

COMPRESION DE VIDEO VS. COMPRESION DE ANCHO DE BANDA

SR. JIMMY RUIZ PROCEL

Egresado, F.I.E. Electrónica y
Telecomunicaciones.
Escuela Politécnica Nacional, 1995

ING. TANIA PEREZ

Profesor Principal F.I.E.
Escuela Politécnica Nacional

RESUMEN

Este trabajo realiza un análisis de cómo las técnicas de compresión de imágenes al actuar sobre una imagen modelo, varían su calidad en función de una reducción de sus características estadísticas y espaciales. La herramienta a utilizar es un determinado Software de compresión, dentro del cual se podrán cambiar algunos de sus parámetros. Se muestra también una comparación de dos técnicas de compresión para imágenes estáticas, tomando en cuenta ciertos criterios tales como la sensibilidad psico-visual humana y la capacidad de canal.

ABSTRACT

This paper presents an analysis of the changes occurred on a model image quality as a function of the statistical and spatial features reduction when image compression techniques are applied. A specific compression Software is employed and changes of some parameters could be permitted. Under certain criterion such as human psychovisual sensibility and the channel capacity, comparison of two static images compression techniques are presented too.

INTRODUCCION

En telecomunicaciones, debido a limitaciones de los sistemas, algunos de los parámetros de la comunicación de datos deben ser restringidos; tal es el caso del ancho de banda de transmisión. Este, en principio, debe ser de una "dimensión" fija tal que el

diseñador que lo vaya a utilizar deba restringirse a esa medida. Existe una gran disparidad entre una tasa de datos digitales (datos MIC) y la capacidad de canal requerida por esta cantidad de datos. Surge entonces la idea de la compresión (reducción) de datos. Existe la desventaja de los múltiples formatos (velocidades de transmisión, anchos de banda, niveles de calidad, etc.) en los que se presentan estos datos. Así para voz: monoaural, estéreo, duplo canal, etc.

En video: Formatos 4:2:2, 4:4:4:, 4:2:0; relación de barrido 2:1, 1:1; resolución, etc. Si lo que se va a transmitir son los pixeles digitalizados de una imagen, ésta puede comprimirse aprovechando la gran redundancia de los datos, su alta correlación, y la baja sensibilidad a las distorsiones del sistema visual humano. Una tasa de datos (comprimidos) transmitidos por unidad de tiempo determina el ancho de banda necesario del canal; entonces se establece un compromiso entre reducción de ancho de banda y tasa de error de los datos (calidad de la imagen): A mayor compresión, menor ancho de banda, pero también menor calidad de imagen. En la codificación de video, los métodos básicos de reducción de datos se pueden representar con el diagrama de flujo que se presenta en la Figura 1.

El interés por la compresión de imágenes data desde hace 25 años. Las investigaciones en este campo se desarrollaron con los métodos analógicos de reducción del ancho de banda de transmisión del video (*compresión de ancho de banda*). La llegada del computador digital y el subsiguiente desa-

Correspondencia a: Ing. TANIA PEREZ,
E.P.N., F.I.E. Tlf. 567750, 507142,
QUITO-ECUADOR.

rollo en los circuitos integrados hizo que el proceso de la compresión vaya de lo analógico a lo digital.

Las técnicas de compresión caen dentro de dos amplias categorías: Las que **preservan la información** y las que **producen pérdidas**. Las primeras son muy útiles en aplicaciones tales como el almacenamiento de archivos médicos o legales. La compresión con pérdidas es útil en radiodifusión de señales de TV, video conferencias y transmisión de facsímil en las cuales es aceptable una cierta cantidad de errores a cambio de una tasa mayor de compresión.

abstracto pero sí una entidad matemáticamente cuantificable. Si n_1 y n_2 denotan el número de unidades portadoras de información en dos arreglos de datos que representan a la misma información, la **Redundancia relativa** R_D de los datos del primer arreglo puede definirse como:[1]

$$R_D = 1 - \frac{1}{C_R} \quad (1)$$

donde C_R , comúnmente llamada la **Tasa de compresión**, es:[1]

$$C_R = \frac{n_1}{n_2} \quad (2)$$

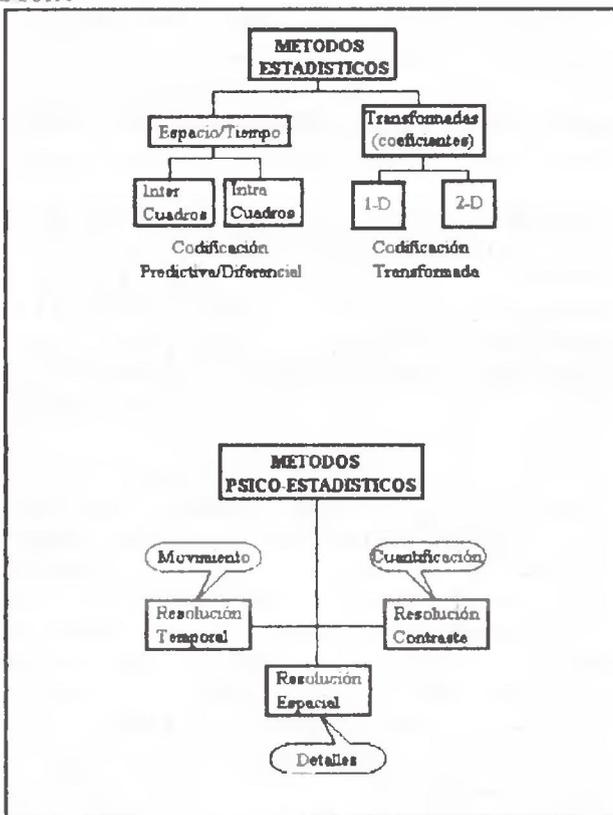


Figura 1: Métodos de Reducción.

FUNDAMENTOS

Se debe dar una clara distinción entre **datos e información**. No son sinónimos, de hecho los datos son los medios en los que se transporta a la información. Se dice que un grupo de datos que no proveen de información relevante o simplemente resultan ser datos ya conocidos de un total de información, son datos **redundantes**.

La redundancia en los datos es el tópico central en la compresión de imágenes digitales. No es un concepto

Para el caso que $n_2 = n_1$, $C_R = 1$ y $R_D = 0$; esto indica que la primera representación de la información no contiene datos redundantes. Cuando $n_2 \ll n_1$, $C_R \rightarrow \infty$ y $R_D \rightarrow 1$, esto implica una significativa compresión y una alta cantidad de datos redundantes. En el caso de que $n_2 \gg n_1$, $C_R \rightarrow 0$ y $R_D \rightarrow -\infty$, el segundo arreglo contiene muchos más datos que el original; este es el caso de expansión de los datos que, por supuesto es el menos deseado. En general, C_R y R_D permanecen dentro de los intervalos abiertos $]0, +\infty[$ y $] -\infty, 1[$, respectivamente. Una tasa de compresión práctica, tal como 10 (o 10:1) significa que el primer arreglo tiene 10 unidades portadoras de información (por ejemplo, bits) por cada unidad en el segundo arreglo de datos comprimidos. La correspondiente redundancia de 0.9 indica que el 90% de los datos del primer arreglo es redundante. En la compresión digital de imágenes, se pueden identificar y explotar tres tipos básicos de redundancias: Redundancia **de codificación**, redundancia **interpíxel**, y la redundancia **psicovisual**. La compresión se alcanza cuando una o más de estas redundancias son eliminadas.

En la primera, la redundancia se presenta cuando al codificar con palabras código de longitud fija un arreglo de datos, los bits componentes de la palabra no se "llenen" de manera uniforme, así por ejemplo para un arreglo de 256 datos que utilizan palabras de 8 bits, puede haber 2 datos con el valor 255 (11111111), 6 datos con el valor

$$SNR_{ms} = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f'(x, y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f'(x, y) - f(x, y)]^2} \quad (6)$$

Aunque este criterio objetivo de fidelidad ofrece un mecanismo simple y conveniente para evaluar la pérdida de información, las imágenes descomprimidas son vistas por seres humanos. Consecuentemente, la calidad de imagen medida por la evaluación subjetiva de un observador humano a menudo es más apropiada. Esto se logra al mostrar una típica imagen descomprimida a un grupo de espectadores y se promedian sus evaluaciones. Las evaluaciones se pueden hacer utilizando una escala absoluta o por medio de comparaciones directas entre $f(x, y)$ y $f'(x, y)$. La Tabla I muestra una posible escala absoluta. Las comparaciones directas se pueden hacer con una escala tal como $\{-3, -2, -1, 0, 1, 2, 3\}$ para representar a las evaluaciones subjetivas $\{\text{deficiente, mala, regular, la misma, un poco mejor, mejor, excelente}\}$, respectivamente.

Valor	Rango	Descripción
1	Excelente	Imagen con calidad extremadamente alta.
2	Fino	Imagen de alta calidad, la interferencia no objetable
3	Pasable	Imagen de calidad aceptable, interferencia no es objetable
4	Marginal	Imagen de pobre calidad, la interferencia objetable
5	Inferior	Una imagen muy pobre, pero que puede distinguirse, una interferencia objetable siempre está presente.
6	No útil	Una imagen tan mala que no se puede distinguir.

TABLA I: Escala de evaluaciones subjetivas.

RELACION COMPRESION-ANCHO DE BANDA

La compresión de video está ya en pleno uso comercial. No solamente que se han establecido estándares (JPEG, MPEG, CCITT, ISO), sino que también los fabricantes han introducido productos de video digital que se encuentran en aplicaciones de muy alta calidad.

El análisis de una imagen se realiza mediante el "barrido" horizontal (línea por línea) y el barrido vertical (cuadro por cuadro). La imagen está dividida en un determinado número de líneas horizontales que contienen información. El barrido de una línea se lo hace punto por punto, en forma discreta, así la imagen se la divide en un grupo de líneas horizontales que contienen información, el barrido de una línea, se lo hace identificando intensidades luminosas, punto por punto, en forma discreta, así un punto tendría un tamaño físico muy reducido, con el fin de captar la mayor cantidad de información de ese pequeño pedazo de imagen que a menudo es llamado "Pixel". Debido a que el barrido es discreto, existe un cierto número de pixeles por línea barrida (L), y un número de M líneas por cuadro, teniéndose por lo tanto [11]

$$\# \text{total de puntos que conforman un cuadro} = (L)(M) = N$$

En realidad no interesa sobremanera la transmisión, o el almacenamiento de un solo cuadro de imagen, sino, de una serie de cuadros, que se representan uno a continuación de otro para dar la sensación de movimiento de la imagen; dicho movimiento continuo de imágenes tiene un ritmo (v) de aparición de cuadro por unidad de tiempo, lo que en conjunto, con el valor de N hallado anteriormente, da un resultado interesante, el "Ritmo de aparición de un pixel", o los pixeles por unidad de tiempo: [11]

$$R = N \cdot v \quad (\text{pixeles/seg}) \quad (7)$$

La necesidad de reducir el ancho de banda de imágenes digitalizadas se debe a los excesivos requerimientos de ancho de banda de este tipo de imágenes y la continua demanda de capacidad de los canales de comunicación. De

240 (11101100), 3 datos con el valor 200 (11001000), 29 datos con el valor 33 (00100001) y 216 datos con el valor 1 (00000001), esto da una codificación ineficiente; por el contrario, si dependiendo de la longitud (número de bits) necesaria para cada dato se realizaría la codificación (Codificación por Longitud Variable o CLV), se lograría optimizar la codificación y reducir la redundancia original.

La redundancia interpixel, llamada también redundancia *geométrica, espacial*, o *intertrama*, aparece luego del proceso de digitalización (en el proceso de Modulación por Impulsos Codificados MIC), y resulta de que en el paso muestreo-cuantificación, se asignan a dos o más muestras el mismo nivel cuando entre estas no existe mucha diferencia de amplitud respecto del tamaño de paso de cuantificación. Se explota esta redundancia al codificarla eficientemente por medio de técnicas diferenciales (que codifican no a las palabras sino a su diferencia con respecto a la adyacente) como MIC Diferencial o la Modulación Delta, seguida de una CLV pues existirán muchas palabras de corta longitud, y muy pocas de longitud extensa.

El ojo humano no responde con igual sensibilidad a toda la información visual. Cierta información es de relativamente menor importancia que otra información en el procesamiento visual normal. Esta información se dice es *Psicovisualmente redundante*, y se la puede eliminar sin una degradación significativa de la calidad de la imagen percibida. En general, un observador busca características distinguibles tales como bordes o regiones de textura uniforme, y mentalmente las combina en agrupamientos reconocibles. El cerebro entonces correlaciona a estos agrupamientos y a priori los reconoce a fin de completar el proceso de interpretación de imágenes. A la reducción de esta redundancia se le conoce muchas veces como *cuantificación* y como este proceso es irreversible (la información visual se pierde), da como resultado una compresión de datos con pérdida. El remover datos psicovisualmente redundantes resulta entonces en una pérdida de información real o visualmente cuantitativa. Debi-

do a que la información de interés se podría perder, es deseable contar con un medio repetidor o reproductor de la naturaleza y extensión de la información perdida. Dos clases generales de criterios se utilizan como las bases de tales valoraciones: (1) un criterio de fidelidad objetiva y (2) un criterio de fidelidad subjetiva. Cuando el nivel de pérdida de información se puede expresar como una función de la imagen original y la imagen de salida comprimida y subsecuentemente descomprimida, se dice que está basado en un *criterio objetivo de fidelidad*. Un buen ejemplo es el *error raíz cuadrático medio* entre una imagen de entrada y una de salida. Sea que $f(x,y)$ represente a un pixel de la imagen de entrada y sea que $f'(x,y)$ denote a una estimación o aproximación de $f(x,y)$ que resulta de la compresión y subsecuente descompresión de la imagen de entrada. Para cualquier valor de x y de y , el error $e(x,y)$ entre $f(x,y)$ y $f'(x,y)$ se puede definir como:[1]

$$e(x,y) = f'(x,y) - f(x,y) \quad (3)$$

así que el error total entre las dos imágenes es:[1]

$$\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f'(x,y) - f(x,y)] \quad (4)$$

donde las imágenes son de tamaño $M \times N$ pixeles. El error raíz-cuadrático-medio, e_{rms} , entre $f(x,y)$ y $f'(x,y)$ entonces es la raíz cuadrada del error cuadrático promedio sobre el arreglo $M \times N$, o:[1]

$$e_{rms} = \sqrt{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f'(x,y) - f(x,y)]^2} \quad (5)$$

Un criterio objetivo de fidelidad es la *relación señal a ruido cuadrática media* de la imagen comprimida-descomprimida. Si se considera a $f'(x,y)$ como la suma de la imagen original $f(x,y)$ y una señal ruido $e(x,y)$, la relación señal a ruido cuadrática media de la imagen de salida, denotada SNR_{dB} es:[1]

otro lado, el procesamiento, el almacenamiento, la transmisión, y la difusión de imágenes en forma digital, ofrece muchas ventajas distintivas con respecto a estas operaciones realizadas en la forma analógica. Estas son la flexibilidad del procesamiento, la transmisión con bajo número de errores, precisión controlada, registro preciso, facilidad de encriptamiento y codificación, facilidad de almacenamiento y recuperación, así como también la compatibilidad con redes y computadores digitales. Muchas veces lo que se hace es buscar la relación señal a ruido de cuantificación y el ancho de banda requerido para una transmisión digital.

Se sabe que un mensaje de ancho de banda W (Hz) requiere de una tasa mínima de muestreo de $2.W$ (Hz). En cada muestra de la señal representada por una palabra código de n -bits, la duración del bit, T tiene un valor de:[3]

$$T = \frac{1}{2 \cdot n \cdot W} \quad (\text{seg}) \quad (8)$$

El ancho de banda del canal requerido B para transmitir un pulso de esta duración es:[3]

$$B = K \cdot n \cdot W, \quad K > 2 \quad (9)$$

Asumiendo que se usa una palabra código binaria de n bits, es posible definir el número de niveles L para un cuantificador del tipo "Mid-Tread":

$$L = 2^n - 1 \quad (10)$$

Ahora, para un sistema MIC es posible obtener que la relación señal a ruido (SNR) es:[3]

$$SNR = \frac{3}{2} L^2 \quad (11)$$

con lo que la SNR en función del número n de bits está dada por la relación (12) para $n \geq 8$. [3]

$$SNR = \frac{3}{2} (2^n)^2 \quad (12)$$

Introduciendo la ecuación (9), la SNR en función del ancho de banda W queda:[3]

$$SNR = \frac{3}{2} \left(4 \frac{B}{K \cdot W} \right) \quad (13)$$

Esta última expresión muestra que un sistema MIC es capaz de mejorar la relación Señal/Ruido exponencialmente con la relación de expansión de ancho de banda, B/W .

El siguiente cuadro ([2]) muestra algunas diferencias entre los parámetros de codificación según el tipo de señal de audio que se use.

	Rango Hz	f _{muestreo} KHz	b/muestra	Kbps
Voz(Telefonía)	300-3400	8	8	64
Voz(Mejorada)	50-7000	16	14	224
Audio	10-20000	48	16	768

Para un medio de almacenamiento como el disco compacto(CD) se tiene por ejemplo una frecuencia de muestreo de 44.1KHz, una tasa de audio de 1.41M-bps, y un "over-head" de 2.91Mbps para bits de control de errores y sincronismo, se obtiene una tasa total de 4.32Mbps. Las capacidades de canal para voz van de 2.4Kbps a 64Kbps; Voz mejorada: 64Kbps; Audio estéreo: 128Kbps, 256Kbps, o 384Kbps.

En lo que respecta a la codificación de material de video la Recomendación 601 del CCIR(Comité Consultivo Internacional de Radio) para EQTV(Televisión de Calidad Mejorada) y EDI(Imágenes Digitales Mejoradas) según el patrón 4:2:2, con 8 bits por muestra y codificación MIC da la siguiente tasa de datos:[2]

$$\text{Tasa de datos} = (13.5 + 2 \times 6.75) \times 8 = 216 \text{Mbps}$$

Sin embargo la capacidad de canal prevista es de:[2]

$$35 \text{Mbps}$$

Por tanto se debe reducir la tasa por lo menos en:

$$6.17 : 1$$

Las tasas de datos(en Mbps) requeridas por algunos formatos de Televisión de Alta Definición(HDTV) pueden tabularse como se muestra a continuación[2]

CCIR-601	Europa	Japón	U.S.A.
216	864,1152,2304	778	864,1205

Y las capacidades de canal suministradas por el estándar de compresión para imágenes en total movimiento MPEG-2 son: [2]

CCIR-601/PAL:	3Mbps-5Mbps
CCIR-601/Componentes:	8Mbps-10Mbps
HDTV:	20Mbps-40Mbps

Por otro lado, las tasas de datos(en Mbps) requeridas por algunos formatos de Televisión de Menor Resolución pueden tabularse también como sigue: [2]

CIF(Formato Intermedio Común):	37.3Mbps
QCIF(Formato Cuádruple Intermedio Común):	9.3Mbps
SIF(Formato Entrante de Fuente):	30.4Mbps

Y las capacidades de canal previstas son: [2]

CIF y QCIF:	64Kbps-1Mbps
SIF:	1.5Mbps

Mientras que el estándar de compresión para imágenes en total movimiento(baja resolución) MPEG-1 podría suministrar: [2]

$$1856\text{Kbps(VIDEO)}+128\text{Kbps(AUDIO)}=31\times 64\text{Kbps} = 1984\text{Kbps}$$

CODIFICACION POR TRUNCAMIENTO DE BLOQUE

Esta técnica realiza una compresión de código al hacer una post-codificación por medio de una técnica de longitud variable. La imagen de trabajo se la divide en bloques de $U \times U$ pixeles y cada bloque se codifica por separado. Este es un método muy interesante de compresión que utiliza una codificación no de los pixeles de una imagen sino de los valores estadísticos de aquella. Básicamente lo que se hace es calcular para cada bloque de pixeles su primer momento (media) y su varianza, seguido luego de una cuantificación de todos los pixeles para obtener un MAPA binario del bloque constituido de 1's y 0's según el valor del ponderado del pixel (que es una medida del valor del pixel según dos parámetros a y b que se explicarán luego)

que pueden estar sobre o debajo del valor medio. En la práctica lo que el receptor-descompresor necesita para reconstruir la imagen original es la palabra código de la varianza, la del primer momento, y el mapa binario. En esta simulación se asignan 8 bits para la varianza, 8 bits para el valor medio, y $U \times U$ bits para el mapa.

Sea $m=U^2$ y sean X_1, X_2, \dots, X_m los valores de los pixeles en un bloque del cuadro original. Entonces el primer y segundo momentos y la varianza de la muestra son, respectivamente: [8]

$$\bar{X} = \frac{1}{m} \sum_{i=1}^m X_i \quad (14)$$

$$\overline{X^2} = \frac{1}{m} \sum_{i=1}^m X_i^2 \quad (15)$$

$$\sigma^2 = \overline{X^2} - (\bar{X})^2 \quad (16)$$

Como con el diseño de cualquier cuantificador binario, se halla un umbral X_u y dos niveles de salida, a y b , tales que: [8]

$$\begin{aligned} \text{Si } X_i \geq X_u & \quad \text{Salida} = b \\ \text{Si } X_i < X_u & \quad \text{Salida} = a \end{aligned} \quad (17)$$

para $i=1, 2, \dots, m$.

Para el primer cuantificador, se coloca a $X_u = \bar{X}$, este razonable criterio se puede modificar para mejorar la eficacia de la codificación. Los niveles de salida a y b se hallan de las siguientes ecuaciones:

Sea q =número de X_i 's mayores que X_u , entonces para conservar \bar{X} y $\overline{X^2}$: [8]

$$m\bar{X} = (m-q)a + qb; \text{ y,}$$

$$m\overline{X^2} = (m-q)a^2 + qb^2$$

Resolviendo para a y b : [8]

$$a = \bar{X} - \sigma \sqrt{\frac{q}{m-q}} \quad (18)$$

$$b = \bar{X} + \sigma \sqrt{\frac{m-q}{q}} \quad (19)$$

Cada bloque es luego descrito por los valores de X , σ y un plano binario UxU consistente de 1's y 0's que indican si los pixeles están sobre o bajo X , respectivamente.

El receptor reconstruirá la imagen bloque al calcular a y b con (18) y (19) y asignando estos valores de pixeles de acuerdo con el código en el plano binario. Debido a que los cálculos son relativamente simples y el dato a almacenar, pequeño; CTB es fácil de implementar en un circuito integrado[15].

DESCRIPCION DEL PROGRAMA COMPUTACIONAL UTILIZADO

Debido a su capacidad de trabajo en forma de "Procedimientos", depurador incorporado; y, por tener una característica de lenguaje "Estructurado", se ha escogido al Lenguaje de Programación PASCAL versión 7.0 como herramienta informática de implementación de los algoritmos de compresión a simularse. Se establecieron cinco bloques auxiliares y de esta forma la ejecución total de la simulación consta del programa principal:

COMPIMAG.PAS

y las seis unidades:

VARIOS.TPU
VARIABLE.TPU
INICIAL.TPU
PROCESO.TPU
CUADROS.TPU
ANALISIS.TPU

El programa trabaja bajo entorno DOS (Disk Operating System) y es capaz de manejar hasta 1Gbyte de memoria en pilas de 64Kbytes para pseudo-código o para datos.

EJECUCION

A fin de utilizar una imagen de interés para ejecutar la aplicación de la técnica CTB de compresión se va a utilizar una imagen tipo de 256 tonalidades de gris que es típica en los anales de la IEEE para procesamiento de imágenes: LENA.

Como se puede notar, la imagen tiene ciertas características importantes: bordes(que significan altas frecuen-

cias), pequeños detalles en donde podría perderse más información, grandes zonas de una tonalidad definida en donde algún error de decodificación se visualizaría como pequeñas manchas, y ciertas zonas de gran intensidad luminosa, lo que significa palabras binarias muy grandes.



Imagen de prueba: LENA

Aunque esta imagen originalmente tiene dimensiones de 256x256 pixeles, para la simulación solo se requiere de 256x240 pixeles, esto para que se ajuste a las medidas de la pantalla de trabajo del programa. Pese a que existen monitores de muy alta resolución y de varias tonalidades de un solo color, los tonos grisáceos de la imagen se re-escalan a fin de poder visualizar más cómodamente a la imagen, sin embargo todos los cálculos que se realizan se los hace en base a la imagen original, es decir con sus 256 tonalidades de gris. La ejecución se realiza variando el tamaño del bloque de cálculo desde 2x2 pixeles hasta 128x128 pixeles los resultados a obtener son:

Tamaño del archivo comprimido: Así se mide el nivel de compresión de una manera aproximada.

Varianza: Un equivalente AC de la potencia de la imagen.

Media: Un equivalente DC de la potencia de la imagen.

Valor cuadrático medio: Un equivalente de la potencia total de la imagen.

Bits/pixel: El nivel de compresión.

Error cuadrático medio: Potencia del ruido de codificación.

Relación señal/ruido(dB): Medida de la calidad "analógica" de la imagen.

BER promedio: Medida de la calidad "digital" de la imagen.

En el ANEXO-I se presentan las gráficas que relacionan los valores estadísticos mas importantes de la codificación de la imagen en función del tamaño del bloque expresado en número total de pixeles tanto para el método CTB como para el método MICD (Técnica predictiva) en presencia de ruido; esta última técnica se la emplea con el propósito de compararla con CTB y así realizar también una confrontación de métodos de compresión. De estos resultados se puede concluir que al incrementar el tamaño del bloque ocurren los siguientes efectos:

- El tamaño del archivo comprimido se reduce.
- La potencia AC y la DC se reducen debido a que los valores de pixeles conformadores de cada bloque de la imagen se cuantifican a un solo valor: la media de cada bloque.
- La tasa bits/pixel se reduce bordeando el 1bit/pixel.
- La potencia de ruido aumenta.
- La relación Señal/Ruido disminuye debido a que cada vez el bloque es más grande y se puede llenar con apenas dos valores (a y b).
- El BER promedio aumenta.

Y lo más notable, la calidad visual de la imagen decrece. Las imágenes obtenidas para bloques de 2x2, 8x8, 16x16, 32x32, 64x64, y 128x128 pixeles se muestran a continuación como imágenes (a), (b), (c), (d), (e), y (f).



(a)



(b)



(c)



(d)



(e)



(f)

CONCLUSIONES

(*) En este estudio, se ha descrito como actúan las técnicas de reducción de la velocidad binaria (bps) y el ancho de banda (Hz). Se ha visto sus aplicaciones actuales y su desempeño. Y se ha aclarado que la mejor forma de llegar a un entendimiento entre la velocidad binaria y la calidad de la imagen se logrará solo al seleccionar con mucho cuidado los parámetros de compresión según los requerimientos de calidad de imagen y de la aplicación particular.

(*) Se ha dividido los sistemas de reducción de redundancia en tres clases: (1) Los que optimizan la longitud de las palabras código, (2) Los que aprovechan el alto grado de correlación entre muestras codificadas, y (3) Los que toman ventaja de la poca sensibilidad visual del ojo humano.

(*) Los métodos de compresión se basan tanto en la redundancia de los datos de la imagen digitalizada, como en la no linealidad de la visión humana. Explotan la correlación en el espacio para imágenes estáticas y la correlación en el espacio y el tiempo para las imágenes en movimiento.

(*) La compresión en el **ESPACIO** se conoce como compresión **INTRA-TRAMA**, mientras que la compresión en el **TIEMPO** se la llama compresión **INTER-TRAMA**.

(*) Generalmente, los métodos que alcanzan altas tasas de compresión (de 10:1 a 50:1 para imágenes estáticas y de 50:1 a 100:1 para imágenes móviles) son con pérdidas, debido a que los datos reconstruidos no son idénticos al original. Los métodos que no producen errores existen, pero sus tasas de compresión son muy bajas, quizá no mayores que 3:1. Tales técnicas son

utilizadas solo en aplicaciones de alta sensibilidad como imágenes médicas o archivos legales.

(*) Los algoritmos con pérdidas generalmente explotan aspectos del sistema visual humano. Por ejemplo, el ojo es mucho más perceptivo a finos detalles en la señal de luminancia (o brillo) que en las señales de crominancia (o color). Consecuentemente, la señal de luminancia se muestrea usualmente a altas resoluciones espaciales (Por ejemplo, en la difusión de televisión de calidad, la resolución digital de la señal de luminancia muestreada es de 720x480 pixeles, mientras que para las señales de color esta puede ser de solamente 360x240 pixeles). Generalmente a la representación codificada (o comprimida) de la señal de luminancia se le asigna más bits (un rango dinámico mayor) que a las señales de crominancia.

(*) La técnica CTB es una opción sencilla que se ha descrito y que se puede modificar para mejorar su desempeño (CTB híbrida, mayor número de momentos). Las imágenes reconstruidas resultantes tienen artefactos muy distintos de los producidos por otras técnicas, pero las imágenes resultantes son comparables y de alguna forma superiores a aquellas producidas por técnicas más sofisticadas. Este método produce una imagen codificada que es más robusta en presencia de errores en el canal y así se requiere muy poca protección contra errores.

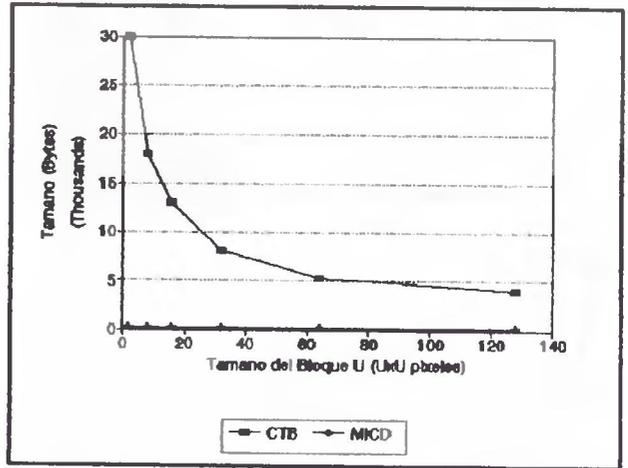
(*) En cuestiones de programación para el campo de la simulación o la implementación circuital, es recomendable trabajar con los procesos de cálculo en forma de bloques. Es decir que si lo que se quiere es construir un circuito impreso que codifique y decodifique imágenes por las técnicas estudiadas (u otras más), se acelera el proceso si el tratamiento se lo hace por secciones de imagen (bloques), mucho mejor si estos son cuadrados ($U \times U$ pixeles).

(*) Con respecto a la ejecución del programa de compresión de imágenes con la técnica CTB es necesario notar que el tiempo de procesamiento es dependiente de los siguientes factores:

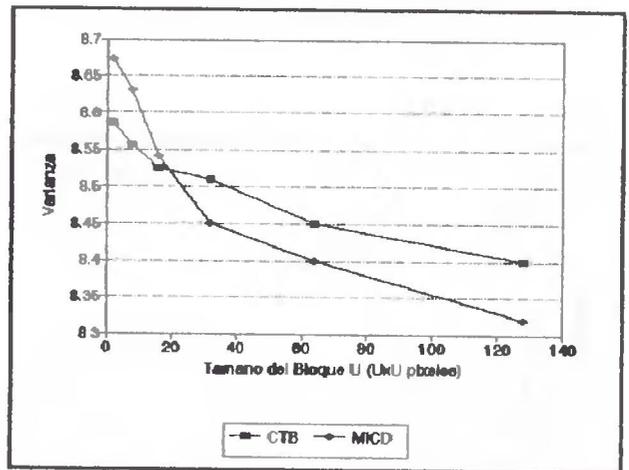
- El tamaño del bloque.
- Los niveles de color de la imagen.
- La incidencia de ruido.

(*) El compromiso entre reducción de datos y reducción de ancho de banda se establece cuando la imagen procesada pierde sus características visuales útiles (cuando una imagen se vuelve menos distinguible) y cuando la capacidad del canal ya está limitada.

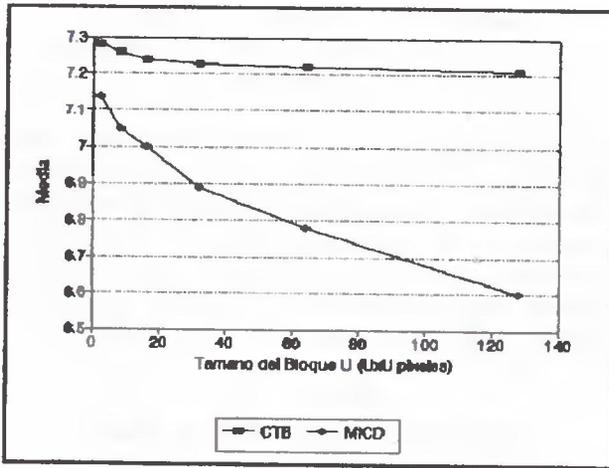
ANEXO - I
Compresiones por CTB y MICD



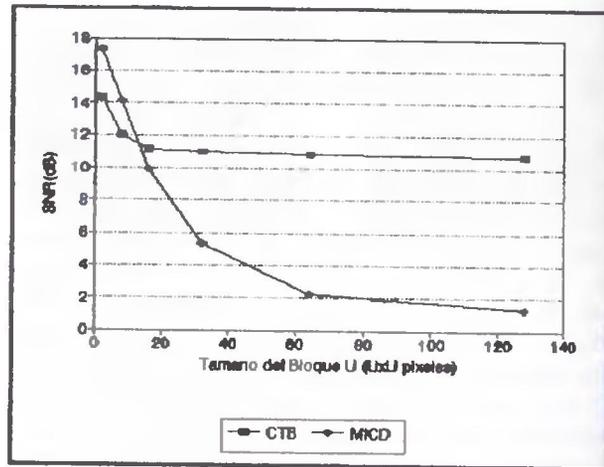
Tamaño del archivo comprimido



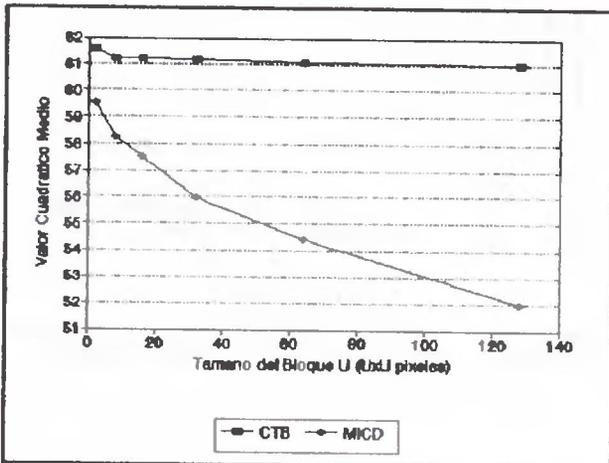
Varianza



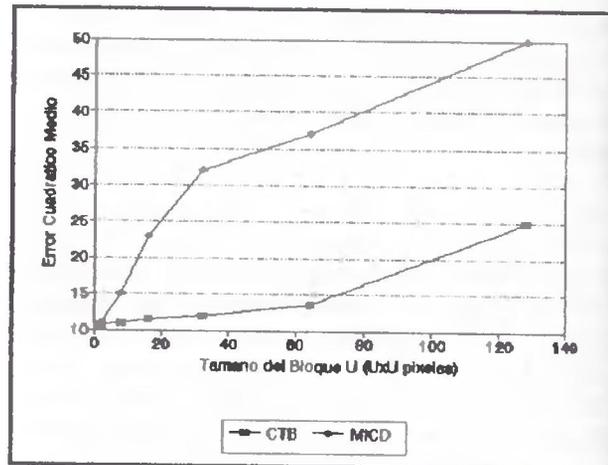
Media



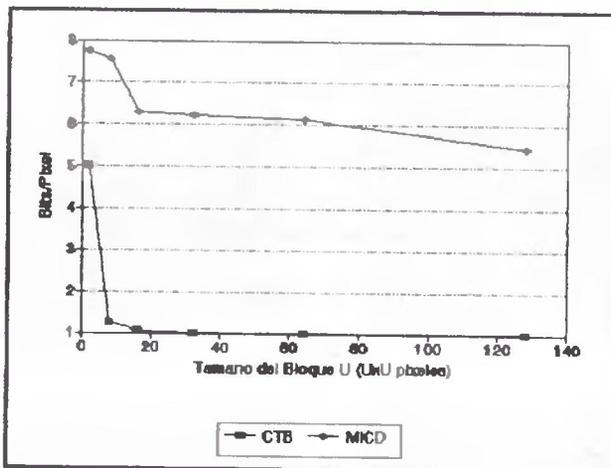
Relación Señal/Ruido



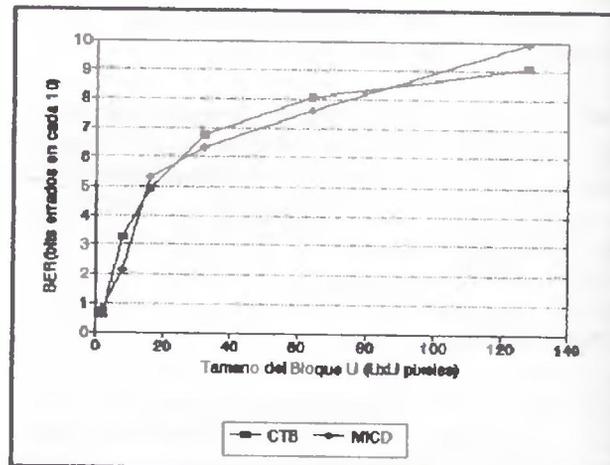
Valor Cuadrático Medio



Error Cuadrático Medio



Bits/Píxel



BER(bits errados en cada 10)

BIBLIOGRAFIA

- [1] GONZALEZ C. R., WOODS R., *Digital Image Processing*, Addison-Wesley Publishing Co., 1992.
- [2] SAKANE T. F., *Codificación Digital de Señales de Voz y Video.*, ESPE, 1995.
- [3] LATHI B. P., *Modern Digital and Analog Communication System*, Prentice-Hall, 1990.
- [4] MARPLE L. Jr., *Digital Spectral Analysis Whith Applications*, Prentice-Hall 1987.
- [5] ANIL J., *Fundaments of Digital Image Processing*, Prentice-Hall, 1989.
- [6] HABIBI A., *Special Issue on Image Bandwith Compression*, IEEE Transactions on Communications, Vol. COM-25, N°11, Noviembre 1977.
- [7] GRISWOLD N. C., HALVERSON D. R., WISE G. L., *A Note on Adaptive Block Truncation Coding for Image Processing*, IEEE Transactions on Acustics, Speech, and Signal Processing, vol. ASSP.35, N°8, Agosto 1987.
- [8] DKLP K. J., MITCHELL R. O., *Image Compression Using Block Truncation Coding*, IEEE Transactions on Communications, Vol. COM-27, N°9, Septiembre 1979.
- [9] HEALY D. J., MITCHELL R. O., *Digital Video Bandwidth Compression Using Block Truncation Coding*, IEEE Transactions on Communications, Vol. COM-29, N°12, Diciembre 1981.
- [10] HOPKINS R., *Advanced Television Systems*, IEEE Transactions on Consumer Electronics, Vol. 34, N°1, Febrero 1988.
- [11] CAGGIONI H., *Video Compression*, SONY Corporation of America, 1994.
- [12] GATICA M. J., *Procesamiento Computacional de Audio y Video*, PC-WORLD Ecuador, Año XV, N° 124, Ed. ECUASIS-TEM, Abril 1993.
- [13] PENG H., RUETZ P., AULD D., LSI LOGIC CORP., *Video compression makes big gains*, IEEE Spectrum, Octubre 1991.
- [14] ESPECIAL REPORT/CONSUMER ELECTRONICS, *An abundance of video formats*, IEEE Spectrum, Marzo 1992.
- [15] EVERSOLE W. L., MAYER D. J., FRAZER F. B. & CREEK T. F., *Investigation of VLSI Technologies for Image Processing*, Proceedings: Image Understanding Work, Pittsburgh, PA, Noviembre 14-15 1978, Sponsored by the Defense Advanced Research Projects Agency.

BIOGRAFIAS

PEREZ, TANIA ING. Ingeniero en Electrónica y Telecomunicaciones graduada en el Instituto Bonch Bruyevich, Leningrado, 1977. Actualmente se desempeña como profesor principal a tiempo completo en la Escuela Politécnica Nacional. Ha realizado estudios de Postgrado en Computación e Informática en la misma Institución. Sus áreas de interés se encaminan al procesamiento de imágenes.



RUIZ, JIMMY.- Nació en Quito-Ecuador el 23 de noviembre de 1970. Obtuvo el título de Bachiller en Humanidades Modernas en el Colegio "PAULO VI" de Quito en 1988. Sus estudios superiores los realizó

en la Escuela Politécnica Nacional y está concluyendo su tesis de grado con el estudio de las técnica de compresión de video. Se ha desenvuelto en las áreas del audio y video y en sistemas de computación. Actualmente está realizando estudios de Postgrado en Ingeniería Electrónica en la Escuela Politécnica del Ejército. Sus áreas de interés son los sistemas de radar, telemetría y biomedicina.