

SECURE TRANSMISION OF CODED DIGITAL IMAGES**TRANSMISION SEGURA DE IMAGENES DIGITALES CODIFICADAS****PhD. Francisco J. García Ugalde****Universidad Nacional Autónoma de México (UNAM)****Facultad de Ingeniería, División de Ingeniería Eléctrica, Edificio “Bernardo Quintana”**

2do piso, Ciudad Universitaria, Circuito Exterior, 04510 México D.F., México.

Tel.: (+52) 55-56223011, Fax: (+52) 55-5616-1855

E-mail: fgarciau@servidor.unam.mx

Abstract: In nowadays telecommunications the information channels bandwidth is an expensive resource that should take care. On the other hand, the digital images to transmit diverse types of information represent big data volumes. Before this alternative, it is required to code this type of information with the purpose of reducing the amount of bits/seg, maintaining at the same time a good quality of the information received.

Resumen: En las telecomunicaciones actuales el ancho de banda de los canales de información es un recurso costoso que se debe cuidar. Por otro lado las imágenes digitales para transmitir diversos tipos de información representan grandes volúmenes de datos. Ante esta disyuntiva se requiere codificar este tipo de información con el propósito de reducir la cantidad de bits/seg manteniendo al mismo tiempo una buena calidad de información recibida.

Keyword: Digital image coding, Digital image compression, Non equal digital image coding.

INTRODUCCION

El auge económico propiciado por la red Internet y las comunicaciones inalámbricas, ha desencadenado una actividad importante de investigación en el área de la codificación de señales multimedia (voz, audio, imagen, video). Uno de los soportes formales de la codificación fuente es la teoría de la tasa-distorsión (R-D). Considerando las propiedades de la percepción humana, la señal original y la señal decodificada mediante una estrategia con pérdidas resultarán casi idénticas para el usuario final. Para una compresión dada, la codificación fuente busca entonces llegar lo más cerca como sea posible a este punto de operación en el plano tasa-distorsión.

Por otro lado, la transmisión de la información codificada se realiza a través de un canal de comunicaciones ruidoso caracterizado por una capacidad de canal que cuantifica la tasa máxima a la cual la información se puede transmitir confiablemente, tarea conocida en comunicaciones digitales como codificación canal. En las comunicaciones punto a punto la codificación fuente y canal pueden teóricamente realizarse de una manera separada sin provocar una baja en el desempeño, para esto la teoría de la información reposa en dos suposiciones: i) el uso de bloques arbitrariamente largos; ii) la disponibilidad de recursos de cómputo arbitrariamente altos. Tales condiciones no se satisfacen en la práctica, entonces se requiere de una interacción estrecha

entre la codificación de fuente y canal. Dependiendo del número de elementos involucrados en el diseño del sistema de comunicaciones global, la codificación conjunta fuente-canal se puede realizar de diferentes maneras. Cuando sólo los dispositivos de los codificadores fuente canal se involucran, el recurso a ser asignado es la tasa total de bits entre la fuente y el codificador de canal. Si además se incluye al modulador-demodulador, entonces la restricción es la potencia de transmisión. Debido a la naturaleza probabilística del canal (en la presencia de ruido), típicamente la función costo evalúa la distorsión extremo a extremo de la copia que se recibe de la fuente, la distorsión es aleatoria y entonces deben considerarse los valores esperados de la misma, utilizando para ello la razón señal pico a ruido (PSNR). Los receptores modernos de comunicaciones están conformados por cascadas de subsistemas que realizan procesamiento de señales de manera intensiva, cada uno optimizado para realizar una tarea única. Convencionalmente, la interfase entre los subsistemas implica un paso de bits resultantes de una toma de decisión llamada “dura”, perdiéndose con esto información que ya no puede utilizarse en los subsistemas subsecuentes. Siendo por lo tanto aconsejable aprovechar esta información haciendo uso de una decisión “suave” en la decodificación. Adicionalmente, las etapas al inicio de la cadena de procesamiento no se benefician de la información resultante de las etapas finales. Para hacer la codificación conjunta fuente-canal se trabajará con un juego de codificadores, acoplados con métodos de control de errores de tasa compatible (RCPC), y así poder hacer una protección desigual contra errores (UEP) en base a la percepción.

2. CODIFICACION FUENTE

La codificación fuente se divide fundamentalmente en dos grupos: compresión sin pérdidas y compresión con pérdidas. Aunque dependiendo de la aplicación, la mayoría de los codificadores actuales son híbridos, es decir, cuentan con una etapa de codificación con pérdidas, seguida de otra sin pérdidas. En la compresión sin pérdidas (*lossless*), las imágenes reconstruidas son numéricamente idénticas a las imágenes originales. Entre los métodos que corresponden a este grupo de compresión podemos destacar los que emplean métodos estadísticos, basados en la teoría de Shannon, como son: la codificación de Huffman, la codificación aritmética y la de Lempel-Ziv. En las

imágenes médicas es indispensable la utilización de métodos de compresión sin pérdidas, con el propósito de no introducir artefactos que pudiesen falsear algún diagnóstico, sin embargo la tasa de compresión obtenida con estos métodos es menor. En la compresión con pérdidas (*lossy*), dependiendo de la tasa de compresión, se produce cierta degradación con respecto a la imagen original, a mayor compresión, mayor degradación. A costa de esta degradación, los métodos de compresión con pérdidas permiten alcanzar tasas de compresión mas elevadas. Entre los métodos que corresponden a este grupo y que permiten una transmisión progresiva de la información codificada, podemos citar: JPEG2000, EZW (*embedded zero tree wavelet*) [1], SPIHT (*set partitioning in hierarchical trees*) [2], etc. La compresión de imágenes se puede abordar también en dos dominios: el espacial de origen, o bien alguno otro, al cual se llega por medio de una transformación reversible. La idea de una transformación reversible es conservar la energía pero distribuida de otra manera, con el propósito de aprovechar esta nueva distribución de la energía para aumentar la compresión, sin por lo tanto aumentar la degradación de las imágenes comprimidas. Tal es el caso de la DCT (*discrete cosine transform*) y de la transformación *wavelet*. Particularmente esta última ha dado muy buenos resultados en el dominio del análisis y procesamiento de imágenes desde su aparición a principios de los años 80's. Su éxito se debe a su analogía con el sistema de percepción visual que trabaja a múltiples resoluciones, y en los últimos años a la obtención de algoritmos rápidos para aplicaciones en tiempo real, como el algoritmo FLWT (*Fast Lifting Wavelet Transform*) [3].

En la figura 1 se tiene el esquema de las principales etapas de un codificador que utiliza una transformación

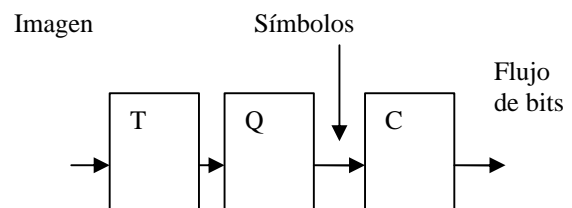


Fig. 1: Esquema de un codificador fuente híbrido

El bloque T representa la transformación que podría ser *wavelet*, Q es un cuantificador y C un codificador de entropía (Huffman). En el avance de

los codificadores por transformación las funciones *wavelet* y el estudio de sus propiedades han aportado una herramienta moderna y fundamental, que ha permitido el desarrollo de algoritmos de compresión eficientes y de aplicación en tiempo real. Estos algoritmos sacan ventajas de la existencia de dependencias significativas espaciales utilizando estructuras de datos tales como los árboles de ceros [1]. El pionero de estos algoritmos ha sido el llamado EZW (*embedded zerotree wavelet*), el cual genera una cadena binaria progresivamente ordenada en base a la importancia relativa de los bits (*embedded chain*). Se aprovecha entonces de la existencia de una correlación entre los coeficientes *wavelet* de sub-bandas diferentes, como se puede apreciar en la figura 2.

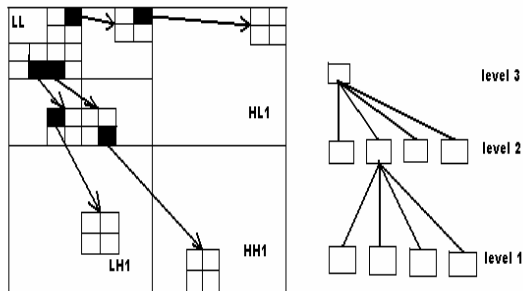


Fig. 2: Relaciones entre los coeficientes wavelet de diferentes sub-bandas y estructura de árbol asociada

Habiendo realizado una transformación *wavelet* de una imagen, se puede verificar que dado un coeficiente $c[k_1/2, k_2/2]_{j+1}$, irrelevante en la escala $j+1$, respecto a un umbral, es altamente probable que el coeficiente $c[k_1, k_2]_j$ de la escala j también sea irrelevante. Siguiendo este principio de ordenamiento, los coeficientes que conllevan la mayor parte de la información tendrán una alta prioridad durante la compresión. En el algoritmo EZW un coeficiente *wavelet* será irrelevante, cuando su valor sea inferior a un umbral predeterminado T . Este comportamiento de coeficientes irrelevantes en diferentes escalas se puede expresar analíticamente con la siguiente probabilidad condicional:

$$P\left(c[k_1, k_2]_j \leq T \mid c\left[\frac{k_1}{2}, \frac{k_2}{2}\right]_{j+1} \leq T\right) \approx 1 \quad (1)$$

Entonces, después de haber transformado la imagen utilizando funciones *wavelet*, el algoritmo EZW y de acuerdo a (1), representa los coeficientes *wavelet* de la imagen con una estructura de árbol,

donde cada raíz corresponde a un coeficiente *wavelet padre* $c[k_1/2, k_2/2]_{j+1}$, con cuatro ramas descendentes $c[k_1, k_2]_j$, como se ilustra en la figura 2. El análisis debe respetar que ningún elemento descendente puede ser analizado antes que su raíz padre. Esta restricción asegura que las sub-bandas de baja frecuencia, serán completamente escaneadas, antes que las de alta frecuencia sean procesadas. El orden de barrido conocido con el nombre de Morton, inicia en la banda de más baja frecuencia LL_N , continua en seguida con las bandas HL_N, LH_N, HH_N , luego se sigue con los coeficientes del nivel $N-1$ y así sucesivamente. En [1] se define un árbol de ceros como un "quad-tree" en el cual todos los nodos descendentes son iguales, o menores que la raíz padre, la cual a su vez es más pequeña que el umbral T . Este árbol es codificado con un solo símbolo y reconstruido por el decodificador como un "quad tree" relleno con ceros. Para que el algoritmo produzca una cadena de bits embebidos, se utiliza junto con la estructura de árbol de ceros una cuantificación por aproximaciones sucesivas (SAQ). Esta cuantificación, se relaciona directamente con la codificación de planos de bits de las magnitudes. El umbral inicial T_0 es escogido tal que, $|x_j| < 2T_0$, para todos los coeficientes transformados x_j . La SAQ aplica sucesivamente una secuencia de umbrales T_0, \dots, T_{N-1} para determinar el grado de importancia de los coeficientes. Dichos umbrales son escogidos de tal manera que $T_i = T_{i-1}/2$. Y el umbral inicial está dado por:

$$T_0 = 2 + \log(\text{MAX} |?(x,y)|) + \quad (2)$$

Donde $+^x +$ representa la parte entera de x , MAX es el valor máximo de los coeficientes $?(x,y)$ y \log es el logaritmo base 2. Dado un umbral, se determina si un coeficiente es significativo respecto a ese umbral. Se dice que un coeficiente es significativo respecto a un umbral dado, si su magnitud excede el valor del umbral. Y por otro lado, se dice que un coeficiente x es un elemento de un árbol de ceros si dicho coeficiente y todos sus descendientes son insignificantes respecto a T . Un elemento de un árbol de ceros, para un umbral T , es una raíz de un árbol de ceros, si este coeficiente no es un descendente de una raíz encontrada previamente para el mismo umbral T . Una raíz del árbol de ceros indica que la detección de coeficientes con valores insignificantes en las sub-bandas de alta frecuencia (escala baja), es completamente predecible. Un mapa que indica el resultado de una decisión ternaria (significativo positivo, significativo negativo, o insignificante), se

le llama mapa de significancia, y puede ser representado por una cadena de símbolos. Los tres elementos usados en esta representación son: a) raíz de *árbol de ceros*, b) cero aislado, c) coeficiente significativo. El algoritmo de codificación realiza dos pasos principales: dominante y subordinado. En el dominante, la imagen es escaneada y se asigna un símbolo para cada coeficiente comparado con el umbral; en el subordinado se cuantifican todos los coeficientes significativos encontrados en el paso dominante. El proceso continúa, alternando los pasos dominante y subordinado. Y el valor del umbral se va haciendo igual a la mitad antes de iniciar el siguiente paso subordinado. De esta manera se obtienen dos listas: la dominante y la subordinada. La primera contiene las coordenadas de los coeficientes que no se encontraron significativos durante un barrido de la imagen. Este escaneo se realiza de tal manera que las sub-bandas son ordenadas, y dentro de cada sub-banda, el conjunto de coeficientes son ordenados también. Este ordenamiento es conocido por el decodificador. En el codificador, la lista subordinada contiene las magnitudes de aquellos coeficientes que fueron clasificados significativos. En el decodificador la lista subordinada guarda las coordenadas de los coeficientes significativos, además de la mejor aproximación de cada coeficiente. Durante un paso dominante, los coeficientes con coordenadas en la lista dominante son comparados con el umbral T_i , con el propósito de determinar su significancia, y en caso de resultar significativo, también se codifica su signo. El algoritmo de codificación *árbol de ceros* produce una secuencia de planos de bits que corresponden a una representación binaria de la imagen previamente transformada. Los bits más significativos son codificados primero, y los bits menos significativos son codificados al final. El proceso se termina cuando el umbral haya alcanzado su valor mínimo, o cuando una función de costo previamente calculada haya sido satisfecha. La función de costo más común es aquella que cuenta el número de bits escritos en la cadena de datos comprimida y termina cuando alcanza un valor predeterminado. La función de costo puede garantizar tasas de compresión exactas. Otro tipo de función de costo se centra en el error máximo absoluto por píxel, que se tiene entre la imagen original y la imagen reconstruida; y termina cuando este error se encuentra por debajo de un margen especificado. La figura 3 muestra el resultado de una transmisión progresiva de la imagen "Lena" utilizando este algoritmo, ejecutado en tiempo real con el FLWT.



Fig. 3: Transmisión progresiva de la imagen "Lena" de izquierda a derecha y de arriba abajo, con PSNR

En procesamiento de imágenes digitales, se considera que una razón señal pico a ruido PSNR superior, o igual, a 30 [dB] es un buen resultado de codificación fuente. Donde:

$$\text{PSNR} = 10 \log_{10} \left(\frac{(255)^2}{\text{MSE}} \right) \quad (3)$$

Siendo MSE el error medio cuadrático entre la imagen original $f(x,y)$ y la imagen decodificada $F(x,y)$:

$$\text{MSE} = \sum_{\forall x,y} \frac{[f(x,y) - F(x,y)]^2}{N^2} \quad (4)$$

3. CODIFICACION CANAL

De acuerdo a la teoría matemática de comunicaciones de Shannon [4], al trabajar por debajo de la capacidad teórica del canal de información $C = W \log_2(S/N+1)$, los sistemas de comunicaciones digitales cuentan con métodos eficientes de control de errores de transmisión que permiten disminuir la probabilidad de error tanto como se desee. En la teoría de códigos, los sistemas de control de errores se dividen en dos grandes grupos, conocidos como: de bloque y convolucionales. Para codificación en tiempo real, los códigos convolucionales son generalmente más adecuados y utilizando matrices de perforación en su salida es posible adaptarlos a las condiciones variantes en ancho de banda y en potencia de ruido de los canales de comunicaciones.

Teniendo entonces un mecanismo de control de errores llamado “códigos convolucionales perforados de tasa compatible” (RCPC) [5]. Así a partir de un código original de tasa $r=1/2$, con las matrices de perforación es posible obtener códigos de tasa $r=(n)/(n+1)$.

En la figura 4 se muestra esquemáticamente un codificador convolucional perforado de tasa final $r=2/3$, obtenido a partir de un código original de tasa $r=1/2$ y una matriz de perforación P , dada por:

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad (5)$$

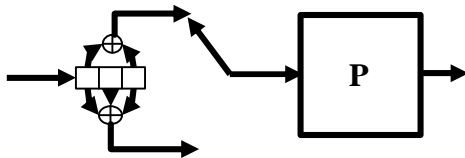


Fig. 4: Codificador convolucional perforado de tasa $r=2/3$

En este ejemplo, los polinomios que definen el código original son:

$$g_1(x) = x^2+1 \text{ y } g_2(x) = x^2+x+1$$

La perforación hace que el elemento nulo de la matriz P cancela el bit de salida del codificador original en esa posición, generando en los tiempos sucesivos t_1 y t_2 , después de la perforación, 3 bits de salida por cada 2 bits de entrada.

La gran ventaja de un código perforado es que el algoritmo de decodificación permanece casi inalterado con respecto al algoritmo de decodificación del código original, y por consecuencia se tiene un tiempo de decodificación igual en ambos casos, resultado de gran interés en aplicaciones en tiempo real.

Las figuras 5 y 6 muestran respectivamente el diagrama de transición de estados de un enrejado “trellis” del código original de tasa $r=1/2$, y del código perforado resultante de tasa $r=2/3$. En la figura 6, debido a la perforación, los símbolos marcados con X no se toman en cuenta en la decodificación para el cálculo de la distancia entre la secuencia recibida y la secuencia generada en el enrejado.

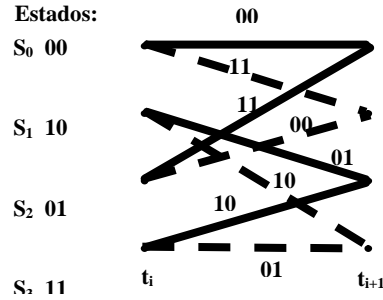


Fig. 5: Diagrama de transición de estados de un código original de tasa $r=1/2$

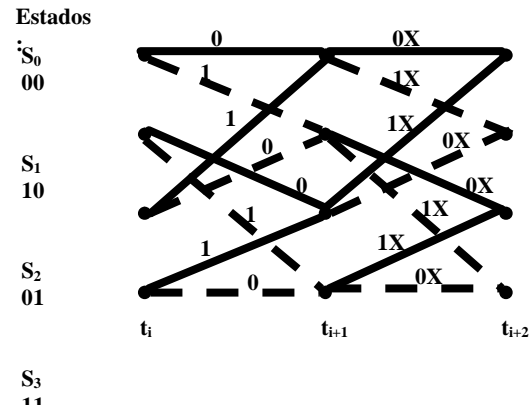


Fig. 6: Diagrama de transición de estados de un código perforado resultante de tasa $r=2/3$

Otra ventaja de los códigos convolucionales es que se puede lograr de manera sencilla un acoplamiento entre el demodulador digital y el decodificador, permitiendo aprovechar la información propia de una decisión llamada suave “soft” del demodulador. En términos generales esto permite ganar 2 [dB] en la razón señal pico a ruido del receptor, utilizando para ello, en el algoritmo de decodificación de Viterbi como medida de comparación entre la secuencia recibida y la secuencia generada en el enrejado, la distancia euclidiana (decisión suave), en vez de la distancia de Hamming (decisión dura). Para una aplicación, considerando la imagen original “Lena” de dimensiones 256x256 y 8 bits/píxel de la figura 7, se aplica una protección desigual (UEP) a las diferentes 8 capas binarias que conforman la imagen, protegiendo de esta manera con mayor capacidad de corrección de errores, las capas binarias de los 4 bits más significativos. La figura 10 muestra el resultado de una decodificación con decisión dura “hard”, en donde para codificar las 4 capas binarias más significativas, se ha utilizado el

código convolucional de tasa $r=1/2$ de polinomios $g_1(x) = x^3+x^2+1$ y $g_2(x) = x^3+x^2+x+1$, con longitud de restricción $K=4$ y distancia libre $d_f=6$, y para codificar las 4 capas binarias menos significativas el código convolucional de tasa $r=1/2$ de polinomios $g_1(x) = x^2+1$ y $g_2(x) = x^2+x+1$, con longitud de restricción $K=3$ y distancia libre $d_f=5$.



Fig. 7: Imagen "Lena" original de dimensiones 256x256 y 8 bits/píxel

Se puede observar una mejora significativa en la figura 10, con respecto a la figura 8, donde no hay control de errores.



Fig. 8: Imagen con ruido aditivo, sin control de errores, PSNR=21.1 [dB]

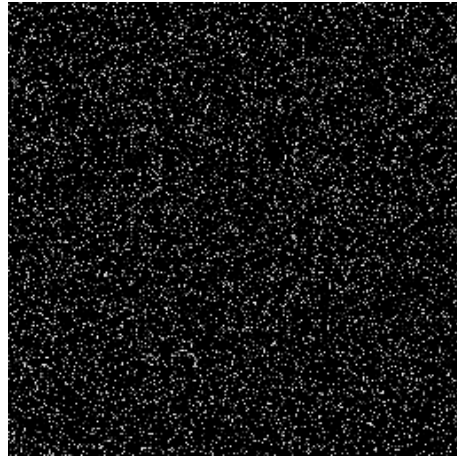


Fig. 9: Diferencia amplificada por un factor de 30 entre la imagen original y la imagen con ruido, sin control de errores



Fig. 10: Imagen decodificada con decisión dura, PSNR=28.35 [dB]

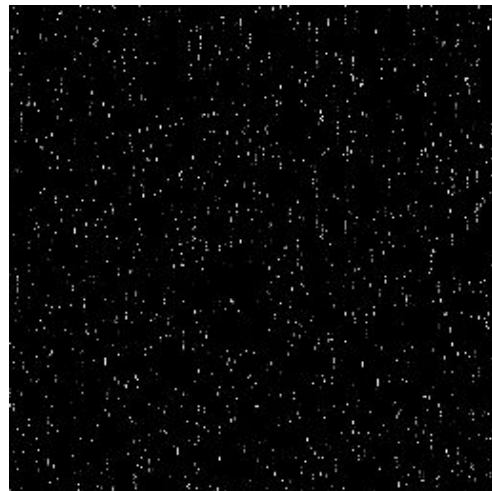


Fig. 11: Diferencia amplificada por un factor de 30 entre la imagen original y la decodificada con decisión dura

Para poder corroborar la mejora obtenida al utilizar una decisión *suave*, se muestra el resultado de la decodificación en la figura 12, donde se han utilizado los mismos códigos que en el caso anterior. La mejora es significativa.



Fig. 12: Imagen decodificada con decisión suave,
PSNR=37.9 [dB]



Fig. 13: Diferencia amplificada por un factor de
30 entre la imagen original y la decodificada con
decisión suave

4. CONCLUSIONES

Considerando la teoría de la codificación de Shannon, en este trabajo se han presentado métodos de codificación fuente y canal aplicados a imágenes digitales. En la codificación fuente se ha utilizado un codificador basado en la transformada *wavelet*, la cual desde su introducción en los años 80's ha demostrado su utilidad en el diseño de codificadores con propiedades similares al

funcionamiento del sistema de percepción visual, permitiendo alcanzar altas tasas de compresión sin degradaciones perceptibles en las imágenes obtenidas después de la descompresión. Explotando además las propiedades de los coeficientes generados por la transformación, con el algoritmo EZW es posible obtener codificadores con transmisión progresiva de información.

Esta característica es importante en aplicaciones donde el ancho de banda asignado para la transmisión, es variable, como es el caso de las redes de tipo Internet. En las simulaciones del codificador se ha utilizado un algoritmo rápido para la transformada *wavelet* llamado FLWT, de manera que se ha podido simular la operación del codificador en tiempo real. Los resultados de la transmisión progresiva se muestran en la figura 3.

En la codificación canal para el control de errores de transmisión se ha utilizado una estrategia de protección desigual UEP con códigos convolucionales, donde los bits más significativos de cada píxel de la imagen se protegen con un código de mayor desempeño en términos de control de errores. También se ha usado la decisión suave "*soft*" del demodulador para optimizar la cantidad de información transferida de éste al decodificador canal y aumentar así el desempeño de la decodificación.

El resultado se puede apreciar comparando las imágenes mostradas en las figuras 11 (decisión dura "*hard*") y 13 (decisión suave "*soft*") donde se han utilizado los mismos códigos, y el control de errores es mejorado cuando se utiliza la decisión suave "*soft*".

En este trabajo se han probado métodos de codificación fuente y canal en imágenes digitales. Mostrando un algoritmo y sus propiedades en la transmisión progresiva. En cuanto a la codificación canal, se ha demostrado la utilidad de una protección desigual UEP y de una decisión suave "*soft*" en la etapa de decodificación.

Los resultados ilustran la conveniencia de las estructuras de los codificadores utilizados.

AGRADECIMIENTOS

La presente investigación fue desarrollada con apoyo de los proyectos CONACyT 14069 y PAPIIT-IN101305.

REFERENCIAS

- Hagenauer, J. “*Rate-compatible punctured convolutional codes (RCPC codes) and their applications*”. IEEE Transactions on Communications, 1988, vol. 36, issue 4, pp. 389-400.
- Said, A. and Pearlman, W.A. “*A new, fast, and efficient image codec based on set partitioning in hierarchical trees*”. IEEE Transactions on Circuits and Systems for Video Technology, 1996, vol. 6, issue 3, pp. 243-250.
- Shannon, C. E. “*A mathematical theory of communication*”. Bell Systems Technical Journal, 1948, vol. 27, issue July, pp. 379-423 and issue Oct., pp. 623-656.
- Shapiro, J.M. “*Embedded image coding using zerotrees of wavelet coefficients*”. IEEE Transactions on Signal Processing, 1993, vol. 41, issue 12, pp. 3445-3462.
- Sweldens, W. “*The lifting scheme: A new philosophy in biorthogonal wavelet constructions*”. Wavelet Applications in Signal and Image Processing III, Laine, A. F. and Unser, M. Eds. New York SPIE, 1995, vol. 2569, pp. 68-79.