



## Criterios físicos para el estudio de la compresión con pérdidas en la imagen médica digital: revisión del tema

J. Paz y M. Pérez

Centro de Estudios de Electrónica y Tecnologías de la Información, Universidad Central “Marta Abreu” de las Villas; jpaz@uclv.edu.cu

Recibido el 15/07/2008. Aprobado en versión final el 03/04/2009.

**Resumen.** Se realiza un estudio de la calidad de imagen que se obtiene tras la compresión con pérdidas. Se enfoca un problema abierto de investigación, dirigido hacia el establecimiento de criterios objetivos para estimar el nivel de pérdidas durante la compresión, utilizando métricas tales como el Error Cuadrático Medio (MSE), la Relación Señal-Ruido máxima (PSNR), la relación Contraste-Ruido (RCR), el Índice de Similitud Estructural (SSIM index), Medidas de Distancia, basadas en la descomposición en valores singulares de la matriz imagen, medidas de correlación y espectrales. Se realiza además una revisión de los enfoques físicos de predicción de calidad de imagen a partir de utilizar observadores matemáticos como el de Hotelling en su forma de ventana (HO), en sus formas canalizadas (CHO y LGHO) y el del filtro acoplado no equalizador (NPWE). Finalmente, se analiza el enfoque subjetivo de análisis de calidad de imagen realizada a partir de análisis ROC basado en Curvas de Rendimiento Diagnóstico.

**Abstract.** The evaluation and description of image quality after applying lossy compression is described. It remains an open research issue in the sense of finding an objective criterion to stop lossy compression at a maximum level of compression with a minimum distortion. This value is found by implementing a set of quantitative bivariate measures such as the Mean Square Error (MSE), the Peak Signal to Noise Ratio (PSNR), the Structural Similarity index (SSIM index) and the Spectral Distance among others. A review is also done about the physical meaning of prediction algorithms for quality comparison such as the Hotelling Observer (HO), the channelized forms (CHO and LGHO), and the Non-Prewhitened Matched Filter (NPWE). Finally, the subjective evaluation of image quality is investigated through ROC analysis based on diagnostic performance curves.

**Palabras clave.** Digital imaging, image processing algorithms, 07.05.Pj

### 1 introducción

Una imagen médica es la representación gráfica de una estructura, región, órgano, o tejido del cuerpo humano que se obtiene a través de procesos físicos y computacionales ordenados para este fin.<sup>1,2</sup> Uno de los factores que incide en la calidad y cantidad del tráfico que se genera en una red de intercambio de imágenes, está relacionado con el tamaño de los ficheros que las

contienen.

Como se observa en la tabla 1, el volumen de estos datos puede complicar las tareas de almacenamiento y/o transmisión de estas imágenes<sup>3,4</sup>. Si se necesita almacenar la mayor cantidad de estudios por el mayor tiempo posible, o transmitir éstos con mayor efectividad y en menor tiempo, se hacen necesarios entonces los algoritmos de compresión o *códecs* y en particular los *códecs* con pérdidas.

La tasa de compresión (TC) que puede lograrse se mide mediante la expresión:

$$TC(\text{veces}) = \frac{T_{\text{Inicial}}}{T_{\text{Final}}}, \quad (1)$$

donde  $T_{\text{Inicial}}$  y  $T_{\text{Final}}$  son los tamaños de la imagen sin compactar y compactada respectivamente. La norma actual para la compresión con pérdidas de imágenes digitales es JPEG 2000, adoptado por DICOM (*Digital Imaging Communications in Medicine*), protocolo que regula el manejo e intercambio de datos e imágenes la medicina.<sup>5,6</sup>

La mayor desventaja de utilizar un códec con pérdidas radica en la cantidad de información que se elimina sin poderse recuperar más tarde de la data compactada<sup>7</sup>. Es de esperar que para cada tipo de imagen médica exista un valor de TC con el que se reduce al mínimo el espacio necesario para almacenarla, manteniendo en ella una calidad necesaria para diagnóstico. Este es un tema abierto de investigación hoy, común tanto para la Física Médica como para el Procesamiento Digital de Imágenes en general. En este caso el problema gira alrededor de la calidad de la imagen posterior a la compresión con pérdidas.

## 2 Calidad de la imagen médica digital

Históricamente, y gracias al surgimiento y desarrollo de otras disciplinas como la microscopía, la astronomía y la fotografía, se han establecido un conjunto de procedimientos, algoritmos y mediciones directas de tipo objetivo y subjetivo sobre las imágenes con el fin de tener un estimado global de su calidad.<sup>8</sup> Ésta se expresa en términos de los parámetros físicos tales como resolución espacial, ruido, contraste así como diferentes tipos de distorsiones. El conjunto de métricas objetivas y subjetivas utilizadas para medirla, se refieren al modo concreto de cuantificar uno o más de estos parámetros. Sin embargo, no se ha logrado definir cuál es la medida o el conjunto de ellas por las cuales juzgar los esquemas de compresión con pérdidas o determinar umbrales apropiados para cada tipo de imagen médica y protocolo de adquisición.<sup>9</sup>

**2.1 Medición de calidad de imagen a partir de medidas objetivas.** El conjunto de medidas de calidad que se presenta a continuación se han estado utilizando para evaluar calidad de imagen en estos casos.<sup>8-11</sup>

Entre las más utilizadas se encuentran las medidas de distancia y entre ellas la Relación Señal-Ruido (RSR) que puede definirse como:

$$RSR = \frac{S_{rms}}{N_{rms}}, \quad (2)$$

donde  $S_{rms}$  y  $N_{rms}$  son la raíz media cuadrática de la señal y el ruido respectivamente. Relacionada con esta magnitud puede estimarse la Ganancia en Relación Señal-Ruido como el incremento en la magnitud medido en *dB*, tras un determinado procesamiento con la imagen.

$$RSR(\text{dB}) = 10 \cdot \log_{10} \left( \frac{\sigma_x^2}{\sigma_e^2} \right), \quad (3)$$

donde  $\sigma_x^2$  es la varianza de la señal original;

$\sigma_x^2 = \frac{1}{N} \cdot \sum_{i=1}^N (x_i - \bar{x})^2$  con  $\bar{x} = \frac{1}{N} \cdot \sum_{i=1}^N x_i$  y  $\sigma_e^2$  es la varianza del error de reconstrucción medido con la diferencia entre la imagen sin compactar y compactada ( $e[n] = x[n] - \hat{x}[n]$ ). Si el error de reconstrucción no tiene valor medio  $e_i = 0$  entonces  $\sigma_e^2 = MSE$  (*Mean Square Error* o Error Cuadrático Medio).

$$\sigma_e^2 = \frac{1}{N} \sum_{i=1}^N (e_i - 0)^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{x}_i)^2 = MSE \quad (4)$$

**Tabla 1**

Modalidades de imagen médica, tamaño de la matriz de imagen, profundidad de bits y espacio que ocupa en memoria.

Modalidad	Matriz	PB*	Tamaño**
Medicina Nuclear	128 x 128	8	16,384
Resonancia Magnética	256 x 256	12	98,304
Ultrasonido	256 x 256	8	65,536
Tomografía Compt.	512 x 512	12	393,216
Fluoroscopia	1024 x 1024	8	1,048,576
Radiografía Digital	2048 x 2048	12	6,291,456
Mamografía	4096 x 5120	12	31,457,280

\*PB : profundidad de bits [bpp], \*\* Tamaño [Bytes].

No obstante, un valor alto de esta magnitud no garantiza una buena percepción de la imagen.<sup>8,9,11</sup> A fin de distinguir los objetos circundantes importa más la diferencia entre sus contrastes que la intensidad propia de cada uno de ellos. Por este motivo se prefiere en muchas ocasiones la magnitud Relación Contraste-Ruido (RCR) en lugar de la RSR.

La expresión para esta magnitud puede definirse como:

$$RCR = 10 \cdot \log_{10} \left( \frac{\bar{x}}{\sigma_x} \right), \quad (5)$$

donde  $\bar{x}$  es el valor medio de las intensidades de píxel en la imagen o en una región de interés (ROI) y  $\sigma_x$  la desviación típica de la misma imagen o ROI.

Las medidas de tipo bi-variadas tales como la *MSE* en la ecuación 4, comparan la imagen sin compactar con la compactada para brindar un estimado de la distorsión que introduce el procesamiento realizado. En general pueden clasificarse en medidas de distancia, de correlación, espectrales y las que tienen en cuenta la forma en que el sistema visual humano (SVH) interpreta

la información que percibe.<sup>9</sup> A continuación se describen algunas de estas medidas:

Considerando  $F(i,j)$  y  $\hat{F}(i,j)$  las matrices que representan la imagen sin compactar y reconstruida luego de haber sido compactada, se puede calcular:

- la Distancia Media Absoluta (*MAD*) como:

$$MAD = \frac{1}{m \cdot n} \sum_{i=1}^m \sum_{j=1}^n |F(i,j) - \hat{F}(i,j)|, \quad (6)$$

donde  $m$  es el número de filas y  $n$  es el número de columnas de la imagen.<sup>8</sup>

- la Relación Señal-Ruido máxima (PSNR):

$$PSNR(dB) = 10 \cdot \log_{10} \left( \frac{MAXp^2}{MSE} \right), \quad (7)$$

donde  $MAXp = 2^B - 1$  y  $B$  es la profundidad de bits de la imagen. Por ejemplo,  $MAXp = 255$  para una imagen de 8 bpp,

- la Ganancia en Relación Contraste-Ruido (*gRCR*) descrita por la expresión:

$$gRCR(dB) = 10 \cdot \log_{10} \left( \frac{RCR}{RCR_0} \right), \quad (8)$$

donde  $RCR$  y  $RCR_0$  son las magnitudes de la relación contraste-ruido de la imagen sin compactar y reconstruida respectivamente estimadas según la ecuación 5,

- la medida basada en la descomposición en valores singulares ( $M_{DVS}$ ) de la matriz de la imagen:

$$M_{DVS} = \frac{\sum_{i=0}^{\left(\frac{k}{n}\right) \times \left(\frac{k}{n}\right)} |D_i - D_{mid}|}{\left(\frac{k}{n}\right) \times \left(\frac{k}{n}\right)}, \quad (9)$$

donde:  $D_i = \sqrt{\sum_{i=1}^n (s_i - \hat{s}_i)^2}$  es la distancia entre los valores singulares de la imagen sin compactar ( $s_i$ ) y compactada ( $\hat{s}_i$ ) respectivamente y  $D_{mid}$  es el valor medio de los valores de  $D_i$ . Si  $n$  es el tamaño del bloque y la imagen es de dimensión  $k$  entonces existirán  $(k/n) \times (k/n)$  bloques<sup>9</sup>,

- las medidas de correlación (*Cor1*, *Cor2* y *Cor3*)<sup>10</sup> descritas por las expresiones:

$$Cor1 = \frac{\sum_{i=1}^m \sum_{j=1}^n F(i,j)^2}{\sum_{i=1}^m \sum_{j=1}^n \hat{F}(i,j)^2}, \quad (10)$$

$$Cor2 = \frac{\sum_{i=1}^m \sum_{j=1}^n F(i,j) \hat{F}(i,j)}{\sum_{i=1}^m \sum_{j=1}^n \hat{F}(i,j)^2}, \quad (11)$$

$$Cor3 = \frac{1}{m \cdot n} \sum_{i=1}^m \sum_{j=1}^n \left( 1 - \frac{2 \cdot \min(F(i,j), \hat{F}(i,j))}{F(i,j) + \hat{F}(i,j)} \right), \quad (12)$$

- la distancia espectral (*SD*) descrita por la expresión:

$$SD = \frac{1}{m \cdot n} \sum_{i=1}^m \sum_{j=1}^n (\varphi(i,j) - \hat{\varphi}(i,j))^2, \quad (13)$$

donde  $\varphi(i,j)$  y  $\hat{\varphi}(i,j)$  son las componentes de fase de las transformadas de Fourier de  $F(i,j)$  y  $\hat{F}(i,j)$  respectivamente.

Por último, y no menos importante, está el índice de similitud estructural (*SSIM*) propuesto por Wang, *et al.*,<sup>11</sup> el que compara los patrones locales de intensidad de los píxeles que han sido normalizados por la luminancia y el contraste. Una forma de esta expresión es:

$$SSIM(x,y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (14)$$

donde  $\mu_i$  es el estimado de luminancia de la imagen  $i$ ,  $\sigma_i$  es el estimado del contraste en la imagen  $i$ ,  $C_1 = (K_1 L)^2$  y  $C_2 = (K_2 L)^2$  son constantes con  $K_1$  y  $K_2 \ll 1$  y  $L$  = rango dinámico de los píxeles. El valor medio sobre toda la imagen se obtiene mediante la expresión:

$$MSSIM(X,Y) = \frac{1}{M} \sum_{j=1}^M SSIM(x_j, y_j), \quad (15)$$

donde  $X$  y  $Y$  son la imagen sin compactar y reconstruida respectivamente,  $x_j$  e  $y_j$  son los contenidos de cada imagen dentro de la ventana local  $j$ -ésima y  $M$  es el número de ventanas locales o bloques considerados<sup>11</sup>.

En relación con este resumen de medidas objetivas se debe enfatizar que históricamente estas se han reportado como poco correlacionadas con las opiniones de observadores humanos sobre la calidad de la imagen<sup>8-11</sup>. Por otro lado, las métricas basadas en SVH son técnicas en desarrollo que no han sido minuciosamente comparadas con el desempeño de un observador en aplicaciones médicas.<sup>10</sup>

En lo referente a la compresión con pérdidas, se estudia cómo el grado de compresión afecta el valor de estas métricas, proponiendo detener la compresión cuando la calidad de imagen (medida en estos términos), comienza a afectarse significativamente. La toma de decisiones sobre el valor de compresión óptimo a utilizar no dependerá en ningún caso de métricas de este tipo. Este valor sólo puede ser estimado mediante el empleo de observadores humanos expertos calificados y certificados para esta tarea específico de cada tipo de imagen médica y cada protocolo de adquisición. La figura 1 muestra un ejemplo de la utilización de este tipo de criterio físico para detener compresión en imágenes de cerebrales de RM de 512 x 512 píxeles 16 bpp.

**2.2 Medición de calidad de imagen a partir de observadores matemáticos.** La calidad de la imagen puede ser definida objetivamente en términos del desempeño de un observador sea este humano o matemático.<sup>12</sup> Ambos toman una decisión calculando una determinada estadística  $\lambda(\bar{g})$  y comparando su valor con un umbral de decisión. Si definimos como clase 1 la de casos positivos o de señal presente, la decisión será "positiva" si la estadística calculada sobrepasa dicho umbral. Si en realidad la señal pertenecía a esta clase se dice que la decisión es un Verdadero Positivo (VP) y si no, entonces es un Falso

Positivo (FP). El umbral de decisión controla el compromiso existente entre los VP y los FP.<sup>13</sup>

Una figura de mérito asociada a esta magnitud es la Relación Señal Ruido asociada a una estadística en particular y que también se denomina Índice de Detectabilidad y tiene la forma:

$$d^2 = \frac{[E(\lambda(g)|2) - E(\lambda(g)|1)]}{\frac{1}{2} \text{var}(\lambda|1) + \frac{1}{2} \text{var}(\lambda|2)}, \quad (16)$$

donde  $E(\lambda(g)|k)$  es la media condicional de la estadística  $\lambda(\bar{g})$  dado que  $\bar{g}$  provenga de la clase  $k$ , mientras que  $\text{var}(\lambda|k)$  es la varianza condicional correspondiente.

En el caso de los observadores matemáticos o bayesianos se trata de algoritmos, de mayor o menor complejidad, que pueden hacer un estimado de la calidad en la imagen, pueden predecir la actuación del observador humano y logran, en algunos casos, hasta reemplazar su labor.<sup>12-14</sup> Estos observadores modelo, son capaces de generar predicciones explícitas sobre la detectabilidad de una señal en un medio ruidoso.<sup>15-20</sup>

Entre los más conocidos destacan el observador de Hotelling (HO) y sus versiones canalizadas CHO (eng. *channelized Hotelling observer*) y LGHO (eng. *Laguerre-Gauss Hotelling observer*) y el filtro complementado no equalizador (NPWE, eng. *Non-PreWhitening matched filter with Eye model*)<sup>15,16</sup>.

La forma general de estos observadores es el producto interno entre una plantilla  $w_{n,j}$  y los datos  $g_{n,i}$  de la señal en cada una de las locaciones posibles.

$$\lambda_{i,j} = \sum_{n=1}^{N^2} w_{n,j} g_{n,i} = \vec{w}'_j \vec{g}_i \quad (17)$$

donde  $\lambda_{i,j}$  es la respuesta escalar del modelo en la locación  $i$  usando la plantilla  $j$ .  $\vec{w}'_j$  es una plantilla bidimensional de dimensión  $N \times N$  representada aquí por un vector columna de dimensión  $N^2 \times 1$ . El superíndice  $t$  se refiere a la transpuesta y el vector  $\vec{g}_i$  contiene los valores de la señal en la  $i$ -ésima locación presentados en forma de vector columna.<sup>12</sup>

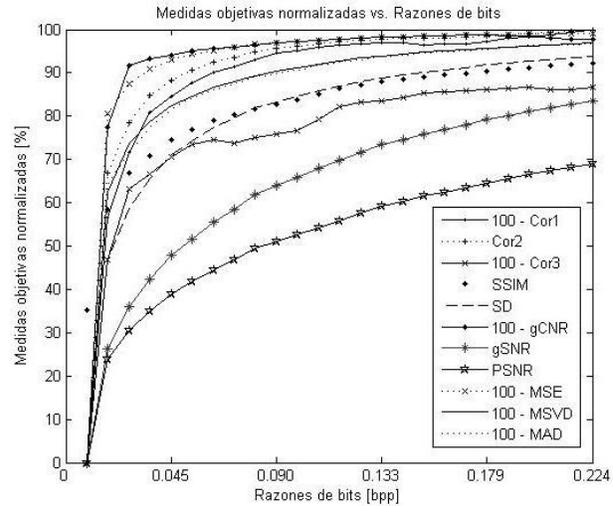
El observador de *Hotelling* utiliza información de la señal y de la correlación espacial del ruido para calcular sus plantillas. Si  $\hat{K}_g$  es la matriz de covarianza,  $\vec{g}_s$  es la media de la señal más ruido y  $\vec{g}_n$  es la media del fondo de la imagen la plantilla toma la forma:

$$\vec{w}_{Hot} = \hat{K}_g^{-1} [\langle \vec{g}_s \rangle - \langle \vec{g}_n \rangle]. \quad (18)$$

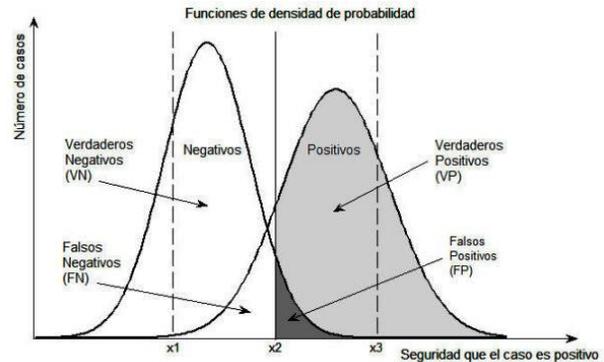
Un gran problema de este observador es la inversión de la matriz de covarianza en la anterior ecuación, a menos que la cantidad de imágenes exceda el número de píxeles en ellas; una condición que es virtualmente imposible de alcanzar con imágenes de tamaños comunes.<sup>12-14</sup>

Para resolver este problema se han planteado dos

soluciones; la primera consiste en restringir la matriz de covarianza a regiones de la imagen donde se encuentran las posibles localizaciones de la señal. A pesar de este recurso las matrices de covarianza continúan siendo cercanas a ser singulares (determinante muy cercano a cero) lo que hace muy difícil la inversión de la matriz.<sup>15</sup>



**Figura 1.** Medidas objetivas normalizadas vs. razones de bits (Inversamente proporcional a TC). Se observan las diferencias entre una métrica y otra en cuanto a sensibilidad ante los cambios de calidad. Por debajo de 0.045 bpp las métricas colapsan relacionado con un deterioro considerable en la calidad.



**Figura 2.** Distribución de probabilidad de señal útil y ruido.

La segunda solución utiliza funciones tales como los polinomios de *Laguerre-Gauss*

$$L_n = \sum_{m=0}^n (-1)^m \binom{n}{m} \left(\frac{x}{m}\right)$$

para extraer rasgos significativos de las imágenes.<sup>15,16</sup> Los canales están definidos en coordenadas polares como:

$$\vec{C}_{LG} = \exp\left(\frac{-\gamma}{2}\right) L_n(\gamma), \quad (19)$$

donde:  $\gamma = 2\pi \left( (x - x_c)^2 / a^2 + (y - y_c)^2 / b^2 \right)$ ,

$x_c$  e  $y_c$  marcan la localización de la señal mientras que  $a$

y  $b$  definen la orientación de éstos.

En este caso se utiliza la matriz de covarianza de la respuesta de los canales en lugar de la matriz de covarianza de los píxeles de la imagen para obtener la plantilla. De esta manera la plantilla queda de la forma:

$$\vec{w}_{LGHO} = \hat{w}_c^t \vec{C}_{LG} \quad (20)$$

donde  $\hat{W} = \hat{K}_v^{-1} \left[ \langle \vec{g}_{c/s} \rangle - \langle \vec{g}_{c/n} \rangle \right]$  es la covarianza de la salida de los canales.

A diferencia del LGHO, el CHO usa un conjunto de canales que pretenden reflejar en la etapa de preprocesamiento en el sistema visual humano<sup>15,16</sup>. Este conjunto de canales reduce la cantidad de información disponible para el modelo. La plantilla de CHO es la más apropiada que puede resultar de una combinación lineal de la frecuencia espacial y los canales de orientación que supuestamente reflejan la respuesta de las neuronas en la corteza visual primaria. Los canales de *Gabor* están definidos por:

$$C_{CHO} = \exp\left(\frac{-4\ln(x^2 + y^2)}{W_s^2}\right) \cdot \cos[2\pi f_c(x\cos\theta + y\sin\theta) + \beta] \quad (21)$$

donde  $f_c$  es la frecuencia espacial,  $\theta$  es la orientación,  $W_s$  es el ancho y  $\beta$  es la fase. La plantilla en este caso toma la misma forma del caso anterior, el LGHO.

Por último, entre los más sencillos de implementar, el NPWE tiene una plantilla de la forma:

$$\tilde{w}_j = \tilde{s}_j | \tilde{E} |^2 \quad (22)$$

donde  $\tilde{s}_j$  y  $\tilde{E}$  son la señal y el filtro visual respectivamente en el espacio de la transformada de *Fourier*. La plantilla se obtiene transformando inversamente  $\tilde{w}$ :

$$E_\rho = \rho^\eta \exp(-c\rho^\gamma) \quad (23)$$

donde  $\rho = \sqrt{u^2 + v^2}$  es la frecuencia especial radial en ciclos por grado. De acuerdo con Chawla *et al.*<sup>20</sup>, el parámetro  $c$  brinda mejor respuesta visual a 4 ciclos/grado,  $\eta = 1.5$  y  $\gamma = 0.68$ .

La figura 3 muestra un ejemplo del desempeño de algunos observadores matemáticos para predecir calidad de imagen a diferentes TC en imágenes de RM.

**2.3 Medición de calidad de imagen a partir de observadores reales.** El análisis ROC, también bajo el nombre de curvas de rendimiento diagnóstico, es una técnica matemática para describir y estudiar las propiedades de las pruebas diagnósticas. Mediante ella es posible tener un estimado de la exactitud diagnóstica de un determinado método, modalidad de imagen o evaluador, midiendo su capacidad de clasificar correctamente en categorías los casos bajo estudio.<sup>21,22</sup>

Cuando se realizan pruebas de tipo dicotómico, los datos de una prueba realizada a un grupo de pacientes se

colocan en una tabla de contingencia por el resultado de la evaluación y el estado real del paciente respecto a la enfermedad.

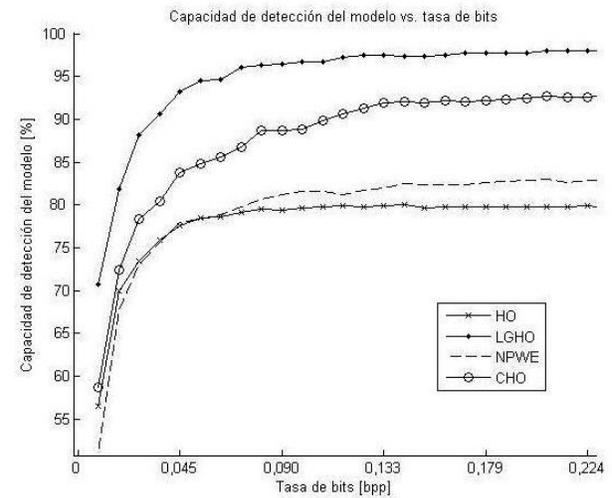
A partir de los datos en esta tabla se puede estimar la exactitud diagnóstica la cual se expresa en términos de la sensibilidad o fracción de verdaderos positivos,

$$\text{Sensibilidad} = FVP = \frac{VP}{VP + FN} \quad (24)$$

que no es más que la probabilidad de que el observador aprecie una señal anómala cuando en realidad ésta se encuentra presente y la especificidad o fracción de verdaderos negativos

$$\text{Especificidad} = FVN = \frac{VN}{VN + FP} \quad (25)$$

que mide la probabilidad de que el observador no aprecie señal anómala cuando realmente ésta no está presente.



**Figura 3.** El comportamiento de los cuatro modelos empleados en el intervalo de razones de bits desde 0 hasta 0.223 *bpp*.

<b>Tabla 2</b>			
Resultados de una prueba y su estado respecto a la enfermedad.			
		Verdadero Diagnóstico	
		Enfermo	Sano
Resultado de la prueba	Prueba positiva	Verdadero positivo (VP)	Falso Positivo (FP)
	Prueba negativa	Falso Negativo (FN)	Verdadero Negativo (VN)
		VP + FN	VN + FP

En el caso de evaluar la calidad de las imágenes médicas, en lugar de una simple evaluación dicotómica, se propone más de un umbral de calidad lo que resulta en varios valores de corte para la prueba determinados por la experiencia del observador, la información previa que éste tiene sobre la patología del paciente, el número de observaciones que realiza, las condiciones de

iluminación de que dispone para visualizar la imagen, la calidad de la imagen, y las características del sistema de detección (sensibilidad, resolución entre otras).

Un gráfico de estos valores a medida que varía el umbral de decisión se denomina curva ROC y el área bajo de ésta, (*AUC*, eng. *Area Under the Curve*) puede ser interpretada como la medida de precisión diagnóstica del observador. Sus valores oscilan entre 0.5 (clasificación regular) y 1.0 (clasificación perfecta) <sup>21,22</sup>.

En relación con la compresión de imágenes se deduce que a mayor compresión existirá mayor superposición en las distribuciones de señal útil y ruido reflejadas en la figura 2, lo cual se traduce en menores niveles de detectabilidad por parte del observador real reduciéndose el área bajo la curva ROC.

Desde el punto de vista de la teoría de las probabilidades, los valores de reales y falsos positivos se calculan según las expresiones:

$$P(S/s) = \int_{x_c}^{\infty} f(x/s) dx \quad (26)$$

$$P(S/r) = \int_{x_c}^{\infty} f(x/r) dx \quad (27)$$

donde  $f(x/s)$  es la distribución de probabilidad normalizada que describe la probabilidad de obtener un suceso particular de la variable aleatoria  $x$  cuando la señal y el ruido están presentes,  $f(x/r)$  es la distribución de probabilidad normalizada que describe la probabilidad de obtener un suceso particular de la variable aleatoria  $x$  cuando solo el ruido está presente y  $x_c$  es el umbral subjetivo planteado por el observador para delimitar la señal útil "S" ( $x > x_c$ ) de lo que es solo ruido "R" ( $x < x_c$ ).

De la normalización al área unitaria queda claro que  $P(S/s) + P(R/s) = 1$  y que  $P(S/r) + P(R/r) = 1$ . En las imágenes médicas esto significa presencia de un sujeto enfermo o no enfermo entre un grupo mixto presentado aleatoriamente a un observador que valora entre 1 - seguro que hay lesión a 4 - seguro que no hay lesión pasando por las categorías intermedias 2 - creo que existe lesión, 3 - creo que no existe lesión.

La relación entre la probabilidad condicional de una respuesta real positiva respecto a la probabilidad condicional de un falso positivo se basa en el criterio de que  $X_c$  viene dada por:

$$\frac{\int_{x_c}^{\infty} f(x/s) dx}{\int_{x_c}^{\infty} f(x/r) dx}, \quad (28)$$

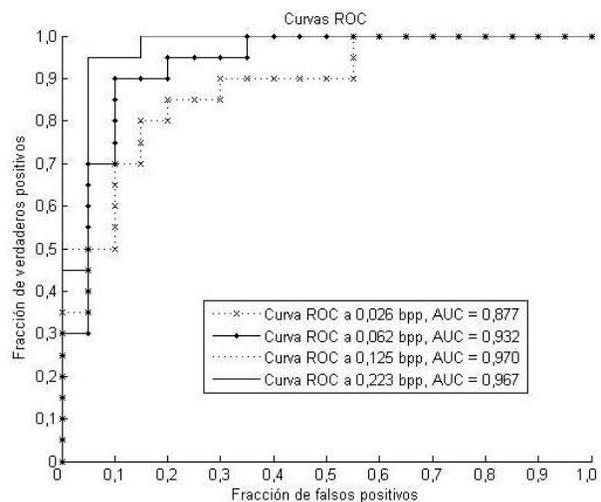
la cual es igual al cociente entre el área bajo la cola derecha de la distribución de señal más ruido, entre el área bajo la cola derecha de la distribución de ruido.

Se ha mostrado cómo un observador experto tiende a generar una única curva ROC en repetidos experimentos, pero diferentes observadores generarán curvas diferentes. Una de las principales limitaciones del método es que el observador se convierte en parte del

"sistema" bajo análisis. Mientras mayor sea el solapamiento de las distribuciones de la señal y del ruido mayor será la probabilidad de que el observador responda incorrectamente. Esto se torna crítico para valores de solapamiento menores que dos desviaciones típicas de las señales en cuestión, producto del umbral limitado de detección del ojo humano.

La figura 4 muestra un ejemplo del empleo del análisis ROC para evaluar calidad de imagen durante la compresión a partir de la evaluación subjetiva de observadores reales de imágenes cerebrales de RM de 512 x 512 píxeles.

El área bajo la curva ROC representa la probabilidad de que un sujeto enfermo elegido aleatoriamente sea correcto o ponderado con mayor sospecha que sujetos no enfermos también elegidos aleatoriamente. Este índice va desde 0.5, (curva 'A' en las figs. 8 y 9), en el caso que no exista una precisión aparente, hasta 1 en un caso ideal donde existe perfecta precisión. Este valor constituye el índice de precisión diagnóstica y puede ser utilizada para comparar dos o más procedimientos diagnósticos, dos modalidades de imágenes o dos evaluadores.



**Figura 4.** Curvas ROC para un observador modelo a 4 niveles diferentes de compresión en imágenes de RM cerebral de 512 x 512 píxeles.

### 3 Conclusiones

A partir del análisis del tema de la calidad de la imagen posterior a la compresión con pérdidas podemos concluir que:

Cuando se utilizan los algoritmos de compresión con pérdidas, se pueden eliminar detalles que representan información útil para el diagnóstico. La cantidad de información que se elimina esta directamente relacionada con el valor de TC que se emplee. Por lo tanto es necesario estimar un valor de TC para cada tipo de imagen médica donde se reduzca al mínimo el espacio necesario para su almacenamiento a la vez que se mantiene la calidad que la hace útil para el diagnóstico.

Este valor puede ser estimado con la ayuda de un grupo de medidas de calidad que van desde las de tipo objetivas, hasta medidas de la precisión diagnóstica de un grupo de especialistas entrenados para este fin (las que se reportan como las mejores pero también las que mas tiempo consumen), pasando por los observadores modelo (algoritmos matemáticos que simulan la tarea de observación de los especialistas).

Las medidas de calidad de tipo objetivas son útiles en sus modos bivariados en el que realizan una comparación con la imagen sin compactar. Hasta el momento no se ha reportado una medida de calidad objetiva correlacionada con el SVH y que pueda utilizarse para estimar el valor de TC óptimo.

Aunque se han hecho diversos esfuerzos por encontrar una medida de calidad unificada, el problema de encontrar un valor numérico para TC que se ajuste a cada modalidad de imagen en particular, y a cada patología además, continúa estando abierto a la investigación.

Los observadores matemáticos, por su parte, son útiles para predecir el comportamiento de los observadores humanos. De esta forma es posible planificar mejor los experimentos ROC los que tienden a ser altos consumidores de tiempo; situación no propicia al emplear especialistas.

El valor de TC óptimo puede ser estimado solamente a partir del análisis estadístico de las observaciones hechas por especialistas quienes son los únicos certificados y calificados para esta labor.

Los resultados que se obtengan para cada tipo de imagen médica en formato digital y para cada protocolo de adquisición podrían pasar a integrar las normas existentes para la manipulación e intercambio de las imágenes.

## Referencias

1. Russ, J.C., The Image Processing Handbook, 2ed. CRC Press, Boca Raton, Florida, (1995).
2. Bankman, I., Medical Imaging Handbook: Processing and Analysis, Academic Press. Vol. 1, San Diego, California, (2000).
3. Kalyanpur, M. A., et al., "Evaluation of JPEG and Wavelet Compression of Body CT Images for Direct Digital Teleradiologic Transmission.", *Radiology*; 217, 772-779,(2000).
4. Lau, C., et al., Telemedicine, Handbook of Medical Imaging, ed. Bellingham. 3, SPIE. 305-331,(2000).
5. Foos, D.H., et al. "JPEG 2000 compression of medical imagery", in *Medical Imaging 2000: Pacs desing and Evaluation: Engineering and Clinical Issues, Image Compression and Presentation*. San Diego, California, USA. (2000)
6. Digital Imaging and Communications in Medicine (DICOM). Available online at:

<http://www.xray.hmc.psu.edu/dicom/>

7. Erickson, B.J., "Irreversible Compression of Medical Images", in *White Paper from Society for Computer Applications in Radiology (SCAR)*, Department of Radiology, Mayo Foundation Rochester, Minnesota. pp. 1-9, (2000).
8. Eskicioglu, A. et al. "Quality measurement for monochrome compressed images in the past 25 years", in *Proceedings of the International Conference on Acoustics Speech (ICASSP)*, (2000).
9. Shnayderman, A., et al., A multidimensional image quality measure using Singular Value Decomposition, Department of Computer and Information Science, CUNY Brooklyn College, 2900 Bedford Avenue, Brooklyn, NY 11210.
10. Delgorge, C., Rosenberger, C., Poison, G., and Vieyres, P., "Towards a New Tool for the Evaluation of the Quality of Ultrasound Compressed Images", *IEEE Trans. on Medical Imaging*, 25 (11), (2006).
11. Wang, Z., et al., "Image Quality Assessment: From Error Visibility to Structural Similarity", *IEEE Transactions on Image Processing*, 13 (4), (2004)
12. Barret, H.H., et al., "Model Observers for assessment of image quality", *Proc. Natl. Acad. Sci. USA*, 90, 9758-9765, (1993).
13. Chakraborty, D. P. and Berbaum, K. S., "Observer studies involving detection and localization: Modeling, analysis, and validation", *Med. Phys.* 31 (8), (2004).
14. Yao J., and Barret, H. H., "Predicting human performance by a channelized Hotelling observer model", *Mathematical Methods in Medical Imaging, SPIE*, 1768, 161 - 168, (1992).
15. Zhang, Y. and Eckstein, M. P., "Evaluation of JPEG 2000 Encoder Options: Human and Model Observer Detection of Variable Signals in X-Ray Coronary Angiograms", *IEEE Trans. on Medical Imaging*, 23(5), (2004).
16. Eckstein, M. P., Bartoff, J. L., Abbey C.K., Whiting, J. S., and Bochud, F. O., "Automated computer evaluation and optimization of image compression of x-ray coronary angiograms for signal known exactly detection tasks", *Optic Express*, 11(5), (2003).
17. Zhang, Y., Pham B., and Eckstein, M. P., "Evaluation of internal noise methods for Hotelling observer models", *Med. Phys.* 34 (8), August (2007).
18. Burgess, E., and Colborne, B., *Visual Signal Detection. IV Observer Inconsistency*, *Opt Soc. Am. A*, Vol. 5, No. 4, May (1988).
19. Eckstein, M. P., Abbey C.K., and Bochud, F. O., "Visual Signal Detection in structured backgrounds. IV Figures of Merit for model performance in multiple-alternative forced-choice detection tasks with correlated responses", *J. Opt. Soc. Am. A*, Vol. 17, No. 2, February (2000).
20. Chawla, A. S., "Analyzing the effect of dose reduction on the detection of mammographic lesions using mathematical observer models", *Med Phys.*, 34 (8), pp.3385-98, (2007).
21. Hanley, J. A., McNeil, B. J., "The Meaning and Use of the Area under a Receiver Operating Characteristic (ROC) Curve". *Radiology* 143, 29-36, April (1982).
22. López de Ullibarri, Galparsoro I, Píta Fernández, S., "Curvas ROC", *Cad. Aten. Primaria*, 5(4), 229-235, (2001).