la derivada de las franjas doblemente desenfocadas para varios valores de la constante aditiva. Cuando el valor de ϕ es cero, la visibilidad es 1, como se puede deducir de la ecuación (5.3). Un pequeño valor de ϕ produce ceros en la imagen derivada resultante dando origen a valores altos de la visibilidad aun cuando la distancia desenfocada no es despreciable. Inversamente, valores altos dan resultados muy bajos de visibilidad haciendo por lo tanto insensible al desenfoque. De todas maneras, el valor de ϕ más adecuado es dependiente de la escena.

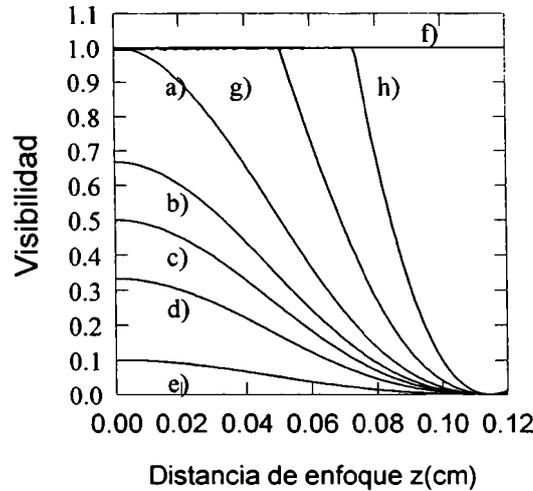


Figura 5.6. Gráfico de la visibilidad en función del desenfoque z (cm) para diferentes valores de la constante ϕ . a) $\phi=I_{\min}$; b) $\phi=1.5 I_{\min}$; c) $\phi=2I_{\min}$; d) $\phi=5I_{\min}$; e) $\phi=I_{\min}$; e) $\phi=I_{\min}$; f) $\phi=I_{\min}$; g) $\phi=I_{\min}$; h) $\phi=I_{\min}$

Las curvas mostradas en la Figura 5.6 han sido calculadas usando: $l=2\text{ cm}$, $\lambda=0.5\mu$, $F=20\text{ cm}$, $f_1/f_2=0.3$, y $NA=0.05$. Se puede observar, que el comportamiento de la visibilidad mostrado en la Figura 5.6 coincide con el de la Figura 5.4.

En todos los casos, la visibilidad fue calculada en regiones que contenían un período completo de las franjas proyectadas. Si ahora usamos regiones más pequeñas para calcular la visibilidad, simulando ventanas más pequeñas, encontramos el comportamiento mostrado en la Figura 5.7, para las franjas derivadas doblemente desenfocadas. La curva 5.7 a) representa el comportamiento de la visibilidad cuando el tamaño de la ventana es tal que contiene un período completo de las franjas. Las curvas 5.7 b) y c) usan ventanas, cuyos dimensiones son 3 y 5 veces menores que en el caso a). Se puede observar que para una cierta región del objeto, a medida que el tamaño de la ventana decrece, el valor máximo de la visibilidad también decrece. Esto es debido al hecho de que los valores máximos y mínimos de la intensidad que están dentro de la ventana, no son los valores extremos absolutos tomados por la intensidad en un período completo. Por lo tanto el tamaño óptimo de la ventana debería ser tal que contenga al menos un período.

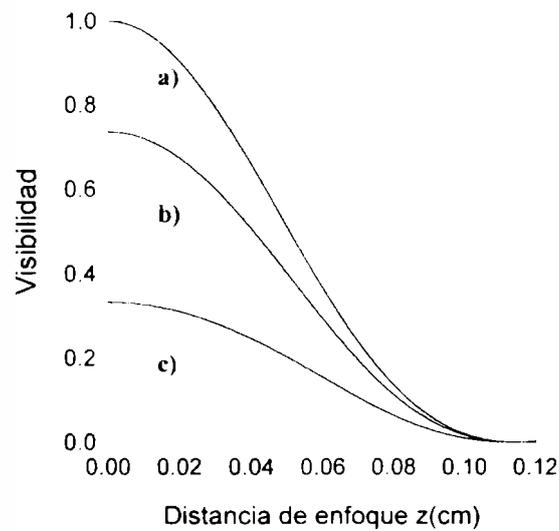


Figura 5.7. Gráfico de la visibilidad en función de la distancia de enfoque para distintos tamaños de ventanas. a) Un período. b) Tres veces menor que a), c) Cinco veces menor que a)

5.5 Corrección del desenfoque usando imágenes de objetos 3-D.

El siguiente conjunto de experiencias fue realizado para construir una imagen de un objeto 3D con el desenfoque parcialmente corregido.

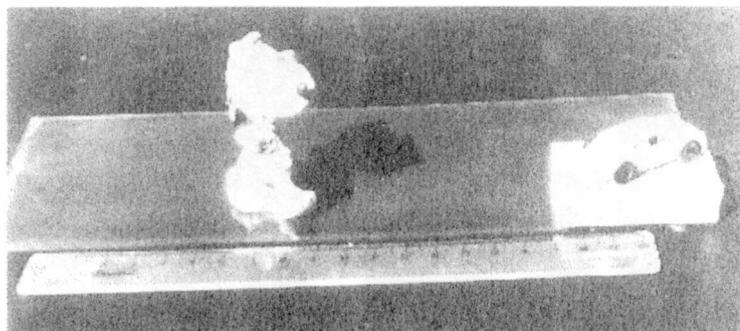


Figura 5.8. Vista lateral del objeto 3-D utilizado en la experiencia.

Comenzamos registrando dos conjuntos de imágenes de un objeto 3D con desenfoque progresivo. Una vista lateral de la escena se muestra en la Figura 5.8. Un conjunto $I(x,y,z)$ fue obtenido con iluminación común y el otro $I_p(x,y,z)$ proyectando ranjas sobre el objeto, bajo el esquema experimental mostrado en la Figura 5.1. Dos ejemplos de

cada conjunto se muestran en la Figura 5.9 y 5.10 respectivamente. A cada imagen del conjunto $I_p(x,y,z)$ se le aplicó el operador derivada, obteniendo como resultado el conjunto $I'(x,y,z)$. Dos ejemplos de estas imágenes se ven en la Figura 5-11.

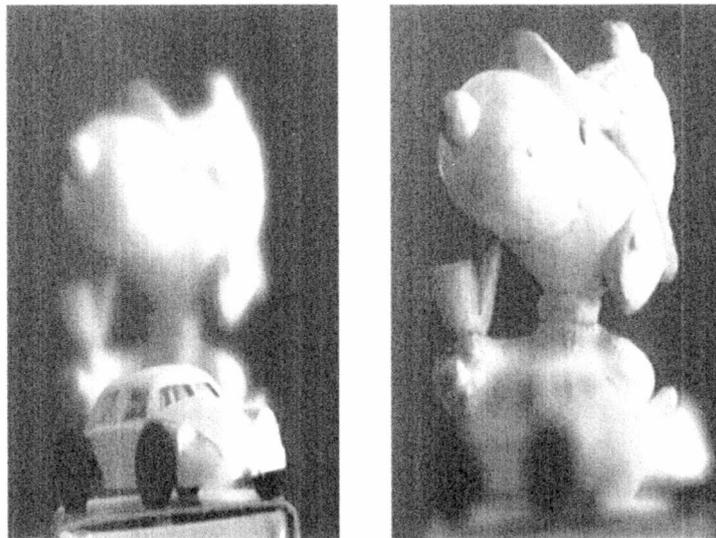


Figura 5.9. Imágenes de dos pequeños juguetes que corresponden a dos posiciones diferentes de enfoque.

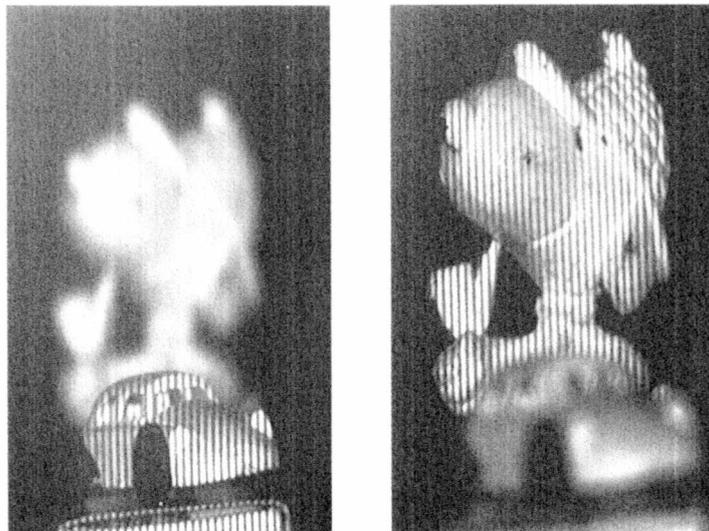


Figura 5.10. Las mismas imágenes que se muestran en la figura anterior (en la misma posición). con franjas proyectadas.

cada conjunto se muestran en la Figura 5.9 y 5.10 respectivamente. A cada imagen del conjunto $I_p(x,y,z)$ se le aplicó el operador derivada, obteniendo como resultado el conjunto $I'(x,y,z)$. Dos ejemplos de estas imágenes se ven en la Figura 5-11.

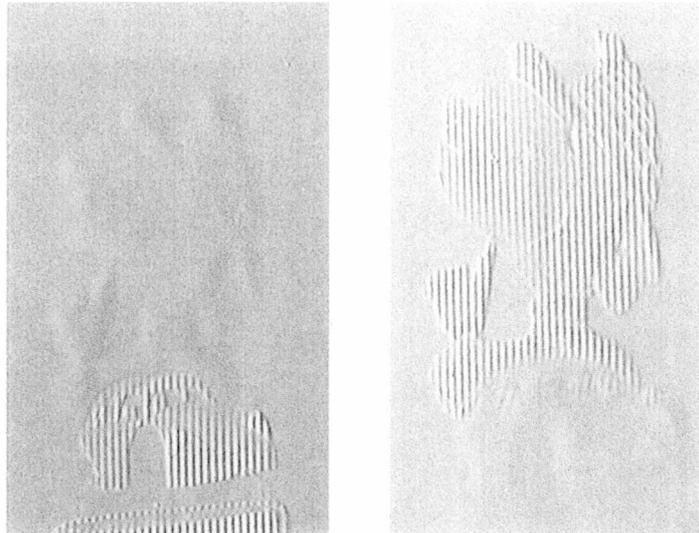


Figura 5.11. Las mismas imágenes de la Figura 5.7, sobre las cuales se les aplicó el operador derivada.

Los conjuntos fueron procesados de la siguiente manera:

- a) El conjunto $I(x,y,z)$ fue procesado usando el criterio de máxima visibilidad [4.1] sin tener en cuenta la iluminación estructurada. La Figura 5.12 muestra la imagen procesada con el fin de comparar los resultados.
- b) Se aplicó el algoritmo que calcula la visibilidad en cada imagen del conjunto $I_p(x,y,z)$. En las imágenes procesadas se buscaron los valores máximos de visibilidad en la dirección del desenfoque (z) para cada posición de pixel. La imagen corregida se construyó usando los valores de intensidad en el conjunto obtenido con iluminación común, correspondientes a los máximos de visibilidad encontrados anteriormente. El resultado se muestra en la Figura 5.13.
- c) Para mejorar el resultado final, se usaron las imágenes obtenidas con el operador derivada. Se repitió el proceso b) usando el conjunto $I'(x,y,z)$ en lugar de $I_p(x,y,z)$. El resultado se muestra en la Figura 5.14.

Comparando las imágenes resultantes corregidas se puede concluir que la adición de

iluminación estructurada sobre la escena reduce y elimina las sombras periféricas.

En todos los casos el tamaño de las ventanas utilizadas fue de 21x21 pixels y el número de imágenes procesadas fue 10.



Figura 5.12. Imagen corregida usando iluminación común.



Figura 5.13. Imagen corregida usando franjas proyectadas.



Figura 5.14. Imagen corregida usando las imágenes con franjas proyectadas, a las cuales se les aplicó el operador derivada.

La resolución longitudinal en la imagen procesada está limitada por el número de imágenes que componen el conjunto. Para mejorar la resolución, se necesita un mayor

número de imágenes incrementándose de esta manera el tiempo computacional.

Usamos también patrones de speckle, redes de Ronchi de distintas frecuencias espaciales y redes rectangulares como medios para introducir la iluminación estructurada en la escena. Todos los resultados obtenidos fueron consistentes.

5.6 Bibliografía

- [5.1] A. E Savakis., H. J. Trusell “On the accuracy of PSF representation in image restoration”. IEEE Trans. Image Process. **2**, 252-259, (1993)..
- [5.2] J. W. Goodman, *Introduction to Fourier Optics*, MacGraw-Hill Book Company Cap 6, 120-125, (1968).
- [5.3] R. W Ditchburn. “Focusing of instrument”, *Optica Acta*, **27**, 713-715, (1980).
- [5.4] G. R Lockheed., M. L Wolbarsht., “Toying with the moon illusion”, *Applied Optics*, **30**, 3504-3507, (1981).
- [5.5] P. Torroba, N. Cap H.. Rabal “Defocus correction using structure illumination”, *Optik*, **108**, 68-77, (1998).
- [5.6] K. Engelhardt, G. Häusler. “Acquisition of 3-D data by focus sensing”, *Applied Optics*, **27**, 4684-4689, (1988)..

Capítulo 6

El modelo visual de Cornsweet y Yellott

6.1 Introducción

En el capítulo 2 estudiamos el problema de disminuir el ruido en imágenes aplicando una serie de filtros no lineales y en éste queremos saber como se comporta el mismo en la naturaleza. Con este motivo empezamos a analizar como influye el ruido en la visión humana. Estudiamos algunos modelos visuales y cómo responden éstos a distintos efectos visuales.

Comenzamos describiendo el modelo de Marr y Hildreth. A continuación, tratamos el comportamiento del sistema visual humano y de los modelos lineales invariantes ante traslaciones cuando se producen cambios en los niveles de iluminación. Luego, se presenta la idea básica del modelo visual de Cornsweet, también llamado Suma Espacial Dependiente de la Intensidad (IDS) u Operadores de volumen constante. Este modelo crea bandas de Mach sin inhibición lateral (con funciones de punto extendido puramente

positivas) y simultáneamente modela un sorprendente rango de fenómenos visuales adicionales. Los operadores de volumen constante son no lineales y muestran que un mecanismo de un solo parámetro fijo puede reproducir muchas propiedades de la visión que son usualmente atribuidas a los procesos retinales, incluyendo los cambios en aquellas propiedades que se producen por cambios en el nivel de iluminación. Además, se analiza la suposición sobre la cual está basado el modelo y se describe su formalismo. Por último se comparan la respuesta del sistema visual y el modelo IDS.

6.2 Modelo visual de Marr y Hildreth

En esta sección se describirá la aproximación computacional de la percepción visual, como fue introducida por David Marr en su trabajo "Vision" [6.1]. La aproximación computacional trata la percepción visual como una operación de procesamiento de la información. Esta operación tiene como entrada la distribución de intensidad sobre la retina y la salida será una descripción simbólica del mundo externo. *Simbólica* significa que aunque todos parecen tener una imagen del mundo inmediatamente después de abrir los ojos, no existen estas imágenes en nuestras cabezas, solamente hay activación de patrones neuronales. Por lo tanto el mundo debe ser representado por símbolos

El propósito de la visión es obtener una descripción de la escena y para su estudio se lo fracciona arbitrariamente en subprocesos. Una de las características más importantes de la teoría de Marr es su énfasis en que deben construirse diferentes representaciones a partir de la información de la imagen retinal. Una primera representación, *the primal sketch*, (bosquejo primal) captura la estructura bidimensional de la imagen. El bosquejo primal es una descripción de los cambios significativos en la imagen, tales como bordes y barras presentes y su arreglo espacial en dos dimensiones. La segunda representación, que es tridimensional, denominada sketch de dimensión 2.5, describe como están orientadas las superficies con respecto al observador y cuan lejos están del mismo. Finalmente, la representación llamada modelo sólido o 3D, es una representación volumétrica de los cuerpos de la escena.

Como la tarea del sistema visual es relacionar “imagen” y “mundo”, el primer paso es localizar estructuras en la imagen que deben reflejar aspectos significativos del “mundo”. Uno de tales aspectos significativos es el contorno de un objeto, o el contorno entre una superficie de una región y otra. Estos contornos producirán discontinuidades de intensidad (bordes) en la imagen. Entonces, un primer paso en el análisis de una imagen debe involucrar la detección de bordes.

Existen muchas maneras diferentes de encontrar bordes en una imagen. Una forma es derivar la función intensidad respecto de la ubicación espacial, de manera de obtener el cambio de la intensidad a través de la imagen. El borde puede entonces, estar localizado donde exista el pico en la derivada primera de la distribución de intensidad. Ahora, si derivamos nuevamente para obtener la derivada segunda, el pico de la derivada primera se convierte en un cruce por cero. Así, los bordes pueden localizarse por los picos de la derivada primera de la distribución de intensidad o por los cruces por cero en la derivada segunda.

Sin embargo, la detección de bordes no se realiza directamente. Naturalmente, los bordes pueden ser abruptos o estar extendidos espacialmente. Esto es desafortunado porque una manera fácil de obtener la operación de derivar es comparar la intensidad de la luz que cae en cada mitad de la máscara mostrada en la Figura 6.1 Si hay alguna diferencia, entonces debe estar presente alguna clase de borde. Sin embargo, la máscara mostrada en la figura 6.1 fallaría en detectar un cambio de intensidad muy gradual y se necesitarían máscaras de mayor tamaño, que a su vez no detectarían bordes agudos.

| | |
|----|---|
| -1 | 1 |
| -1 | 1 |

Figura 6.1. Máscara que podría detectar un borde vertical.

Marr and Hildreth sugirieron [6.2] que este problema podría ser resuelto si los bordes fueran detectados simultáneamente en diferentes escalas. Esto se obtendría borroneando la imagen con distintos grados y luego buscando los bordes independientemente dentro de estos canales con distinto borreo. El borreo de una imagen involucra dejar pasar solamente las componentes de bajas frecuencias espaciales. Una manera de obtener esto es dividir la imagen original en pequeñas áreas y reemplazar la intensidad original por la intensidad promedio dentro de cada una de estas pequeñas áreas. El problema es que este proceso puede introducir altas frecuencias espaciales espúreas o ruido que no está presente en la imagen original. Marr y Hildreth sugirieron que el uso de una función Gaussiana (normal) bidimensional minimizaría este problema. Esta función trabaja reagrupando las intensidades dentro de una región circular de la imagen, de modo que las intensidades en el centro del área circular tienen un peso mayor que aquellas que están sobre la periferia, de acuerdo con la función Gaussiana. El tamaño del círculo se especifica por la desviación estándar de dicha función. Los diferentes canales de borreo son creados por medio de la convolución de la imagen original con funciones Gaussianas de diferentes anchos. Para detectar los cambios de intensidad dentro de cada canal, Marr y Hildreth toman la derivada segunda de la imagen borreada con el operador Laplaciano (∇^2). El operador Laplaciano fue elegido para la diferenciación porque es no direccional, y puede ser usado para buscar los cambios de intensidad en todas las direcciones. Entonces, los primeros pasos en la detección de bordes, de acuerdo a Marr y Hildreth involucra la localización de los cruces por cero independientemente en diferentes escalas espaciales convolucionando la imagen I con operadores que tienen diferente scaleo ($\nabla^2 G$).

Los cruces por cero proveen una manera natural de moverse desde una representación continua, como son los valores de intensidad de una imagen bidimensional, a una representación simbólica discreta. Un hecho interesante sobre esta transformación es que probablemente no se incurra en pérdida de información. El argumento que soporta esta idea todavía no es seguro y está expresado en el teorema de Logan [6.3]. Este teorema establece que, siempre que ciertas condiciones sean satisfechas, una señal pasabanda de una octava de ancho puede ser completamente reconstruida (a menos de una constante multiplicativa) a partir de sus cruces por cero. La demostración del teorema es dificultosa,

pero consiste esencialmente en mostrar que si la señal tiene un ancho de banda menor que una octava, entonces debe cruzar el eje x al menos tantas veces como requiere el teorema del muestreo.

Lamentablemente, el teorema de Logan no es suficientemente fuerte. Primero, la aplicación visual de los cruces por cero es bidimensional y es difícil extender los argumentos de muestreo del caso unidimensional al bidimensional. Segundo, el operador $\nabla^2 G$ es un filtro pasabanda con ancho mayor de una octava. Por otro lado, se tiene información extra, es decir, los valores de las pendientes de las curvas cuando ellas cruzan por cero, puesto que éstas corresponden al contraste de los bordes presentes de la imagen. Una aproximación analítica del problema parece ser muy complicada, pero en una investigación empírica, se encontraron evidencias que avalan el punto de vista de que una imagen filtrada bidimensional puede ser reconstruida a partir de sus cruces por cero y por sus pendientes [6.4]

6.3 Modelos visuales y operadores lineales

Varias teorías actuales sobre modelos visuales están basadas en operadores que transforman la imagen de entrada $I(x,y)$ en una de salida $O[I](x,y)$ por medio de una convolución de I con alguna función de punto extendido; en otras palabras, operadores lineales invariantes frente a traslaciones. En particular, una forma estándar para modelar los primeros pasos del modelo visual retinal es un operador lineal, circularmente simétrico de la forma:

$$O[I](x,y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x',y') S[(x-x')^2 + (y-y')^2] dx' dy' \quad (6.1)$$

donde la respuesta impulsiva tiene un perfil parecido a un sombrero mejicano, S es igual a la diferencia de dos Gaussianas, o S es el Laplaciano negativo de una Gaussiana como el modelo de Marr y Hildreth [6.2]. Para cada respuesta impulsiva el operador lineal representado en la ecuación (6.1) produce un corte en las bajas frecuencias en el módulo de

la función de transferencia y como consecuencia un realce de bordes (la respuesta a un borde contiene un mínimo local sobre el lado de intensidad más baja y un máximo local del lado de mayor intensidad). Ambos efectos se parecen a importantes propiedades de la visión humana que son: las bandas de Mach que se observan en los bordes [6.5] y la caída de las bajas frecuencias que presenta la curva de sensibilidad al contraste obtenida psicofísicamente por Campbell y Robson [6.6].

Estos fenómenos perceptivos han sido generalmente atribuidos a procesos de inhibición lateral en la retina. Estos procesos están representados en el modelo lineal dado en la ecuación (6.1) por la presencia de lóbulos negativos en la respuesta impulsiva. Dentro del marco de la teoría de los sistemas lineales estos lóbulos negativos son la manera natural de crear las bandas de Mach y otros efectos de filtros pasa-altos, porque la respuesta a los bordes de los operadores lineales invariantes frente a traslaciones, debe ser monótona si su respuesta impulsiva es enteramente positiva.

Por lo tanto cuando el proceso retinal es modelado por operadores lineales, las bandas de Mach y la inhibición lateral están inseparablemente ligados en una relación sí-sólo-sí (si existe uno debe existir el otro y reciprocamente). Pero este tipo de relación existe para operadores lineales. A continuación comenzaremos a tratar con una clase simple de operadores no lineales (operadores de volumen constante) que crea bandas de Mach sin inhibición (con funciones de punto extendido puramente positivas) y simultáneamente modela un sorprendente rango de fenómenos visuales adicionales.[6.7]

6.3.1 Motivación psicofísica

Desde el punto de vista psicofísico, los operadores de volumen constante son interesantes porque ellos muestran que un mecanismo de un solo parámetro fijo puede reproducir muchas propiedades de la visión que son usualmente atribuidas a los procesos retinales, incluyendo los cambios en aquellas propiedades que se producen por cambios en el nivel de iluminación. Cuando estos niveles crecen desde luz muy baja hasta luz normal, la visión humana sufre tres cambios importantes:

El primero, mejora la resolución espacial: la función de sensibilidad al contraste (CSF) se corre hacia la derecha a lo largo del eje de las frecuencias espaciales [6.8], de modo que se incrementa la mayor frecuencia que puede ser resuelta (acuidad visual), con la raíz cuadrada de la iluminancia retinal.

Segundo, la forma de la CSF cambia desde un filtro pasabajo a un filtro pasabanda [6.8].

Tercero, la visibilidad de los objetos se hace independiente de su luminancia absoluta y comienza a depender solamente de su contraste con el fondo. Esto es, la detectabilidad de los cambios de luminancia de I a $I+cI$ en algunas regiones del campo visual depende eventualmente solamente del contraste c , una vez que I sea suficientemente alta. Este tercer hecho de la visión es la ley de Weber [6.9].

Los operadores lineales no tienen en cuenta estos cambios relacionados con las variaciones de la iluminación. Ellos predicen el crecimiento en la resolución espacial con la iluminación retinal, simplemente porque incrementar la iluminación mejora la relación señal ruido de los cuantos atrapados por los fotorreceptores. Pero esta mejora sólo produce que la CSF se traslade hacia arriba y no predice el corrimiento horizontal a lo largo del eje de las frecuencias espaciales. Para obtener dicho corrimiento en el modelo lineal, el ancho de la respuesta impulsiva debe decrecer cuando la iluminación retinal crece y como los operadores lineales no se ajustan automáticamente de esta manera, debe por lo tanto postularse otro mecanismo adicional para que sea posible dicho comportamiento. Lo mismo vale para el cambio de un filtro pasabajo a otro pasabanda: para un operador lineal, este cambio implica que la forma de la respuesta impulsiva sea diferente en los distintos niveles de iluminación. En otras palabras, un operador lineal con parámetro fijo puede sólo modelar la curva CSF para un solo nivel de iluminación y un mecanismo no lineal debe ser propuesto para tener en cuenta la forma en que la CSF cambia desde un nivel a otro. La misma situación se presenta con la ley de Weber.

La visión espacial analizada desde este punto de vista de los sistemas lineales, parecería estar relacionada con un conjunto complejo de cambios a medida que los niveles de iluminación crecen desde muy bajos a niveles de luz diurna. Los operadores de volumen constante (CV) pueden ser vistos como consecuencia de una simple operación cuyo propósito real es maximizar la resolución espacial en vista del ruido fotónico. Esencialmente, un operador de volumen constante es un mecanismo que ajusta el área de la función de punto extendido de acuerdo al nivel de luz existente y produce que la resolución espacial crezca con la iluminancia retinal. Intuitivamente, uno esperaría que la única consecuencia fuera el aumento de la acuidad visual porque así fue construido. Sin embargo, esta misma operación da origen al realce de bordes, a otros efectos de filtro pasa altos y también a la ley de Weber. En otras palabras, todas las propiedades básicas de la visión pueden ser reproducidas por un operador de volumen constante de un solo parámetro fijo, motivado por el ruido intrínseco de la luz.

La idea básica de los operadores de volumen constante es que cada punto de la imagen de entrada crea una función de punto extendido cuya área varía inversamente con la intensidad de la imagen de entrada. Este mecanismo fue sugerido originalmente por el hecho psicofísico de que el tamaño del área de la suma espacial en la retina humana se angosta a medida que crece el nivel de luminancia media [6.10]. Consecuentemente, estos operadores inicialmente fueron denominados operadores de “Suma espacial dependiente de la intensidad”. Sin embargo, posteriormente se adoptó el nombre de operador de volumen constante porque contiene la diferencia fundamental entre los operadores lineales y los no lineales.

6.3.2 Ruido fotónico y operadores de volumen constante

Muchos experimentos psicofísicos y fisiológicos pueden ser interpretados de la siguiente manera: la luz que cae sobre cualquier punto de la retina crea un efecto excitatorio en los puntos vecinos y esta excitación lateral se combina aditivamente con la excitación directa producida por la luz [6.10]. Evidencias psicofísicas indican también que la extensión de la excitación lateral (el tamaño del área de la suma espacial) se incrementa cuando la

iluminancia retinal decrece y disminuye cuando la iluminancia aumenta [6.11]. Un motivo para que se produzcan estos resultados podría ser que el sistema visual tuviera una respuesta adaptativa al ruido intrínseco de la luz. Cuando los niveles de luz son muy bajos la suma sobre grandes áreas ayudaría a la captura fotónica, permitiendo que la detección del contraste sea más confiable. Por otro lado cuando los niveles de luz crecen, la resolución de los detalles finos se hace físicamente posible y las áreas no necesitan ser tan grandes para tener un buen contraste. Por lo tanto, el sistema ideal es aquel para el cual, cuando el nivel de luz es bajo un gran grupo de receptores de la retina suman sus señales y cuando el nivel de luz se incrementa, el tamaño de estos grupos se hace menor, en correspondencia con las propiedades estadísticas de la luz.

Rose [6.12] y de Vries [6.13] señalaron varios años atrás que la naturaleza aleatoria de la absorción de fotones impone un límite en la detección del contraste en el sistema visual. Si la iluminancia (valor medio de cuantos absorbidos por receptor por unidad de tiempo) en alguna región de una imagen es I , la cantidad total de cuantos absorbidos en un área que contiene R receptores para cualquier intervalo de tiempo T es una variable aleatoria de Poisson con media y varianza igual a IRT . Esta relación estadística impone una restricción en la detección del contraste. Si suponemos que quiere ser detectado un cambio de iluminancia de I a $I+cI$ ($c>0$) entonces, usando la aproximación normal de la distribución de Poisson, la probabilidad de detección es $N\left\{c\left[\frac{RI}{2+c}\right]^{1/2}\right\}$, donde N es la función de distribución normal dada por:

$$N(z) = \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx$$

Se puede demostrar que para que sea satisfecha esta condición con un error del orden del 1%, suponiendo que el sistema visual es un detector perfecto limitado solamente por las fluctuaciones de los cuantos, se debe verificar que $N\left\{c\left[\frac{RI}{2+c}\right]^{1/2}\right\} \geq 2.3$ o sea que uno

necesita que la cantidad media de cuantos absorbidos por unidad de tiempo satisfaga una relación del orden de

$$IR > 10/c^2 \quad (6-2)$$

De modo que una detección confiable del 100% ($c=1$) de contraste requiere una cantidad media total del orden de 10 cuantos, 10% de contraste requiere 1000, etc. Algunas estimaciones [6.14] indican que la cantidad media de cuantos absorbidos por un solo fotorreceptor es alrededor de 4 fotones por segundo por troland (nivel de iluminancia retinal que resulta cuando una superficie con una luminancia de 1 candela/m² es vista a través de una pupila de un área de 1 mm²) de iluminancia retinal. Visualmente, un segundo completo es un intervalo un poco extenso: el período de integración de los fotorreceptores es más parecido a un cuarto de segundo. Se puede decir que un receptor captura un promedio de un cuanto por “unidad de tiempo visual” por cada troland de iluminancia retinal. El rango de trabajo del sistema visual se expande alrededor de 10 unidades logarítmicas, desde 10⁻⁴ a 10⁶ troland. Sobre este rango, de acuerdo con la relación (6.2), el contraste umbral nunca excede el 100% y generalmente es mucho menor. Cuando estos parámetros se combinan en la ecuación anterior, es obvio que sobre su rango operacional el sistema visual debe detectar contraste sobre la base de sumar los cuantos capturados por muchos receptores: debe activar al menos 10⁵ receptores para detectar 100% de contraste a 10⁻⁴ trolands, al menos 25000 para detectar 20% a 10⁻² trolands, etc. Es decir solamente cuando la iluminancia media alcanza 1000 trolands, puede ser detectado el 1% de contraste sobre la base de los cuantos capturados por un solo detector.

La suma espacial puede ser vista como un mecanismo que debe realizar el sistema visual, forzado por la estadística de la luz. Este mecanismo produce un crecimiento en los cuantos capturados hasta un nivel aceptable aumentando el área efectiva de recolección. Al mismo tiempo, sin embargo, la suma espacial limita la resolución, puesto que dos puntos no pueden ser resueltos si ambos caen dentro de la misma área. Entonces, las consideraciones de sensibilidad al contraste determinan que las áreas a sumar deben ser extensas, mientras que las de resolución espacial conducen a sumar áreas pequeñas. El

requerimiento de detección de la señal expresado en la ecuación anterior, sugiere cómo debe ser resuelto este conflicto: para maximizar la resolución espacial a través de diferentes niveles de luz manteniendo una detección confiable constante para algún nivel de contraste dado, el área de recolección de cuantos R , debe variar inversamente con el nivel de luz I . Los operadores de volumen constante son propuestos para producir este efecto automáticamente.

Un punto interesante es que este mecanismo produce realce de bordes y la ley de Weber que son efectos que no parecen estar relacionados usualmente con el ruido fotónico.

6.4 Modelo visual de Cornsweet y Yellott

El modelo visual de Cornsweet y Yellott llamado "Intensity dependent spatial summation" (suma espacial dependiente de la intensidad o IDSS) responde a este tipo de cambio en la suma espacial y permite con un mínimo de suposiciones establecer un posible mecanismo para las primeras etapas del sistema visual humano.

Este modelo explica un gran número de efectos visuales bien conocidos como son la ley de Weber-Fechner [6.8], las bandas de Mach [6.15], la ley de Ricco [6.7], la sensibilidad del sistema visual humano a la detección de frecuencias espaciales [6.6], su variación con la luminancia [6-8] y la adaptación a la luz y a la oscuridad [6.16] como se describirá a continuación

Primero se va a describir el modelo visual mencionado y cómo responde a este tipo de cambio en la suma espacial y por último se comparará la respuesta del sistema visual humano con las del modelo IDS.

La figura 6.2 ilustra la idea básica de este modelo. Una imagen bidimensional, que es alguna distribución de luz (en este caso un borde) es recibida por un arreglo de fotorreceptores y estos envían una señal a una red que realiza la operación IDS. La operación consiste en dos pasos. Primero, cada receptor da lugar a una función de punto

extendido no negativa cuya altura en su centro es proporcional a la intensidad de la imagen de entrada en ese receptor y cuyo volumen es constante. Segundo, estas funciones de punto extendido se suman para crear la imagen de salida y esta imagen es luego leída por un arreglo de canales de salida, uno por cada receptor.

Las respuestas de este modelo son esencialmente las mismas, independientemente de cuál sea la forma de la función de punto extendido y sólo dependen de la suposición fundamental de que el área bajo la función es inversamente proporcional a la intensidad local de la imagen. En este ejemplo se usa como función de punto extendido una Gaussiana.

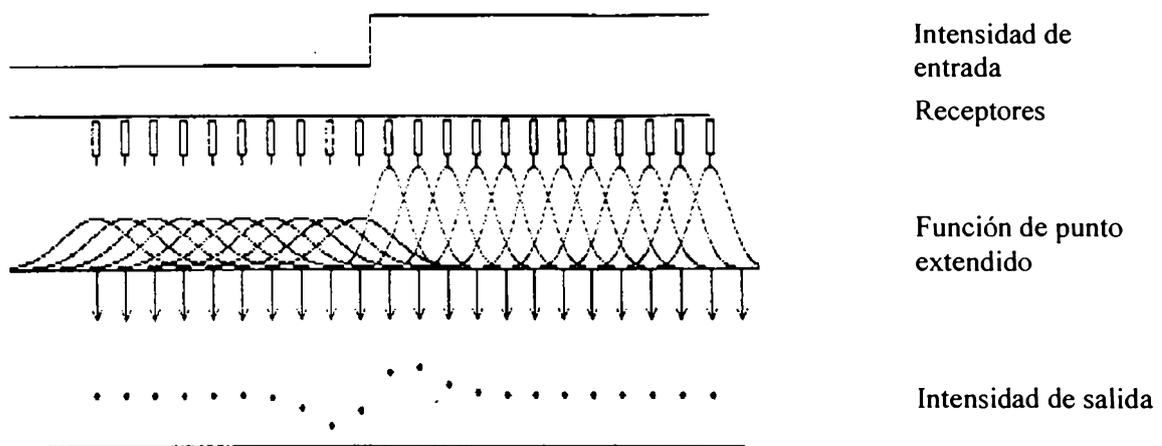


Figura 6.2. Diagrama esquemático del modelo IDSS. En la parte superior se indica el perfil de la imagen de entrada que en este caso es un escalón. A continuación están los fotorreceptores; luego la función de punto extendido (caso Gaussiano); siguen las flechas que indican los canales de salida y por último el perfil de la imagen de salida representado por puntos.

6.4.1 Formalismo del modelo

Por una conveniencia matemática, el tratamiento analítico supone que los fotorreceptores son infinitamente pequeños con respecto al tamaño de las imágenes de entrada y de salida. Esto es, se tratará el caso continuo, de la misma manera que lo hacen las teorías que tratan

el modelo retinal por medio de una convolución de imágenes retinales continuas con respuestas impulsivas continuas. Esta aproximación continua de la naturaleza discreta de la retina produce resultados realistas hasta un cierto nivel de intensidad máximo de la imagen de entrada. Por encima de este nivel se produciría una función de punto extendido más angosta que un solo receptor. Se designa con $I(x,y)$ a la intensidad de la imagen de entrada en el punto (x,y) y con $O[I(x,y)](p,q)$ a la intensidad respuesta de la imagen de salida en el punto (p,q) cuando la entrada es $I(x,y)$.

La idea básica del modelo es que cada punto (x,y) contribuye con un valor a la función de punto extendido no negativa en cada punto de salida (p,q) , dependiendo el tamaño de la contribución del valor de la intensidad de entrada $I(x,y)$ y de la distancia entre (x,y) y (p,q) . Por lo tanto se necesita especificar la forma general de dicha función $S[(x,y),(p,q),I]$ que da la contribución de (x,y) a (p,q) cuando la intensidad de entrada es I . Se supone que:

- 1) La función de punto extendido S es no negativa.
- 2) S es espacialmente homogénea y circularmente simétrica. O sea, S es de la forma:

$$S\left\{\left[(x-p)^2 + (y-p)^2\right], I\right\}.$$

- 3) La altura del centro $S(0,I)$ es directamente proporcional a la intensidad de entrada, mientras que el volumen bajo S permanece constante para todo valor de I distinto de cero:

$$S\left\{\left[(x-p)^2 + (y-p)^2\right], I\right\} = I \times S\left\{I \times \left[(x-p)^2 + (y-p)^2\right], 1\right\}$$

Para cualquier función S , integrar el lado derecho de la suposición (3) sobre p y q debe dar un valor constante V_S que es independiente de I , mientras que la altura del centro es $S(0,I) = I \times S(0,1)$. Entonces el área bajo la función de punto extendido alrededor de cualquier punto de entrada es:

$$A = \frac{V_S}{I \times S(0,1)}$$

La elección del volumen es arbitraria y se elige $v_s = 1$, que es la respuesta del modelo a campos uniformes como se verá más adelante.

4) Por lo expresado en (3) la integral de S sobre el plano p, q es igual a 1.

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S \left\{ (x-p)^2 + (y-p)^2 \right\} I \, dpdq = 1$$

$S(0, I)$ determina el valor numérico de las áreas de la función de punto extendido y debe ser obtenida a partir de experiencias psicofísicas. Este valor va a depender de las unidades usadas para medir las áreas retinales e intensidades de luz.

La suposición fundamental del modelo es la siguiente:

La dispersión del punto desde el punto de entrada (x, y) al punto de salida (p, q) es:

$$I(x, y) \times S \left[I(x, y) \times \left[(x-p)^2 + (y-q)^2 \right] \right]$$

donde $I(x, y)$ es la intensidad de la imagen de entrada en (x, y) y S es una función real no negativa para la cual se verifica

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(p^2 + q^2) \, dpdq = 1$$

Diferentes casos del modelo pueden ser creados a partir de distintas elecciones de la función de punto extendido S . Sin embargo, la forma exacta de dicha función sólo produce ligeras diferencias en los resultados.

5) Por último se supone que la imagen de salida es la suma de las funciones de punto extendido.

$$O[I(x, y)](p, q) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(x, y) \times S \left\{ I(x, y) \times \left[(x-p)^2 + (y-p)^2 \right] \right\} dx dy$$

Esta última suposición contiene la idea fundamental del operador IDS.

Los siguientes teoremas ponen de manifiesto dos importantes características del modelo; la no linealidad y el resultado de multiplicar la intensidad de entrada por un factor constante.

Es importante destacar que este modelo es no lineal. Una manera de ver esta propiedad es la siguiente:

Teorema 1:

La salida a cualquier campo uniforme distinto de cero es el campo uniforme con valor 1. Esto es, cuando $I(x, y) \equiv I > 0$, $O(p, q) \equiv 1$.

El significado físico de este teorema se puede entender de la siguiente manera. A causa de que el volumen bajo la función de punto extendido en cada punto es constante e independiente de la intensidad de entrada, la salida total del sistema es independiente de su entrada. Puesto que una imagen de entrada espacialmente uniforme genera una imagen de salida uniforme, se sigue que las amplitudes de salida correspondientes a todas las imágenes uniformes de entrada deben ser idénticas. El modelo es ciego a variaciones uniformes.

El siguiente teorema describe el efecto de multiplicar todas las intensidades de la imagen de entrada por un factor común, como sucedería con la imagen retinal de una escena real si cambia la iluminación sobre la misma. Este simple teorema es realmente el corazón matemático del modelo: A través del mismo se puede demostrar que la ley de Weber vale en los bordes, que la ley de Ricco vale para puntos brillantes sobre fondos oscuros y que la acuidad visual se incrementa en proporción a la raíz cuadrada del nivel de luminancia media.

Teorema 2

Para toda constante positiva c y toda imagen de entrada $I(x, y)$

$$O[cI(x, y)](p, q) = O\left[I\left(\frac{x}{\sqrt{c}}, \frac{y}{\sqrt{c}}\right)\right](p\sqrt{c}, q\sqrt{c})$$

Esto es, cuando la intensidad de entrada se incrementa en un factor c , la altura de la función de punto extendido también se incrementa en el mismo factor c , mientras que el ancho decrece en un factor de $1/\sqrt{c}$. En palabras, la ecuación anterior significa que el efecto de multiplicar todas las intensidades de la imagen de entrada por una constante c es lo mismo que: primero expandir la imagen original a lo largo de ambos ejes por un factor \sqrt{c} , luego aplicar el operador suma (dado en la suposición 5) a la imagen y finalmente achicar la imagen de salida al tamaño original. Por ejemplo, cada frecuencia espacial f en la imagen $cI(x,y)$, es tratada como una frecuencia f/\sqrt{c} en la imagen $I(x,y)$.

Las demostraciones de ambos teoremas pueden ser vistas en la referencia [6.7].

6.5 Comparación entre las respuestas del sistema visual y el modelo IDS.

Ahora se analizará cómo responde el modelo y se comparará con la respuesta del sistema visual.

Supongamos primero que el ojo observa un campo iluminado uniformemente de modo que la imagen retinal tiene la misma intensidad en todas partes. Entonces, la imagen neuronal o el patrón de salida será el mismo en todas partes. Ahora se superpone sobre este campo un punto brillante. La respuesta del modelo IDS se muestra en la figura 6.3. Se observa que cuando se suma un punto brillante a un campo uniforme la intensidad en el centro crece pero, la intensidad de salida alrededor del punto brillante es menor que el campo uniforme. Entonces, sumar luz a la retina produce en algunos lugares una intensidad menor en la salida.

Otra forma de mostrar el patrón de salida para este estímulo es indicar la actividad neuronal como se observa en la figura 6.4. En el centro la actividad crece, este hecho está indicado por un signo más; y alrededor decrece, indicado con el signo menos. Esta clase de dibujo es el campo receptivo usado por primera vez por Kuffler, en 1953 para describir los

patrones de actividad en las células ganglionales retinales de los gatos [6.17]. El campo receptivo de las células ganglionares retinales es un área particular cuya característica es agrupar las respuestas que reciben las células receptoras. Cuando un campo receptivo se presenta como el mostrado en dicha figura, se dice que muestra antagonismo de centro-alrededor, la actividad en el centro parece oponerse a la actividad del entorno. Este fenómeno, donde sumar luz a una parte de la retina produce que se reduzca la señal en otra parte se ha encontrado en la retina de los mamíferos y se dice que la actividad en el centro brillante debe inhibir la actividad de su alrededor y por lo tanto que este fenómeno demuestra la acción de inhibición lateral en la retina. En realidad, esta clase de antagonismo de centro-alrededor es usualmente usado como evidencia para demostrar y probar la existencia de la inhibición lateral en la retina. La respuesta del modelo IDS produce el mismo efecto y no involucra un proceso de inhibición o sustracción en la retina, sólo realiza sumas.

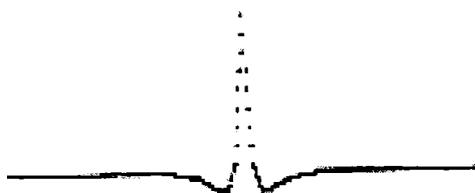


Figura 6.3. Respuesta del modelo IDS cuando la entrada es un punto brillante sobre un fondo menos intenso.

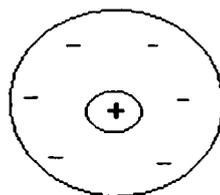


Figura 6.4. Esquema de una vista bidimensional llamada campo receptivo. El signo positivo de la región central está indicando que la actividad en ella aumenta cuando el punto brillante se suma a una campo uniforme. Los signos negativos de los alrededores indican que la actividad decrece. De esta manera el modelo IDS pone de manifiesto lo que se llama antagonismo de centro alrededor, aunque no está involucrado ningún proceso de inhibición.

La figura 6.5 muestra la respuesta del modelo IDS cuando existe un punto brillante en un fondo uniforme menos intenso. En la parte superior de la figura se muestra el patrón de entrada. Cada receptor expande su señal y esta expansión se muestra debajo de cada receptor (en este caso es un triángulo). Los canales de salida suman toda la actividad que les llega y el resultado de cada canal de salida se muestra como un punto en la parte inferior de la figura 6.5. Se concluye que el cambio en el ancho de la función de punto extendido hace que la respuesta del modelo produzca un efecto que se parece a la inhibición lateral, pero el modelo no realiza ninguna sustracción.

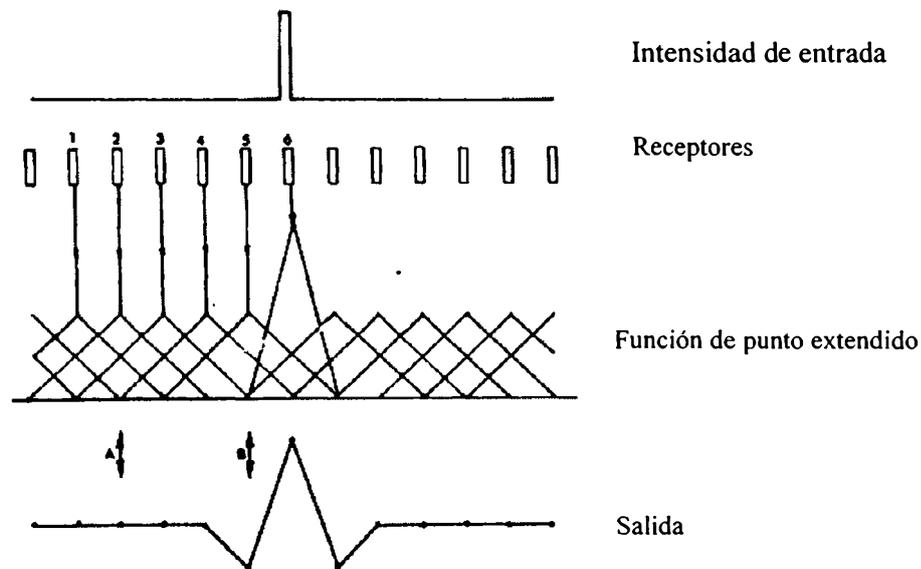


Figura 6.5. Descripción diagramática de cómo la adición de un punto sobre un campo uniforme puede reducir la actividad de algunos canales de salida.

6.5.1 Bandas de Mach

Las Bandas de Mach son un fenómeno visual que fue informado por primera vez cien años atrás por el físico austriaco Ernst Mach. Dichas bandas dependen estrictamente de la distribución de la iluminación. Mach, formuló un criterio simple para su efecto: “Siempre que la curva de intensidad de luz de una superficie iluminada (cuya intensidad de luz varíe en una sola dirección) tenga una flexión cóncava o convexa con respecto a la abscisa, el lugar aparecerá más brillante o más oscuro que sus alrededores”. Este fenómeno es

puramente subjetivo porque medidas físicas de la luminancia muestran que no existen valores mayores o menores que los del entorno. Las bandas de Mach pueden observarse en la figura 6.6.

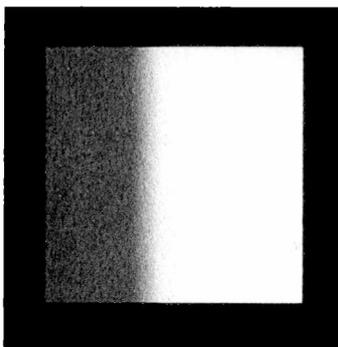


Figura 6.6. Imagen de un borde desenfocado. En la parte izquierda del borde, se puede observar una luminancia subjetiva más oscura que en los alrededores y en la parte derecha del mismo una luminancia más brillante que en las cercanías.

La importancia de las bandas de Mach en la percepción ha sido señalada por Stockham [6.18], quien construyó una imagen donde era suprimida la ilusión. Esta imagen era percibida como una imagen degradada de la original.

La figura 6.7 muestra la salida del modelo IDS cuando el perfil de intensidad de entrada es un escalón. Se supone que la entrada es un borde unidimensional $I(x)$ de la forma $I(x)=I$ para $x \leq 0$, $I(x)=I+D$, para $x > 0$ como el esquematizado en la parte superior de la figura 6.2 . Haciendo esta sustitución en la suposición 5 y luego de un poco de álgebra se obtiene la respuesta mostrada en dicha figura. En este caso, la función de punto extendido utilizada es una Gaussiana y se consideraron diferentes bordes. Estos bordes difieren en la luminancia, pero la relación del lado más brillante al más oscuro y consecuentemente la fracción de Weber D/I , es constante. En dicha figura se puede observar que el brillo que una persona ve en cada punto del patrón de salida del modelo IDS, muestra una región oscura donde la luminancia empieza a incrementarse y una región más brillante cuando la luminancia decrece, llamadas bandas de Mach. Esto es exactamente lo que sucede en la visión humana.

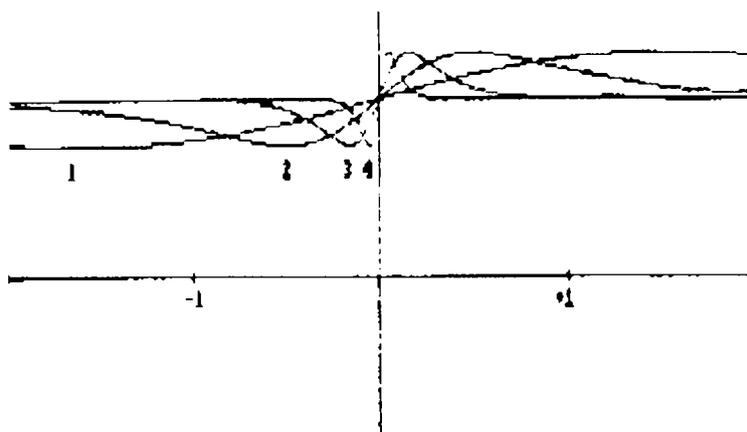


Figura 6.7. Respuesta del modelo IDS de un borde. La imagen de entrada es un escalón en cero desde una intensidad I a $I+D$. $I/D=10$ en todos los casos. En la curva 1, $I=0.1$; en la curva 2; $I=1$; en la curva 3, $I=10$, en la curva 4, $I=100$.

Para entender intuitivamente como se crean las bandas de Mach por medio de un mecanismo de funciones de punto extendido puramente positivas puede ser útil la figura 6.2. A medida que se aproxima el borde desde la izquierda, la salida decrece por debajo del nivel base porque existe una menor excitación que proviene de los receptores del lado derecho del borde, los cuales tienen una función de punto extendido mucho más angosta. Inversamente, a medida que se aproxima el borde desde el lado derecho, la salida crece respecto del nivel base debido a una mayor excitación que proviene de los receptores de la izquierda cuya función de punto extendido es más ancha.

Una segunda característica importante del perfil de la respuesta es que el efecto de incrementar I es mover los picos y los valles de las bandas de Mach hacia el borde mismo. Esto es una consecuencia del hecho de que los bordes de entrada tienen todos la misma fracción de Weber D/I .

Este modelo se comporta como un filtro pasabanda y es usado como un primer paso hacia la detección de bordes. Este método conduce a una descripción más rica de la escena en aplicaciones de visión por computadora y a un nuevo esquema de codificación para

aplicaciones que requieren gran compresión de datos. Alter-Gatenberg et al. [6.4] consideran dos filtros pasabanda diferentes basados en modelos visuales humanos. Uno es el modelo de Marr [6.1] y el otro es el IDS. Ambos filtros preservan la información significativa más allá de detectar los bordes. El modelo de Marr da información de la cual se puede extraer los cambios de intensidad a través del borde. El filtro IDS contiene información de la cual se puede obtener los cambios de intensidad en los bordes y, a menudo más importante, la relación de reflectancia a través el borde. En el trabajo mencionado en la referencia se presenta un método para recuperar la intensidad y la reflectancia aplicando ambos modelos. Ellos demuestran que muchas características de una escena están contenidas en las bandas de Mach.

En otro trabajo Reese [6.19] describe un método para realzar bordes basado en el modelo IDS. El autor muestra que el método es particularmente útil en imágenes que tienen regiones de bajo contraste, tales como áreas muy brillantes o aquellas cubiertas por sombras.

6.5.2 Función de Sensibilidad al Contraste

La sensibilidad del sistema visual humano a las frecuencias espaciales se estudia midiendo el contraste umbral. El contraste se define como :

$$C = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} \quad (6.3)$$

donde I_{max} e I_{min} son los valores máximos y mínimos de luminancia de una red sinusoidal [6.20]

El contraste umbral es el mínimo necesario para percibir una red de cierta frecuencia y la sensibilidad al contraste se define como la recíproca del contraste umbral. Campbell y Robson [6.6] midieron la Función de Sensibilidad al contraste en función de la frecuencia espacial para una red sinusoidal (Figura 6.8).

La figura 6.9 es una gráfica de la misma función, pero obtenida con el modelo IDSS. Los resultados muestran que ambas curvas presentan un comportamiento similar. Se observa que en la parte media, la sensibilidad es alta es decir, patrones que tienen frecuencias espaciales medias y que tienen bajo contraste pueden ser vistos fácilmente. Si se observa la parte derecha, se encuentra que la sensibilidad cae, cuando los patrones son más apretados, (frecuencias espaciales altas) y estos patrones no pueden ser resueltos a pesar de que el contraste sea muy alto. El extremo derecho de la curva tiene cierto interés porque representa el límite de la acuidad visual. En la parte izquierda, la sensibilidad también cae, es decir es difícil ver patrones que son demasiado apretados o demasiado anchos.

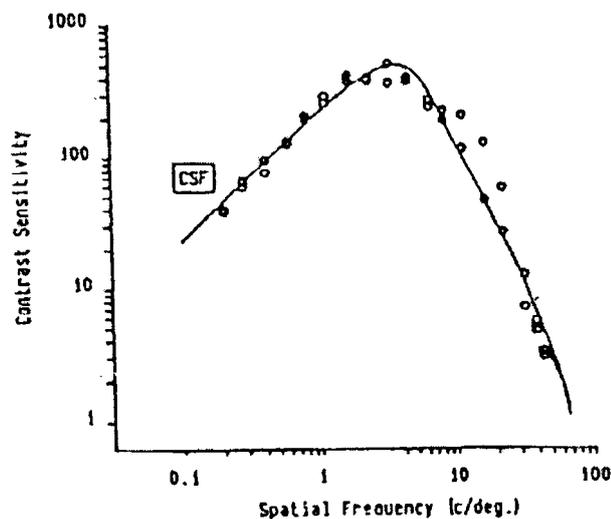


Figura 6.8. Datos de Campbell y Robson que muestran la curva de sensibilidad al contraste para observadores de redes sinusoidales de diferentes frecuencias espaciales.

Debido a que esta curva tiene mucha información, es usada en ambientes clínicos para la evaluación y descripción de la visión. Por ejemplo, se aplica al diagnóstico de las patologías del nervio óptico y para definir la naturaleza de la pérdida de la visión en ambliopía (oscurecimiento de la visión por sensibilidad imperfecta de la retina). Cuando esta curva es usada para ello, cualquier pérdida que muestre usualmente tiene una interpretación fisiológica. Por ejemplo, el hecho de que en un ojo normal la sensibilidad caiga hacia la parte izquierda de la curva, es decir hacia las bajas frecuencias espaciales, ha

sido atribuido a la inhibición lateral. Por lo tanto, cuando alguna patología produce cambios en esa parte de la curva, la patología es usualmente interpretada como la que causa cambios en los mecanismos de inhibición. El hecho de que el modelo IDS (sin inhibición lateral) dé la misma curva de Sensibilidad al Contraste que la obtenida psicofísicamente, sugiere la revisión de algunas de las interpretaciones anteriormente mencionadas.

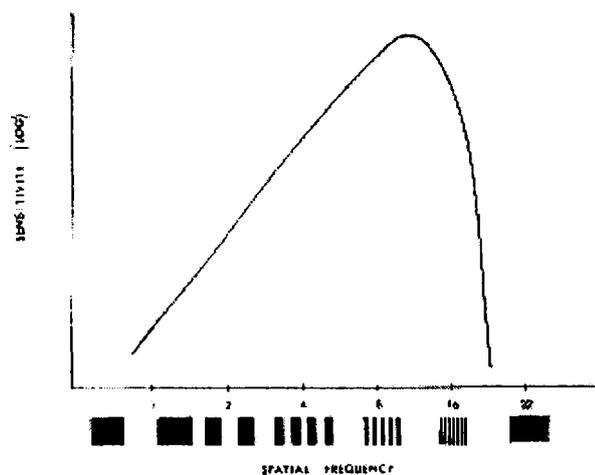


Figura 6.9. La función de sensibilidad al contraste para el modelo IDS. Las redes que están a lo largo del eje horizontal indican aproximadamente la relativa fineza de las frecuencias espaciales. Cuando cada una de estas funciones es medida psicofísicamente, el estímulo tiene un perfil de luminancia sinusoidal. La función de punto extendido usada es una Gaussiana. Nótese que las escalas son logarítmicas

La figura 6.10 es una gráfica de la Función de Transferencia (relación entre el contraste de salida y el contraste de entrada) del modelo IDS para diferentes niveles promedio de iluminación. Cuando ésta se incrementa, la curva se corre hacia las frecuencias espaciales altas y cuando el eje horizontal es graficado como el logaritmo de la frecuencia, la forma de la curva no cambia, solo se traslada en una manera perfectamente ordenada. Este corrimiento es una consecuencia de la no linealidad del modelo, indica un incremento de la acuidad visual y varía con la raíz cuadrada de la luminancia media.

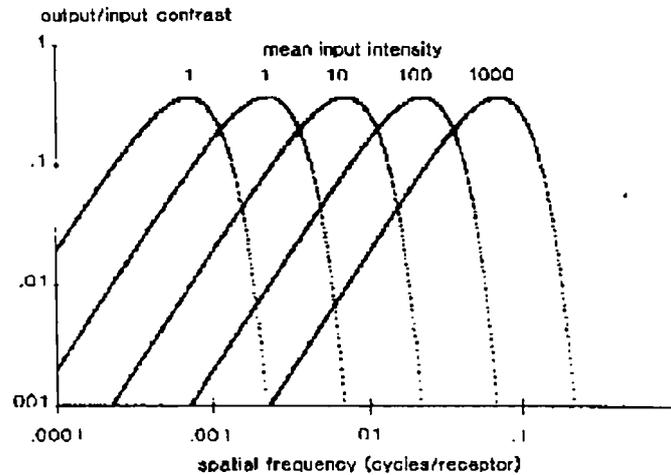


Figura 6.10. Módulo de la Función de Transferencia del modelo IDS para una función de punto extendido con forma Gaussiana.

Los antagonismos de centro- alrededor, las bandas de Mach, el realce de bordes y esta clase de curva de Sensibilidad al contraste pueden ser explicadas tanto por el modelo IDS o por el modelo de inhibición lateral. De hecho, si se calcula la curva de Sensibilidad al contraste con el modelo IDS usando una forma particular de la función de punto extendido y luego se calcula la correspondiente curva con el modelo de inhibición lateral con la misma forma de la función de punto extendido, las dos curvas son idénticas.

Con el modelo de inhibición lateral de Marr y Hildreth se obtiene una curva que coincide con el modelo IDS, pero dicha curva no cambia para diferentes niveles de iluminación. Las curvas que corresponden al sistema visual humano fueron obtenidas por Van Ness y Bouman (figura 6.11) y difieren de las del modelo IDS en dos aspectos importantes: éstas últimas no exhiben el cambio característico de filtro pasabajo a pasabanda a medida que la luminancia crece y el valor pico de la Función de Transferencia es el mismo para todas la luminancias, mientras que el pico de la sensibilidad en la curva del sistema visual crece de alrededor de 10 (contraste umbral del 10%) a más de 100 (contraste umbral < 1%) a medida que la luminancia media crece desde el umbral absoluto hasta 100 Trolands. En la figura 6.11 se grafica el contraste umbral en función de la frecuencia espacial mientras que en la figura 6.9 se grafica la función recíproca.

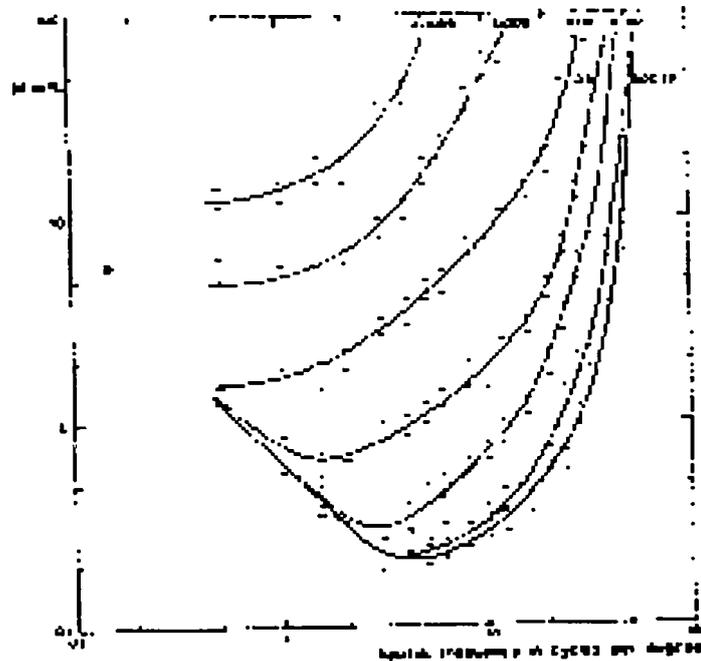


Figura 6.11. Contraste en función de la frecuencia espacial para diferentes niveles de iluminación. Las líneas horizontales indican los umbrales superiores e inferiores de la modulación para cada frecuencia espacial. Las curvas para 5900td y 900td son idénticas.

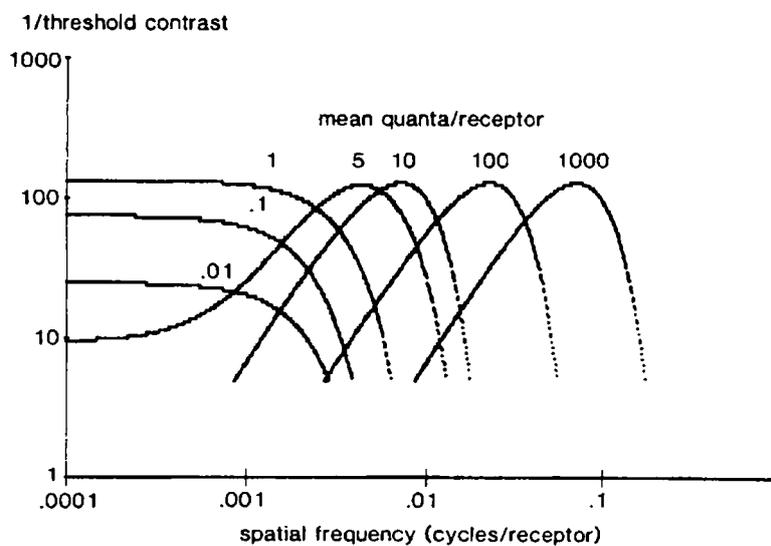


Figura 6-12. Función Sensibilidad al Contraste espacial del operador Gaussiano (CV) aplicado a redes sinusoidales con ruido de Poisson y con diferentes niveles de iluminación.

En un trabajo posterior, Yellott [6.21] analiza las consecuencias de aplicar los operadores de volumen constante a imágenes que contienen ruido de Poisson. Una comparación detallada entre datos psicofísicos y predicciones de los operadores de volumen constante aplicados a imágenes de entrada con ruido de Poisson muestra que el modelo Gaussiano reproduce cualitativamente las principales propiedades de la visión espacial pero no coincide cuantitativamente con el comportamiento humano en ciertos aspectos críticos. En particular, la función de sensibilidad al contraste que se obtiene es la mostrada en la figura 6.12. En este gráfico se puede observar que el comportamiento de dicha función no coincide con los datos obtenidos empíricamente.

6.5.3 Ley de Ricco

La ley de Ricco establece que la detectabilidad de un spot de luz depende sólo del producto de su área por la intensidad. Experimentalmente, en el sistema visual humano esta ley es válida hasta cierta área crítica del spot (figura 6.13). Esta área decrece cuando la intensidad del fondo se incrementa. El modelo IDS obedece esta ley de una manera muy parecida a la forma en que lo hace el sistema visual [6.7] y se muestra en la figura 6.14.

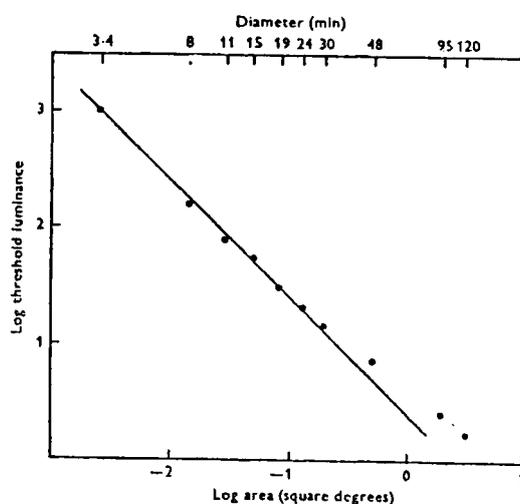


Figura 6.13. Gráfico del logaritmo de la luminancia umbral absoluta en función del logaritmo del área de testeo para un sujeto B. S. El test fue circular, de 32 mseg. de duración y la imagen del filamento fue de 2 mm en la pupila. Cada punto del gráfico representa el promedio de ocho

lecturas. La desviación estandar de los valores medios fue menor que 0.04 unidades logarítmicas. La línea recta tiene una pendiente de -1 .

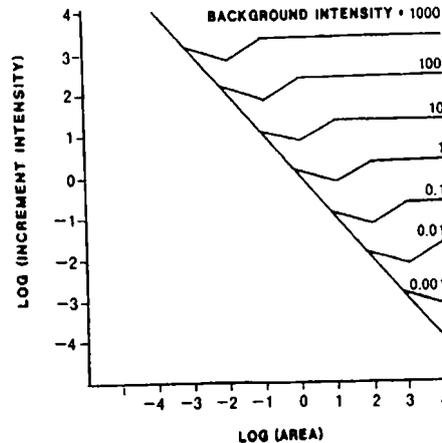


Figura 6.14. El modelo IDS obedece la ley de Ricco. La entrada al modelo IDS es un spot de luz al que se le puede modificar el tamaño y su luminancia sobre un fondo con una luminancia también ajustable. La luminancia de este spot se modifica hasta que produce una salida de una determinada amplitud. En la correspondiente experiencia psicofísica, la luminancia del spot se ajusta hasta que alcance la detección umbral subjetiva. Cada punto de estas curvas representa un spot de alguna luminancia y tamaño que justamente produce esta respuesta. Para las menores luminancias de fondo (0.001), la intensidad del spot (el incremento) que se requiere para su detección es inversamente proporcional al área del spot hasta un área de alrededor 3 unidades logarítmicas. Esto significa que, para todas las áreas menores que esta, el producto del área por la luminancia es constante. Esto es, la energía total en el umbral es constante. Esto es la ley de Ricco. Como se puede ver el sistema visual presenta un comportamiento similar.

6.5.4 Ley de Weber

La parte superior de la figura 6.15 muestra la respuesta del modelo IDS a una serie de bordes. En la fila inferior de la figura, la luminancia de las franjas y la del fondo se va aumentando manteniéndose la relación 2:1. Como puede verse en la parte superior de la misma figura, las respuestas del modelo se hacen más angostas pero las amplitudes son idénticas. Esto significa que el modelo IDS obedece la ley psicofísica de Weber: si I es la intensidad para la cual el ojo está adaptado y ΔI , el incremento mínimo de la intensidad que puede ser percibido por el ojo, la relación $\Delta I/I$, puede ser considerada como el poder discriminante del ojo y es constante dentro de amplios límites de intensidad. En el modelo de inhibición lateral esta ley no es válida. Otro fenómeno psicofísico que se produce es la

no validez de la ley de Weber para objetos muy pequeños. Esto mismo predice el modelo IDS.

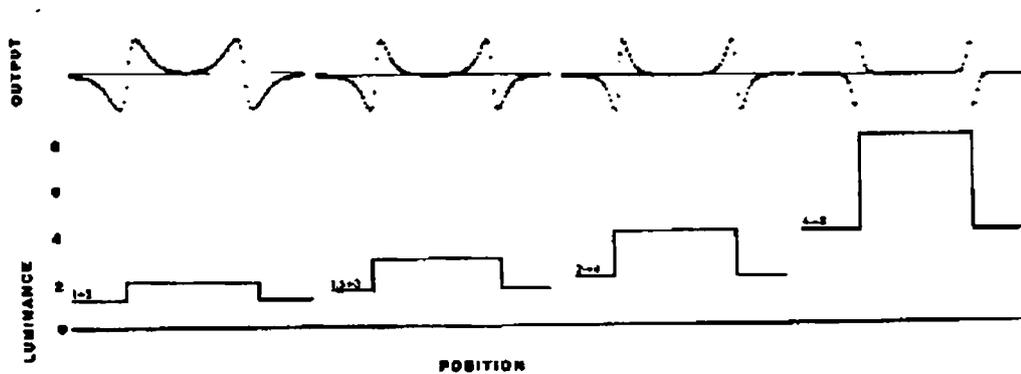


Figura 6.15. En la fila inferior están graficados los perfiles de diferentes spots o franjas y en la superior sus correspondientes salidas. Las luminancias de varios patrones son diferentes, pero la relación de las luminancias de las franjas con respecto a los fondos es la misma, 2:1. Los patrones de salida se hacen más angostos para los estímulos de luminancia mayor, pero las amplitudes de los picos y de los valles son todas idénticas.

6.5.5 Adaptación

El último fenómeno que se comentará es la adaptación a la luz y a la oscuridad. Estos términos se refieren al hecho de que la sensibilidad visual cambia con el tiempo después de haber cambiado el nivel de luz. El resultado del proceso de adaptación es que cuando se está en presencia de baja iluminación, el sistema visual es más sensible que con una iluminación intensa. Es decir, que si se mira un área brillantemente iluminada y luego se le suma un spot de luz, se puede detectar el spot sólo si éste es lo suficientemente brillante; mientras que si se mira un área pobremente iluminada se puede detectar un spot iluminado con una intensidad mucho más reducida. Cualquier sistema visual que obedece la ley de Weber hace exactamente esto, y cualquier teoría que explica este proceso bajo la ley de Weber también explica la mayor parte de los procesos de adaptación a la luz y la oscuridad. Como el modelo IDS cumple con la ley de Weber, explica totalmente estos fenómenos.

6.6 Bibliografía

- [6.1] David Marr. Vision. A Computacional Investigation into the Human Representation and Processing of Visual Information, W. H. Freeman and Company. San Francisco. (1982).
- [6.2] Marr, D. Y. and Hildreth, E.: Theory of edge detection. Proc. R. Soc. London Ser. **B 207**, 187-217, (1980).
- [6.3] B. F. Logan Jr. Information in the Zero Crossing of Bandpass Signals. The Bell System Technical Journal, **56**, 487-510, (1977).
- [6.4] Rachel Alter-Gatenberg, Friedrich Huck, Ramkumar Narayanswamy: Image recovery from edge primitives, J. Opt. Soc. Am. A, **7**, Nro 5, 898-911, (1990).
- [6.5] F. Ratliff. Contour and contrast. Scien. Am. **226**, 6 , 91-101, (1972).
- [6.6] Campbell, F. W. and Robson, J. G. Application of Fourier analysis to the modulation response of the eye. J. Opt. Soc. Am., **54**, 581, (1964).
- [6.7] T. Cornsweet and J. Yellott: Intensity Dependent spatial summation. J. Opt. Soc. Am, **2**, 1769-1786, (1985).
- [6.8] F. Van Nes and M. Bouman: Spatial modulation transfer in the human eye. J. Opt. Soc. Am. **57** 401-406, (1967).
- [6.9] S. Hecht: The visual discrimination of intensity and the Weber-Fechner law. J. Gen. Physiol. **7** 235-265, (1924).
- [6.10] Las Demostraciones psicofísicas de la suma espacial comenzaron con A. Ricco “Relazione fra il minimo angolo visuale e l’intensita luminoso,” Ann, Ottalmol. **6**,

373-479 (1877); la literatura fue revisada por P. E. Hallett, "Spatial summation," *Vision Res.* **3**, 9-24, (1963) y por B. Sakitt, "Configuration dependence of scotopic spatial summation," *J. Physiol. (London)* **216**, 513-529 (1971); demostraciones fisiológicas de suma espacial en retinas de vertebrados comenzaron con H. K. Hartline, "The receptive fields of optic nerve fibers," *Am. J. Physiol.* **130**, 690-699, (1940) una posterior revisión de esta literatura fue realizada por P. O. Bishop, "Processing of visual information within the retinostriate system," en *Volumen III of the Handbook of Physiology: The Nervous System*, I. Darian Smith, ed. (American Physiological Society, Bethesda, Md., 1984); la suma espacial en el nivel de fotorreceptores fue primeramente informada por D. A. Baylor, M. G. F. Fourtes and P. M. O'Bryan, "Receptive fields of cones in the retina of the turtle," *J. Physiol. (London)* **214**, 265-294 (1971).

- [6.11] H. B. Barlow, "Temporal and spatial summation in human vision at different background intensities". *J. Physiol. (London)* **141**, 337-350, (1958)[V. D. Glezer, "The receptive field of the retina," *Vision Res.* **5**, 497-525 (1965).
- [6.12] Rose, A The sensitivity performance of the human eye on an absolute scale. *Journal of the Opt. Soc. of Am.* **38**, 196-208.
- [6.13] DeVries, H.. The quantum character of light and its bearing upon threshold of vision, the differential sensitivity and visual acuity of the eye. *Physica*, **11**, 179-189, (1943)
- [6.14] Baylor, D. A., Nunn, B. J., & Schnaff, J. L.. The photocurrent, noise and spectral sensitivity of rods of the monkey macaca fascicularis, *Journal of Physiology*, **357**, 575-607, (1984).
- [6.15] J. Ross, M. Morrone and D. Burr: The conditions under which Mach bands are visible, *Vision Res.* **29** N0 6, 699-715, (1989).

- [6.16] Barlow, Hb.: Dark and light adaptation: psychophysics. In: Jamenson D., Hurvich L., eds. Handbook of sensory Physiology, vol 7, pt 4, Visual Psychophysics. Berlín: Springer Verlag, 1972: 1-28.

- [6.17] Kuffler S. W. Discharge patterns and functional organization of mammalian retina. J. Neurophysiol., 16, 37-68, (1953).

- [6.18] T. Stockham: Image processing in the context of a visual model. Proc. of the IEEE **60** N^o 7 (1972) 828-842.

- [6.19] Greg Reese: Image enhancement by Intensity-Dependent Spread Function, Graphical models and image processing, **54**, Nro 1, 45-55, (1992).

- [6.20] Max Born and E. Wolf, *Principles of Optics*, sixth edition, New York, Pergamon Press, p. 267, (1980).

- [6.21] Yellot, J.: Photon noise and constant-volume operators, J. Opt. Soc. Am., **4**, 2418-2446, (1987).

Capítulo 7

El modelo de volumen aproximadamente constante

7.1 Introducción

En el capítulo anterior estudiamos el modelo IDS y en éste lo aplicaremos, incluyendo en él una modificación, de manera tal de obtener una medida útil que indique cuáles son las áreas correctamente enfocadas, en un conjunto de imágenes con desenfoque progresivo. Ésto no significa que el modelo modificado tenga valor en cuanto a que describa las operaciones que efectivamente hace el sistema visual.

Este trabajo [7.1] está basado en la suposición de que las bandas de Mach juegan un rol importante en la percepción del desenfoque y que éstas proveen una herramienta para discriminar la posición del mejor foco en un conjunto de imágenes de una escena 3D.

Vamos a sugerir una manera empírica de medir el área bajo las bandas de Mach usando una versión modificada del modelo IDS. Este valor llamado por nosotros *señal* del

modelo modificado, va a ser utilizado como criterio para determinar y corregir el desenfoque.

Como se ha explicado anteriormente, se necesita incluir detalles de alta frecuencia en la escena para una mejor elección del foco correcto. Por lo tanto los mismos se introducen como iluminación estructurada a través de franjas proyectadas sobre el objeto que va a ser registrado.

Calcularemos la *señal* para perfiles de intensidades teóricos de redes sinusoidales proyectadas sobre un objeto plano y observada a través de un sistema formador de imágenes. Se va a considerar el desenfoque del proyector y el del sistema formador de imágenes. Luego, la medida propuesta va a ser aplicada a imágenes experimentales de objetos planos y sus valores van a ser graficados en función de la distancia de enfoque para verificar su comportamiento. Para incrementar aun más el contenido relativo de las frecuencias altas, será aplicado a dichas imágenes el operador derivada antes de calcular la *señal*. Finalmente, la medida propuesta se aplicará para seleccionar las regiones de enfoque correcto en un conjunto de imágenes 3D con desenfoque progresivo y construir una imagen procesada donde este defecto será parcialmente corregido.

Este método es no lineal, es adaptativo al ruido y toma en cuenta todas las áreas de los alrededores del pixel considerado.

7.2 Formalismo del modelo de volumen aproximadamente constante

Cuando el modelo IDS se implementa teniendo en cuenta el tamaño finito de los receptores, como es el caso del procesamiento digital de imágenes, el tamaño de los mismos debe ser incluido en el cálculo. Este modelo considera que los receptores que intervienen son sólo bastones. Estos detectores están agrupados densamente en la retina, la cantidad de bastones varía entre $110 \cdot 10^6$ y $125 \cdot 10^6$, tienen un diámetro aproximado de 2μ y un área media denominada por nosotros A_m , que varía con la ubicación en la retina [7.2]. Cuando los niveles de iluminación son muy altos, el área de integración no puede achicarse

a dimensiones menores que un solo detector por lo tanto, la suposición de volumen constante no puede seguir valiendo. Además, las integrales deben ser reemplazadas por sumatorias.

Para tener en cuenta el tamaño finito de los receptores sugerimos restar al área bajo la función de punto extendido el tamaño de un receptor, tratando de conservar las propiedades del operador de volumen constante. Por lo tanto la segunda ecuación de la suposición 3 de la sección 6.2.1 se convierte en:

$$I(A - A_m) = C$$

donde A_m es el tamaño de un receptor y C es una constante. Ahora el volumen IA no es más constante y muestra una pequeña variación con I . Por este motivo proponemos llamarlo “Operador de volumen aproximadamente constante” (ACVO). Este modelo es también no lineal.

El área puede ser expresada como:

$$A = \frac{C}{I} + A_m \quad (7.1)$$

El efecto de la corrección será notable en el límite de altas intensidades como se puede deducir de la ecuación anterior. En este caso se obtiene que $A = A_m$ y la intensidad I es asignada a un sólo receptor. Por lo tanto la imagen resultante es una copia de la imagen de entrada y no se producen ilusiones tales como las Bandas de Mach. En el límite opuesto de bajas intensidades, están involucrados un gran número de receptores y el tamaño de uno solo puede ser despreciado. Entonces, el resultado tiende a coincidir con el del modelo de volumen constante como se deduce de la ecuación anterior de la cual se obtiene:

$$A = \frac{C}{I}$$

El modelo ACVO puede ser descripto de la siguiente manera: Una imagen de entrada incide sobre un arreglo cuadrado de fotorreceptores, cada uno de los cuales tiene un tamaño de 1×1 , (de modo que la distancia unidad en el plano imagen de entrada es el diámetro de un receptor) coordenadas (x,y) donde $x = \pm 0,1,2,..$ e $y = \pm 0,1,2,..$ e intensidad $I(x,y)$ la cual es considerada constante dentro de cada receptor. El plano imagen contiene otro arreglo cuadrado similar al de entrada. Las salidas tiene coordenadas (p,q) y tamaño 1×1 . Cuando el receptor (x,y) recibe una intensidad $I(x,y)$ se crea una función de punto extendido sobre el plano imagen con un valor en cada punto (p,q) dado por:

$$I(x,y)S \left\{ \frac{I(x,y)}{1 + I(x,y)A_m} [(x-p)^2 + (y-q)^2] \right\} \quad (7.2)$$

donde A_m es el tamaño general de un receptor, en este caso de 1×1 . Suponemos que el valor obtenido en la ecuación anterior es constante sobre cada receptor. La contribución total a la función de punto extendido de todos los receptores (x,y) a la salida (p,q) , es la suma de la ecuación anterior sobre x,y y ese número es el valor de la imagen de salida en la coordenada (p,q) . La imagen discreta de salida está dada por:

$$D[I(x,y)](p,q) = \sum_x \sum_y I(x,y)S \left\{ \frac{I(x,y)}{1 + I(x,y)A_m} [(x-p)^2 + (y-q)^2] \right\} \quad (7.3)$$

Con la suposición anterior podemos aproximar la versión discreta de este operador con la continua y toma la forma general

$$D[I(x,y)](p,q) = \iint I(x,y)S \left\{ \frac{I(x,y)}{1 + I(x,y)A_m} [(x-p)^2 + (y-q)^2] \right\} dx dy \quad (7.4)$$

Se puede observar que cuando $A_m = 0$ se obtiene el operador de volumen constante. Cuando la intensidad es suficientemente alta de modo que el área se reduce a un solo receptor, la ecuación anterior puede ser escrita como:

$$D[I(x,y)](p,q) = \iint I(x,y) S \left\{ \frac{[(x-p)^2 + (y-q)^2]}{A_m} \right\} dx dy \quad (7.5)$$

Esta expresión es la convolución entre la imagen de entrada y una función de punto extendido (cuyo ancho es constante) y es una versión pasa bajo del objeto. El sistema simplemente copia la entrada en la salida. Entre estos dos extremos, es decir entre el operador de volumen constante y la copia existe un continuo de valores de intensidades donde se combinan ambos efectos.

De la misma manera que antes, la función de punto extendido del ACVO debe cumplir las condiciones de simetría establecidas para el CVO. Para describir las propiedades generales que cumple el ACVO elegimos la forma Gaussiana.

7.3 Las bandas de Mach y el modelo ACVO

El fenómeno de las Bandas de Mach fue explicado en la sección 6.2.1. En ésta vamos a comparar la respuesta del sistema visual a una rampa de luminancia con la correspondiente al ACVO.

Se realizó una experiencia en la intervinieron 15 observadores a los cuales se les mostró una rampa de luminancia. Cada participante debía indicar el valor del brillo percibido. Este valor debía ser elegido por el método de estimación de magnitud [7.3] a partir de dos valores fijos que correspondían a los valores objetivos mínimos y máximos de luminancia de la rampa. El resultado obtenido se muestra en la figura 7.1. La curva (a) de la figura es la intensidad de la rampa y la (b) es el valor medio de los valores de brillo observados por los participantes. El modelo CVO no predice este comportamiento. En la figura 6.5 se puede observar la salida del modelo CVO cuando la entrada es un borde. La respuesta muestra que la intensidad de la izquierda del borde, más allá del valle, toma el mismo valor que la intensidad de la derecha, más allá de la cima. Esta respuesta no coincide con los datos psicofísicos. La respuesta del CVO cuando la entrada es una rampa se puede observar en la referencia [7.4] y presenta un comportamiento similar al borde en

cuanto a que los valores de intensidad a ambos lados de la rampa son los mismos. A continuación, analizaremos la respuesta del modelo ACVO cuando la entrada es una rampa y la función de punto extendido es una Gaussiana. Para ello, comenzamos haciendo una generalización cuando la entrada es una función unidimensional de la forma: $I(x, y) = I(x)$.

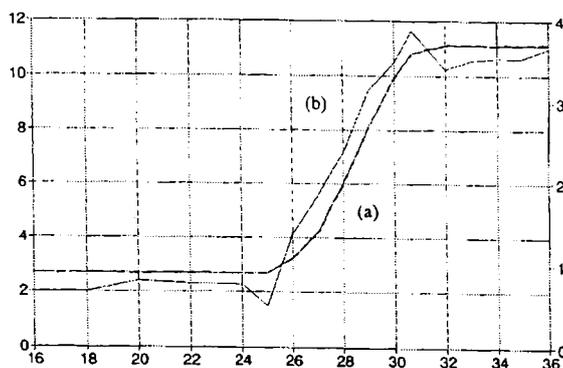


Figura 7.1. Bandas de Mach experimentales. a) Perfil de intensidad de la rampa. b) Respuesta a este perfil de intensidad obtenida por medio de datos psicofísicos.

Analíticamente, si consideramos una imagen de entrada unidimensional de la forma: $I(x, y) = I(x)$ y se la sustituye en la ecuación (7.4), se obtiene:

$$D[I(x)(p,q)] = \iint \sqrt{I(x)} \sqrt{I(x)} S \left\{ \left[(x-p) \sqrt{\frac{I(x)}{1+I(x)A_m}} \right]^2 + \left[(y-q) \sqrt{\frac{I(x)}{1+I(x)A_m}} \right]^2 \right\} dy dx$$

Si ahora en la integral sobre y , hacemos el siguiente cambio de variables

$$v = (y-q) \sqrt{\frac{I(x)}{1+I(x)A_m}}, \text{ entonces } dv = \sqrt{\frac{I(x)}{1+I(x)A_m}} dy$$

se obtiene:

$$D[I(x)(p,q)] = \int \sqrt{I(x)} \sqrt{1+I(x)A_m} \int S \left\{ \left[(x-p) \sqrt{\frac{I(x)}{1+I(x)A_m}} \right]^2 \right\} dv dx$$

Si S es una Guassiana de la forma

$$S(x, y) = \frac{1}{2\pi} \exp \left[-\frac{1}{2} (x^2 + y^2) \right]$$

que corresponde a la función densidad de probabilidad de dos variables aleatorias independientes, cada una con media cero y varianza uno, la salida del ACVO es:

$$D[I(x)(p)] = \frac{1}{\sqrt{2\pi}} \int \sqrt{I(x)} \sqrt{1 + I(x) \cdot A_m} \exp \left[-\frac{1}{2} \left[(x-p)^2 \frac{I(x)}{1 + I(x) \cdot A_m} \right] \right] dx \quad (7.6)$$

o en su versión discreta

$$D[I(x)(p)] = \frac{1}{\sqrt{2\pi}} \sum_x \sqrt{I(x)} \sqrt{1 + I(x) \cdot A_m} \exp \left[-\frac{1}{2} \left[(x-p)^2 \frac{I(x)}{1 + I(x) \cdot A_m} \right] \right] \quad (7.7)$$

En el límite de altas luminancias, la ecuación (7.7) puede ser escrita como:

$$D[I(x)(p)] = \frac{1}{\sqrt{2\pi}} \sum_x I(x) \sqrt{A_m} \exp \left[-\frac{1}{2} \left[\frac{(x-p)^2}{A_m} \right] \right] \quad (7.8)$$

Si suponemos que $I(x)$ (para intensidades intermedias) es una rampa de la forma

$$I(x) = \begin{cases} I_{min} & x \leq x_1 \\ I_{min} + m(x - x_1) & x_1 < x < x_2 \\ I = I_{max} & x \geq x_2 \end{cases}$$

donde I_{min} , I_{max} son las luminancias mínimas y máximas, m la pendiente y x_1 , x_2 las posiciones iniciales y finales de la rampa. La respuesta del ACVO (usando la ecuación (7.7)) está dada por

$$\begin{aligned}
 D[I(x)(p)] = & \frac{1}{\sqrt{2\pi}} \sum_{x < x_1} \sqrt{I_{min}} \sqrt{1 + I_{min} A_m} \exp \left[-\frac{1}{2} \left((x-p)^2 \frac{I_{min}}{1 + I_{min} A_m} \right) \right] + \\
 & + \frac{1}{\sqrt{2\pi}} \sum_{x_1 < x < x_2} \sqrt{I_{min} + m(x-x_1)} \sqrt{1 + (I_{min} + m(x-x_1)) A_m} \exp \left[-\frac{1}{2} \left((x-p)^2 \frac{I_{min} + m(x-x_1)}{1 + (I_{min} + m(x-x_1)) A_m} \right) \right] + \\
 & + \frac{1}{\sqrt{2\pi}} \sum_{x_2 < x} \sqrt{I_{max}} \sqrt{1 + I_{max} A_m} \exp \left[-\frac{1}{2} \left((x-p)^2 \frac{I_{max}}{1 + I_{max} A_m} \right) \right] \quad (7.9)
 \end{aligned}$$

La figura 7.2 muestra la salida discreta del ACVO a una rampa, donde A_m fue elegido igual a 1. Se puede ver que los valores bajos y altos de luminancias de los costados de la rampa se perciben diferentes y esta gráfica reproduce cualitativamente los resultados experimentales mostrados en la literatura [7.5] y en la figura 7.1.

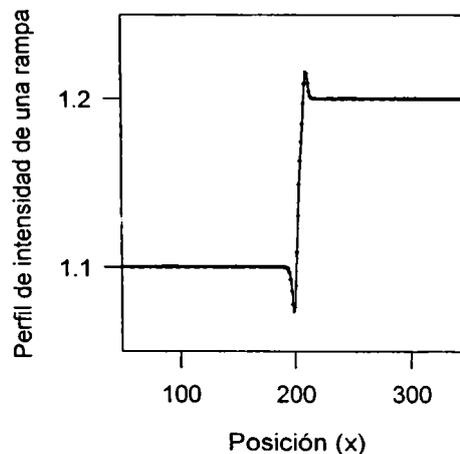


Figura 7-2. Respuesta discreta del ACVO a una rampa de intensidad. $I_{min}=0.1$, $I_{max}=0.2$, ancho=10.

7.4 Respuestas del ACVO a barras verticales

Ahora vamos a considerar las respuestas del ACVO a barras verticales. Suponemos que $I(x, y) = I(x) = I$ (un valor constante) para $|x| > b/2$; $I(x) = I + \Delta I$ para $|x| \leq b/2$ de modo que la entrada consiste en una barra de ancho b , de intensidad $I + \Delta I$ superpuesta sobre un campo uniforme de intensidad I . Por lo tanto dicha entrada puede ser escrita como:

$$I(x, y) = I(x) = \begin{cases} I & |x| > \frac{b}{2} \\ I + \Delta I & |x| \leq \frac{b}{2} \end{cases}$$

Entonces, la salida analítica del ACVO usando la ecuación (7.6) es:

$$\begin{aligned} D[I(x)(p)] = & \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{b}{2}} \sqrt{I} \sqrt{I + I.A_{min}} \exp - \left[\frac{1}{2}(x-p)^2 \frac{I}{1 + I.A_{min}} \right] dx + \\ & + \frac{1}{\sqrt{2\pi}} \int_{\frac{b}{2}}^{-\frac{b}{2}} \sqrt{I + \Delta I} \sqrt{1 + (I + \Delta I).A_{min}} \exp - \left[\frac{1}{2}(x-p)^2 \frac{I + D}{1 + (I + D).A_{min}} \right] dx + \\ & + \frac{1}{\sqrt{2\pi}} \int_{\frac{b}{2}}^{\infty} \sqrt{I} \sqrt{I + I.A_{min}} \exp - \left[\frac{1}{2}(x-p)^2 \frac{I}{1 + I.A_{min}} \right] dx \quad (7.10) \end{aligned}$$

Si hacemos, en la primera y en la tercera integral, un cambio de variables de la forma:

$$u = (x - p) \sqrt{\frac{I}{1 + I.A_{min}}}$$

y en la segunda otro definido por:

$$u_1 = (x - p) \sqrt{\frac{I + \Delta I}{1 + (I + \Delta I) A_{min}}}$$

la respuesta está dada por:

$$\begin{aligned} D[I(x)(p)] = & (1 + IA_{min})N\left(-\sqrt{\frac{I}{1 + IA_{min}}}\left(\frac{b}{2} + p\right)\right) + (1 + IA_{min})N\left(\sqrt{\frac{I}{1 + IA_{min}}}\left(\frac{b}{2} - p\right)\right) + \\ & + (1 + (I + \Delta I)A_{min})N\left(\sqrt{\frac{I + \Delta I}{1 + (I + \Delta I)A_{min}}}\left(\frac{b}{2} - p\right)\right) - \\ & - (1 + (I + \Delta I)A_{min})N\left(\sqrt{\frac{I + \Delta I}{1 + (I + \Delta I)A_{min}}}\left(\frac{b}{2} + p\right)\right) \end{aligned} \quad (7.11)$$

donde N es la Función de Distribución Normal acumulativa dada por:

$$N(p) = \int_{-\infty}^p \frac{1}{\sqrt{2\pi}} \exp(-1/2 x^2) dx$$

Como se puede ver en la ecuación (7.11) la salida depende del ancho b , de la intensidad del fondo I y del incremento de intensidad ΔI . La figura 7.3 muestra, usando la versión discreta (ecuación (7.11)), la respuesta del modelo para $b=I$ y para la fracción de Weber $\Delta I/I$ constante. Cuando la intensidad I se incrementa, la barra angosta produce una respuesta cuyo perfil tiene la forma de un sombrero. Este resultado es el mismo que el que se obtiene usando tanto el modelo CVO como la función de dispersión lineal del modelo lineal de inhibición lateral, basado en la diferencia de dos Gaussianas o el Laplaciano negativo de una Gaussiana.

7.5 La Función de Sensibilidad al Contraste y el ACVO

La Función de Sensibilidad al Contraste (CFS) es una medida de la modulación de la función de Transferencia (relación entre el contraste de salida y el contraste de entrada) del

sistema visual humano cuando la entrada es un conjunto de sinusoides de diferente frecuencia espacial. Esta Función de Transferencia puede definirse sólo en sistemas lineales y el sistema visual es no lineal, lo mismo que el CVO y el ACVO. Sin embargo, cuando la entrada es una senoide de muy bajo contraste ambos modelos son aproximadamente lineales [7.6] y se puede definir la Función de Transferencia.

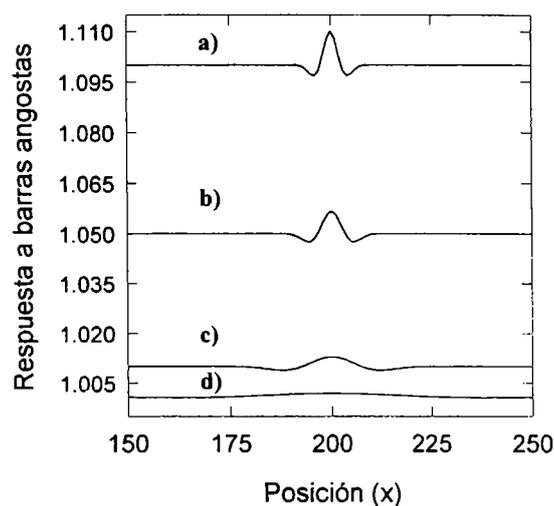


Figura 7.3. Respuesta del modelo ACVO a barras angostas. En este caso se eligió: $b=1$ y $\Delta I/I=0.1$.
 a) $I=0.1$, $\Delta I=0.01$. b) $I=0.05$, $\Delta I=0.005$. c) $I=0.01$, $\Delta I=0.001$. d) $I=0.001$, $\Delta I=0.0001$.

Para medir la CFS, se muestra al observador un campo uniformemente iluminado donde él es capaz de cambiar el contraste de la senoide. Cuando ese contraste, definido por la ecuación (6.3) alcanza la detección umbral, esa medida es almacenada y el experimento comienza nuevamente con una frecuencia espacial diferente. La sensibilidad al contraste se define como la inversa de este contraste umbral para cada frecuencia.

En la formulación del CVO el cálculo de la sensibilidad al contraste se realiza analíticamente y tiene la forma dada por la siguiente ecuación (referencia [7.7]) :

$$G(f, L) = 2\pi^2 \left(\frac{f}{L} \right)^2 \exp \left(-2\pi^2 \left(\frac{f}{L} \right)^2 \right)$$

donde L es el nivel de intensidad media y f es la frecuencia espacial. Se puede observar el comportamiento de esta función en la figura 6.8.

Para calcular la CFS con el ACVO comenzamos usando un valor de contraste umbral externo (es decir, objetivo) obtenido mediante experimentos psicofísicos. Umbral externo significa que es el menor valor del contraste objetivo que induce la detección del observador cuando mira un patrón generado. Entonces, se supone que una señal interna (es decir, subjetiva) es inducida por dicho valor de contraste, que es la necesaria para alcanzar la detección. Este valor fue calculado usando el modelo. Luego, supusimos que ese contraste interno era el mismo para todas las frecuencias y calculamos cual era el contraste externo necesario para alcanzar la detección en todas las frecuencias restantes. Repetimos los cálculos para distintas luminancias medias y los resultados se pueden ver en la figura 7.4

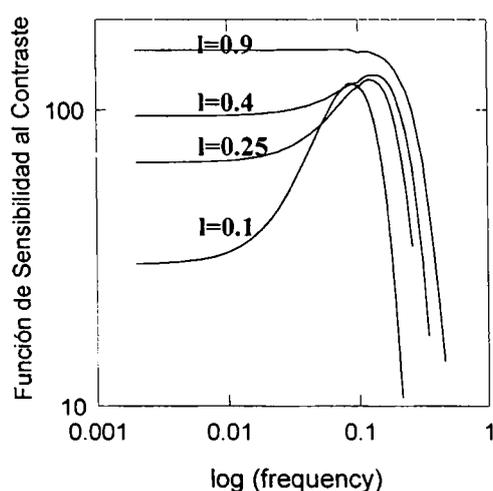


Figura 7.4. Curva de la Función de Sensibilidad al Contraste para varias luminancias medias, en función de la frecuencia espacial en escala logarítmica medidas en ciclos por receptor.

En dicha figura se puede ver que a medida que la luminancia media aumenta, la curva se corre hacia las altas frecuencias y la frecuencia pico se incrementa levemente. Pero este comportamiento se produce en un rango de variación de la luminancia que es menor que 1 unidad logarítmica, mientras que en los datos experimentales mostrados en la figura 6.9, esto ocurre sobre un rango de 6 unidades. Otra discrepancia que surge es que las curvas de la figura 7.4 son pasabandas para bajas luminancias medias y pasabajos para valores altos

de la luminancia media, mientras que los datos experimentales presentan el comportamiento opuesto.

7.6 La ley de Weber y el ACVO

La ley de Weber surge en los operadores de volumen constante como una consecuencia natural del teorema de cambio de escala. Este teorema lamentablemente, no es válido cuando se considera el tamaño finito de los receptores. Si aplicamos esta corrección a un borde se obtiene a la salida que la parte superior de las Bandas de Mach no obedece la ley de Weber y ambos lados de dichas bandas no son más simétricos. Cuando la luminancia superior de un escalón es suficientemente alta de tal manera que estén activos pocos o un sólo receptor, la banda de Mach correspondiente tiende a desaparecer. Esto es debido, como mencionamos antes, a que la respuesta del modelo modificado tiende a copiar la distribución de luminancia de entrada.

Deseamos mantener válida la ley de Weber en este modelo. Para recuperar esta ley debemos restar precisamente lo que hemos sumado, esto es las partes laterales de la respuesta del escalón de luminancia; las cuales son efectivamente percibidas. Para obtener la ley de Weber sustraemos de la respuesta del modelo el resultado que se obtendría si no hubiera dispersión en la suma. Esto es como si los detectores mediaran entre dos sistemas. En uno de ellos se aplica el modelo de volumen aproximadamente constante y en el otro, cada detector transmite la información sin modificarla como si estuviera en el rango de altas luminancias. A este modelo originado mediante la respuesta del ACVO para altas luminancias lo denominaremos ACVO corregido. Cuando ambas respuestas se sustraen, se obtienen los valores correspondientes a los del modelo de volumen constante y con ellos una aproximación a la validez de la ley de Weber.

Aplicamos esta idea a un escalón de entrada. La figura 7.5 muestra la relación $\Delta I/I$ requerida para alcanzar un cierto umbral de detección interno, calculada con la diferencia entre la respuesta del ACVO y el ACVO corregido, como función del logaritmo de la intensidad I , donde ΔI es el incremento de intensidad umbral objetivo necesario para

detectar el escalón. Se puede observar que sobre un cierto rango de intensidades, $\Delta I/I$ puede ser considerado aproximadamente constante. Este resultado coincide con los datos psicofísicos obtenidos por tres grupos de experimentadores pero sólo en una zona restringida como se puede ver en la figura 7.6 , y en la referencia.[7.8].

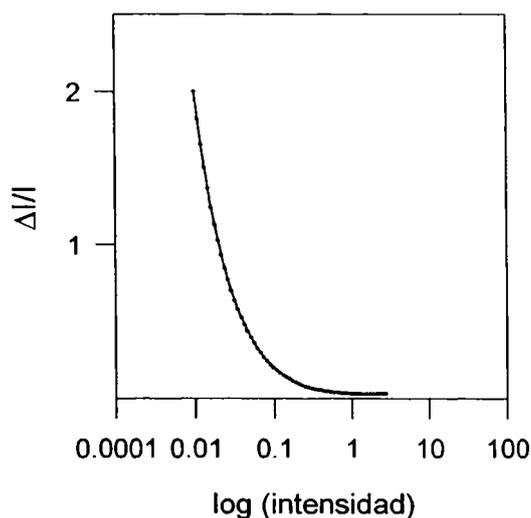


Figura 7.5. Ley de Weber obtenida con el modelo ACVO corregido. La relación $\Delta I/I$ en función del logaritmo de la intensidad I .

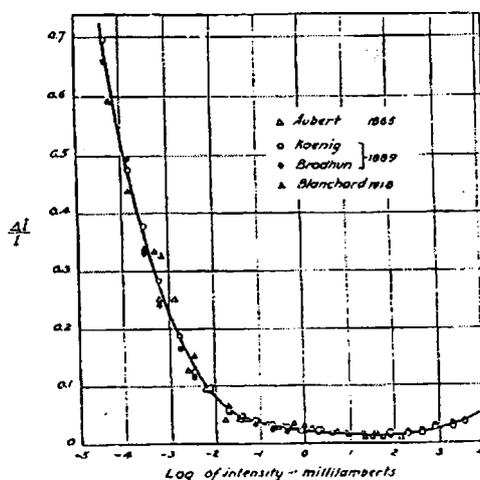


Figura 7.6. Discriminación de intensidades sobre el rango total de iluminación visible del ojo. Los datos de Blanchard están dados en mililamberts. Las intensidades de Koenig y Brodhun han sido divididas por 250 para convertirlas en mililamberts. De la misma manera las de Aubert han sido divididas por 10.000. Estos tres grupos de experimentadores que han trabajado independientemente han arribado esencialmente a los mismos resultados.

7.7 La señal. Aplicación de las bandas de Mach

Vamos a definir ahora una medida asociada con la magnitud del realce de bordes introducida por el modelo. Será útil para esta descripción considerar que la imagen de entrada es sólo un borde como en la figura 7.7 a). La salida del modelo ACVO a este borde se muestra en la figura 7.7 b), la cual presenta bordes de Mach.

Para dar cuenta sólo del realce introducido por el modelo no lineal podemos sustraer la respuesta que se obtendría si el modelo usara sólo un receptor como área de integración (modelo ACVO corregido). Ésta (la imagen obtenida de un solo receptor) es una versión como la que se obtendría a través de un filtro pasabajo del objeto. Su imagen sería vista como si usara la resolución espacial completa, compatible con el tamaño finito de los receptores. La salida es una copia del objeto de entrada y no se producen ilusiones. En el ejemplo considerado, el resultado de la sustracción se muestra en la figura 7-7 c).

El valor medio de la señal resultante es 1 como en el CVO. Si se sustrae 1 de este resultado se obtiene una función de valor medio cero quedando solamente las bandas de Mach. Entonces, podemos usar el valor absoluto o la energía de esta función y sumar todos los valores para obtener un resultado que sea proporcional al área bajo las bandas de Mach. El resultado final (figura 7-7 d) depende del borde considerado y su valor va a ser testeado como una salida cuantitativa del modelo.

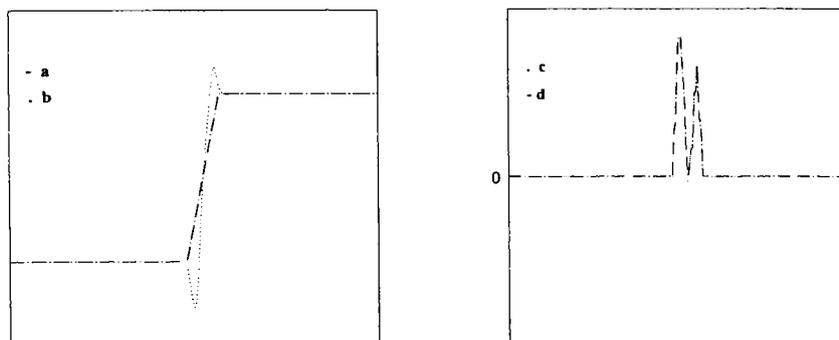


Figura 7.7. La *señal* de un borde. a) Borde de entrada. b) Salida del modelo ACVO que corresponde a intensidades intermedias. c) Respuesta de modelo cuando se ha sustraído a b) la respuesta del mismo a intensidades muy altas. d) Resultado final cuando el valor medio (1) es sustraído a c) y luego se ha tomado el valor absoluto.

La señal resultante de cualquier imagen puede ser considerada como un conjunto de Bandas de Mach correspondientes a todos los bordes presentes en la imagen, contribuyendo en la salida, en principio, en forma no lineal. Podemos decir que esta *señal* da cuenta del “contenido de las bandas de Mach de la imagen”, para mantener en mente qué es lo que se supone que se está midiendo, pero debe recordarse que el modelo es no-lineal. Es sólo aproximadamente lineal cuando el contraste es muy bajo [7.6]. La no linealidad del modelo hace muy difícil demostrar que el valor sugerido de la señal es máximo para un caso general en las cercanías del enfoque correcto. De todas maneras, como la operación de enfocar una imagen se facilita en la presencia de iluminación estructurada, vamos a introducirla a través de franjas proyectadas sobre la escena. Por lo tanto, vamos a calcular ahora la respuesta *señal* del ACVO sobre un conjunto de franjas sinusoidales proyectadas sobre un objeto plano y que sufren desenfoque progresivo.

En resumen, el procedimiento es como sigue:

- a) se calcula la salida del modelo visual ACVO (ecuación 7.7).
- b) se calcula la respuesta del modelo ACVO corregido (ecuación 7.8).
- c) se restan los resultado a)-b).
- d) se resta 1 a d) para obtener una función con valor medio cero.
- e) Se integra el valor absoluto sobre el área de interés.

7.8 El valor de la señal para perfiles sinusoidales teóricos

En esta sección, para respaldar este criterio, vamos a calcular el valor de la *señal* para perfiles de intensidades teóricos de entradas sinusoidales desenfocadas. Para ello, en primer lugar utilizamos la distribución de intensidad $I_1(x)$ observada por un sistema formador de imágenes, de una entrada unidimensional sinusoidal desenfocada, dada por la ecuación (4.10). A continuación, suponemos que esta intensidad se proyectó sobre un plano, que observamos su imagen (produciéndose un doble desenfoque) y que obtuvimos la distribución de intensidad $I_2(x)$ dada por la ecuación (5.1). Luego, para incrementar el contenido de las altas frecuencias en la imagen resultante aplicamos el operador derivada a dicha ecuación. Este resultado está dado por la ecuación (5.2). Por último, calculamos

numéricamente la *señal* en la forma descrita en la sección anterior, para las intensidades $I_1(x)$, $I_2(x)$, $I_3(x)$ usando las ecuaciones (7.7) y (7.8). En la figura 7.8 se muestran los resultados correspondientes a $I_1(x)$, e $I_2(x)$ en función del desenfoque z y en la figura 7.9, los de $I_3(x)$.

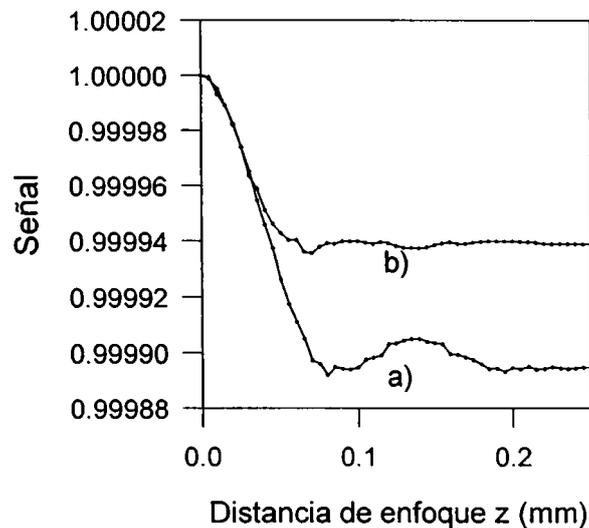


Figura 7.8. La respuesta teórica de la señal en función del desenfoque z (mm) para: a) desenfoque simple, b) doble desenfoque. La posición del foco correcto corresponde a $z=0$.

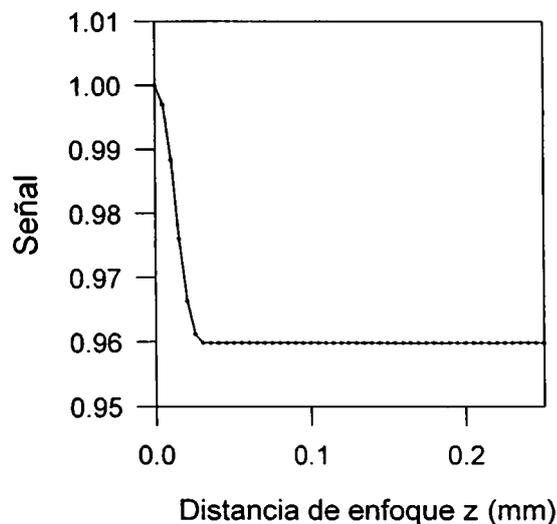


Figura 7.9. La respuesta teórica de la señal en función del desenfoque z (mm) para la derivada de franjas doblemente desenfocadas. La posición del foco correcto corresponde a $z=0$.

Se puede ver en estas figuras que el máximo se encuentra en la posición que corresponde al mejor foco ($z=0$). Los lóbulos laterales, si están presentes, son tan pequeños que no

pueden ser percibidos en el dibujo. También, se puede observar que la parte de la variación de la señal sin derivada es baja, aparece en el quinto decimal.

7.9 Experiencia.

Realizamos una serie de experiencias para chequear el comportamiento del método como un criterio para construir una imagen con el desenfoque corregido.

Comenzamos trabajando con objetos planos con desenfoque progresivo para determinar la posición del enfoque correcto. A efectos de poder comparar estos resultados con los obtenidos usando el criterio de máxima visibilidad, utilizamos las imágenes de un objeto plano, mostradas en la figura 5.2. Medimos el valor de la señal, como describimos antes, para cada pixel de cada imagen, para todas las imágenes de cada conjunto y graficamos la respuesta del modelo en función del número de imagen. La imagen 10 fue considerada por el observador como la mejor enfocada. Los resultados que corresponden al objeto plano con iluminación común (figura 5.2 a)) se muestran en la figura 7.10 a), los obtenidos con proyección de franjas (figura 5.2 b)) en la figura 7.10 b) y los hallados con la derivada de las franjas proyectadas, en la figura 7.10 c).

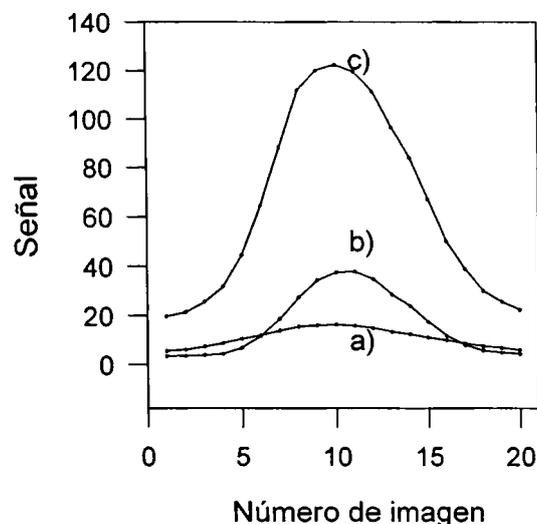


Figura 7.10. Resultados experimentales de la respuesta señal en función del número de imagen para los tres casos. a) simple desenfoque con iluminación común. b) doble desenfoque con franjas proyectadas. c) doble desenfoque con la derivada de las franjas proyectadas. La imagen 10 corresponde a la imagen con mejor foco.

Se puede ver en dicha figura que los valores del criterio son mayores para las imágenes a las cuales se les aplicó la derivada a las franjas proyectadas.

Por último, para construir una imagen resultante, a partir de un conjunto de imágenes de un objeto tridimensional, usamos las imágenes de las figuras 5.9, 5.10 y 5.11. Cada conjunto fue procesado de la siguiente manera:

- a) El criterio se aplicó a cada imagen.
- b) Se buscó en la imagen resultante procesada, el valor máximo de intensidad en la dirección del desenfoque para cada valor de pixel.
- c) La imagen corregida fue construida usando las correspondientes intensidades en el conjunto obtenido usando iluminación común, correspondientes a las posiciones de los máximos encontrados en b).

Los resultados correspondientes a las imágenes con iluminación común se muestran en la figura 7.11.a), los obtenidos con iluminación estructurada en la figura 7.11 b) y con la derivada de las franjas proyectadas en la 7.11 c).



a)



b)

Figura 7.11. Imágenes resultantes con el desenfoque parcialmente corregido. a) usando iluminación común. b) usando franjas proyectadas.



c)

Figura 7.11 c) usando la derivada de las franjas proyectadas.

La figura 7.11 a) muestra un sombreado que es debido a la falta de detalles y al viaje radial que realiza la función de punto extendido durante el desenfoque. Se puede ver en la figura 7.11 b) que el método mejora con la presencia de iluminación estructurada, como se puede notar en las regiones del objeto que no están cercanas al borde con el fondo. La oscuridad de este fondo carece de detalles e induce errores. Esos errores son severos debido a que las zonas de baja intensidad requieren grandes áreas de integración. Esta situación podría mejorar si se incluye un fondo más brillante y si es posible también franjas proyectadas sobre él. En la figura 7.11 c) el resultado es mejor que en la figura anterior porque las franjas han sido realizadas mediante el operador derivada.

Como en la práctica existen pequeños errores de alineación, las franjas proyectadas no sólo cambian de contraste sino sufren pequeños corrimientos de fase. Estos corrimientos inducen errores en la elección de la imagen que muestra máxima señal para cada pixel. Estos errores aparecen en forma de franjas en la imagen corregida. Para evitar estos errores, dos conjuntos de franjas complementarias o franjas desfasadas en forma escalonada podrían usarse, pero se incrementaría sustancialmente tanto el tiempo de registro como el de procesamiento.

Encontramos que un paso de filtrado pasabajo en la señal de las imágenes disminuye este efecto obteniéndose resultados aceptables como puede verse en la figura 7.12. Un filtro pasabajo se aplicó a cada imagen señal con franjas proyectadas derivadas y luego se construyó la imagen corregida usando las correspondientes intensidades en el conjunto con iluminación común, correspondientes a la posición de los máximos valores encontrados en cada imagen señal. Los resultados pueden ser comparados con los obtenidos con el mismo objeto y con la maximización de la visibilidad local en las figuras 5.12, 5.13 y 5.14; esto es son de calidad similar.

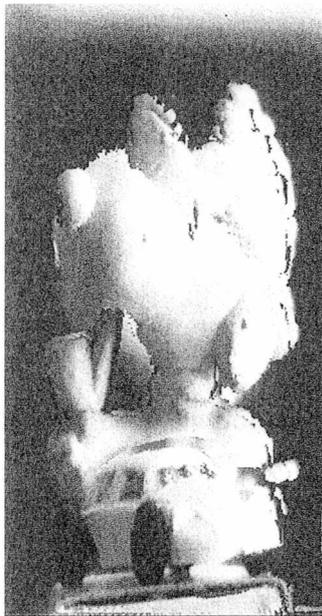


Figura 7.12. Imagen corregida de un objeto 3D usando en un paso intermedio un filtro pasabajos

7.10 Bibliografía

- [7.1] P. Torroba, N. Cap, H. Rabal. "Correction of defocusing using a modified visual model", *Optik*, **110**, N^o 7, 341-349, (1999).
- [7.2] Günter Wyszecki & W. S. Stiles, *Color Science, Concepts and Methods, Quantitative and Formulas*, John Wiley & Sons.

- [7.3] Peter Lindsay y Donald Norman Ed.: Human Information Processing.
- [7.4] Cornsweet, T. N. Prentice Award Lecture: A Simple Retinal Mechanism That Has Complex and Profound Effects on Perception, *Am. J. of Optom. & Physiol. Optics*, vol. 62, N^o 7, pp.427-438, 1985
- [7.5] Lowry, E. And Palma, J. :Sine-Wave Response of the Visual System. I. The Mach Phenomenon, *J. Opt. Soc. Am* , 51,740,746, (1960).
- [7.6] T. Cornsweet and J. Yellott: Intensity Dependent spatial summation. *J. Opt. Soc. Am*, 2, 1769-1786, (1985).
- [7.7] Yellot, J. Jr.: Constant Volume Operators and Lateral Inhibition, *J. Of Mathematical Psychology*, 33, N^o1, (1989).
- [7.8] Hecht,S.:The Visual Dicrimination of intensity and theWeber-Fechner law, *J. Gen. Physiol.*, 7, 235-265, (1924).

Capítulo 8

Corrección del desenfoque usando la matriz de Co-ocurrencia y la Función Estructura

8.1 Introducción

En el capítulo 4 propusimos el uso de la maximización de la función visibilidad para identificar áreas correctamente enfocadas en un conjunto de imágenes de escenas 3D con desenfoque progresivo. Este criterio que tiene un buen comportamiento, utiliza para los cálculos solamente dos valores de intensidad dentro de cada ventana (el máximo y el mínimo). Este hecho tiene como consecuencia que la visibilidad sea sensible al ruido. Entonces, es necesario aplicar previamente a la corrección del desenfoque el filtro mediana para reducir el ruido impulsivo. A continuación, estudiamos el uso de una versión modificada del modelo visual de volumen constante de Cornsweet y Yellott para introducir una medida del contenido de las Bandas de Mach presentes en una imagen. Esta medida es también sensible al desenfoque y usa valores locales de intensidad en una ventana cuyo

tamaño depende los valores de intensidad de la imagen. Esta herramienta da buenos resultados también aunque lleva tiempos computacionales más largos.

En este capítulo presentamos otras dos aproximaciones que hemos desarrollado de la corrección del desenfoque en un conjunto de imágenes de escenas 3D, usando el momento de segundo orden de la matriz de Co-ocurrencia de los valores de intensidad de la imagen y la Función Estructural [8.1]. Ambos criterios eligen zonas enfocadas dentro de un conjunto de imágenes de escenas 3D. Luego de esta selección, se puede construir una imagen con el desenfoque parcialmente corregido. Estas dos herramientas alternativas usan todos los puntos que están dentro de la ventana de modo que las diferencias entre dos pixels contiguos contribuyen al resultado final.

Comenzamos definiendo la matriz de Co-ocurrencia modificada (COM) de los niveles de intensidad por valores normalizados, dividiendo cada uno de ellos por el número de veces que aparece el primer nivel de gris. Luego, definimos el momento de Inercia de esta matriz con respecto a la diagonal principal, como una medida de la importancia de los cambios de intensidad de la imagen. Por último, este valor y la Función Estructural se aplican a simulaciones, a imágenes de objetos planos y a objetos 3D para construir una imagen en donde el desenfoque está parcialmente corregido.

8.2 La matriz de Co-ocurrencia y su Momento de Inercia

La matriz de Co-ocurrencia (COM) de los niveles de intensidad de una imagen ha sido usada anteriormente en el análisis de texturas [8.2] y en la caracterización del fenómeno de speckle dinámico [8.3]. Esta matriz se define como:

$$COM = [N_{ij}] \quad (8.1)$$

Esto es, la matriz tiene como entradas el número N de ocurrencias de un cierto valor de intensidad i , que es inmediatamente seguido por un valor de intensidad j . Éste es un caso particular de lo que se llama “dependencia matricial de niveles de gris” [8.4]. Con fines de

normalización, es conveniente dividir cada fila por el número de veces que aparece el primer nivel de gris.

$$M_{ij} = \frac{N_{ij}}{\sum_i N_{ij}} \quad (8.2)$$

Esta matriz provee varias medidas que pueden ser usadas en la caracterización de texturas. En particular, el momento de segundo orden de su versión modificada, es una medida de cuan abruptos son los cambios de los niveles de gris que contiene una imagen.

La intensidad de una imagen es el resultado de la convolución entre la intensidad de un objeto y la función de punto extendido no negativa, la cual se ensancha a medida que el desenfoque se hace más severo. Esto es, la imagen desenfocada es una versión cada vez más acentuada de un filtro pasabajos del objeto, a medida que el desenfoque aumenta. Entonces, los cambios más abruptos de la intensidad se esperan en las regiones mejor enfocadas. Vamos a usar el momento de segundo orden de la COM como una medida de cuan abruptos son estos cambios.

El momento de segundo orden, llamado acá Momento de Inercia (IM), se define como:

$$IM = \sum_{ij} M_{ij} (i - j)^2 \quad (8.3)$$

Este es un momento particular de segundo orden de la matriz y lo vamos a llamar Momento de Inercia (IM) con respecto a la diagonal principal en la dirección de la fila (se adoptó este nombre en base a la analogía mecánica de esta operación). Vamos a ver que el IM presenta un máximo en la imagen mejor enfocada cuando se lo grafica en función de la distancia de enfoque. La ubicación de este máximo es entonces usada como criterio para la elección de la mejor posición de enfoque.

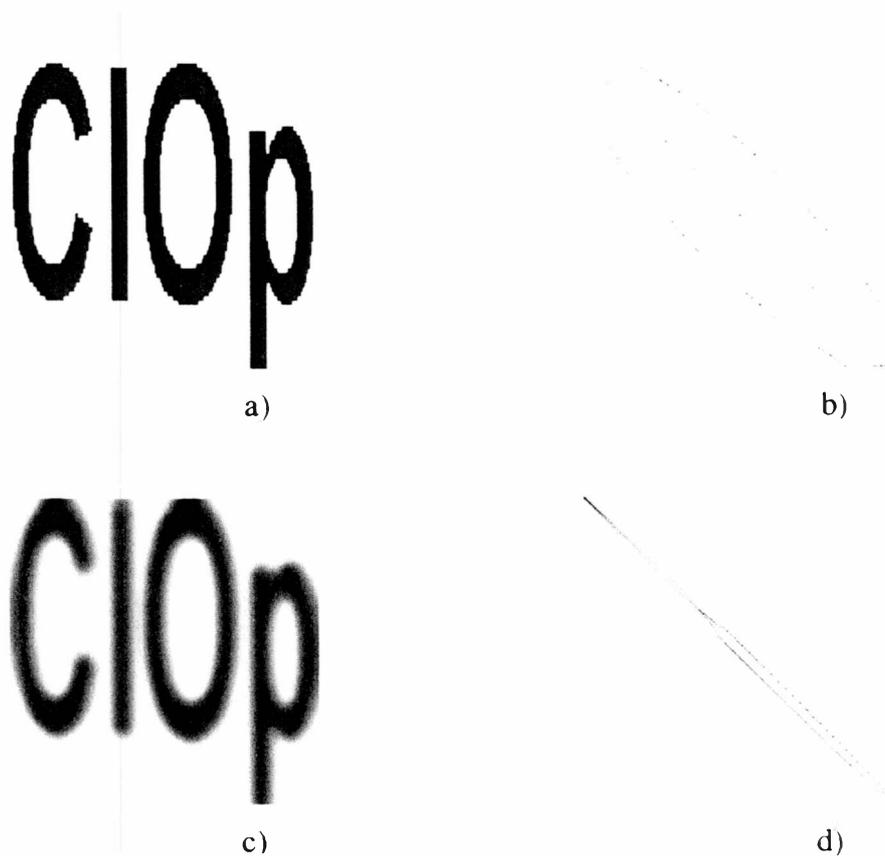


Figura 8.1. Simulaciones por computadora. a) Imagen enfocada. b) Matriz de co-ocurrencia correspondiente a la imagen enfocada. c) Imagen desenfocada. d) Matriz de co-ocurrencia correspondiente a la imagen desenfocada.

La figura 8.1 a) y c) muestra dos imágenes desenfocadas obtenidas por simulación por computadora. Esta simulación es producida mediante una convolución con una Gaussiana. Este tipo de PSF es a veces usado para simular el efecto de desenfoque [8.5]. Las COM correspondientes a las dos posiciones de enfoque se muestran como imágenes en la figura 8.1 b) y d). Se puede observar que las entradas de la COM están más dispersas en el caso enfocado que en el desenfocado. El correspondiente Momento de Inercia está graficado en la figura 8.2 en función de la desviación estándar de la Gaussiana.

La medida del IM de la dispersión de la COM es notable en dicho gráfico. La ubicación del máximo indica la posición correcta de enfoque. El valor del IM cae rápidamente, mostrando una gran sensibilidad de esta medida con respecto al desenfoque,

que es proporcional a la desviación standard. Además, la curva no muestra lóbulos laterales que es una característica importante para su uso en enfoque automático.

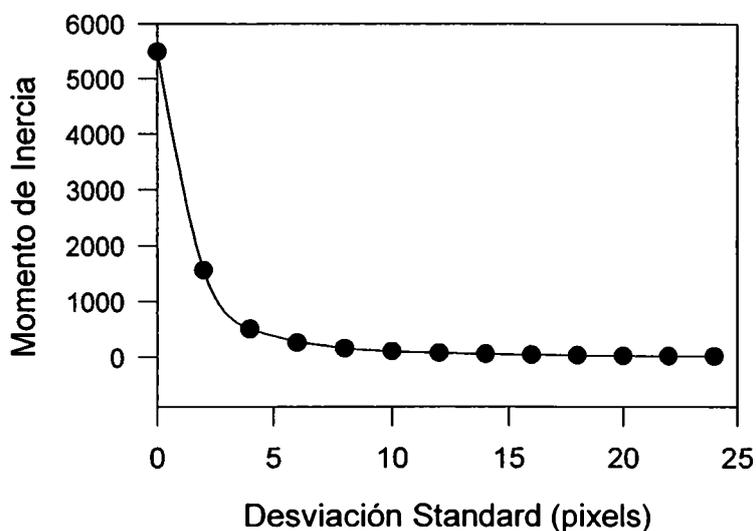


Figura 8-2. Simulación del momento de Inercia de imágenes desenfocadas, generadas computacionalmente.

Como pueden existir regiones en el objeto donde no hay demasiada energía en las zonas de alta frecuencia de su espectro, el método se beneficia con la proyección sobre el objeto de detalles de alto contraste, usando iluminación estructurada. Estos detalles son introducidos por medio de franjas proyectadas sobre el objeto en la misma dirección de observación. El esquema experimental es el mostrado en la figura 5.1.

A continuación, vamos a verificar el procedimiento con una distribución de intensidad teórica que es obtenida tomando en cuenta la óptica física, cuando un conjunto de franjas sinusoidales es proyectado sobre una superficie plana y observado por un sistema formador de imágenes en una situación de doble desenfoco experimental.

8.3 El Momento de Inercia de la COM de perfiles de intensidades teóricos desenfocados

Cuando una estructura sinusoidal es proyectada sobre una superficie plana y el resultado es observado con un sistema óptico, entonces, tanto la óptica de proyección como la de observación sufren desenfoque. En el capítulo 5 calculamos la distribución de intensidad de la imagen de un patrón de intensidad sinusoidal, proyectado sobre un plano en el cual se produce un doble desenfoque. Esta distribución de intensidad está dada por la ecuación (5.1). Usando esta expresión generamos imágenes con desenfoque simulado y calculamos la COM para cada una de ellas. La figura 8.3 muestra dos casos de la COM correspondientes a imágenes con diferente posición de enfoque. La COM, usualmente con una apariencia de nube, está en este ejemplo muy bien localizada. Las entradas distintas de cero están más dispersas y ocupan un rango mayor en el caso enfocado, mostrando el correspondiente alto contraste. Luego medimos los valores del IM en función del desenfoque y los resultados se muestran en la figura 8.4.

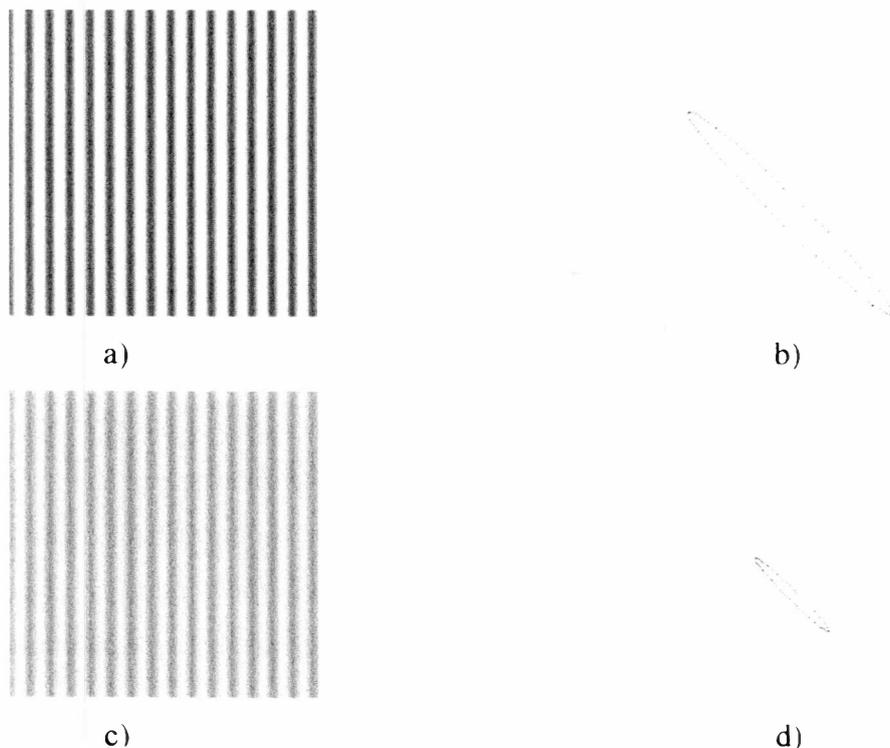


Figura 8.3. La matriz de co-ocurrencia de patrones de franjas sinusoidales desenfocadas generadas computacionalmente. a) Imagen enfocada. b) Matriz de co-ocurrencia correspondiente a la imagen enfocada. c) Imagen desenfocada. d) Matriz de co-ocurrencia correspondiente a la imagen desenfocada.

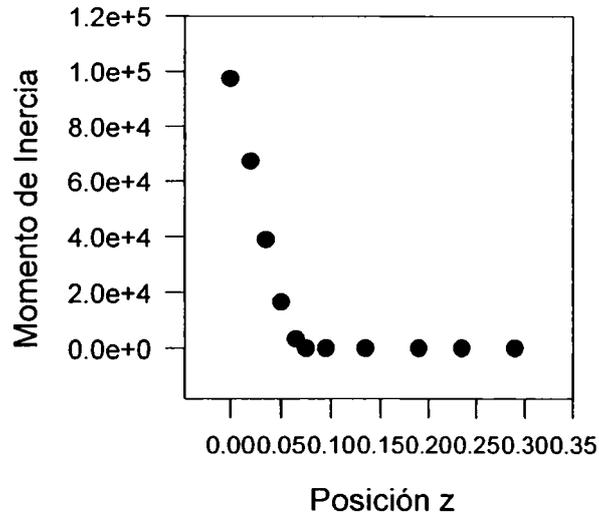


Figura 8.4. El momento de inercia de patrones de franjas desenfocados generados computacionalmente.

Como era de esperar, el IM es máximo en este caso y decrece a medida que el desenfoco aumenta. El IM parece ser entonces, una medida apropiada para elegir el mejor enfoque.

8.4 La Función Estructura

La Función Estructura (SF) de una señal de intensidad variable, la cual es usada en el estudio de fenómenos de scattering y en el movimiento Browniano [8.6], se define como:

$$SF = \frac{1}{N} \langle [n_i - n_{i+1}]^2 \rangle \quad (8.4)$$

en donde el paréntesis quebrado indica el promedio de ensamble, n_i es el número de fotones detectados en el intervalo i . Un estimador usual de este valor es el que vamos a aplicar a las imágenes y lo definimos como:

$$SF = \frac{1}{N} \sum [n_i - n_{i+1}]^2 \quad (8.5)$$

en donde, ahora n_i son los valores de intensidad en cada pixel y N es el número de pixels dentro de la ventana donde se evalúa la Función Estructura.

Para estimar el desenfoque cuando hay franjas proyectadas sobre el objeto, es suficiente calcular esta función solamente en la dirección perpendicular a las franjas, que es la dirección donde se espera a priori que las variaciones de intensidad sean más abruptas. Esta Función Estructura está cercanamente relacionada con el IM ya definido y muestra evidentes ventajas computacionales (dos órdenes de magnitud más rápido en el cálculo).

La figura 8.5 muestra la SF en función de la distancia de enfoque para un patrón sinusoidal proyectado sobre un plano que sufre doble desenfoque. Se puede observar un comportamiento similar a los valores del IM (Figura 8.4).

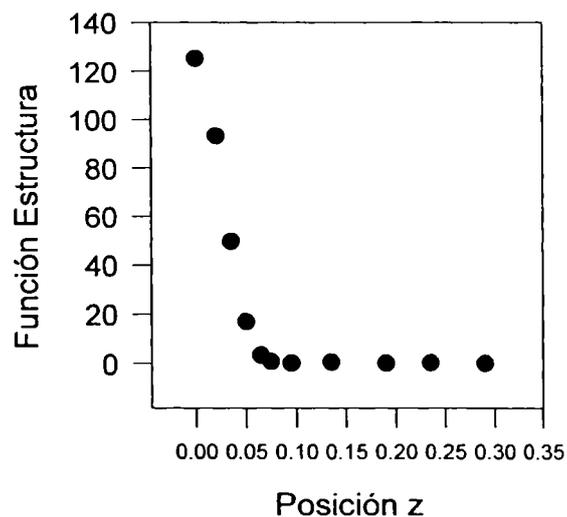


Figura 8.5. La función estructura de patrones sinusoidales desenfocados generados computacionalmente.

8.5 Experiencia

Realizamos varias experiencias para verificar el comportamiento local de las medidas del IM y de la SF con el objeto de utilizarlas como criterios para construir una imagen con el desenfoque corregido.

8.5.1 Objetos 2D

Comenzamos registrando con una cámara CCD treinta imágenes de un objeto plano con desenfoque progresivo, en ambas direcciones con respecto al foco correcto. Luego se calcularon el IM y la SF para cada pixel usando una ventana de 21x21 pixels centrada en él. En cada imagen se sumaron los resultados y se repitió este cálculo en cada una de las imágenes.

Debido a la presencia de ruido, las mismas medidas realizadas en iguales condiciones experimentales no dan los mismos resultados. Por lo tanto, repetimos las medidas diez veces y calculamos el valor medio y la desviación standard de los resultados. La figura 8.6 muestra los resultados obtenidos aplicando el valor medio de la SF en función de la distancia de enfoque. Se puede observar un máximo cerca de la imagen que fue considerada por el observador como la mejor enfocada.

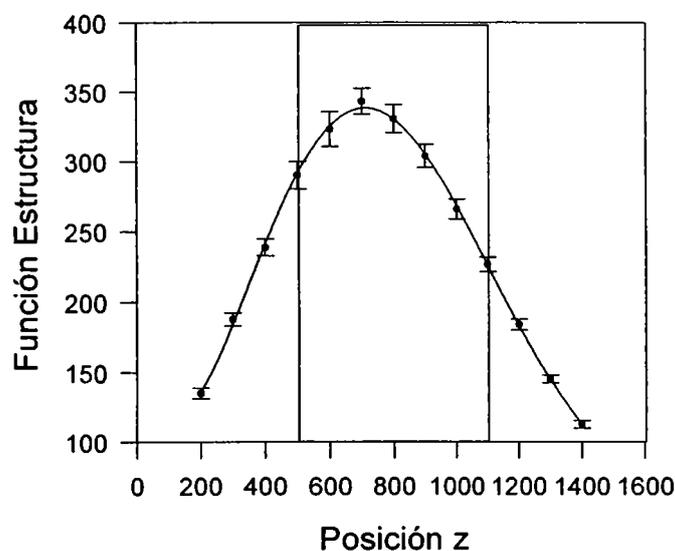


Figura 8.6. El valor medio de la SF de imágenes de objetos planos, tomadas experimentalmente.

En las mismas condiciones experimentales, un observador indicó mediante una inspección visual la mejor posición de enfoque y el menor intervalo en el cual era percibida la diferencia entre una y otra imagen. Esto es, el observador modificaba la distancia objeto hasta que notaba una variación en el enfoque. Este intervalo está indicado en la figura 8.6 como barras verticales y se puede ver que las medidas de la SF resuelven varios puntos dentro de dicha zona. El mismo comportamiento se obtuvo con el IM como se puede ver en la figura 8.7.

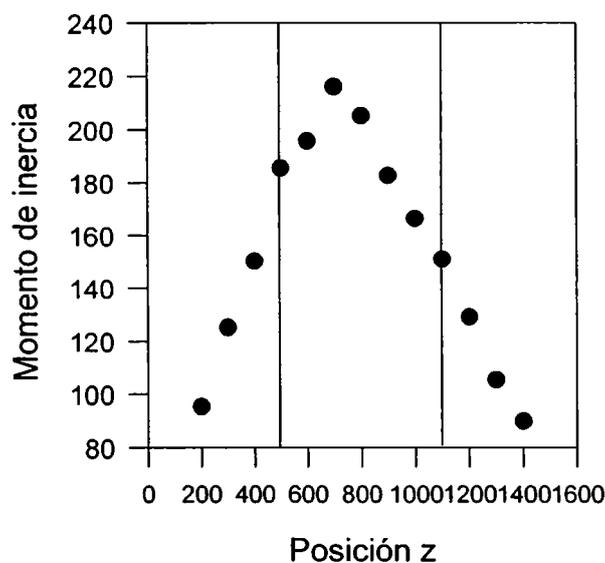


Figura 8.7. El valor medio del IM de imágenes de objetos planos, tomadas experimentalmente.

8.5.2 Objetos 3D

A continuación utilizamos los conjuntos de imágenes mencionadas en la sección 5-5 para construir las imágenes con el desenfoque parcialmente corregido. Los conjuntos se procesaron de la siguiente manera:

- a) Con el propósito de poder comparar los resultados, primero sólo se procesó el conjunto de imágenes obtenido con iluminación común. En este grupo, se midieron el IM y la SF en cada pixel de cada imagen en las diferentes posiciones de enfoque y la imagen resultante fue construida usando los valores de intensidad de las imágenes que mostraban un valor máximo de dichas medidas. La figura 8.8 muestra el resultado obtenido usando la SF.

b) Luego, el algoritmo usa el conjunto obtenido con franjas proyectadas para determinar la posición z que maximiza el IM y la SF. A continuación, selecciona los valores de intensidad que se corresponden con dichas posiciones, en el conjunto obtenido sin proyección de franjas, y construye la imagen corregida con estos valores de intensidad. La figura 8.9 es la imagen resultante usando el IM y la figura 8.10, usando la SF.

c) Por último, se repitió el inciso b), pero usando el conjunto obtenido con franjas proyectadas derivadas para determinar la posición que maximiza el IM y la SF. La figura 8.11 muestra el resultado usando la SF.

Si el objeto ya tiene detalles de alta frecuencia, el resultado no mejora si se agrega iluminación estructurada a las imágenes; pero si el objeto tiene bajo contraste, la mejora es importante.

En todos los casos el tamaño de ventana utilizado fue de 21x21 pixels y el número de imágenes fue 10.

Las imágenes en la figuras 8.9 y 8.10 son muy similares y compiten favorablemente con aquellas obtenidas previamente en los capítulos 5 y 7. Se puede ver que, excepto por algunos defectos (*artifacts*) en los bordes, los resultados son buenos. El procedimiento “eligió” la mayoría de las veces la imagen mejor enfocada. Los resultados obtenidos usando la SF son ligeramente mejores, esto es, el método aplicado comete menos “errores” en las elecciones que realiza. Se nota también, que el resultado en la figura 8.11 es sutilmente mejor que en la figura 8-10.

Se puede observar que la ausencia de franjas proyectadas en el fondo (figura 8-8), induce a elecciones equivocadas cerca de los bordes. Este inconveniente no es debido al algoritmo y se puede corregir introduciendo iluminación estructurada en el fondo.

El método no selecciona adecuadamente, cerca de los lugares donde se producen cambios abruptos en el objeto (límite entre el auto y el Snoopy). Estos errores son debidos

al tamaño finito de la ventana usada en los cálculos y son notablemente menores que en las aproximaciones usadas en los capítulos 5 y 7.



Figura 8.8. Imagen construida usando iluminación común y la SF



Figura 8.9. Imagen corregida usando el momento de inercia.



Figura 9.10. Imagen corregida usando la Función estructura.



Figura 8.11 Imagen resultante obtenida usando la SF y el operador derivada.

8.6 Bibliografía

- [8.1] P. Torroba, H. Rabal and R. Arizaga. "Correction of defocusing using the co-occurrence matrix and the structure function", *J Mod Optics*, vol 47, 69-80, (2000).
- [8.2] S. Allam, M. Adel, and P. Réfrégier. "Fast Algorithm for Texture Discrimination by use a Separable Orthonormal Decomposition of the Co-occurrence matrix", *appl. opt.* 36 , 8313-8321,(1997).
- [8.3] R. Arizaga, M. Trivi and H. Rabal. "Speckle time evolution characterization by the co-occurrence matrix analysis". *Optics Laser Technology* **31**, 163-169 (1999).
- [8.4] A. L. Zobrist and William Thompson. *IEEE Transactions on Computers*, 718-728 (1975)..
- [8.5] X. Roca, X. Binefa and J. Vitria, "New autofocusing algorithm for cytological tissue in a microscopy environment". *Opt. Eng.* **37** (2) 635-641 (1998).

- [8.6] Benjamin Chu, *Laser Light Scattering, Basic Principles and Practice*, Academic Press Inc., Second edition, Boston, pp 129-130, (1991).

Capítulo 9

Conclusiones y perspectivas

En el marco de esta Tesis estudiamos dos fuentes de degradación ligadas a la formación de imágenes, el ruido y el desenfoco. Comenzamos tratando el ruido e implementamos la aplicación de una familia de filtros para disminuirlo. Luego, realizamos una descripción matemática del desenfoco y propusimos criterios nuevos para su corrección en imágenes de escenas 3D. Posteriormente, analizamos cómo soluciona la naturaleza ambos defectos a través del sistema visual. A partir del estudio de modelos visuales definimos otro criterio original para corregir el desenfoco en imágenes de objetos 3D. A continuación se sintetizan los resultados más significativos y se bosquejan algunas extensiones del presente trabajo.

En el capítulo 2 estudiamos el problema de disminuir el ruido en imágenes utilizando una familia de filtros correspondiente a la Media generalizada de orden fraccionario, con el orden h en el intervalo $[0,2]$. Estos filtros se aplicaron a imágenes que contenían ruido sal y pimienta y ruido gaussiano. Se utilizaron ventanas de 9 pixels y definimos otra ventana original de 8 pixels. Evaluamos el comportamiento de estos filtros en la reducción de

ruido, aplicando criterios de calidad objetivos ECM, CC y PL y uno subjetivo el ERM. Los criterios objetivos están relacionados con medidas de magnitudes físicas (distancia, energía) y el criterio subjetivo, propuesto por nosotros, tiene en cuenta las características promedio del sistema visual humano y lo denominamos ERM. Analizamos también el comportamiento de estos filtros en preservar bordes.

Los resultados obtenidos mostraron que todos los criterios objetivos dan aproximadamente la misma información y elegimos el ECM para describir el comportamiento de aplicar estos filtros a sinusoides. La calidad de las imágenes filtradas, desde el punto de vista del ECM, fue muy buena cuando se usaron ventanas de 9 pixels y valores de h cercanos a 1.3, en sinusoides de altas frecuencias con ruido sal y pimienta. Cuando se consideran las bajas frecuencias, el menor ECM se corre hasta $h=1.2$. Para estos valores de h la mejora con respecto a $h=1$ (la mediana) es de 20%. Entonces, en condiciones similares a las anteriormente mencionadas, es conveniente aplicar este filtro fraccionario con $h=1.3$ o $h=1.2$ en lugar de la mediana, a pesar del tiempo computacional que emplea en procesar la imagen. En otras situaciones, podría ser más eficiente usar la mediana porque consume menor tiempo de procesamiento.

Observamos también que las frecuencias espaciales bajas se corrigen mejor que las altas, esto es una consecuencia del tamaño de la ventana. Los resultados mostraron que para valores bajos de h la ventana de 8 pixels no es tan buena como la de 9 para la reducción del ruido.

Encontramos que el ECM usando ventana de 9 pixels es sistemáticamente menor que con ventanas de 8 pixels. Obtuvimos que el valor mínimo del ERM para ambos tipos de ventanas está en $h = 1$, esto coincide con el buen comportamiento del conocido filtro mediana.

Hallamos también excelentes resultados con respecto a la habilidad que tienen estos filtros en preservar bordes en imágenes con ruido Gaussiano. Sobre la imagen original libre de ruido, agregamos hasta 55% de ruido Gaussiano aditivo y luego aplicamos los filtros

sobre una fila de 14 pixels. No encontramos ninguna diferencia entre la imagen original y la filtrada usando ventanas de 9 pixels y todos los valores de h . Obtuvimos los mismos resultados aplicando la varianza local que es una medida de la difusión de los bordes. Por otro lado, todos los filtros tanto los de 8 como los de 9 pixels no lograron resolver una línea de 1 pixel de ancho.

A su vez, realizamos una evaluación visual preliminar de las imágenes filtradas y encontramos algunas evidencias psicofísicas que dirían que el ERM se relaciona mejor con la calidad de la imagen percibida que el ECM. Además, las ventanas de 8 pixels mostraron menor ERM que las de 9. Por lo tanto, esta evidencia parecería sugerir que cuando el aspecto visual es el motivo del filtrado, esto es, cuando no está orientado hacia aplicaciones metrológicas, las ventanas de 8 pixels son las más adecuadas para la reducción de ruido.

Está fuera del propósito de este trabajo, pero sería interesante aplicar estos filtros en imágenes con speckle para analizar franjas.

En el capítulo 3 comenzamos a estudiar el efecto del desenfoque. Se desarrolló el análisis de Fourier en la teoría de formación de imágenes para obtener la respuesta en frecuencia de un sistema óptico desenfocado iluminado incoherentemente. Se analizó el comportamiento de la FTO en el caso particular de una pupila rectangular. Una de las características de estas curvas es el rápido deterioro de la función respuesta de la lente para altas frecuencias aún para pequeños desenfoques. En el gráfico de la FTO como función de la frecuencia s , se produce un cambio de signo que da cuenta de una inversión del contraste, llamada resolución espúrea. A partir de la representación de la FTO en función del desenfoque z se puede observar que esta inversión del contraste se produce en zonas cada vez más cercanas a la posición de enfoque, a medida que aumenta la apertura numérica. Este hecho se acentúa cuando aumenta la frecuencia espacial s .

Se concluye de los gráficos que cuando el sistema óptico está fuera de foco se produce una atenuación del contraste y el número de veces que se invierte aumenta para las

altas frecuencias. Surgió también del análisis que el sistema óptico se hace más sensible al desenfoque a medida que aumenta la apertura numérica.

En el capítulo 4 comenzamos con el estudio de distintos criterios que ayudan a determinar cual es la posición de la imagen plana mejor enfocada, dentro de un conjunto de imágenes planas con desenfoque progresivo. Primero se analizaron algunos criterios conocidos y luego se propusieron otros nuevos. Los parámetros físicos nuevos que se ensayaron fueron: el gradiente, el desvío medio aritmético local y la visibilidad. A continuación, se realizaron una serie de experiencias con objetos planos con desenfoque progresivo y se calcularon todos los criterios tratados. A partir del análisis de los gráficos, se observó que la visibilidad, la varianza y el desvío medio aritmético presentaban un máximo en la imagen mejor enfocada. Además, los resultados mostraron que el máximo alcanzado por la visibilidad, como consecuencia de que la intensidad es un extremo en dicha posición, era el más agudo de los tres criterios mencionados y fue elegido como el mejor parámetro entre todos los propuestos. Estas experiencias se repitieron con diversos objetos y se obtuvieron resultados consistentes con todos los parámetros.

También notamos que los resultados eran sensibles a la presencia de ruido. Para disminuirlo, aplicamos a las imágenes el filtro mediana y un filtro promediador en el tiempo antes de calcular la visibilidad. Bajo estas condiciones, el gráfico de la visibilidad en función del desenfoque mostró en la imagen de mejor foco un máximo mucho más destacado que cuando no se aplica el promediador y el filtro mediana.

En este contexto propusimos usar la visibilidad para seleccionar la intensidad de la imagen mejor enfocada para cada posición de pixel y construir con ella una nueva imagen en la cual el desenfoque esté parcialmente corregido. Entonces, realizamos una experiencia utilizando imágenes de un objeto 3D, en la que la profundidad variaba de forma continua. En este caso el objeto empleado era una regla inclinada. En cada imagen había sólo una zona enfocada correctamente y ésta se paseaba desde un extremo a hacia el otro de la regla a medida que se variaba el desenfoque. Se calculó la visibilidad para cada posición de pixel de cada imagen y luego se construyó la imagen corregida a partir de los valores de

intensidad correspondientes a las posiciones de los máximos de la visibilidad de cada imagen. En la imagen resultante se pudo observar que todos los números de la regla así como sus divisiones estaban simultáneamente en foco.

También se notaba que la calidad de la imagen construida disminuía desde el centro hacia los extremos. Este resultado, se debe a que la función de punto extendido viaja en dirección radial a medida que el desenfoque progresa. Por lo tanto, si un detalle de alto contraste está cerca de una zona de intensidad constante (baja visibilidad), a medida que se vaya desenfocando invadirá zonas uniformes enfocadas. Entonces, la visibilidad resultará mayor en las áreas en las cuales está el detalle de alto contraste desenfocado que en la zona uniforme enfocada. Por este motivo, sugerimos que debían introducirse en las imágenes detalles de alto contraste que ayuden al método a “elegir mejor” la zona enfocada.

Por otro lado, hacemos notar que una optimización del rango dinámico mejoraría la imagen resultante, ya que la probabilidad de encontrar dos valores de visibilidad iguales en la misma posición de pixel pero en distintas imágenes sería despreciable.

Las imágenes corregidas fueron procesadas usando ventanas de 21x21, de 9x9 y de 3x3 pixels. Observamos que la calidad de la imagen resultante mejora cuando el tamaño de la ventana aumenta, pero tiene la desventaja que disminuye el tamaño de la imagen final. Consideramos que las más adecuadas fueron las ventanas de 21x21 pixels.

Estudiamos, desde la perspectiva de la óptica de Fourier, la visibilidad en función del desenfoque considerando como objeto una fuente sinusoidal unidimensional. Obtuvimos, usando la Función de Transferencia Óptica calculada en el capítulo 3, que la visibilidad en función del desenfoque se comporta como un seno cardinal. Entonces, demostramos teóricamente que la visibilidad es máxima en la posición del foco correcto que se corresponde con el máximo del seno cardinal.

Cuando trabajamos con imágenes de objetos que contienen amplias zonas uniformes, (comparadas con el tamaño de la ventana) el criterio de máxima visibilidad no da muy

buenos resultados en zonas alejadas al centro de la escena. Estos “errores” de enfoque que también los comete el sistema visual humano, se deberían a la ausencia de detalles de alto contraste. En el capítulo 5 introducimos iluminación estructurada sobre la escena para mejorar la calidad de la imagen resultante en donde el desenfoque está parcialmente corregido. Para ello, realizamos una serie de experiencias en las que proyectamos franjas sobre los objetos bajo estudio para incorporar la iluminación estructurada. En el esquema experimental usado se produce un efecto de doble desenfoque, uno en el sistema de proyección y el otro en el sistema formador de imágenes de la cámara CCD. Este esquema experimental produce un rápido desenfoque y un rápido decrecimiento de la visibilidad. A continuación, para realzar aun más las componentes de alta frecuencia le aplicamos a dichas imágenes el operador derivada y comparamos los valores de la visibilidad de ambos grupos con los valores obtenidos bajo iluminación común.

Inicialmente, trabajamos con imágenes de objetos planos y obtuvimos que los valores de la visibilidad, en promedio, son más grandes en las imágenes a las cuales le aplicamos el operador derivada a las franjas proyectadas. Los resultados indicaron también que el procedimiento es sensible a las frecuencias altas.

Desarrollamos también un estudio teórico de la visibilidad en función del desenfoque para una fuente sinusoidal doblemente desenfocada. Los resultados obtenidos teóricamente avalaron los hallados experimentalmente.

Además simulamos el efecto que produce en el resultado el hecho de variar el tamaño de la ventana. Concluimos que el tamaño óptimo de la ventana debería ser tal que contenga al menos un período de las franjas.

Posteriormente, trabajamos con imágenes 3-D con franjas proyectadas derivadas y sin derivar y corregimos el desenfoque aplicando el criterio de máxima visibilidad teniendo en cuenta las imágenes obtenidas con iluminación común. Las imágenes corregidas mostraron que la adición de iluminación estructurada al método de corrección de desenfoque usando el criterio de visibilidad, ayuda a reducir la sombra en la zona

periférica debido a la ausencia de detalles de alto contraste.

Cuando el objeto contiene detalles de alta frecuencia la mejora en el resultado final no es significativa, en cambio cuando el objeto tiene bajo contraste la mejora que produce este método es muy importante.

Cualquier ruido impulsivo produce valores altos de visibilidad en lugares inesperados, lo cual deteriora la imagen final. El efecto de este tipo de ruido es aún más notable cuando se aplica este método en las imágenes derivadas debido a que la derivada realza los valores de los cambios abruptos de intensidad. Estos valores extremos inducen a “elecciones erróneas” del algoritmo en regiones comparables al tamaño de la ventana. Para disminuir este efecto, cada imagen fue filtrada con un filtro mediana de 3x3 pixels. Este filtro fue usado sobre las imágenes antes de aplicar el operador derivada.

Los defectos que quedan en la imagen resultante que son pocos, se deben a que las imágenes que componen los distintos grupos no tenían un fondo sobre el cual hubiera detalles.

En el capítulo 6 comenzamos con el estudio de modelos visuales y analizamos detalladamente los operadores de volumen constante (CVO). En el capítulo 7 propusimos introducir una modificación al modelo visual de Cornsweet y Yellott o modelo IDS, que tiene en cuenta el tamaño finito de los receptores. Este modelo corregido lo denominamos ACVO. Obtuvimos que las salidas del ACVO, cuando las entradas fueron una rampa de luminancia y barras verticales, simularon cualitativamente las respuestas visuales humanas. También, aplicamos el ACVO a un conjunto de franjas sinusoidales de diferentes frecuencias para construir la Curva de Sensibilidad al Contraste. La respuesta del ACVO a estos estímulos reprodujo en sólo un aspecto el comportamiento de la curva experimental. Además, si en el modelo se introduce un término relacionado con las luminancias altas, entonces el ACVO corregido reproduce la ley de Weber pero sólo en una zona restringida.

Luego, en este capítulo propusimos el uso de las áreas bajo las Bandas de Mach para

estimar el desenfoque en un conjunto de imágenes. Usando el modelo ACVO definimos una medida del área bajo las Bandas de Mach. Demostramos que esta magnitud, denominada por nosotros *señal* obtenida con este modelo modificado tiene un máximo en la imagen correctamente enfocada y no presenta lóbulos laterales apreciables. Esta característica fue utilizada para procesar una imagen en donde el desenfoque está parcialmente corregido.

Calculamos el valor de la *señal* para perfiles de intensidades teóricos de entradas sinusoidales desenfocadas proyectadas sobre un plano objeto y obtuvimos que el máximo de la curva estaba en la posición de mejor enfoque. Para objetos arbitrarios que no sean planos, este desarrollo no necesariamente debe cumplirse. De todos modos, las imágenes de las franjas sinusoidales proyectadas pueden hacerse aproximadamente independientes de la reflectividad del objeto dividiendo las imágenes de las franjas por una imagen con iluminación uniforme. Hemos encontrado experimentalmente que este paso no es requerido usualmente pero podría ser implementado si es necesario hacer cambios muy grandes de reflectividad sobre el objeto.

Hacemos notar que nosotros no estamos implicando que este modelo visual modificado simula la manera que emplea el sistema visual para enfocar. Hemos analizado algunas razones por las cuales el modelo ACVO no es plausible para realizar esta operación.

Otro rasgo para destacar es que el área del objeto usada para determinar el mejor enfoque se adapta, al valor local de la Intensidad y como consecuencia disminuye el ruido.

Nuestro modelo requiere, de la misma manera que el sistema visual, la presencia de detalles de alta frecuencia y lo incluimos por medio de iluminación estructurada. Eligiendo el valor del máximo de la respuesta local en las imágenes con desenfoque progresivo, construimos las imágenes con la corrección parcial del desenfoque.

Encontramos que un paso de filtrado pasabajo en la *señal* de las imágenes

disminuye el efecto ocasionado por el corrimiento de fase de las franjas, obteniéndose resultados aceptables.

En el capítulo 8 propusimos y demostramos el uso del Momento de Inercia de la Matriz de Co-ocurrencia y de la Función Estructura como criterios para la selección del mejor enfoque. Usamos simulaciones por computadora, la expresión teórica de un patrón de intensidad sinusoidal desenfocado y un objeto plano para demostrar la factibilidad del método. Una característica importante que muestran las curvas del Momento de Inercia de la Matriz de Co-ocurrencia y la Función Estructura en función del desenfoco es que no muestran lóbulos laterales. Este hecho es un rasgo importante para su uso en enfoque automático. Además, en estas experiencias notamos que cuando el observador variaba la posición de enfoque no detectaba cambios en las imágenes y sí eran percibidos por ambos criterios. Por lo tanto, podemos decir que la Función Estructura y El Momento de Inercia resuelven varios puntos que el observador no distingue.

Luego, se aplicaron los criterios a la corrección del desenfoco en el caso de escenas de objetos 3-D. Como sucede en esta clase de métodos se requiere introducir sobre el objeto detalles de alta frecuencia. Estos fueron provistos por medio de iluminación estructurada.

La Función Estructura fue usada como una alternativa y se obtuvieron buenos resultados con tiempos computacionales sustancialmente menores.

Si comparamos las imágenes corregidas con estos criterios con las construidas en los capítulos 5 y 7 surge que los resultados obtenidos con la Función Estructura son ligeramente los mejores de todos los criterios aplicados.

Por último, para mejorar aún más la calidad de las imágenes corregidas sería conveniente agregar franjas proyectadas sobre el fondo de la escena. Su ausencia es la responsable de las “elecciones erróneas” de los algoritmos en las zonas cercanas a los bordes.

Se propone continuar con la aplicación del modelo ACVO a la solución de otros problemas específicos no considerados en el contexto de este trabajo. Así por ejemplo, se podría emplear para procesar algunas ilusiones de contraste. En particular, sería interesante buscar una similitud cualitativa entre las salidas del modelo y los efectos visuales para chequear la confiabilidad del mismo. Además, este modelo puede ser utilizado como una herramienta de procesamiento para disminuir o realzar algunas de las características de una imagen.

Iniciamos esta tarea generando por computadora la imagen que corresponde a la ilusión de la grilla de Hermann Hering [9.1]. Esta grilla consiste en una imagen de alto contraste de bloques cuadrados oscuros sobre calles blancas como la mostrada en la figura 9.1.

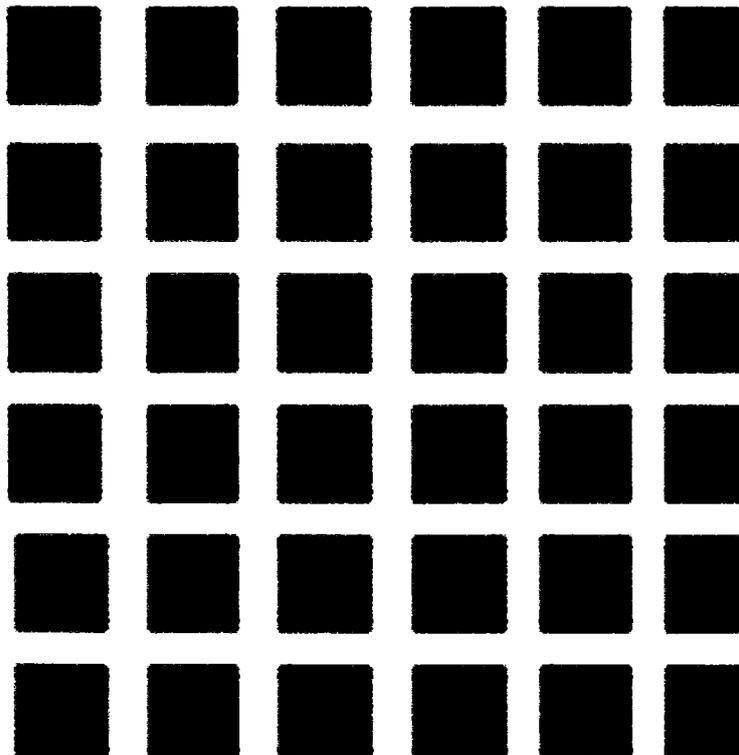


Figura 9.1. Grilla de Herman Hering

La ilusión consiste en la percepción de niveles de gris en las intersecciones de las calles. Es particularmente notable en las intersecciones en las que el observador no mira directamente y es más difícil de percibir en los cruces mirados directamente. Esta ilusión es usualmente atribuida a la inhibición lateral. La respuesta del modelo ACVO a una grilla Herman Hering se puede observar en la figura 9.2. Como se puede apreciar, el resultados hallado simula en forma bastante aproximada la ilusión percibida por el ojo humano. Sugerimos seguir con el estudio de la ilusión de la pared del café de Bristol [9.2] la de Helmholtz [9.3] y el efecto Craik-O'Brien-Cornsweet [9.4]

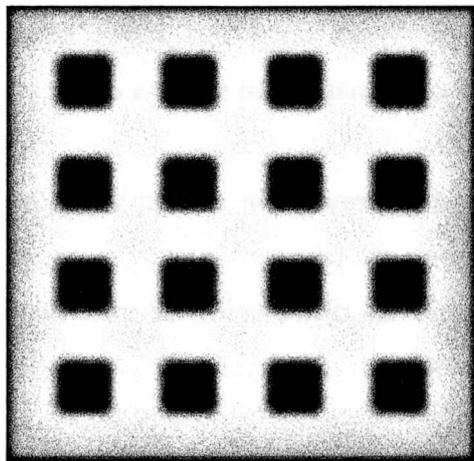


Figura 9.2. Respuesta del ACVO corregido a la Grilla de Herman Hering.

Otra posible aplicación del modelo ACVO sería procesar discontinuidades de gradientes de intensidad (llamadas crestas o líneas) y cambios discontinuos de luminancia (bordes) en presencia de ruido. Se pretende detectar bordes y crestas en imágenes que contienen ruido cuando no es posible hacerlo mediante una observación visual humana. Ambos tipos de discontinuidades son características importantes en una imagen y por lo tanto deben ser detectadas y distinguidas.

Se propone también, continuar con la corrección del desenfoque en imágenes de objetos 3D con desenfoque progresivo usando la función de contraste local Gamma [9.5]. Para aplicar este criterio se utiliza el esquema experimental de doble desenfoque [9.6]. En dicho arreglo experimental, se introduce en cada posición z de desenfoque, un corrimiento

de fase de $2\pi/3$ y otro de $-2\pi/3$. Con estas dos imágenes y la imagen sin desfazar se calcula Gamma, que es similar a la visibilidad, para cada posición (x,y) de la imagen. Gamma es una función de las intensidades I_1 , I_2 , I_3 de las imágenes desfasadas y en fase. Este parámetro tiene la propiedad de no poseer ventana.

Finalmente, se propone continuar con el estudio realizado sobre el desenfoque aplicado a los Diagramas de Proyección. [9.7]. Estos diagramas pueden ser usados para la interpretación de los experimentos de proyección de franjas, de manera similar a como es usado el Holodiagrama en holografía y en otras ramas de la óptica. Este diagrama es obtenido como un patrón de Moiré entre dos spoke target que simulan proyecciones centrales como aquellas que se forman en un sistema de proyección y observación o en dos sistemas de proyección. Usando este diagrama, son realizadas varias características interesantes relacionadas con la proyección de franjas. En el trabajo citado se muestra como se modifica un diagrama de proyección cuando es afectado por el desenfoque en ambos sistemas de proyección. La distribución de intensidad en un sistema desenfocado tratada en el capítulo 5 se incluye en el diagrama de proyección.

Bibliografía

- [9.1] Bela Julesz. Foundations of Cyclopean Perception. The University of Chicago Press. Chicago and London (1971).
- [9.2] Se pueden ver las distintas ilusiones ópticas, conteniendo además una explicación desde el punto de vista fisiológico en <http://www.illusionworks.com/html>.
- [9.3] S. Tolansky. Optical illusions. Pergamon Press. London. (1964).
- [9.4] F. Ratliff. Countour and Contrast. Scientific American., 91-101, Junio de 1972.
- [9.5] P. Huang, Qingying Hu, Feng Jin, Fu-Pen Chiang, Color-encoded digital fringe projection technique for high-speed three-dimensional surface countouring, Opt.

Eng., **38**, p.1065-1065, (1999).

[9.6] P. Torroba, N. Cap H., Rabal “Defocus correction using structure illumination”,
Optik, **108**, 68-77, (1998).

[9.7] M. Muramatsu, P. Torroba, N. Cap, and H. Rabal. The Projection diagram. Appl.
Opt., **vol 42**, N 23, 4765-4771, (2003).

Lista de Publicaciones

Algunos de los aspectos originales de este trabajo forman parte de las siguientes publicaciones:

P. Torroba, N. Cap, H. Rabal, W. Furlan. "Fractional order mean in image processing", *Opt. Eng.*, **33**, N° 2, 528-534, (1994).

P. Torroba, N. Cap, Rabal H. "Defocus correction using visibility criterion", *Journal of Modern Optics*, **41**, 111-117, (1994).

P. Torroba, N. Cap H. Rabal "Defocus correction using structure illumination", *Optik*, **108**, 68-77, (1998).

P. Torroba, N. Cap, H. Rabal. "Correction of defocusing using a modified visual model", *Optik*, **110**, N° 7, 341-349, (1999).

P. Torroba, H. Rabal and R. Arizaga. "Correction of defocusing using the co-occurrence matrix and the structure function", *J Mod Optics*, vol 47, 69-80, (2000).

M. Muramatsu, P. Torroba, N. Cap, and H. Rabal. The Projection diagram. *Appl. Opt.*, **42**, N 23, 4765-4771, (2003).