

01170
2013

COMPRESION DE IMAGENES RADIOLOGICAS UTILIZANDO LA TRANSFORMADA
COSENO DISCRETA (DCT).

Créditos asignados a la tesis DIEZ (10)

APROBADO POR EL JURADO:

Presidente:	<u>DR. FEDERICO KUHLMANN RODRIGUEZ</u> <i>F Kuhlmann</i>
Vocal:	<u>DR. FRANCISCO GARCIA UGALDE</u> <i>[Signature]</i>
Secretario:	<u>DR. JORGE LIRA CHAVEZ</u> <i>[Signature]</i>
Suplente:	<u>DR. ROGELIO ALCANTARA SILVA</u> <i>R Alcantara</i>
Suplente:	<u>DR. ROMEO ORTEGA MARTINEZ.</u> <i>[Signature]</i>

TESIS CON
FALLA DE ORIGEN



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

CONTENIDO

CAPITULO I.-	Introducción.....	1
CAPITULO II.-	Compresión de Imágenes por Transformada.	
II.1.-	Introducción.....	7
II.2.-	Estructura General de la Codificación por Transformada.....	8
II.2.1.-	Cálculo de la Transformada.....	8
II.2.2.-	Cuantización.....	11
II.3.-	La Transformada Optima para Compresión de Imágenes.....	13
II.4.-	La Transformada Coseno Discreta.....	14
II.5.-	Esquema de Compresión Utilizado.....	17
II.5.1.-	Características de las Imágenes Fuente Empleadas.....	17
II.5.2.-	Descripción del Esquema de Codificación....	17
II.5.3.-	Esquema de Decodificación.....	20
CAPITULO III.-	Resultados Experimentales.	
III.1.-	Criterio de Distorsión.....	22
III.2.-	Imágenes y Gráficas de Error.....	23
III.3.-	Observaciones Generales.....	26
CAPITULO IV.-	Conclusiones.....	48
BIBLIOGRAFIA.....		51

CAPITULO I

INTRODUCCION

El rápido desarrollo de nuevas tecnologías en las áreas de electrónica y comunicaciones ha ido cambiando gradualmente los métodos de adquisición, de almacenamiento, de despliegue y de transmisión de imágenes radiológicas, empleadas como medio auxiliar de diagnóstico. Son cada vez más las técnicas digitales empleadas en los equipos médicos, y esto ha permitido la introducción de la radiología digital como algo común en la práctica diaria de la radiología general.

Junto con las ventajas de contar con imágenes radiológicas digitales, como es el poder emplear las técnicas de procesamiento digital de señales en el apoyo al diagnóstico, ha surgido la necesidad de contar con métodos eficientes para el almacenamiento y la transmisión de éste tipo de imágenes, ya que, en general, cuando se trabaja con imágenes digitalizadas, se requiere de una gran cantidad de memoria para su almacenamiento y demasiado tiempo para su transmisión. Por ejemplo, una tomografía computarizada, con una resolución de 512 x 512 "píxeles", cuantizada con 8 bits por píxel, constituye más de dos millones de bits. Existe, por lo tanto, un creciente interés en comprimir este tipo de imágenes, más aun cuando en un Hospital su número va en

aumento y se archivan, a veces, durante varios años.

Los métodos empleados para efectuar la compresión de datos, o codificación de fuente, pueden ser técnicas encaminadas a la reducción de redundancia o a la reducción de entropía, aunque muchos métodos producen simultáneamente ambos tipos de reducciones.

Puesto que la entropía se define como la información promedio, la reducción de entropía produce una reducción de información. Esta información perdida no es posible recuperarla y por tanto, el empleo de la operación de reducción de entropía es irreversible.

A éste tipo de compresión se le llama también codificación reductora de fidelidad, o simplemente, compresión irreversible. Dentro de esta categoría encontramos por ejemplo, los métodos de codificación predictiva como ADPCM (Adaptive Differential Pulse Code Modulation) y LPC (Linear Predictive Coding), y los métodos de codificación basados en las transformadas Karhunen-Loève, Fourier, coseno, Hadamard y Haar.

Por otro lado, la reducción de redundancia elimina o al menos, reduce la redundancia que existe en los datos, de tal manera que pueda reinsertarse posteriormente. Se trata entonces de un proceso reversible, es decir, de una compre-

sión sin error. Como ejemplo de este tipo de codificación se tienen los códigos de Huffman y los "Run Length".

En este trabajo se analiza el comportamiento de la transformada discreta coseno cuando se emplea como herramienta para la compresión de imágenes de tipo radiológico: imágenes digitales como radiografías, tomografías computarizadas, imágenes de resonancia magnética y angiografías. Todas las imágenes empleadas aquí son en blanco y negro, es decir, mono-espectrales. Por tanto, solamente se trabaja con los niveles de intensidad de las mismas.

En general una imagen natural puede ser representada mediante una función continua en la cual el nivel de intensidad o los diferentes tonos de gris que presenta, son una función de la posición.

Antes de procesar una imagen en una computadora, es necesario digitalizarla, es decir, muestrearla de manera que quede representada por un conjunto finito de números reales, llamados muestras o píxeles. Posteriormente, cuantizar las muestras dividiendo la escala de tonos de gris en intervalos y asignar a cada muestra uno de estos intervalos, el cual representará su intensidad. El muestreo y la cuantización son en su forma más simple uniformes: las muestras se toman en incrementos constantes, y los intervalos que forman la

escala de tonos de gris de cada muestra son todos del mismo tamaño. De esta manera, podemos considerar una imagen digitalizada como un arreglo o una matriz, en donde cada uno de sus elementos nos representa la intensidad de radiación de campo instantáneo de vista de la imagen.

La calidad de una imagen digitalizada depende del número de niveles empleados en la cuantización de cada muestra y del número de muestras tomadas en la dirección horizontal y en la dirección vertical del arreglo bidimensional. Entre más grandes sean estos parámetros, se tendrá una mejor reproducción de la imagen original cuando se despliega en un monitor en forma digital. También son importantes en la calidad de una imagen los factores de contraste, el nivel de ruido y la agudeza de los bordes y líneas.

Facilmente se puede observar que la mayoría de las imágenes digitalizadas contienen gran cantidad de redundancia, es decir, los valores de los tonos de gris en algunas zonas de las imágenes están correlacionados, o dicho de otra manera, muchos valores de tonos de gris se pueden predecir a partir de los valores de sus puntos vecinos.

La idea fundamental de la compresión de imágenes es entonces tratar de representar tales imágenes mediante un conjunto de datos no correlacionados entre sí. De manera que cada uno

de estos datos nos proporcione información sobre alguna propiedad única de la imagen, única en el sentido de que esa propiedad no se puede predecir a partir de los otros datos.

Como se mencionó anteriormente, la compresión por reducción de entropía, introduce una cierta distorsión al realizar la descompresión o decodificación de la imagen. Sin embargo, mediante este tipo de codificación se pueden alcanzar tasas de compresión mayores a las que se obtienen con los métodos que no introducen error. Tal es el caso de la compresión por transformada.

Los inconvenientes de codificar imágenes empleando métodos por transformada han sido principalmente, la gran cantidad de cálculos necesarios. Sin embargo, recientemente se han introducido algoritmos rápidos que permiten efectuar tales cálculos más eficientemente. Además, es posible contar ahora con "hardware" especializado que hace factible realizar estas transformadas en un tiempo reducido. De esta manera las tasas de compresión que se obtienen con estos métodos, se muestran accesibles en la práctica.

Al seleccionar una técnica de codificación de fuente para imágenes, hay que tomar en cuenta tres parámetros que la caracterizarán y nos ayudarán a evaluar su desempeño y su utilidad práctica en una aplicación determinada. En primer

lugar el grado o tasa de compresión que se obtiene con la codificación. Esta tasa (o índice de compresión) se define como el cociente del volumen ocupado por la imagen original entre el volumen obtenido después de codificarla. En segundo lugar, el error o la distorsión que se introduce al decodificar la imagen y recuperar la original. Por último, la complejidad computacional del método empleado, que se traduce en tiempo de procesamiento y en costo del equipo empleado para la codificación.

CAPITULO II

COMPRESION DE IMAGENES POR TRANSFORMADA.

II.1.- Introducción.

En este capítulo se describe brevemente la manera de emplear la transformada coseno en la codificación de imágenes. Se muestra cómo, al obtener la transformada de una imagen digital, se trabaja con los coeficientes de esta transformada de manera que aquéllos que son más significativos, es decir que su contribución es mayor al contenido de información y a la calidad subjetiva de la imagen, se cuantizan con más precisión, mientras que aquéllos que son poco importantes desde el punto de vista de la escala de gris y de la resolución espacial son despreciados, lográndose de esta manera una reducción en el número de datos.

Como se mencionó anteriormente, se busca representar la imagen con el mínimo número de bits posible, dada una distorsión máxima permitida. Para esto, por medio de la transformada, buscamos representar las imágenes mediante un conjunto de datos no correlacionados [5,14,16]. Para tipos generales, o clases generales de imágenes, puede resultar que el número de bits necesarios para codificar estos datos ya no correlacionados sea demasiado grande, en consecuencia se

deben ordenar de acuerdo a su importancia, y codificar solo aquellos que contribuyen en mayor grado a la calidad de la imagen. También cabe hacer notar que se deben de tomar en cuenta las características del receptor durante la codificación, pues si éste no es capaz de emplear la resolución o la calidad de las imágenes transmitidas, los coeficientes o los parámetros que nos proporcionan esta calidad pueden ser despreciados. Así por ejemplo, si tenemos una imagen en la cual se tienen 256x256 pixeles, con 12 bits por pixel, pero el receptor solo puede desplegar 256 diferentes tonos de gris, es decir, trabaja únicamente con 8 bits por pixel, no es necesario enviarle los 12 bits por muestra sino solamente 8, teniéndose entonces un ahorro de un tercio en el volumen de los datos.

II.2.- Estructura general de la codificación por transformada.

La codificación de imágenes digitales basada en transformadas consta de dos etapas principales: la obtención de la transformada de la imagen y la cuantización de los coeficientes de la transformada.

II.2.1.- Cálculo de la transformada.

Para efectuar la transformada de una imagen digital generalmente se divide ésta en pequeños bloques de $N \times N$ como se indica en la figura 1 y posteriormente se obtiene la trans-

formada bidimensional de tamaño N de cada uno de los bloques. Esta operación de subdivisión se hace para reducir los cálculos, ya que el número de operaciones se incrementa cuadráticamente con el tamaño de la transformada. El tamaño de los bloques es generalmente una potencia de dos y no hay traslape entre éstos.

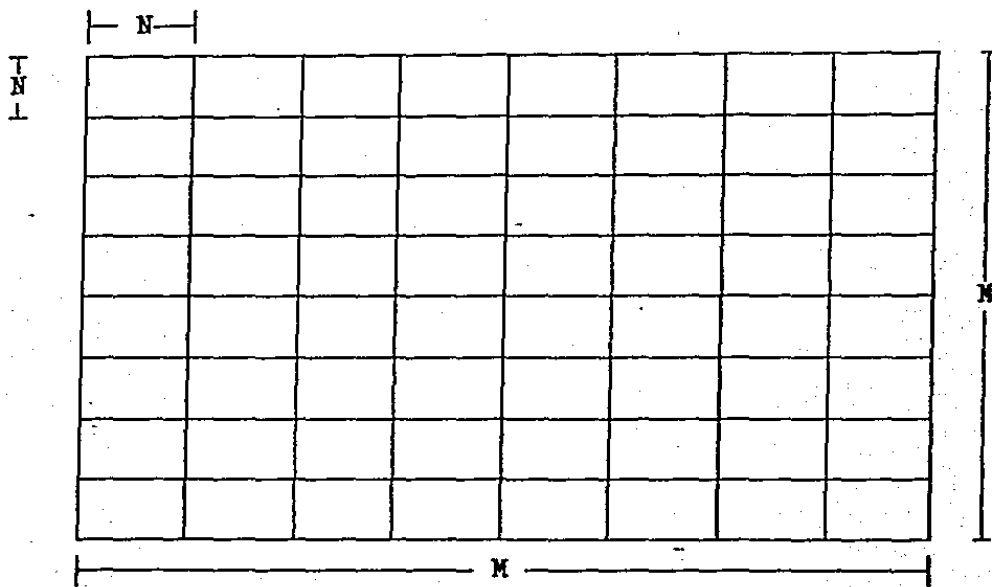


Fig. 1 División de una imagen con una resolución de $M \times M$ "píxeles" en bloques de $N \times N$ "píxeles".

En este trabajo los coeficientes de la transformada bidimensional de un bloque j (alguno de los $(M/N)^2$ bloques que han sido formados con la división mencionada

anteriormente) se representan por $F_j(u,v)$ y forman un arreglo matricial como se indica en la figura 2.

Las componentes de frecuencia que contiene un bloque de la imagen original están representadas por los coeficientes de su transformada, ya que ésta es en realidad una descomposición del bloque en frecuencias generalizadas.

$F_j(0,0)$	$F_j(0,1)$	$F_j(0,2)$. . .	$F_j(0,N-1)$
$F_j(1,0)$	$F_j(1,1)$	$F_j(1,2)$. . .	$F_j(1,N-1)$
.
.
.
$F_j(N-1,0)$	$F_j(N-1,1)$	$F_j(N-1,2)$. . .	$F_j(N-1,N-1)$

Fig. 2 Coeficientes de la transformada del bloque j .

Los coeficientes $F_j(u,v)$ con valores pequeños de u y de v representan los componentes de frecuencias bajas en la imagen, mientras que los efectos de alta resolución espacial, contenidos en las frecuencias altas, están representados por los coeficientes $F_j(u,v)$ con valores grandes de u y de v . Generalmente la mayor parte de la energía de la imagen está

representada en los coeficientes $F_j(u,v)$, que caracterizan los componentes de frecuencias bajas [16].

Al obtener la transformada, queda entonces la imagen representada en el dominio de la frecuencia mediante el conjunto de coeficientes:

$$F_j(u,v) \quad \text{con } u,v = 0,1,2,\dots,N-1 \\ \text{para } j = 1,2,\dots,(M/N)^2$$

donde j nos indica el número de bloque.

Se denota por $\sigma_{u,v}^2$ a la varianza del conjunto de coeficientes que en cada uno de los bloques dan los mismos componentes de frecuencias, es decir, aquellos coeficientes $F_j(u,v)$, para $j = 1,2,\dots,(M/N)$, con los índices u y v iguales.

II.2.2.- Cuantización [14,16].

El procedimiento empleado comunmente para cuantizar los coeficientes de la transformada (y lograr una compresión de la imagen) es ordenar los coeficientes $F_j(u,v)$ de acuerdo al valor de su varianza y retener solamente un determinado número de aquellos cuya varianza sea grande, ya que ésta es una medida del contenido de información.

Debido a que las varianzas de los coeficientes varían mucho de uno a otro, no resultaría eficiente emplear el

mismo cuantizador para todos los coeficientes que se han retenido. Puesto que si se ajustan los niveles de salida del cuantizador para abarcar el rango del coeficiente con la varianza más grande, muchos de los coeficientes con varianzas pequeñas caerán en los rangos más pequeños y en consecuencia la mayoría de los niveles del cuantizador no serán empleados.

Por tal motivo, antes de cuantizar los coeficientes que se han retenido, se normalizan dividiéndolos entre su desviación estandar y de esta manera se emplea un solo cuantizador para todos los coeficientes.

Finalmente a cada uno de los niveles del cuantizador se le asigna un código binario de manera que cada coeficiente cuantizado sea representado por una secuencia de bits de una misma longitud.

Existen sin embargo algunas desventajas de utilizar un solo cuantizador para todos los coeficientes de la transformada y representarlos con palabras de código de longitud constante. Por una parte se introduce el mismo error medio cuadrático en todos los coeficientes normalizados y puesto que en el proceso de decodificación al desnormalizar tales coeficientes cuantizados se deben multiplicar por su desviación estandar, el error medio cuadrático de cada coeficiente se hace proporcional a su varianza. Es decir los

coeficientes con varianzas más grandes, los cuales son más importantes para la reconstrucción de la imagen, son los que resultan con mayor error de cuantización. Por otro lado, al emplear el mismo número de bits para cuantizar a todos los coeficientes que han sido retenidos, no se está tomando en cuenta la cantidad de información que contiene cada uno de ellos. Por lo tanto resulta más eficiente la codificación de los coeficientes de la transformada cuando se cuantiza cada uno de ellos con un número de bits proporcional a su contenido de información.

II.3.-La transformada óptima para compresión de imágenes.

La transformada Karhunen-Loève es la transformada óptima en el sentido de que nos proporciona datos no correlacionados y simultáneamente la máxima concentración de energía en promedio en un número reducido de coeficientes [5,14,16]. Esta transformada es computacionalmente difícil de calcular y por tal motivo, cuando se trabaja con secuencias aleatorias estacionarias, generalmente se emplean otras transformadas unitarias subóptimas que se le aproximan en su eficiencia de compactación de energía, como son la transformada coseno y la de Fourier. Son subóptimas para la compresión de imágenes en el sentido de que sus coeficientes están aún correlacionados

entre si, sin embargo, esta correlación tiende a disminuir conforme se incrementa el tamaño de la transformada.

II.4.- La transformada coseno discreta.

La codificación realizada en este trabajo esta basada en la transformada coseno discreta (DCT) la cual se define de la siguiente manera [2]:

Sea $f(m)$ con $m = 0, 1, 2, \dots, N-1$ (1)

una secuencia de tamaño N .

Su transformada, $F(u)$, es

$$F(u) = 2 / N C(u) \sum_{m=0}^{N-1} f(m) \cos \left[\frac{(2m+1) u \pi}{2N} \right]$$

para $u = 0, 1, 2, \dots, N-1$ (2)

donde $C(u) = \begin{cases} 1/\sqrt{2} & \text{para } u = 0 \\ 1 & \text{en caso contrario.} \end{cases}$ (3)

y la transformada inversa se define por

$$f(m) = \sum_{u=0}^{N-1} C(u) F(u) \cos \left[\frac{(2m+1) u \pi}{2N} \right]$$

para $m = 0, 1, 2, \dots, N-1$ (4)

Para el caso de dos dimensiones, la transformada coseno se define por:

$$F(u,v) = \frac{4}{N} C(u) C(v) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} .$$

$$f(m,n) \cos \left[\frac{(2m+1)u\pi}{2N} \right] \cos \left[\frac{(2n+1)v\pi}{2N} \right]$$

$$\text{con } u, v = 0, 1, 2, \dots, N-1 \quad (5)$$

donde las constantes $C(u)$ y $C(v)$ estan definidas como en (3)

La inversa de (5) es

$$f(m,n) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} C(u) C(v) F(u,v)$$

$$\cos \left[\frac{(2m+1)u\pi}{2N} \right] \cos \left[\frac{(2n+1)v\pi}{2N} \right]$$

$$\text{con } m, n = 0, 1 \dots N-1 \quad (6)$$

Ultimamente, en el procesamiento de imágenes digitales se ha dado especial importancia a ésta transformada y se han desarrollado algunos algoritmos rápidos para su cálculo [4, 6, 11, 17, 18].

Cuando una imagen se puede modelar como un proceso de Markov de bajo orden con grande correlación entre sus pixeles, es posible hacer una aproximación de la transformada Karhunen-Loève mediante la coseno en lo que se refiere a sus propiedades de compactación de energía. Por este motivo, la transformada coseno discreta ha resultado atractiva en la compresión de imágenes [9,16].

Además de su gran capacidad de compactación de energía, la transformada coseno introduce menos falsos contornos ("blocking artifact") en los subbloques que delimitan la transformada, comparados con aquellos que introduce la transformada de Fourier, principalmente cuando la tasa de compresión es grande.

Los coeficientes de la transformada coseno tienden a seguir una distribución Gaussiana, a excepción del coeficiente de d.c. $F_j(0,0)$, el cual es modelado de manera más apropiada por una función de densidad de Rayleigh. Tomando en cuenta que una variable aleatoria Gaussiana con una varianza igual a σ^2 , al ser cuantizada con una distorsión máxima D , necesita por lo menos $1/2 \log(\sigma^2/D)$ bits [3], se observa que si el número de bits asignado a un coeficiente $F_j(u,v)$ debe indicar que tanta es la contribución de dicho coeficiente a la representación total de la

información contenida en la imagen, entonces tal número de bits debe hacerse proporcional al logaritmo de la varianza del coeficiente.

II.5.- Esquema de compresión utilizado.

II.5.1.- Características de las imágenes fuente empleadas.

Las imágenes con que se cuenta son de $256 \times 256 \times 7$, es decir, están representadas por una matriz de 256 renglones y 256 columnas donde cada elemento de la matriz es un número entero en el rango de 0 a 127 el cual nos indica el nivel de intensidad de cada elemento de la imagen (pixel), esto es, se tienen 7 bits por pixel.

II.5.2.- Descripción del esquema de codificación:

a) Cálculo de la transformada.

Inicialmente se divide la imagen fuente en 256 bloques de 16×16 pixeles como se indicó en la sección II.2.1. (figura 1 con $M=256$ y $N=16$), a continuación, mediante (5) y (3), se calcula la transformada coseno bidimensional de cada uno de los bloques que han sido formados.

b) Normalización de la transformada.

Los coeficientes $F_j(u,v)$, para $u,v = 0,1,\dots,15$ y para $j = 1,2,\dots,256$ obtenidos en el inciso anterior, se normalizan con respecto a su desviación estandar de manera que resulten con varianza unitaria.

La varianza $\sigma_{u,v}^2$ de cada coeficiente se estima promediando los cuadrados de los elementos que representan una misma frecuencia a lo largo de los 256 bloques, es decir:

$$\sigma_{u,v}^2 = E \{ |F(u,v)|^2 \} = \frac{1}{N} \sum_{j=0}^{N-1} |F(u,v)|^2$$

$$\text{para } u,v = 0,1,\dots,15 \quad (7)$$

Los coeficientes normalizados $F_{nj}(u,v)$ son entonces:

$$F_{nj}(u,v) = \frac{F_j(u,v)}{\sigma_{u,v}}$$

$$\text{para } u,v = 0,1,\dots,15$$

$$\text{y } j = 1,2,\dots,256 \quad (8)$$

c) Cálculo de la tabla de asignación de bits.

Esta tabla, de 16 x 16 elementos, nos indica el número de bits con el que se van a cuantizar los coeficientes normalizados $F_{nj}(u,v)$. La tabla se calculó mediante la siguiente expresión:

$$B(u,v) = b_{\text{prom}} + \frac{2}{\ln(10)} \ln(\sigma_{u,v}^2) - \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \ln(\sigma_{u,v}^2)$$

$$\text{para } u,v = 0,1,\dots,N-1$$

$$\text{con } N = 16 \quad (9)$$

donde "bprom" es el número de bits promedio esperado para representar los coeficientes de un bloque, y $B(u,v)$ es el número de bits con los que se va a representar a los coeficientes:

$$F_j(u,v) \quad \text{para } j = 1, \dots, 256 \quad (10)$$

La ecuación (9) minimiza el error de cuantización total promedio en los bloques [16], cuando se emplean cuantizadores uniformes. Ha sido obtenida empíricamente por Winz y Kurtenbach [19]. Las diferentes tasas de compresión se obtienen haciendo variar el número de bits promedio "bprom".

Puesto que, en general, $B(u,v)$ es un número real, su valor se redondea al entero próximo y en caso de resultar negativo se le asigna el valor de cero, ya que no es posible tener un número de bits negativo.

d) Cuantización.

Los coeficientes de la transformada, $F_{nj}(u,v)$, se normalizan nuevamente antes de ser cuantizados, de manera que tomen valores entre -1 y 1.

Para esto, se obtienen los máximos $m(u,v)$:

$$m(u,v) = \max \{ |F_{nj}(u,v)| \}$$

$$\text{con } u, v = 0, 1, \dots, 15$$

$$\text{para } j = 1 \dots 256 \quad (11)$$

y los coeficientes normalizados $F'_{nj}(u,v)$, quedan definidos

por:

$$F'_{nj}(u,v) = \frac{F_{nj}(u,v)}{m(u,v)}$$
$$u,v = 0,1,\dots,15$$
$$j = 1,2,\dots,256 \quad (12)$$

Posteriormente, cada coeficiente $F'_{nj}(u,v)$ se representa por un número entero $S_j(u,v)$, mediante:

$$S_j(u,v) = [F'_{nj}(u,v)(2^{B(u,v)} - 1)] \quad (13)$$

donde $S_j(u,v)$ se redondea al entero próximo.

Finalmente, la representación binaria en complemento a dos es el valor codificado del coeficiente $F_j(u,v)$.

Cabe hacer notar que para reconstruir la imagen además de contar con los valores $S_j(u,v)$, es necesario conocer los máximos $m(u,v)$, las varianzas $\sigma_{u,v}^2$ y la tabla de asignación de bits $B(u,v)$.

II.5.3.-Esquema de decodificación.

La descompresión se efectúa realizando el proceso inverso. Primero, se calcula el valor reconstruido de los coeficientes normalizados $F'_{nj}(u,v)$ haciendo uso de la tabla de asignación de bits $B(u,v)$:

$$F'_{nj}(u,v) = \frac{S_j(u,v)}{2^{B(u,v)-1}} \quad (14)$$

posteriormente se desnormaliza con respecto a los máximos y con respecto a las desviaciones estandar mediante la siguiente expresión:

$$F_j(u,v) = \sigma_{u,v}^2 m(u,v) F'_{nj}(u,v)$$

para $u,v = 0,1,\dots,15$
y $j = 1,2,\dots,256$ (15)

una vez obtenidos los coeficientes reconstruidos $F_j(u,v)$, se calcula la transformada inversa de cada uno de los 256 bloques mediante (6) y (3) y, de esta manera, se obtiene finalmente la imagen reconstruida.

CAPITULO III

RESULTADOS EXPERIMENTALES

III.1.-Criterio de distorsión.

Cuando se recupera una imagen, que ha sido comprimida, existen muchas maneras de evaluar la fidelidad con que ésta ha sido reconstruida. Se puede hacer uso, por ejemplo, de métodos que nos proporcionan medidas simples de similitud [16]. Uno de ellos, ampliamente empleado, es el error medio cuadrático. En este trabajo se usa como parámetro de comparación el error medio cuadrático normalizado (NMSE) el cual se define por:

$$\text{NMSE} = \frac{\sum_{u=0}^{T-1} \sum_{v=0}^{T-1} (f(u,v) - \hat{f}(u,v))^2}{\sum_{u=0}^{T-1} \sum_{v=0}^{T-1} (f(u,v))^2} \quad (16)$$

donde $f(u,v)$ y $\hat{f}(u,v)$ son las muestras que representan a la imagen original y a la reconstruida respectivamente. T es el tamaño de los arreglos matriciales que forman las imágenes.

A pesar de que no es muy adecuado el empleo del NMSE cuando se va a evaluar visualmente la calidad de una imagen, debido a que es un parámetro que nos indica la cantidad de

error distribuido a lo largo de toda la imagen (no puede hacer distinción entre unos cuantos errores grandes y muchos pequeños), nos proporciona una idea de la cantidad de distorsión que se ha introducido al codificar la imagen.

A fin de cuentas, la evaluación de la calidad de una imagen reconstruida, es en gran medida subjetiva, debido a que dicha calidad depende de la utilidad que se le dará a la imagen, es decir, depende de si en ella se conservaron o no las características que son de nuestro interés en una aplicación determinada.

III.2.- Imágenes y gráficas de error.

Se ha graficado el número de bits por pixel promedio al que se han comprimido las imágenes fuente, contra el porcentaje de error cuadrático medio normalizado obtenido al reconstruirlas.

La compresión se realizó usando el método descrito en el capítulo anterior y con imágenes radiológicas representadas originalmente con 7 bits por pixel.

En las figuras 3, 5, 7, y 9 se muestran imágenes de resonancia magnética de cortes transversales de la cabeza a diferentes alturas. En cada una de estas figuras se tiene la imagen original y 4 reconstrucciones de ella para diferente

número de bits por pixel (3.5, 3.0, 2.5 y 2.0 bits/pixel), también se muestra una imagen que representa el error entre la imagen original y la que ha sido reconstruida después de haber sido comprimida a 3.5 bits/pixel. Las gráficas de bits por pixel contra por ciento de error en las figuras 4, 6, 8 y 10 corresponden a las imágenes fuente de las figuras 3, 5, 7 y 9 respectivamente.

Las figuras 11, 13, 15 y 17 muestran en este mismo orden, una célula, un tejido, una angiografía y unas vértebras. Para cada imagen fuente se tienen 4 reconstrucciones de la imagen con 2.5, 2.0, 1.5 y 1.0 bits/pixel. El error que se muestra corresponde ahora a la imagen reconstruida con 2.5 bits/pixel. Las gráficas de las figuras 12, 14, 16 y 18 corresponden a las imágenes de las figuras 11, 13, 15 y 17 respectivamente.

Se puede observar que las imágenes de resonancia magnética (figuras 3, 5, 7, y 9) al ser comprimidas a 2 bits/pixel, resultando con un error cuadrático medio normalizado en el rango de 0.75 a 1.0%, no presentan una distorsión notable. Para este tipo de imágenes el NMSE se comporta cae exponencialmente y es bastante predecible su tendencia al disminuir el número de bits por pixel, sin embargo, en las imágenes restantes, conforme se aumenta la tasa de

compresión, llega un momento en que el error se incrementa de manera notoria, como se puede apreciar, por ejemplo, en la figura 12 al codificar la célula con 1.5 bits por pixel. Este cambio abrupto se hace notar de igual manera en la imagen correspondiente, figura 11 d, donde aparecen ya muy marcados los bloques de la transformada para esta tasa de compresión.

La imagen que presentó menor distorsión al comprimirse fue la de las vértebras (figura 17) pues aún codificada con poco menos de 1 bit/pixel se tiene un NMSE menor de 0.1%. Esto se debe a que es una imagen con pocos componentes de alta frecuencia: no presenta contornos con grandes contrastes de tonos de gris.

El efecto de las componentes de alta frecuencia en la compresión con la transformada coseno se puede apreciar en la figura 19, donde se muestra una imagen de resonancia magnética como la de la figura 3 pero con un contraste mayor, con dos reconstrucciones, para 3.5 y 3.0 bits/pixel. En esa figura también se muestra la imagen de error de la reconstrucción de 3.5 bits/pixel.

Se observa que la distorsión introducida al comprimir este tipo de imagen es mucho mayor para la de la figura 19 que para la de la figura 3. Véase también el comportamiento del

NMSE de la figura 20, que corresponde a la imagen de la figura 19.

En la figura 21 se tiene un tejido, y tres de sus reconstrucciones (para 2.0, 1.5 y 1.0 bits/pixel) y en la figura 22 se encuentra su gráfica del error. Se observa en esta imagen un error mayor al obtenido con el tejido de la figura 13, la cual posee menores cambios bruscos de intensidad. La imagen original de la figura 21 presenta contornos más definidos y zonas casi completamente oscuras.

III.3.- Observaciones Generales.

Las imágenes que presentan contornos con grandes contrastes sufren mayor degradación al ser comprimidas, esto se debe a que la transformada coseno concentra la mayor cantidad de energía en los coeficientes que representan bajas frecuencias. Al tenerse una imagen suave (una imagen con pocas frecuencias altas) casi toda su energía se conserva en aquellos coeficientes en los cuales se introduce menor error durante la cuantización, resultando en una mejor calidad de la imagen reconstruida.

Al mismo tiempo, este comportamiento está relacionado con la frecuencia espacial a la que ha sido muestreada la imagen original. Ya que, si se tiene un conjunto de imágenes con

diferentes características y se representan todas con arreglos matriciales de una misma dimensión, es decir, se muestrean todas a una misma frecuencia espacial pero respetando para cada una de ellas el teorema de muestreo, algunas imágenes estarán menos sobremuestreadas que otras y, en consecuencia, será posible comprimir en mayor proporción aquellas que estén más sobremuestreadas.

El NMSE, como parámetro de comparación de la calidad de una imagen, es aplicable solamente a imágenes de la misma clase o que presentan características similares, es decir, no podemos, por ejemplo, comparar el NMSE de una imagen clara con el de otra que presenta grandes zonas oscuras, pues para una misma cantidad de distorsión el NMSE será mayor en una imagen oscura, debido a la normalización que se hace del error.

En el presente trabajo, lo ideal hubiera sido contar con algún especialista en radiología, quien nos pudiera dar su opinión acerca del grado de distorsión que se ha introducido a las imágenes después de haberlas comprimido, pues la evaluación visual que hemos realizado, simplemente se basa en analizar que contornos y formas se mantienen, y que nuevos o falsos contornos se introducen, así como si la cantidad de errores o ruido introducido ha hecho perder determinadas

formas, texturas y contrastes de la imagen original.

Las imágenes fuente han sido proporcionadas por la Universidad Tecnológica de Compiègne y se han procesado en una computadora personal AT. No ha sido el objetivo de este trabajo optimizar los tiempos de ejecución los cuales dependen principalmente de la manera de evaluar las transformadas. Sin emplear un algoritmo rápido para el cálculo de la transformada coseno, el tiempo de ejecución necesario para comprimir cada una de las imágenes con las que se trabajó es aproximadamente de doce minutos.

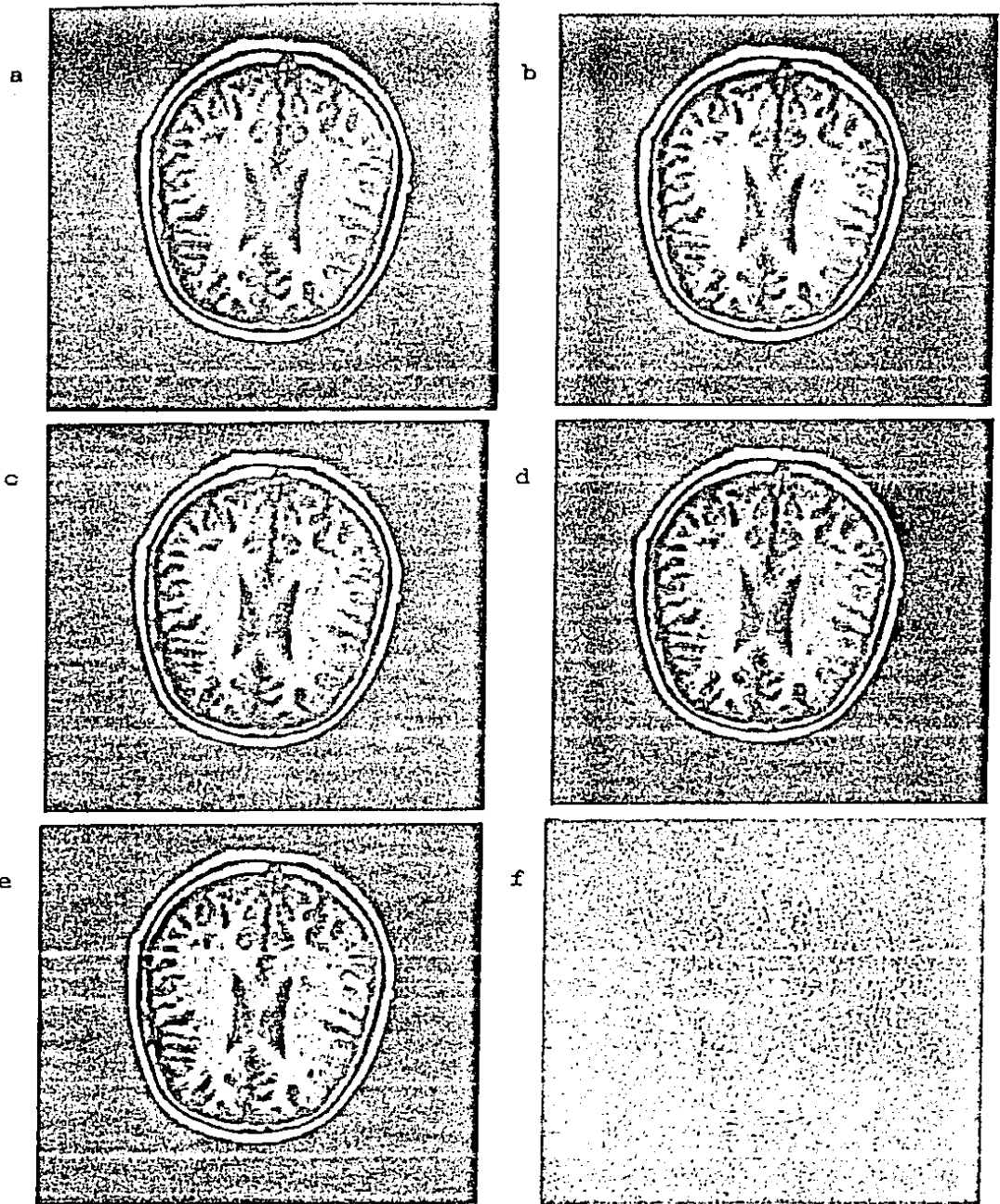


Fig. 3.a) imagen original e imágenes reconstruidas con b) 3.5, c) 3, d) 2.5 y e) 2 bits por pixel respectivamente. f) es la diferencia entre la imagen a) y la b).

NMSE. %

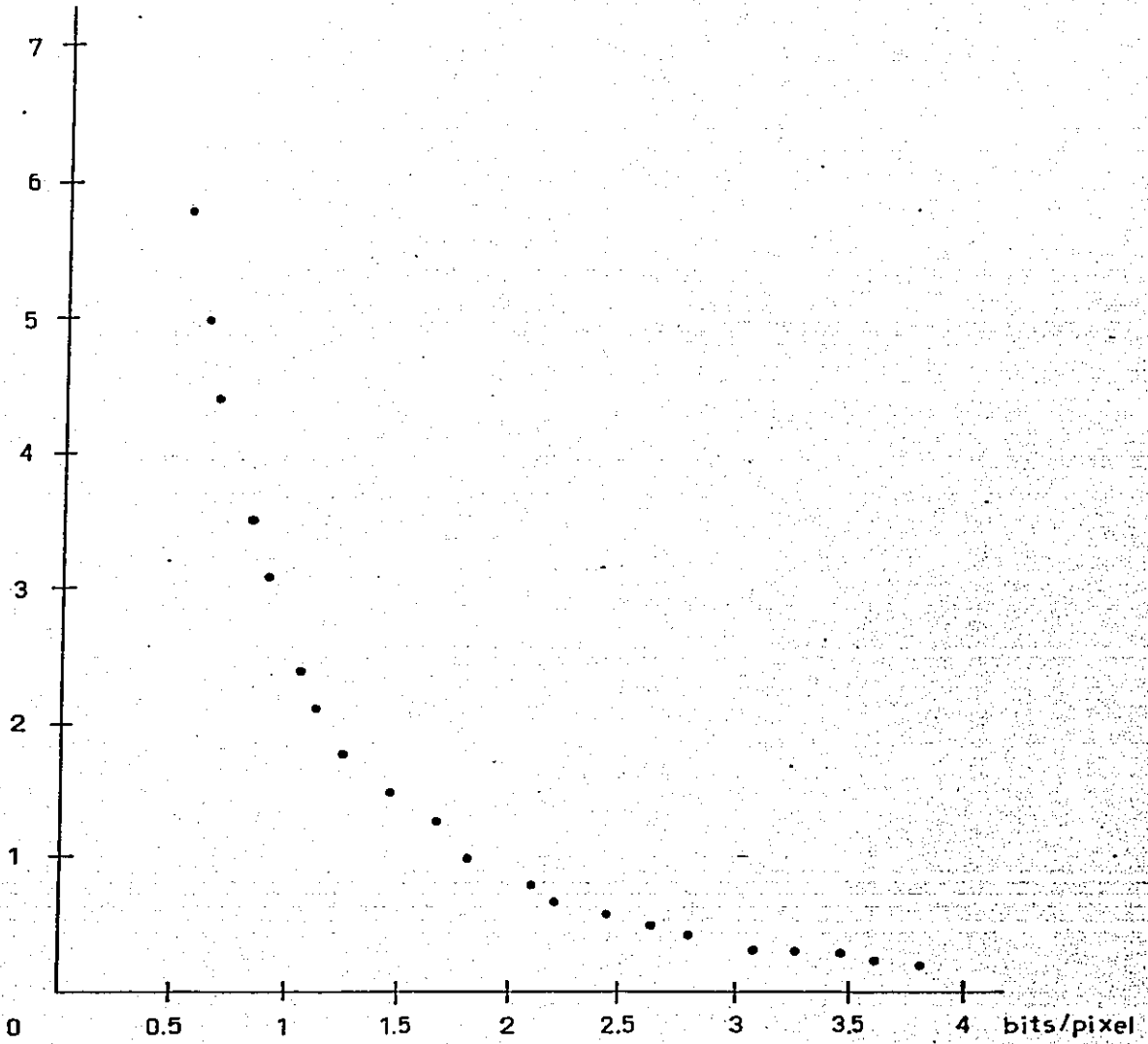


Figura 4.

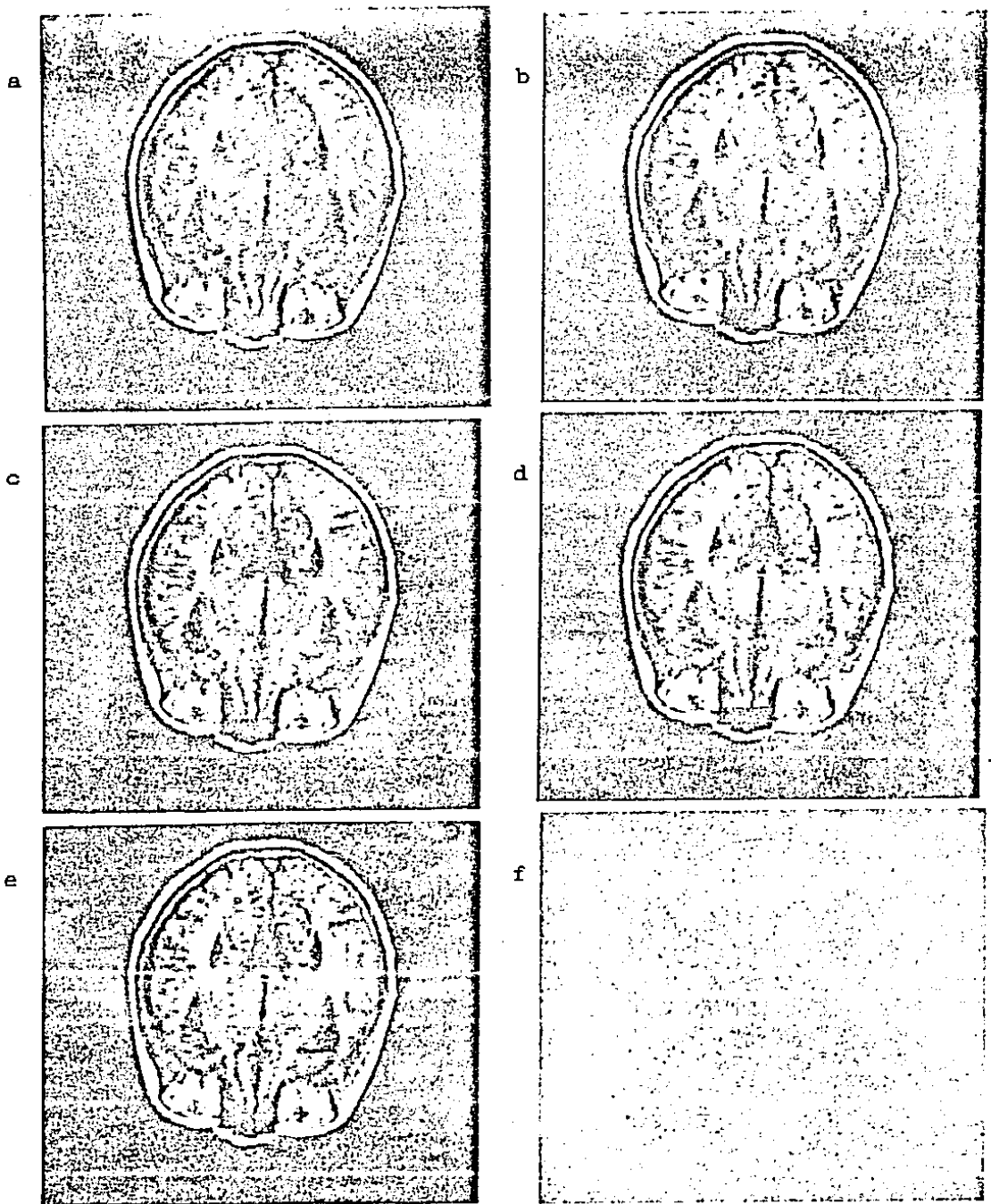


Fig. 5. a) imagen original e imágenes reconstruidas con b) 3.5, c) 3, d) 2.5 y e) 2 bits por pixel respectivamente. f) es la diferencia entre la imagen a) y la b).

NMSE %

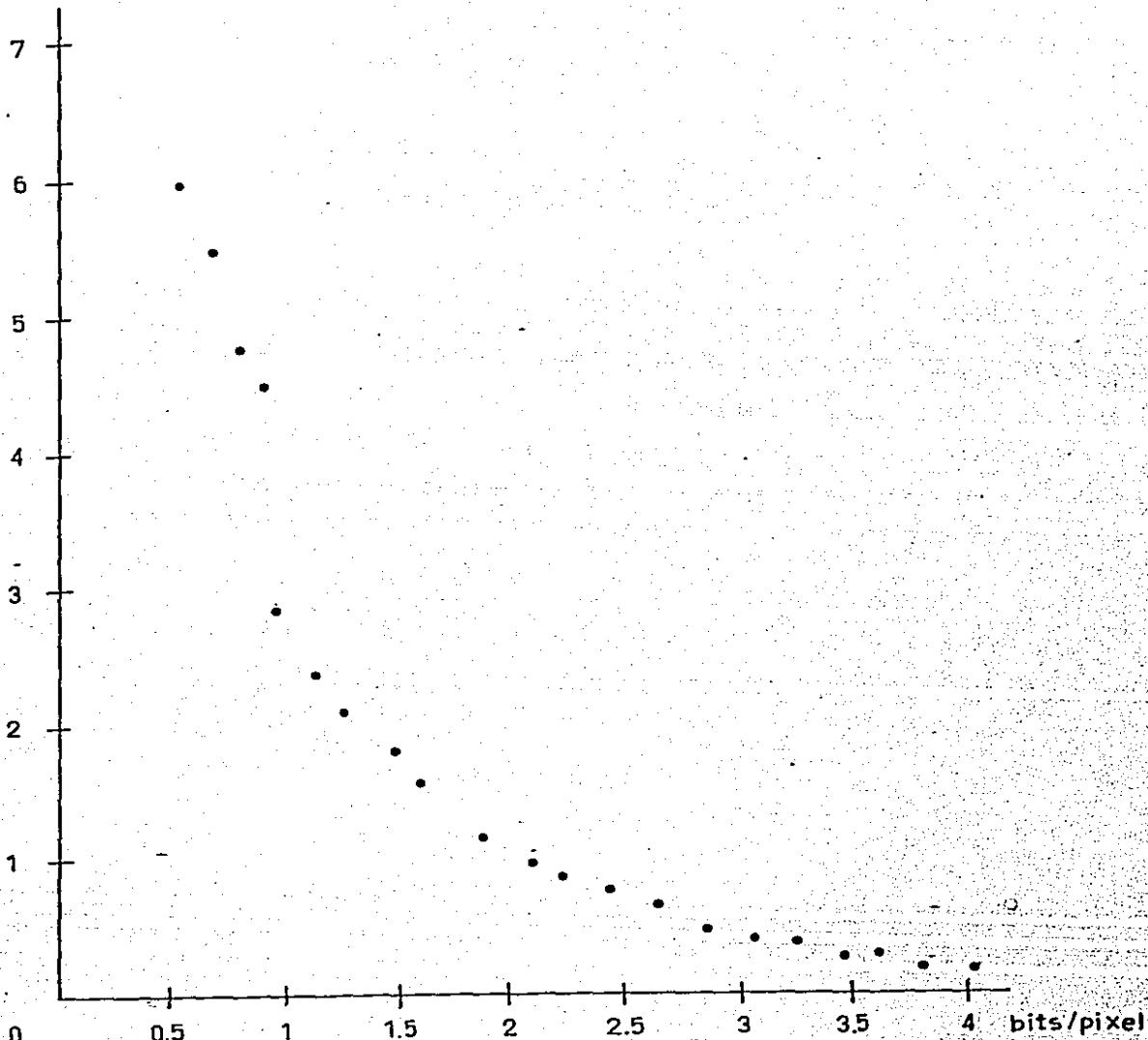


Figura 6.

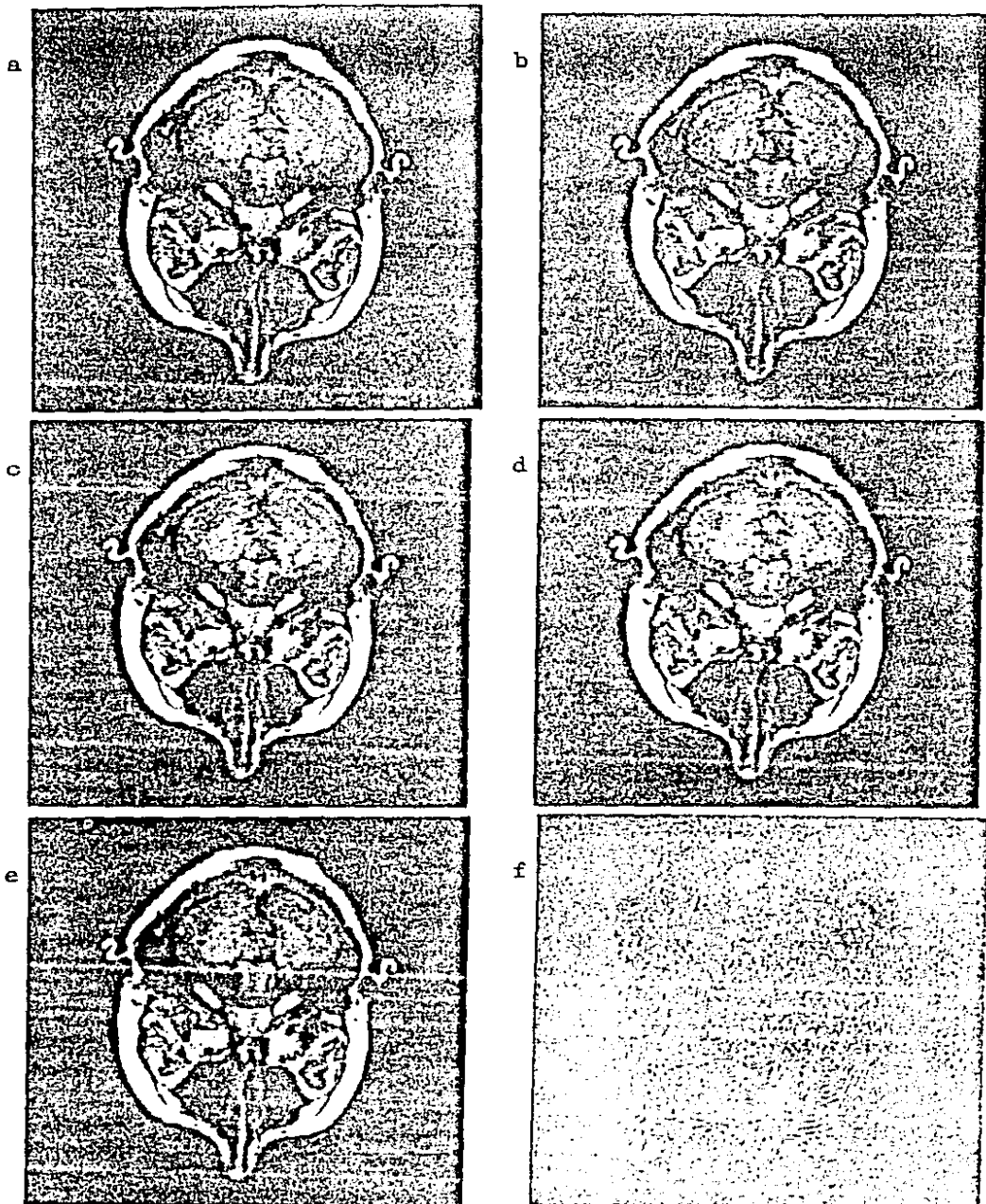


Fig. 7.a) imagen original e imágenes reconstruidas con b) 3.5, c) 3, d) 2.5 y e) 2 bits por pixel respectivamente. f) es la diferencia entre la imagen a) y la b).

NMSE %

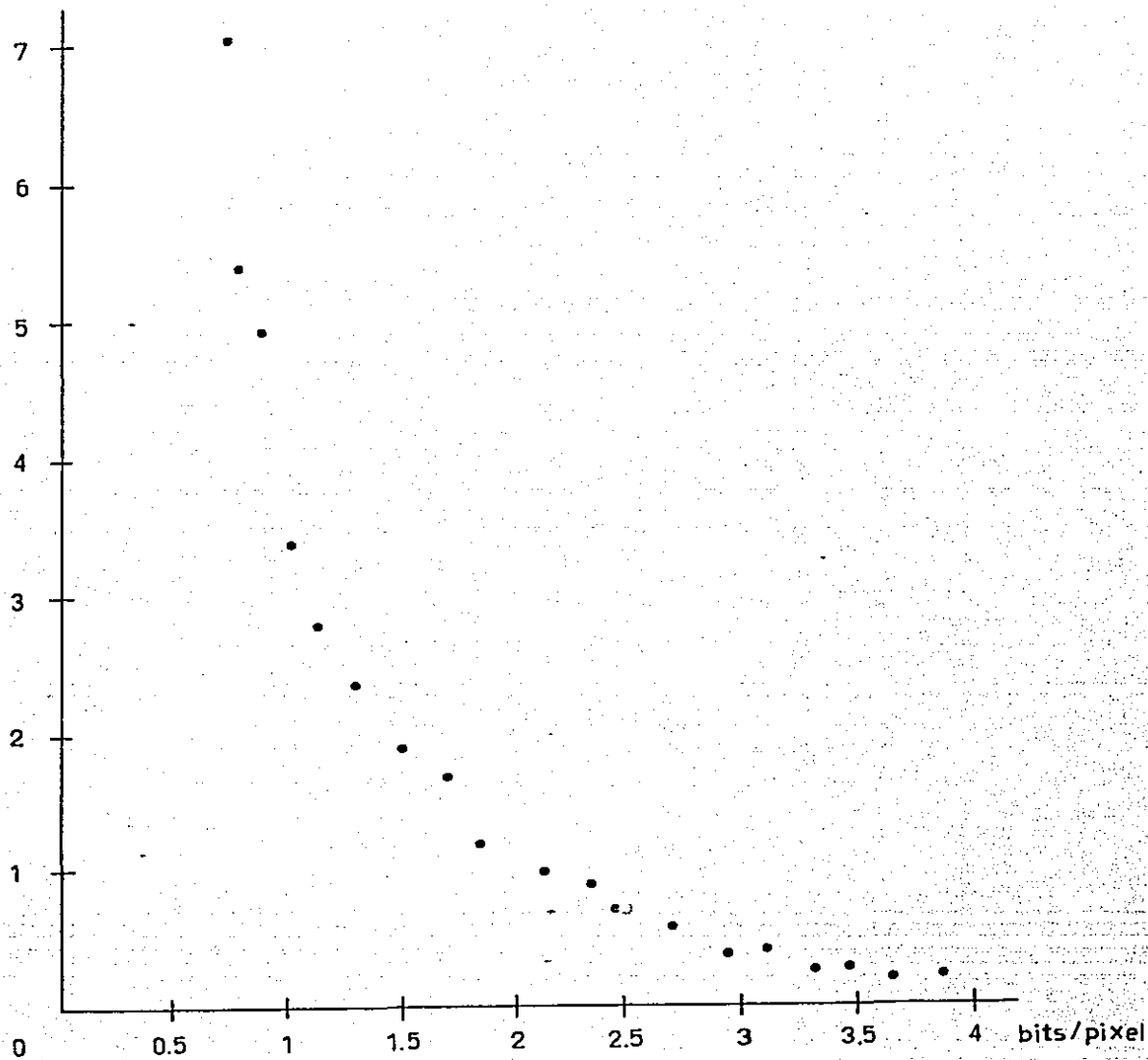


Figura 8.

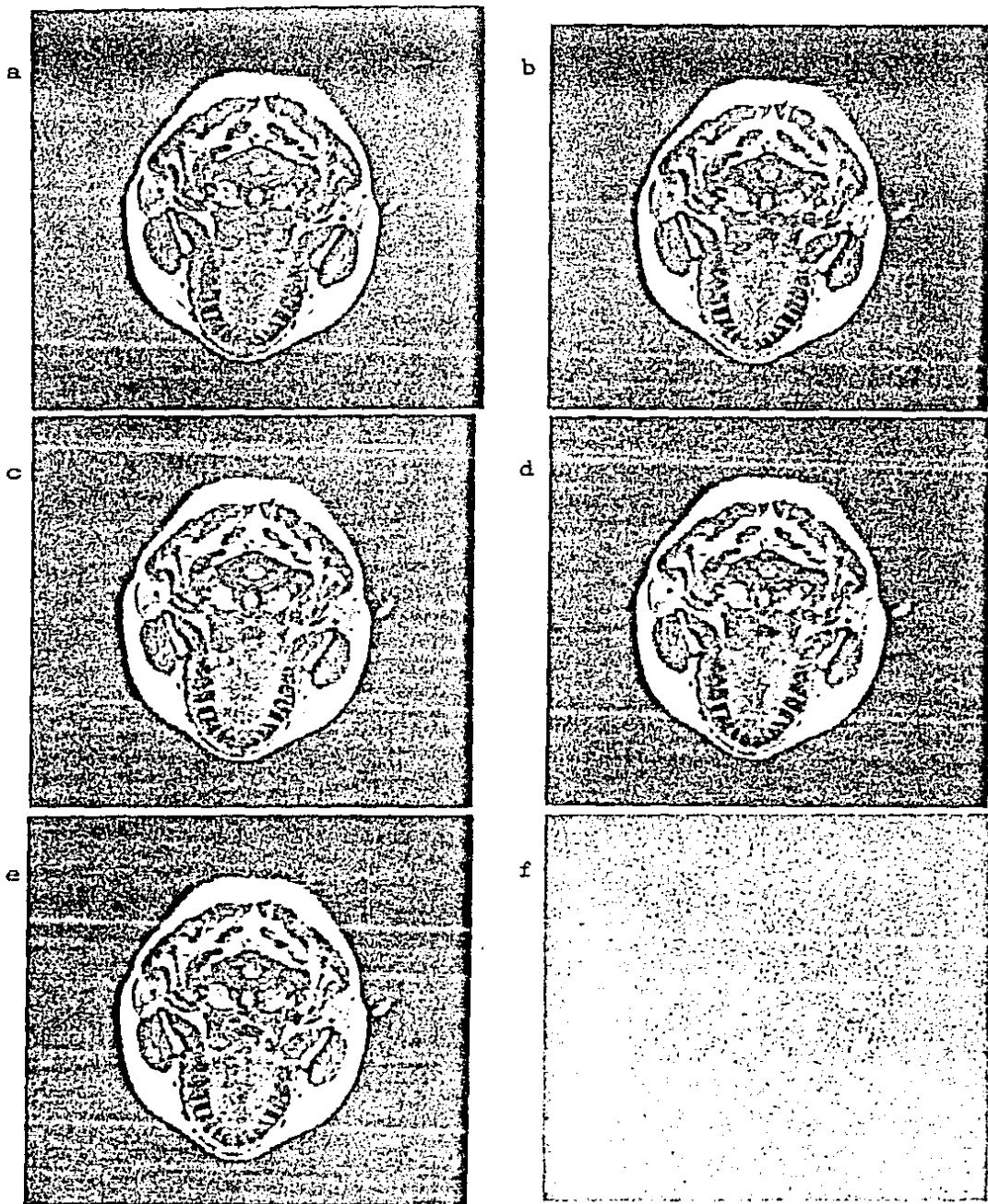


Fig. 9.a) imagen original e imágenes reconstruidas con b) 3.5, c) 3, d) 2.5 y e) 2 bits por pixel respectivamente. f) es la diferencia entre la imagen a) y la b).

NMSE %

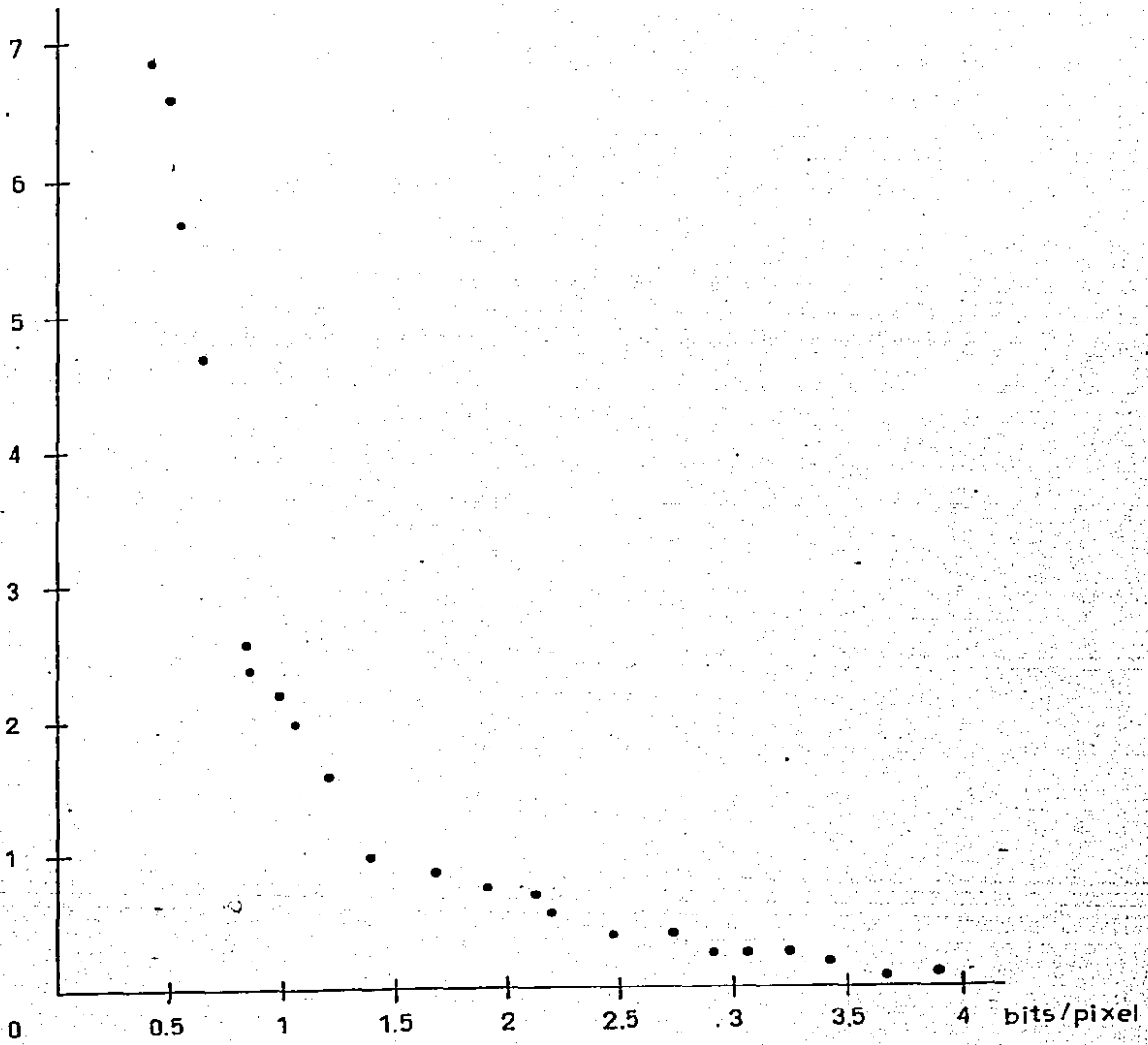


Figura 10.

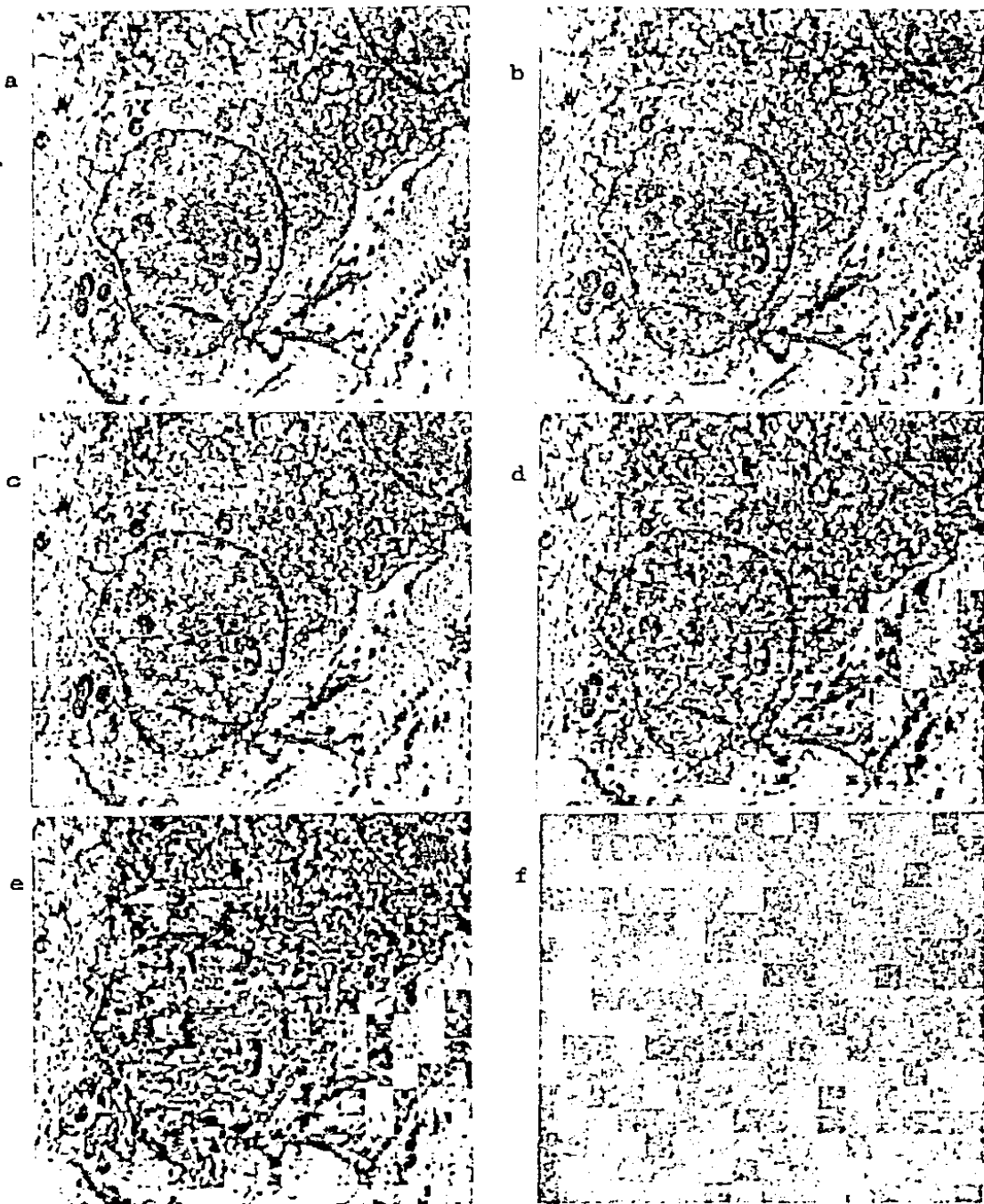


Fig. 11. a) imagen original e imágenes reconstruidas con b) 2.5, c) 2, d) 1.5 y e) 1 bits por pixel respectivamente. f) es la diferencia entre la imagen a) y la b).

NMSE %

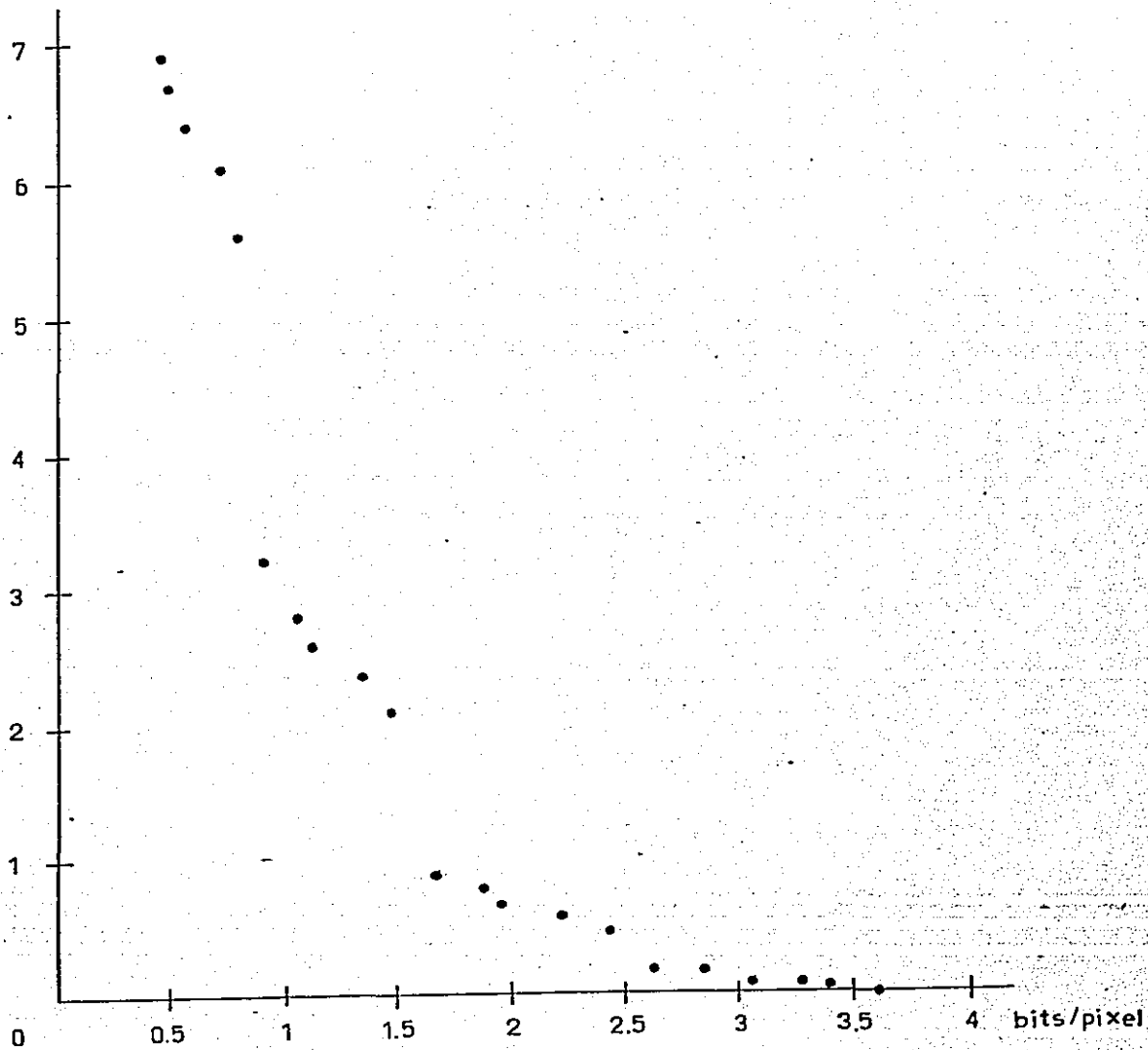


Figura 12.

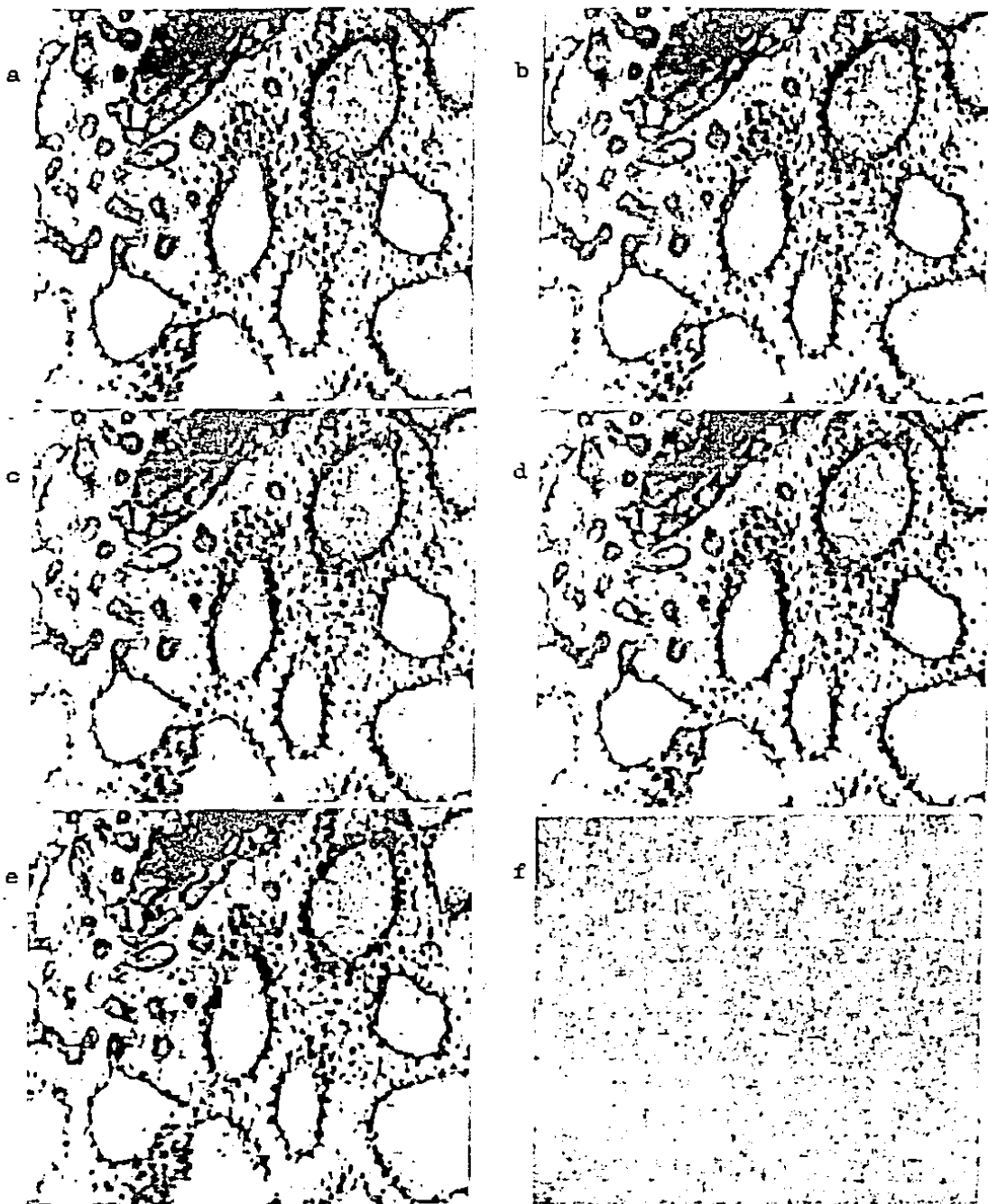


Fig. 13.a) imagen original e imágenes reconstruidas con b) 2.5, c) 2, d) 1.5 y e) 1 bits por pixel respectivamente. f) es la diferencia entre la imagen a) y la b).

NMSE %

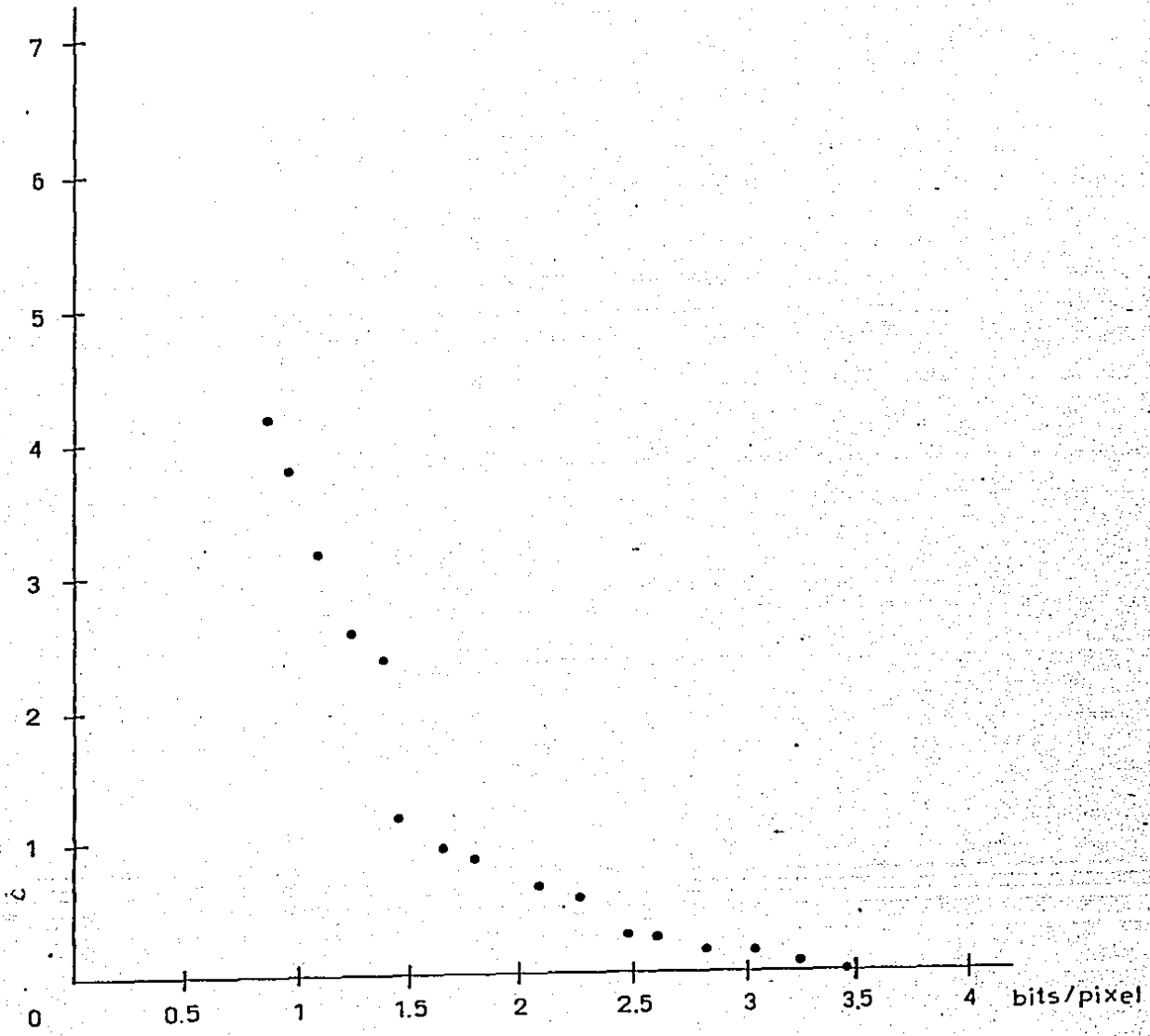


Figura 14.

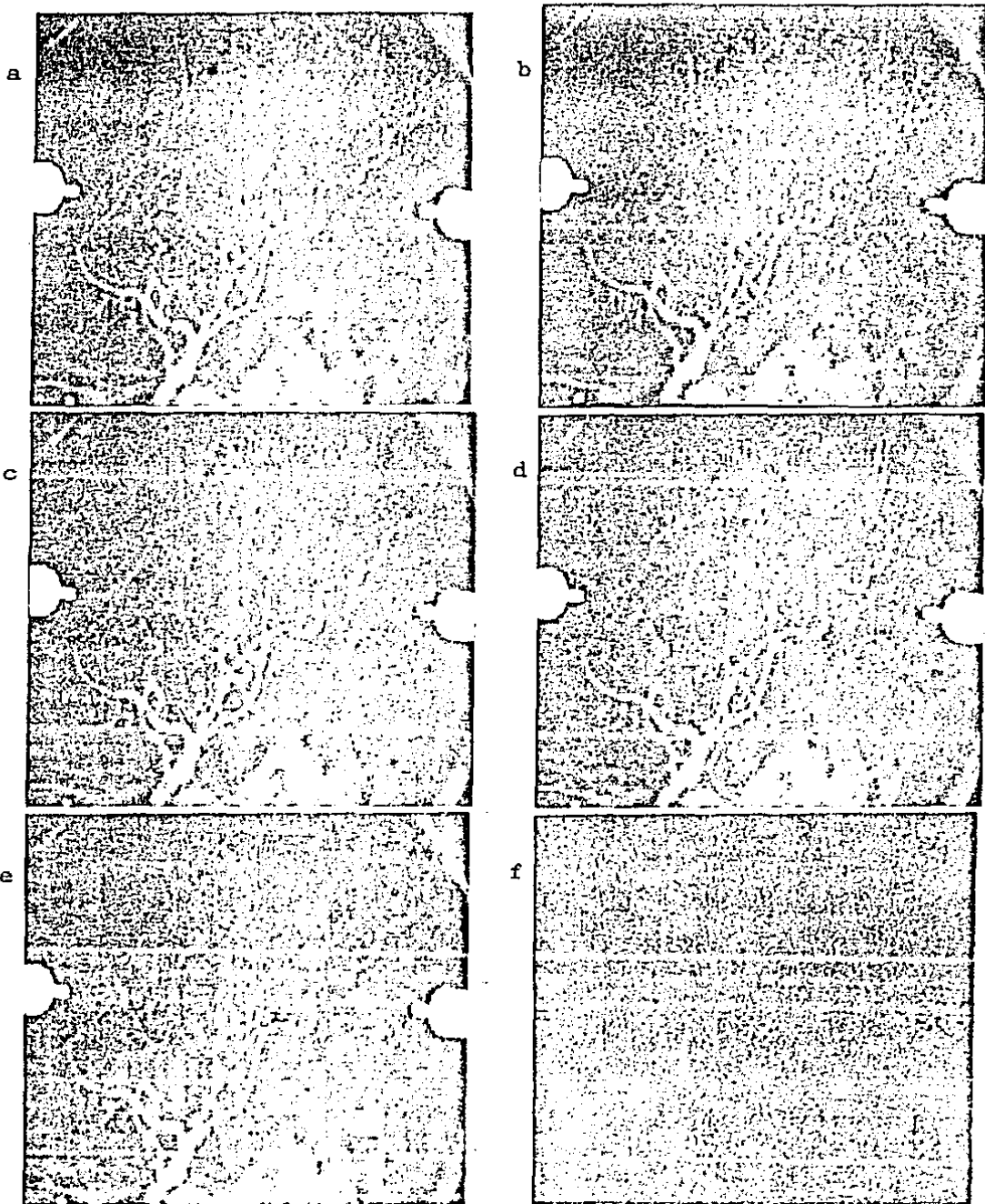


Fig. 15. a) imagen original e imágenes reconstruidas con b) 2.5, c) 2, d) 1.5 y e) 1 bits por pixel respectivamente. f) es la diferencia entre la imagen a) y la b).

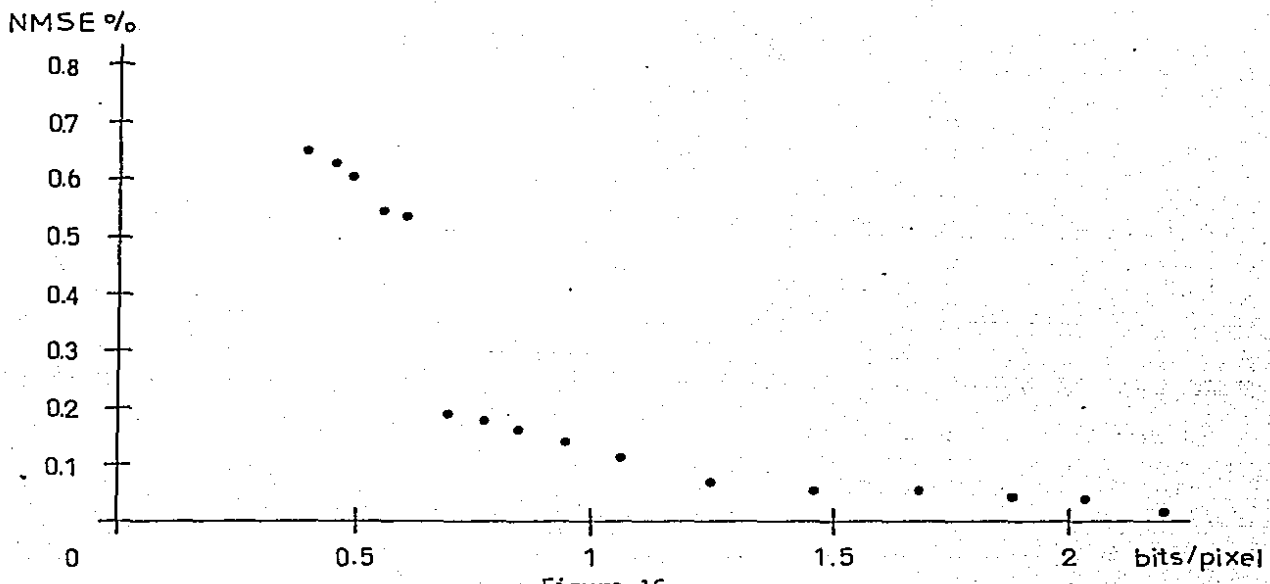


Figura 16.

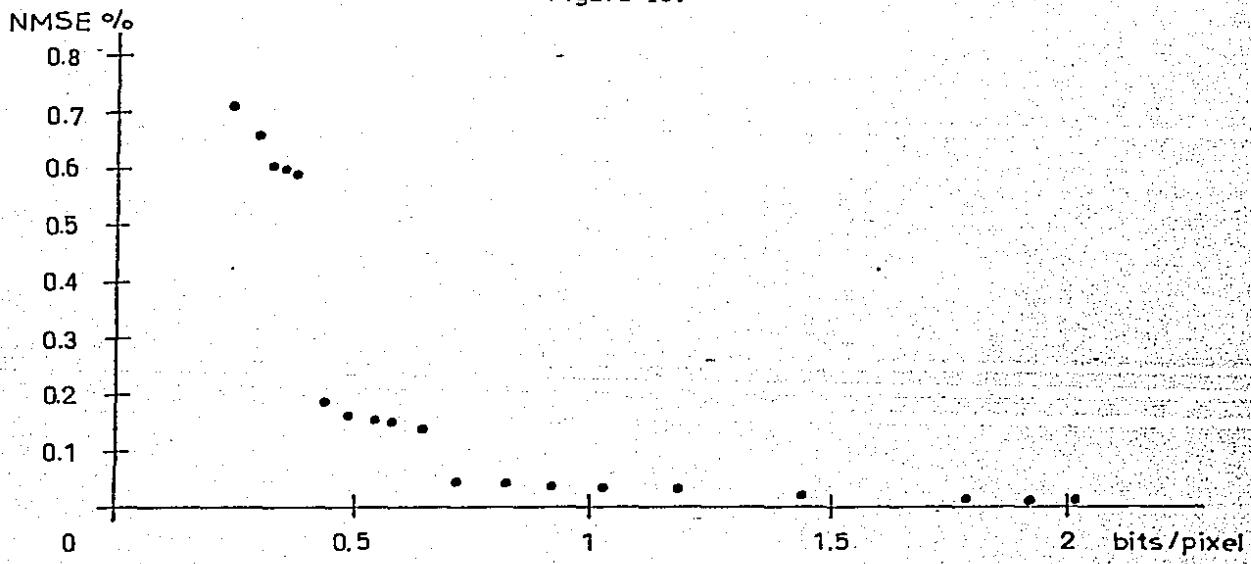


Figura 18.

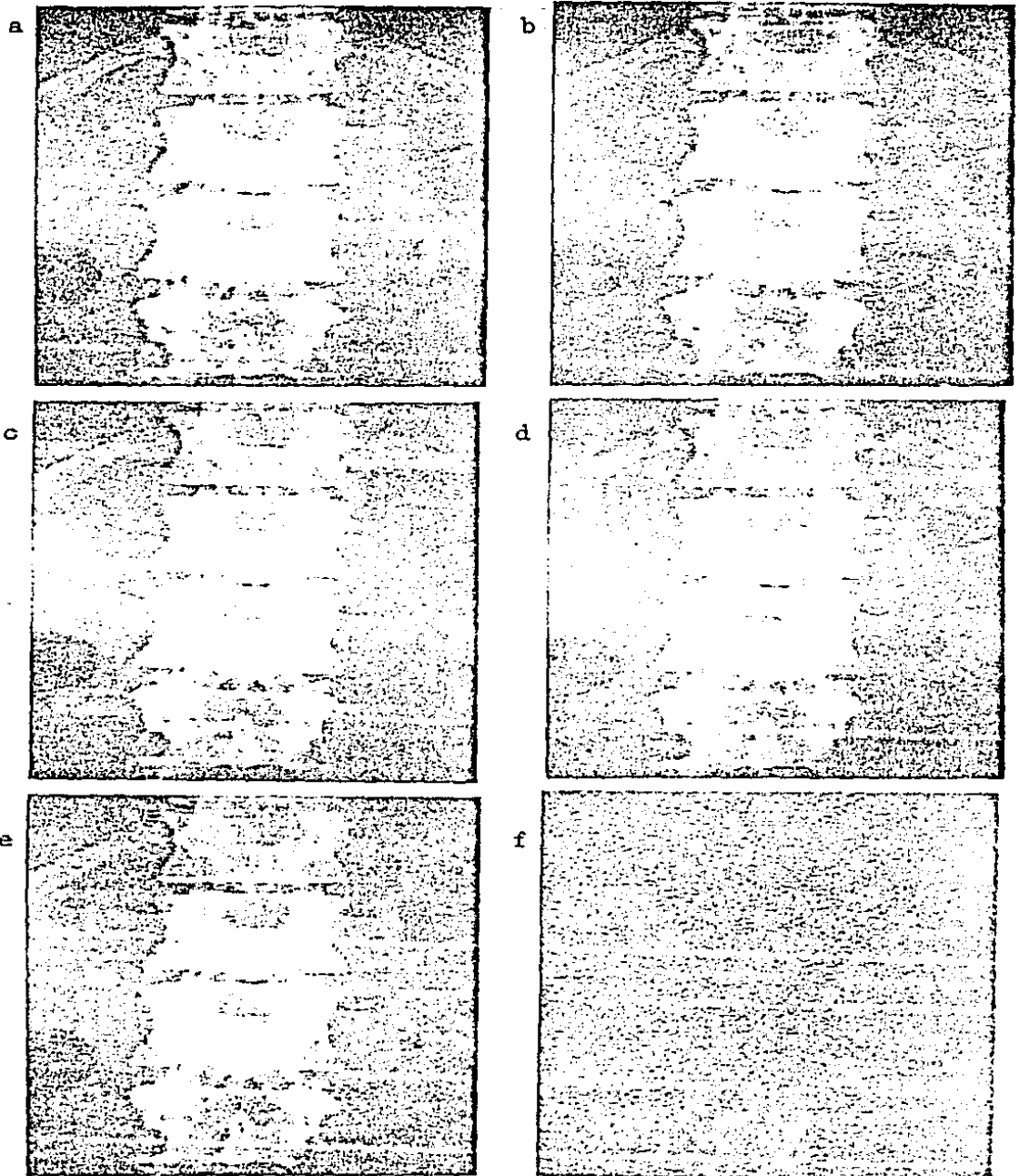
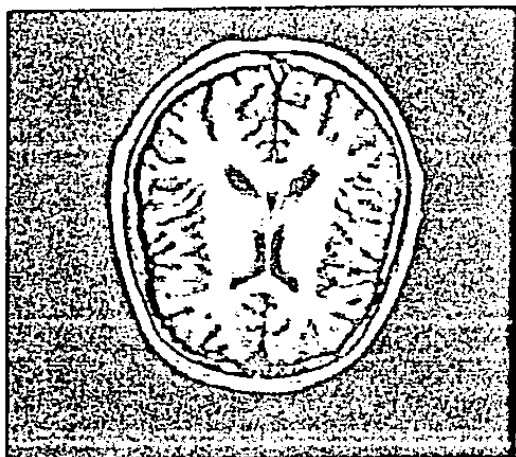
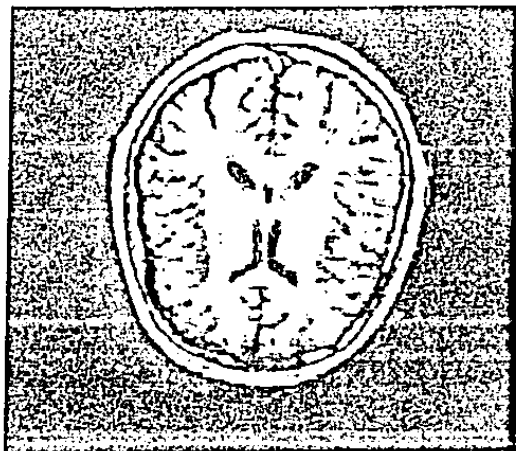


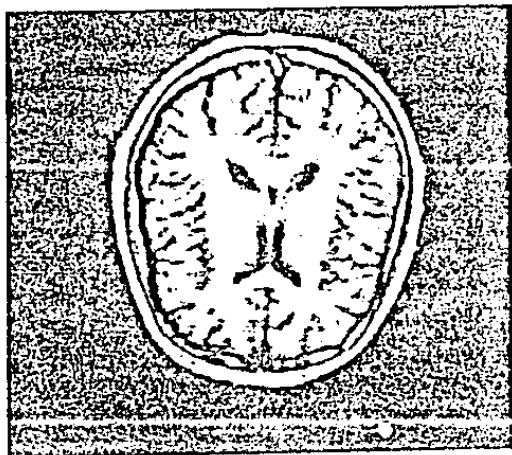
Fig. 17. a) imagen original e imágenes reconstruidas con b) 2.5, c) 2, d) 1.5 y e) 1 bits por pixel respectivamente. f) es la diferencia entre la imagen a) y la b).



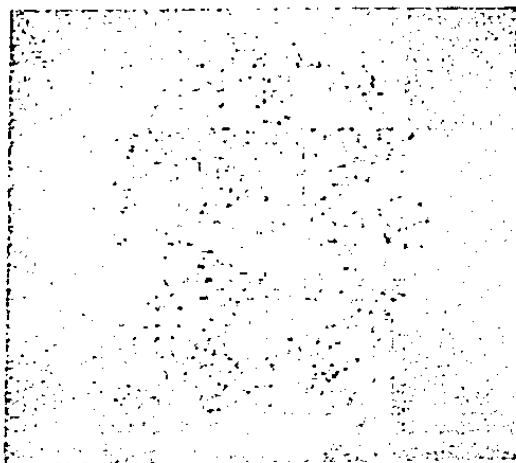
a



b



c



d

Fig. 19. a) imagen original e imágenes reconstruidas con b) 3.5 y c) 3 bits por pixel respectivamente. d) es la diferencia entre la imagen a) y la b).

NMSE %

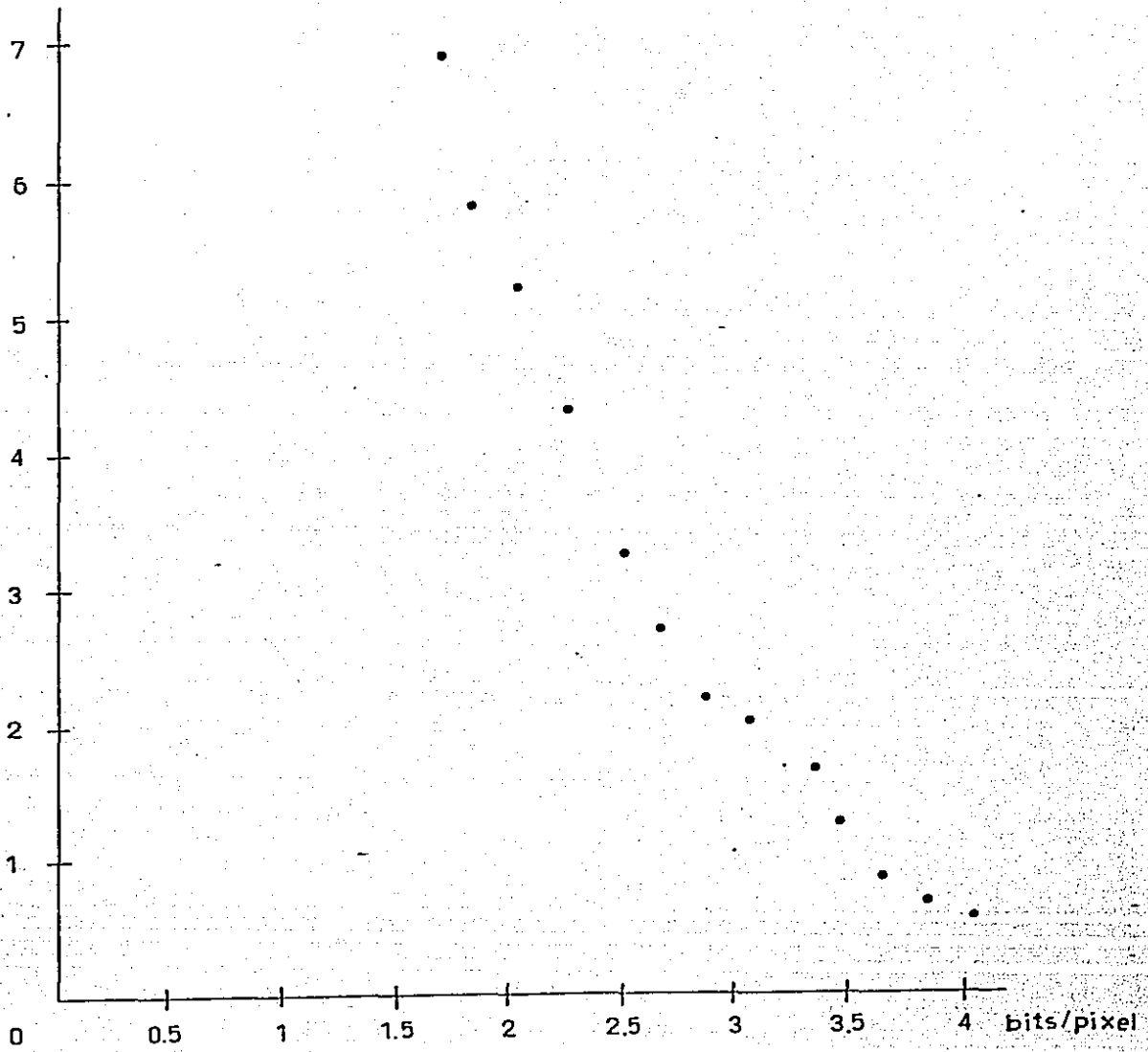
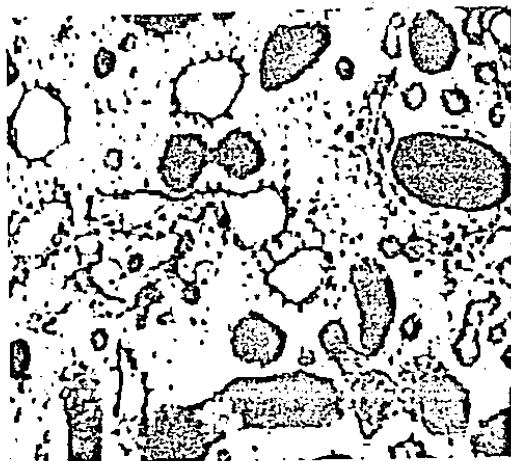


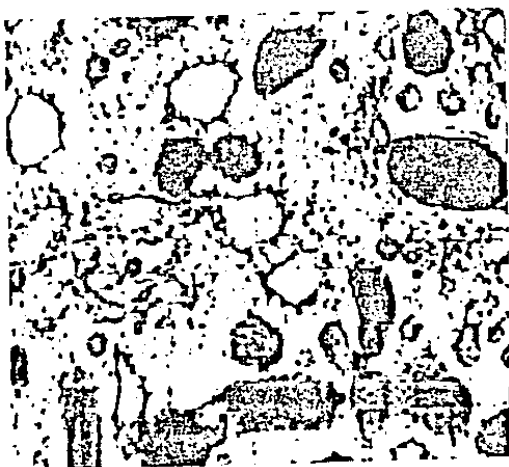
Figura 20.



a



b



c



d

Fig. 21. a) imagen original e imágenes reconstruidas con b) 2, c) 1.5 y d) 1 bits por pixel respectivamente.

NMSE %

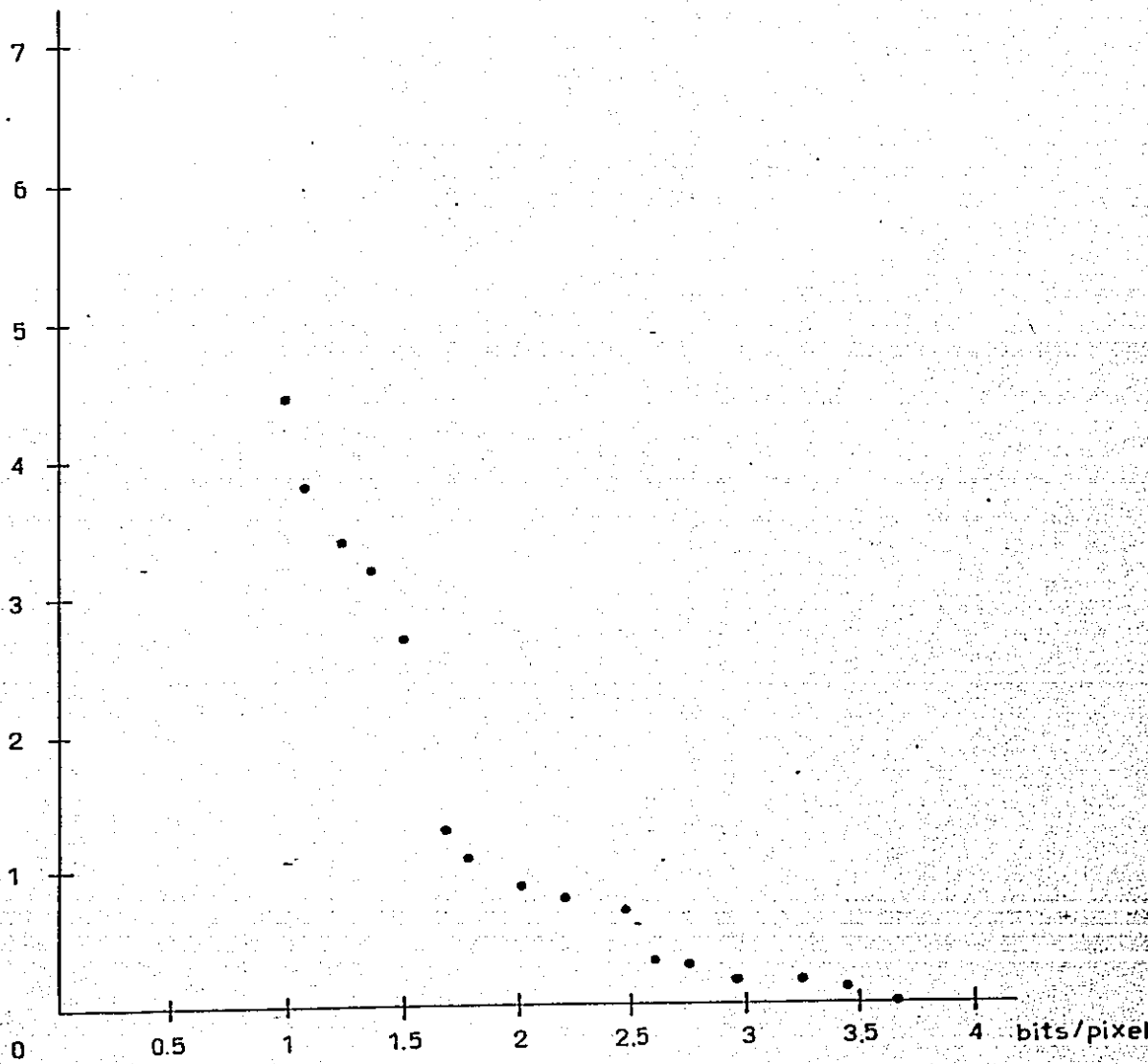


Figura 22.

CAPITULO IV

CONCLUSIONES

En este trabajo se ha evaluado el empleo de la transformada coseno discreta como herramienta para codificar imágenes radiológicas con el objeto de reducir el número de bits por pixel necesario, en promedio, para poder representarlas. En las imágenes con que se trabajó, al hacer una evaluación visual de su calidad, se puede establecer que un error cuadrático medio normalizado menor de 1% es aceptable, ya que, en este caso, no se aprecia la distorsión introducida durante su compresión.

Antes de realizar la codificación se ha dividido la imagen fuente en pequeños bloques, lo que ha permitido reducir el número de operaciones necesarias para el cálculo de la transformada de la imagen. Esta reducción se obtiene debido a que el tamaño empleado de la transformada es igual a la longitud de cada bloque. Y no al tamaño de la imagen completa, tal como se ha realizado durante la compresión de imágenes radiológicas en trabajos anteriores [7,12,13]. Sin embargo, cuando la tasa de compresión es grande, se pueden apreciar en las imágenes reconstruidas falsos contornos introducidos por la división mencionada. Este inconveniente

se puede aminorar, por ejemplo, permitiendo un pequeño traslape entre bloques adyacentes durante la segmentación de la imagen y posteriormente promediando los valores reconstruidos de cada pixel que haya sido codificado más de una vez. Otra manera de reducir los falsos contornos mencionados es empleando filtros paso-bajas solamente donde ocurren las divisiones de los bloques. Estas técnicas se describen con detalle en [15].

Las imágenes radiológicas en general, no pueden clasificarse como un solo tipo de imágenes, pues entre ellas existen grandes diferencias como son, entre otras, sus contenidos de frecuencia, sus factores de contraste y la agudeza de sus bordes y líneas. Por tal motivo, es necesario buscar alternativas en los métodos empleados para codificar estas imágenes (como es el emplear esquemas adaptivos o híbridos de reducción de redundancia y de entropía), de manera que se puedan alcanzar resultados óptimos aprovechando las características particulares de cada clase de imágenes radiológicas.

Para lograr mayores tasas de compresión cuando las imágenes procesadas deban ser evaluadas por el sistema de visión humano, es muy conveniente aprovechar las características del ojo y adecuar a él la codificación. Para esto, se

debe estudiar mas profundamente el sistema visual para tomar en cuenta qué es lo que el ojo ve y cómo lo ve. Existen algunos métodos [10], que toman en cuenta lo anterior y están basados en modelos de contorno-textura, en los cuales, la imagen se segmenta en regiones con determinadas texturas rodeadas por contornos, de tal forma que éstos correspondan, tanto como sea posible, a aquellos que aparecen en la imagen original y, posteriormente, la información de los contornos y de las texturas se codifican por separado.

Las ideas expuestas en estos últimos dos párrafos dan la pauta par el desarrollo de trabajos posteriores en la compresión de imágenes radiológicas.

BIBLIOGRAFIA

- [1] G. B. Anderson and T. S. Huang,
"Picture Bandwidth Compression by Piecewise Fourier
Transformation",
IEEE Trans. Commun. Technol. vol. COM-19, pp. 133-140,
1971.
- [2] N. Ahmed, T. Natarajan, and K.R. Rao,
"Discrete Cosine Transform",
IEEE Trans. Comput., vol. C-23, pp. 90-93, Jan 1974.
- [3] T. Berger
Rate Distortion Theory
Englewood Cliffs, NJ. Prentice-Hall, 1971.
- [4] W. H. Chen, C. H. Smith, and S. C. Fralick,
"A Fast Computational Algorithm For The Discrete
Cosine Transform",
IEEE Trans. Commun., vol COM-25, pp. 1004-1008,
Sept. 1977.
- [5] Rafael C. González and Paul Wintz,
Digital Image Processing,
Addison-Wesley Publishing Company, Inc., 1987.
Second Edition.

- [6] H. S. Hou,
"A Fast Recursive Algorithm For Computing The Discrete
Cosine Transform",
IEEE Trans. Acoust., Speech, Signal Processing,
vol. ASSP-35, pp. 1455-1461, Oct. 1987.
- [7] H. K. Huang,
Elements of Digital Radiology,
Prentice-Hall, Inc., 1987.
- [8] Huffman, D. A.
"A Method For The Construction Of Minimum-Redundancy
Codes",
Proc. IRE, vol. 40, pp. 1098-1101, September 1952.
- [9] A. K. Jain,
"Image Data Compression: a Review",
Proc. IREE, vol. 69, pp. 349-389, March 1981.
- [10] M. Kunt and A. Ikononopoulos,
"Second-Generation Image-Codig Techniques",
Proc. IEEE, vol. 73, pp. 543-574, April 1985.
- [11] B. G. Lee,
"A New Algorithm To Compute The Discrete Cosine
Transform",
IEEE Trans. Acoust., Speech, Signal Processing,
vol. ASSP-32, pp. 1243-1245, Dec. 1986.

- [12] Shih-Chung Lo and H. K. Huang,
"Compression of Radiological Images with 512, 1024,
and 2084 Matrices",
Radiology, vol. 161, pp. 519-525, November 1986.
- [13] Shih-Chung Lo, and H. K. Huang,
"Radiological Image Compression: Full-Frame
Bit-Allocation Technique",
Radiology, vol. 155, pp. 811-817, June 1985.
- [14] T. J. Lynch,
Data Compression, Techniques and Applications,
Van Nostrand Reinhold Company Inc., New York, 1985.
- [15] Reeve III, H.C., and J.S. Lim,
"Reduction of Blocking Effects in Image Coding",
Optical Engineering, vol. 23, pp. 34-37, Jan.-Feb.
1984.
- [16] A. Rosenfeld and A. C. Kak,
Digital Picture Processing,
Second Edition, Vol. 1, Academic Press, Inc., 1982.
- [17] N. Sueniro and M. Hatori,
"Fast Algorithms For The DFT and Other Sinusoidal
Transforms",
IEEE Trans. Acoust., Speech, Signal Processing,
vol. ASSP-34, pp. 642-644, June 1986.

[18] M. Vetterli and H. Nussbaumer,
"Simple FFT and DCT Algorithms With Reduced Of
Operations",
Signal Processing, Vol. 6, no.4, pp. 267-278,
Aug. 1984.

[19] P. A. Wintz and A. J. Kurtenbach,
Waveform error control in PCM telemetry,
IEEE Trans. Informat. Theory,
Vol. IT-14, pp. 650-651, 1968.