

## DETECCIÓN DE CARACTERÍSTICAS EN IMÁGENES DE TELEDETECCIÓN MEDIANTE EL MODELO DE CONGRUENCIA DE FASE

Xose R. Fdez-Vidal\*, C. García-Cuntín\*\*, C. Hernández\*\*

\*Departamento de Física Aplicada, Facultad de Físicas, Santiago de Compostela, España

\*\*Departamento de Electrónica y Computación, Facultad de Físicas, Santiago de Compostela, España

**Palabras Clave:** Características, Congruencia de Fase, Filtros en Cuadratura, Energía Local.

**Abstract.** En este trabajo se presenta una evaluación experimental entre dos enfoques distintos de detección de características en imágenes de teledetección. El enfoque clásico representado por el operador de Marr&Hildreth (operador diferencial) y un enfoque frecuencial, el modelo de congruencia de fase. Para obtener los valores de congruencia de fase en cada punto de una imagen, se ha diseñado un algoritmo basado en un banco de filtros de Gabor. Los resultados finales obtenidos por este algoritmo demuestran que la detección de características mediante congruencia de fase son más acordes con las características reales que percibe un humano. Como conclusión final, cabe destacar que el modelo frecuencial es una potente herramienta de obtención de características en cualquier tipo de imagen.

### I. Introducción

En los procesos iniciales de la visión, parece inevitable buscar algún tipo de estructura en la función luminancia que uno cree que tiene conexión con el objetivo específico de la tarea de inspección que estemos desarrollando en ese momento. Evidentemente, un proceso de este estilo no puede ser entendido sin asumir ningún tipo de conocimiento *a priori* o suposición acerca de los modelos de las características que deseamos detectar.

Existe una gran cantidad de factores que influyen en la función luminancia. Sin embargo, la mayoría de los autores asumen ciertas simplificaciones que afectan, tanto a las superficies, como a las fuentes de iluminación involucradas en la escena visual que deseamos analizar. Bajo estas suposiciones, proponen distintos métodos para detectar los patrones deseados en la imagen. El modelo más usual es aquel que identifica una característica con un cambio brusco en la función luminancia. Sobre esta base se han desarrollado gran cantidad de operadores, llamados operadores gradiente [1], [2].

A pesar de la gran cantidad de operadores diferenciales existentes, cabe destacar fundamentalmente dos: el operador de Marr&Hildreth [3], [4] y el de Canny [5].

El detector de Marr&Hildreth [3] combina el operador Laplaciana con un suavizado mediante una Gaussiana en una única operación (LoG). No obstante, este operador es similar al de Diferencias de Gaussianas (DoG), propuesto por Wilson & Giese [12], cuando la relación entre las sigmas es aproximadamente 1.6. El operador de

Marr&Hildreth alcanzó gran popularidad por varias razones: (i) buenos resultados alcanzados en todo tipo de aplicaciones; (ii) motivaciones biológicas: la forma de este operador coincide bastante bien con los campos receptivos de las células neuronales ganglionares; (iii) es rotacionalmente simétrico (no selectivo a orientaciones); (iv) facilidad en la detección de las características (cruces por cero).

Tratando el problema de detección de características como un problema de optimización, Canny [5] propuso un conjunto de operadores que optimizan tres criterios. Primero, el detector debe tener baja probabilidad de error. Segundo, que los puntos marcados deben estar lo más próximo posible al centro del borde real. Y tercero, deben tener baja probabilidad de mostrar múltiples respuestas ante un borde simple. Esta aproximación supone que los modelos de las características son funciones escalón o rampa de la función luminancia.

Todos los métodos de detección de características basados en gradiente presentan serios problemas. En primer lugar, presuponen un modelo de la característica a detectar. Sin embargo, los humanos son capaces de percibir más patrones de interés que los modelos simples supuestos por estas aproximaciones. En segundo lugar, el gradiente de una imagen depende de varios factores como iluminación, suavizado, ampliaciones, etc. que son desconocidos de antemano. Esto hará dependiente la respuesta del operador al nivel de gris de cada característica. Como consecuencia inmediata, el proceso de umbralización será extremadamente dependiente de cada caso concreto imposibilitando así una estrategia más genérica.

A pesar de los problemas anteriormente mencionados, los métodos basados en gradiente siguen siendo los más utilizados en todo tipo de tareas de procesado de imagen, incluido el campo de la teledetección.

Morrone et al. [6],[7],[8] han propuesto un modelo de detección de características que postula que la estructuras de interés en una imagen se perciben en aquellos puntos en la fase de las componentes de Fourier están máximamente ordenadas. Esta idea tiene sus orígenes en los experimentos llevados a cabo por Oppenheim&Lim [9]. Estos autores demuestran la importancia que tiene la fase en la percepción de una imagen. Morrone et al. [9] demuestran que este modelo explica con éxito un número de efectos psicofísicos en la percepción visual humana.

En este trabajo, presentamos una evaluación comparativa del modelo de congruencia de fase obtenido por nuestro algoritmo frente a los operadores clásicos, representados por el modelo Marr&Hildreth [3]. En los experimentos diseñados se trata de demostrar que el detector de características basado en la congruencia de fase supera claramente al modelos clásicos en la detección de características en imágenes de teledetección. Para la implementación, se ha utilizado una técnica de filtrado multicanal, en donde cada canal se caracteriza por una función Gabor Compleja [11]. En el diseño del banco de filtros de Gabor hemos utilizado cuatro bandas de orientación (ancho de banda angular cte. de 45°) y cuatro escalas radiales (ancho de banda cte. e igual a 1 octava) para recubrir totalmente el espectro. Cada Gabor compleja constituye un par de filtros en cuadratura. La energía local, para cada canal, se obtiene mediante el módulo de la respuesta de cada filtro complejo al convolucionarlo con la imagen. Owens [10] demuestra que esta magnitud está directamente relacionada con la congruencia de fase.

Este artículo se estructura de la siguiente forma. En la sección II abordamos la parte teórica de los modelos computacionales de detección de características empleados. El diseño de los experimentos, así como los resultados experimentales son expuestos en la sección III. Por último, las conclusiones se presentan en la sección IV.

## II. Modelos Computacionales para Detección de Características

El sistema visual humano aporta diferentes capacidades de procesado en los distintos puntos del campo visual. Esta inhomogeneidad constituye un balance entre recursos de procesado y tiempo empleado en la interpretación. Por otro lado, la atención puede desplazarse a diferentes puntos del campo visual que atraigan nuestra atención mediante movimientos oculares. Estos puntos constituyen características visuales que somos capaces de percibir. Luego, los modelos computacionales deben ser capaces de detectar la mayor parte de estos eventos. El objetivo de esta sección es describir los modelos de características empleados. El primero de ellos, está basado en los extremos de la derivada espacial de la imagen suavizada y el segundo es el modelo de congruencia de fase (PC).

### II.1. Características desde Cruces por ceros de la Laplaciana (LoG)

Marr [3],[4] aporta la evidencia que sugiere que las operaciones retinales pueden ser descritas analíticamente como la convolución de la imagen 2-D con el operador Laplaciano de la Gaussiana (LoG),  $\nabla^2 G_\sigma$ , de varios tamaños  $\sigma$ . En este modelo de representación temprana de la visión, las características más significativas son los cruces por cero del operador LoG multiescala, esto es, en los puntos (x,y) en donde se cumple:

$$\nabla^2 G_\sigma(x, y) * I(x, y) = 0 \quad (1)$$

donde  $G_\sigma$  es una Gaussiana 2D que suprime la influencia de los pixel situados a una distancia superior a  $3\sigma$  y  $\nabla^2$  es la Laplaciana, un operador derivada de segundo orden isotrópico. Los cruces por cero son utilizados para detectar y localizar las características más significativas de la imagen.

### II.2. Características desde el Modelo de Congruencia de Fase (PC)

Morrone et al. [6],[7],[8] proponen un modelo de detección de características basados en la Energía local. Este modelo postula que las características se perciben en puntos de la imagen en donde las componentes de Fourier están máximamente ordenadas. Bajo este punto de vista, es interesante resaltar que las bandas de Mach se hacen visibles en aquellos puntos en donde hay mayor acuerdo en la fase.

La función de Congruencia de Fase, PC(x) es definida de forma precisa por Owens [10]. En cada punto x de una señal la función PC(x) es:

$$PC(x) = \max_{\theta \in (0, 2\pi)} \frac{\int a_\omega \cos(\omega x + \phi_\omega - \theta) d\omega}{\int a_\omega d\omega} \quad (2)$$

Aunque la ec. (2) expresa de forma precisa la congruencia de fase, es una función difícil de implementar. Mediante simples manipulaciones trigonométricas Owens [10] demuestra la relación existente entre PC(x) y la Energía local es:

$$E(x) = \int a_\omega d\omega \cdot PC(x) \quad (3)$$

La magnitud E(x) es obtenida como el módulo de la salida de dos filtros en cuadratura de fase. Es ampliamente aceptada que el sistema visual humano utiliza la energía local obtenida de alguna forma por las células simples y complejas del córtex visual. Daugman [13] demostró que los campos receptivos de estas células se ajustan con filtros Gabor complejos.

Para la implementación de E(x) hemos diseñado un banco de filtros Gabor complejos. Estos han sido distribuidos con anchos de banda radiales y angulares constantes (1 octava y 45° grados, respectivamente). Esta partición está inspirada en que el ser humano posee cuatro canales de orientación privilegiadas (aproximadamente 45°). La respuesta espectral de todos los filtros (distintas escalas y orientaciones) se muestra en Fig.1.

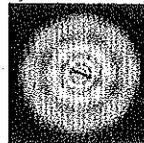


Fig. 1: Conjunto de filtros en el dominio frecuencial (256x256).

Finalmente, la  $E(x)$  total en cada punto de la imagen se obtiene como la suma de la energía local de cada filtro sintonizado a una determinada escala y orientación (ec. 4).

$$E(x) = \sum_i \sum_j E_{ij}(x) \quad (4)$$

### III. Resultados Experimentales

La imagen elegida para nuestros experimentos se muestra en la Fig 2 Su tamaño es de 256x256 pixels y 256 niveles de gris, es un corte selectivo de una imagen LANDSAT, correspondientes a un aeropuerto.

Las imágenes de teledetección pueden tener contraste diferentes en diferentes zonas de la imagen. Esto producirá que los detectores de características respondan de diferente forma en cada una de estas zonas. Esta idea nos sugirió el diseño del experimento para evaluar la robustez del modelo de congruencia de fase frente al modelo diferencial (LoG). Consideremos de nuevo la imagen original de la Fig. 2. Para producir diferentes versiones distorsionadas de ésta, las fases de todos los armónicos en la imagen original son aleatorizados a diferentes grados de acuerdo a la siguiente regla:

$$fase \leftarrow fase + f \times \pi \times random(-1,1) \quad (5)$$

donde  $random(-1,1)$  es un número aleatorio en el intervalo (-1,1) y  $f \in \{0.5, 0.67, 0.83\}$ , respectivamente.

Un diagrama esquemático se muestra en la Fig. 2. La impresión de las imágenes resultantes se muestra en la misma figura. A pesar de los fuertes grados de aleatorización introducidos, debemos notar que las imágenes todavía pueden ser reconocidas puesto que la estructura de las líneas de costa y aeropuerto no han sido totalmente destruidas; excepto para  $f=0.83$  donde las fases de los armónicos ha sido totalmente aleatorizados. En este caso, no somos capaces de percibir ninguna estructura relevante en la imagen. Para estas tres imágenes degradadas, se ha hecho una encuesta a diez sujetos. Como era de preveer, se concentraron en las características marcadas (línea de costa y aeropuerto) para reconocer la imagen en cuestión. En la más aleatorizada, ningún de ellos consiguió percibir nada.

Cuando las imágenes aleatorizadas con distintos valores de  $f$ , mostradas en la Fig. 2, se convolucionaron con operadores Laplaciana de la Gaussiana (LoG) de distintas escalas  $\sigma$ , aparecen un gran número de cruces por cero (Fig. 3) y no todos se corresponden con características visibles en la imagen. Solamente se puede apreciar contornos pertenecientes al aeropuerto y línea de costa para escalas bajas del operador ( $\sigma=1$ ). No obstante, son bordes discontinuos y sólo la capacidad integradora del ojo a una escala superior nos hace percibir dichas estructuras. A medida que se aumenta la aleatorización, el problema se agudiza y los cruces por cero dejan ya de tener relación con las características percibidas en la imagen. Por el contrario,

cuando las características visuales de las imágenes degradadas son encontradas en aquellos puntos en donde hay un gran acuerdo en las fases de las componentes individuales de Fourier, obtenemos un conjunto altamente selectivo de contornos consistentes con las líneas de costa y aeropuerto todavía perceptibles en la imagen. De hecho, estas coinciden con los únicos patrones visuales que los observadores interrogados eran capaces de apreciar. Las imágenes de puntos de máxima congruencia de fase (MPC) se muestran en la Fig. 3. Estos son obtenidos mediante una supresión no máxima. La operación empleada ha sido una histerisis con umbral inferior  $U_i=0.3$  y superior  $U_r=0.5$ , respectivamente.

Si observamos con detenimiento las imágenes producidas por los modelos empleados (Fig. 3), se pone de manifiesto otra diferencia sustancial. Ésta es relativa a la calidad de las características detectadas. El modelo de congruencia de fase produce contornos únicos y máximamente conectados. A medida que las imágenes se van degradando sólo detecta estructuras muy marcadas en la imagen. Si se persiste en el proceso de degradación el algoritmo no detecta ninguna estructura. En esto, coincide con el comportamiento de un observador humano. Por el contrario, el método basado en gradiente produce contornos altamente desconectados entre sí y dobles. Por otro lado, con independencia de la degradación de la imagen obtiene estructuras inexistentes en la imagen para un humano, comportamiento claramente indeseable.

### IV. Conclusiones

La congruencia de Fase es una medida adimensional y normalizada de las características presentes en una imagen. Ello, implica dos ventajas muy deseables para el proceso de umbralización necesario en todo proceso de detección de contornos. La primera, es una medida independiente del nivel de gris que posea el patrón en cuestión, esto es, el grado de orden en la fase es independiente del contraste. Segundo, al estar normalizada es posible obtener valores para los umbrales por anticipado de forma que sean válidos para un gran abanico de imágenes de distinta naturaleza.

En este trabajo, se demuestra que el detector de características implementado con filtros Gabor (4 escalas y 4 orientaciones) para obtener los valores de congruencia de fase obtiene resultados muy superiores al operador de Marr & Hildreth basado en gradientes. Es por tanto una herramienta eficaz para la detección de características en cualquier imagen. Aunque en este trabajo se han utilizado filtros Gabor, es posible utilizar cualquier tipo de filtros en cuadratura. Posiblemente se puedan alcanzar mejoras en los resultados, puesto que los filtros Gabor obtienen resultados satisfactorios en un amplio abanico de tareas pero son subóptimos en todos los casos. Es posible, por tanto, encontrar filtros más específicos para esta tarea que mejoren los resultados obtenidos por los filtros de Gabor.

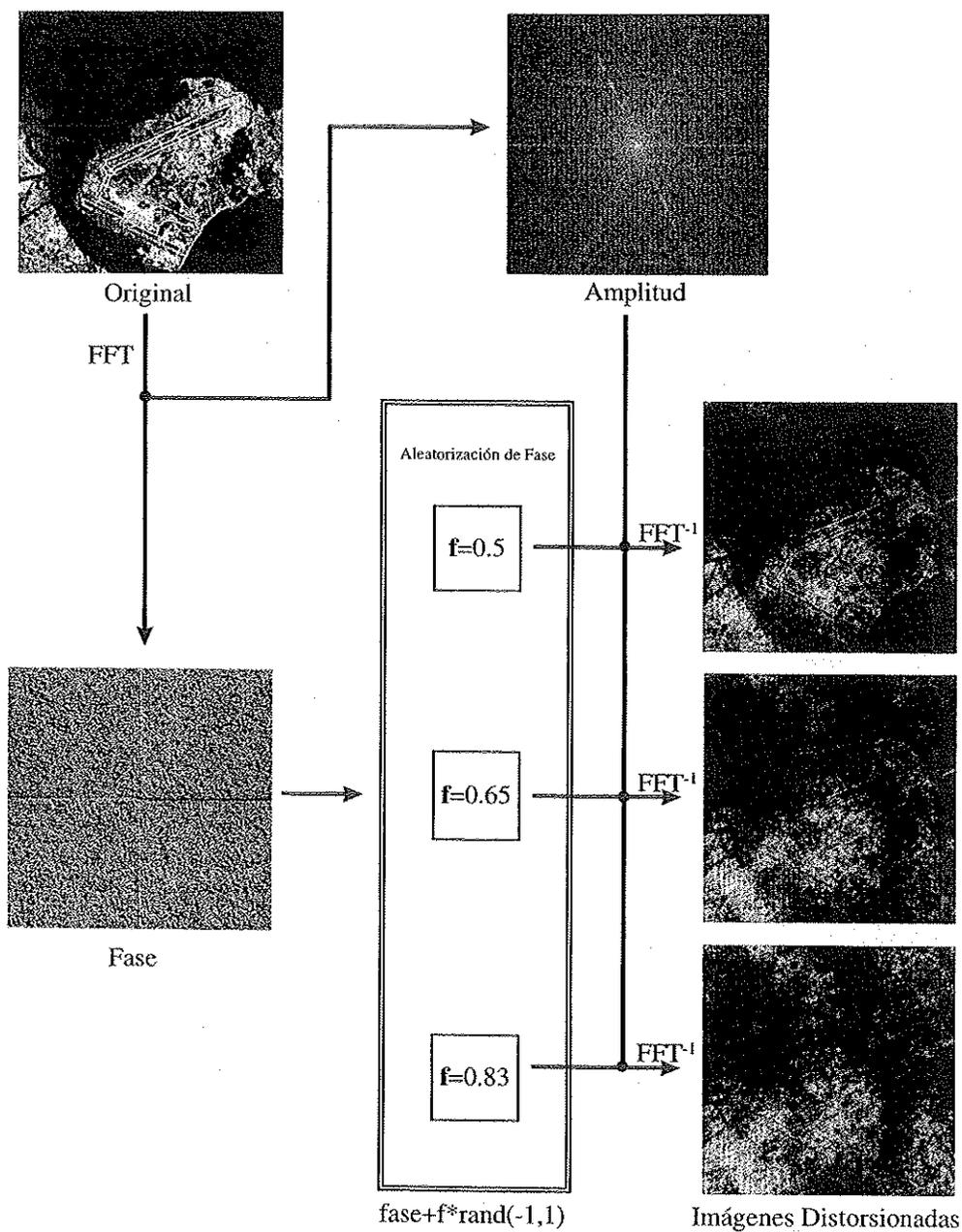


Fig. 2: Para producir tres versiones de diferentes de imágenes degradadas de la original, las fases de todos los armónicos en la original deben ser aleatorizados de acorde con la regla (5). La imagen menos degradada está situada arriba ( $f=0.5$ ) y la más degradada abajo ( $f=0.83$ ).

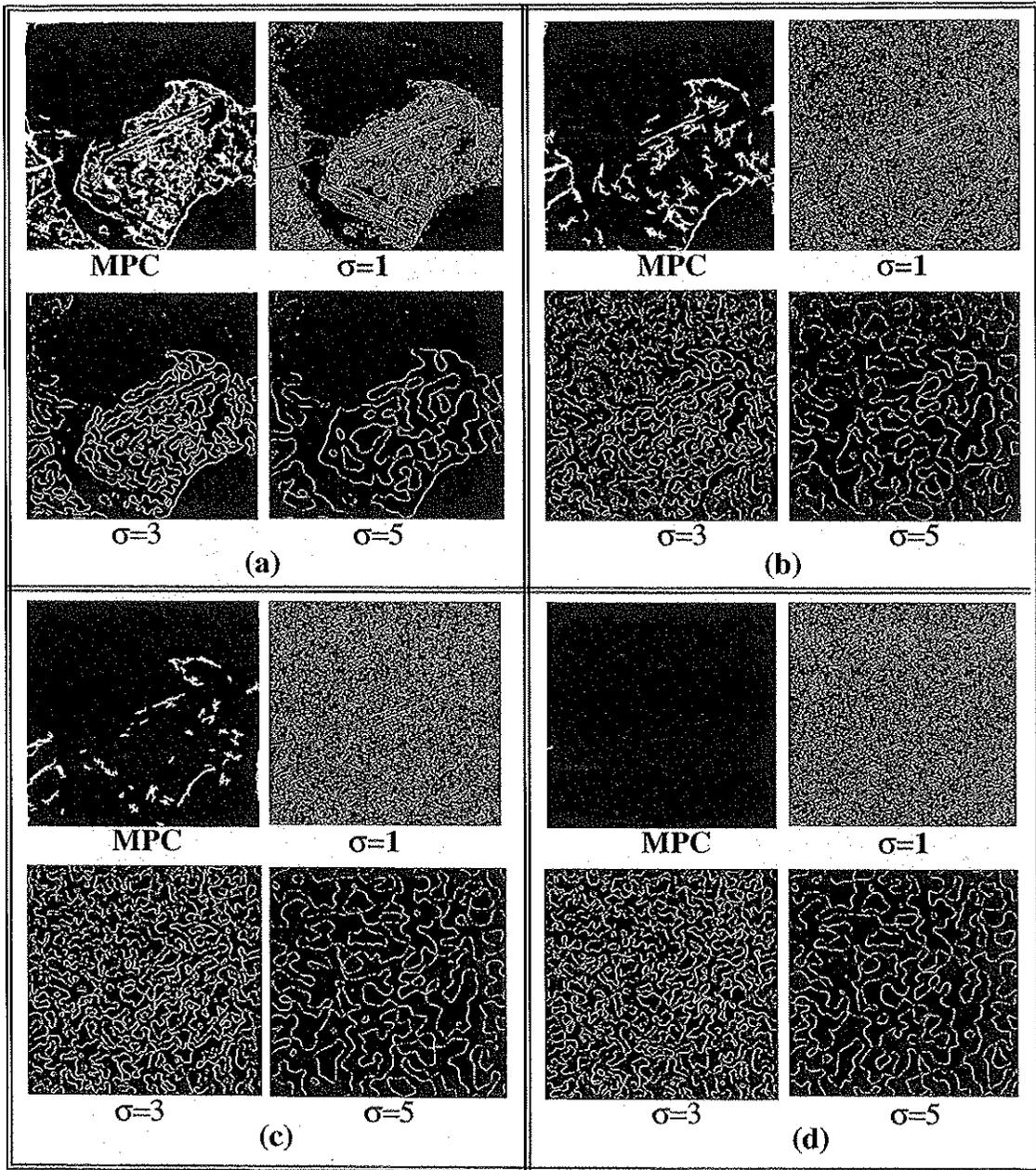


Fig.3: Características obtenidas por los modelos de Congruencia de fase (MPC) y Marr&Hildreth con diferentes  $\sigma$  para las imágenes de entrada : (a) Original, (b) Imagen degradada con  $f=0.5$ , (c) Imagen degradada con  $f=0.67$ , (d) Imagen degradada con  $f=0.83$ .

## V. Bibliografía

- [1] Roberts, L. G. Machine perception of three dimensional solids. Optical and Electro Optical Information Processing. Tipett J. Et al. Eds, 159-196, MIT Press Cambridge, M. A. 1965.
- [2] Prewitt, J.M.S. Object enhancement and extraction, Picture Processing and Psychopictorics. 75-149. Lipkin and Rosenfeld Eds. New York: Academic Press, 1970.
- [3] Marr, D. Vision. 41-43, W.H. Freeman & Co, San Francisco, 1982.
- [4] Marr,D. Hildreth, E. Theory of edge detection. Proc. R. Soc. London. B207, 187-217, 1980.
- [5] Canny, J.F. A computational approach to edge detection. IEEE Trans. On PAMI. Vol. 8, No. 6, 679-697, 1986.
- [6] Morrone, C. Owens, R.A. Features detection from local energy. Pattern Recognition Letters. 6, 303-313, 1987.
- [7] Morrone, M.C. Ross, J., Burr, D.C., Owens, R. A phase-dependent energy model. Proc. R. Soc. London. B 235, 221-245, 1988.
- [8] Morrone, M.C., Ross,J.,Burr, D.C., Owens, R. Mach bands are phase dependent. Nature. Vol. 324, No 6094, 250-253, 1986.
- [9] Openheim,A.L., Lim, S.L. The importance of phase in signals. Proceedings of IEEE, Vol. 69, No 5, 1981.
- [10] Owens,R. Features-free images. Patt. Rec. Letters, 15, 35-44, 1994.
- [11] Jain, A.K., Farokhinia, F. Unsupervised texture segmentation using Gabor filters. Patt. Recognition, Vol 24, No 12, 1167-1186, 1991.
- [12] Wilson,H. R., Giese, S.C., Threshold visibility of frequency gradient patterns. Vision Research, 7, 1177-1190, 1977.
- [13] Daugman, J.G. Two-dimensional spectral analysis of cortical receptive field profiles. Vision Research. Vol. 20, 847-856, 1990.