

UNIVERSIDAD NACIONAL DE INGENIERÍA

FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA



**“ANÁLISIS Y DISEÑO DE UN SISTEMA DE CODIFICACIÓN ADPCM APLICADO A LA
COMPRESIÓN DE IMÁGENES PARA OPTIMIZAR SU TRANSMISIÓN O
ALMACENAMIENTO”**

TESINA

PARA OPTAR EL TÍTULO DE SEGUNDA ESPECIALIZACIÓN PROFESIONAL EN
LA ESPECIALIDAD DE TELECOMUNICACIONES

ELABORADO POR:

OSCAR ANDRÉS CARRIÓN PILLACA

ASESOR:

M. Sc. FERNANDO SALDAÑA HERMOZA

LIMA – PERÚ

2013

AGRADECIMIENTOS

A Dios, por permitirme realizar un aporte más en beneficio de la sociedad y la ingeniería.

A mis padres por su invaluable apoyo en la elaboración de la tesina.

De manera especial, a mi asesor M. Sc. Fernando Saldaña por su dedicación y sus valiosos consejos y aportes que sirvieron para enriquecer el contenido de esta investigación.

A los jurados revisores Dr. Jorge del Carpio y Dr. Guillermo Kemper por su tiempo dedicado y por sus aportes que fueron tomados en cuenta en la redacción de la presente tesina.

ÍNDICE DE CONTENIDOS

INTRODUCCIÓN	1
CAPÍTULO 1	
EXPOSICIÓN DE MOTIVOS	3
1.1. Planteamiento del problema.	3
1.2. Objetivo general.	4
1.3. Objetivos específicos.	4
1.4. Hipótesis.	4
1.5. Area de estudio.	4
1.6. Técnicas de codificación de imágenes fijas y video.	4
CAPÍTULO 2	
SISTEMAS Y TÉCNICAS DE CODIFICACIÓN	6
2.1. Sistema de codificación MPEG-2.	6
2.2. Codificación de la fuente.	7
2.2.1. Codificación entrópica.	7
2.2.2. Codificación predictiva.	9
2.2.3. Predicción por movimiento compensado.	11
2.3. Correlación.	12
2.4. Cuantización.	12
2.4.1. Diseño del cuantizador.	13
2.4.2. Resolución de la intensidad de una imagen.	14
2.5. Sistemas de codificación adaptivos.	16
2.5.1. DPCM con control de ganancia adaptivo.	17
2.5.2. DPCM con clasificación adaptiva.	17
2.5.3. Diferencias entre el codificador propuesto y el codificador de la figura 2.7.	18
2.6. Evaluación de la calidad de las imágenes.	20
2.6.1. Evaluación subjetiva.	20
2.6.2. Evaluación objetiva.	20

CAPÍTULO 3	
MEDICIONES Y SIMULACIONES DEL SISTEMA	22
3.1. Datos del codificador.	22
3.2. Mediciones cuantitativas de las imágenes.	23
3.3. Análisis de la codificación entrópica.	25
3.4. Tasa de datos.	27
3.4.1. Tasa de datos de entrada.	28
3.4.2. Tasa de datos de salida.	28
3.5. Análisis de la codificación predictiva.	28
3.6. Análisis de la codificación con cuantización.	31
CAPÍTULO 4	
ANÁLISIS DE LOS RESULTADOS OBTENIDOS	35
4.1. Correlación.	35
4.2. Codificación entrópica.	35
4.3. Codificación predictiva y entrópica.	36
4.4. Codificación usando cuantización.	37
4.5. Evaluación de tasa de bits y distorsión.	38
4.6. Reducción entrópica de la información.	38
CAPÍTULO 5	
DISEÑO DEL SISTEMA PROPUESTO	40
5.1. Diseño de los parámetros del sistema.	41
5.2. Descripción matemática de la Codificación-Decodificación.	42
5.3. Sistema de codificación sin cuantización adaptiva.	43
5.3.1. Evaluación de los sistemas no adaptivos (Sistemas A y B).	44
5.3.2. Inconvenientes de los sistemas no adaptivos (Sistemas A y B).	45
5.4. Sistemas propuestos de codificación adaptiva.	46
5.4.1. Resultados y evaluación del sistema adaptivo C.	47
5.4.2. Optimización del sistema adaptivo C (Sistemas D y E).	49
5.5. Comparación entre sistemas de codificación.	54
5.6. Interpretación y conclusiones de los sistemas estudiados.	54
5.6.1. Tasa de bits de salida del sistema adaptivo EC.	56

5.7. Análisis de la carga computacional de los sistemas A – E.	57
5.7.1. Cálculo de FLOPS en los sistemas ADPCM diseñados.	59
5.7.2. Comparación de los sistemas A – E considerando las variables de salida.	61
5.7.3. Relación entre las variables de entrada y salida del sistema.	62
CAPÍTULO 6	
EVALUACIÓN DE RESULTADOS	64
6.1. Evaluación y conclusiones de las simulaciones.	64
6.2. Evaluación subjetiva de las imágenes.	65
6.2.1. Fundamentación y descripción de la metodología	65
6.2.2. DSCS – Double Stimulus Comparative Scale (3 escalas)	66
6.2.3. DSCS – Double Stimulus Comparative Scale (9 escalas)	68
6.2.4. Conclusiones de los métodos de evaluación subjetiva empleados	71
6.3. Recomendaciones.	72
6.4. Futuras investigaciones.	73
CONCLUSIONES	75
BIBLIOGRAFÍA	76
ANEXOS	78
ANEXO A: Código de programas en MATLAB.	78
ANEXO B: Código del módulo de evaluación.	86
ANEXO C: Hoja de datos de la evaluación subjetiva.	90
ANEXO D: Imágenes originales y decodificadas del sistema EC.	91

ABSTRACT

The present work studies mainly MPEG encoding and their techniques employed to compress video data. It will analyze several of these techniques used in the coding process of the video signal.

Based on this study, we research how to improve the trade-off between the coding system efficiency and the bit rates generated at the encoder output. To reach that target, we will make quantitative measures from video images, as the correlation between pixels and their entropy. Once we have got the data, we will proceed to classify the images into common features (eg. high correlation, uniform probability of bits or equal information content), we will apply several compression techniques and we will make measures in each stage of coding system to optimize its features and to get enhancements, such as reduce the bandwidth, the processing time of the encoder or the quality of images.

RESUMEN

El presente trabajo estudia básicamente la codificación MPEG y a sus técnicas empleadas para la compresión de los datos de video. Se analizará varias de esas técnicas usadas en el proceso de codificación de la señal de video.

En base a ese estudio, se busca mejorar la eficiencia del sistema de codificación. Para cumplir este objetivo, se realizará mediciones cuantitativas de las imágenes de video, como la correlación entre píxeles y la entropía de las mismas. Una vez obtenidos los datos, se procederá a clasificar las imágenes que reúnan datos comunes (Ej. alta correlación, probabilidad de bits uniforme, similar contenido de información), se aplicará diversas técnicas de compresión y se efectuará mediciones en cada etapa del sistema de codificación para optimizar las características del mismo, y obtener mejoras, como la reducción del ancho de banda, el tiempo de procesamiento del codificador o la calidad de las imágenes.

INTRODUCCIÓN

La SDTV (televisión de definición estandar) que apareció en los años 50's es aún la principal norma para la televisión analógica y digital en casi todos los países del mundo^[1]. La SDTV 4:3, la cual aplicaremos en la investigación, se muestra en la primera fila de la tabla 0.1.

El incremento del tamaño de los televisores implica un incremento en la resolución del video, y por consecuente, un aumento en la cantidad de información o tasa de bits (ver última columna de la tabla 0.1).

Para transmitir esa tasa de bits generada a través de RF (como en el caso de la Televisión digital terrestre), se requerirá de un ancho de banda demasiado grande, por lo tanto, será necesario reducir la información a transmitir. Para lograr una reducción de datos se aplican técnicas de codificación de la fuente^[5].

Para la codificación de la fuente se usan técnicas de compresión como la codificación predictiva y la codificación de transformación^[2]. Existe otra clasificación en la cual se consideran las técnicas de compresión **sin pérdidas** (lossless) y **con pérdidas** (lossy)^[4].

Tabla 0.1. Clasificación de la resolución en sistemas de televisión digital.
(considerando el formato 4:2:2)

Servicio	Relación de aspecto	Resolución	Tasa de bits (bruta)
SDTV*	4:3	640 x 480 (i ó p)	147 Mbps
SDTV	16:9	720 x 480 (i ó p)	165 Mbps
HDTV**	16:9	1280 x 720 p	440 Mbps
HDTV	16:9	1920 x 1080 (i ó p)	≈ 1 Gbps

*SD: (Standar definition)

**HD: (High definition)

Las técnicas de compresión de video requieren un balance entre la tasa de datos

generados y la calidad de las imágenes. Este tema de estudio se abordará en esta investigación.

CAPÍTULO 1

EXPOSICIÓN DE MOTIVOS

1.1 Planteamiento del problema

Codificación en un sistema de comunicación digital.

La codificación esta presente en las distintas etapas de un sistema de comunicaciones y se utiliza para mejorar diversas características del sistema, tales como el ancho de banda, la relación señal a ruido, la tasa de error de bit (BER), la potencia de transmisión, etc.

El estudio de dicha codificación, permite crear sistemas de transmisión digitales más eficientes, con características que se adecuan a las necesidades de las empresas de telecomunicaciones, y que gracias a ello, permiten que los usuarios finales dispongan de una eficiente calidad de servicio.

Con la llegada de la televisión digital y el inicio de las transmisiones en el Perú a mediados del 2011, los requerimientos de sistemas cada vez más óptimos son necesarios porque permiten contar con nuevas alternativas de solución a problemas como la mejora de la calidad de las imágenes y la generación de tasas de bits óptimas que se adecuen al ancho de banda asignado. También será posible ahorrar una parte del ancho de banda asignado para el envío de información adicional (close caption o streaming), manteniendo una calidad adecuada de las imágenes transmitidas.

Para solucionar esos problemas, este proyecto se dedica al estudio de la codificación MPEG. Dentro del sistema de codificación, se ajustarán parámetros tales como los bits de cuantización, los cuales permiten regular la entropía de fuente (tasa de bits) y también, permiten regular y mantener la calidad subjetiva de las imágenes analizadas. (en este caso los bits de cuantización se podrán variar con el coeficiente de correlación espacial de las imágenes fijas).

1.2 Objetivo general

Optimizar la eficiencia del sistema de codificación MPEG usado en la compresión de imágenes.

1.3 Objetivos específicos

- Estudiar y entender la codificación MPEG para la compresión de video.
- Obtener medidas cuantitativas de imágenes reales.
- Realizar simulaciones y mediciones con imágenes fijas reales, en las etapas de codificación y decodificación.
- Diseñar un sistema de codificación que se adecue a las restricciones de diseño: tasa de bits de salida máxima y pérdidas en la calidad de las imágenes mínima.

1.4 Hipótesis

“El análisis y estudio de los parámetros y estructura del sistema de codificación MPEG, permite encontrar un mejor balance entre la calidad de las imágenes de video y la generación de tasa bits a la salida del codificador; reduciendo así, el ancho de banda de la fuente y manteniendo la calidad subjetiva de las imágenes.”

1.5 Área de estudio

Como existen múltiples tipos de codificación, se realizará la siguiente clasificación con la finalidad de ubicar el área de estudio específico para el tema de investigación.

La **codificación de la fuente** generalmente se utiliza para reducir la cantidad de datos a transmitir. Por otra parte, la **codificación de canal** se usa para la detección y corrección de errores (reducir el BER)^[3].

De acuerdo a la tabla 1.1, los temas de investigación abordados serán: codificación de la fuente, codificación de video, norma MPEG 2 y codificación entrópica y predictiva.

De acuerdo a la bibliografía, la codificación entrópica ejecuta una compresión de 2:1, mientras que la codificación predictiva DPCM tiene una tasa de compresión de 4:1^[4].

1.6 Técnicas de codificación de imágenes fijas y video

La tasa de datos generada en la digitalización del video analógico es demasiado alta y

por ello, ocupa un gran ancho de banda.

Tabla 1.1. Tipos de codificación empleados.

Codificación de fuente	Datos y Audio	PCM, DPCM, MP3, AAC, Dolby AC-3		
	Video	MPEG-1 MPEG-2 MPEG-4	Codificación entrópica	VLC (Codificación Huffman) RLC (Run length codes) Codificación Aritmética
			Codificación Predictiva	DPCM (codificación diferencial) MCP (Motion Compensated Prediction)
			Codificación de transformación	DCT (Discrete Cosine Transform) Hadamard Karhunen-Loeve Wavelet
	DVIX			
Codificación de canal	Códigos lineales			
	Códigos convolucionales			
	Turbo códigos			

Según la recomendación de la ITU-R 601-5, la tasa generada de la digitalización de video va de 216 a 270 Mbps.

Con calidad HD, la tasa bruta de bits es aproximadamente mayor a 1 Gbps (1000 Mbps).

Estas tasas de video, deben encajar en un ancho de banda de 6 MHz cuando se usa el estandar ISDB-T para su difusión. Para cumplir este requerimiento se usan diversas técnicas de codificación (que están agrupadas en la norma MPEG-2), de las cuales analizaremos las siguientes:

- Codificación predictiva DPCM (Differential pulse code modulation).
- Cuantización y cuantización adaptiva ADPCM (Adaptive DPCM).
- Codificación entrópica.

En cada una de estas etapas de codificación, los datos de la imagen se comprimen, de modo que a la salida del codificador de video, la tasa de bits generados equivalgan al ancho de banda establecido (6 MHz) para su transmisión por RF.

Seguidamente, se describe los conceptos básicos de cada una de las técnicas arriba listadas.

CAPÍTULO 2

SISTEMAS Y TÉCNICAS DE CODIFICACIÓN

2.1 Sistema de codificación MPEG-2

El diagrama de bloques de un sistema de transmisión digital de TV consta de varias partes, de las cuales, el sistema MPEG-2 usa un sistema **codificador** antes del transmisor y un **decodificador** posterior al receptor. El sistema completo se muestra en la figura 2.1.

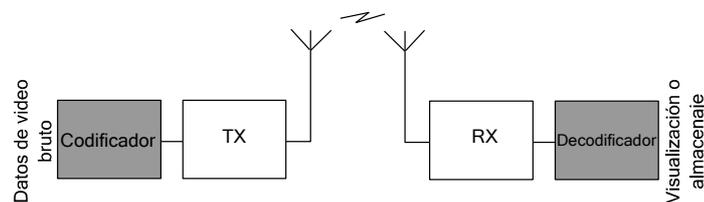


Figura 2.1. Sistema de transmisión de video.

El bloque analizado en el presente trabajo es el codificador mostrado en la figura 2.2^[1]. El mismo consta de varias etapas que son enumeradas seguidamente.

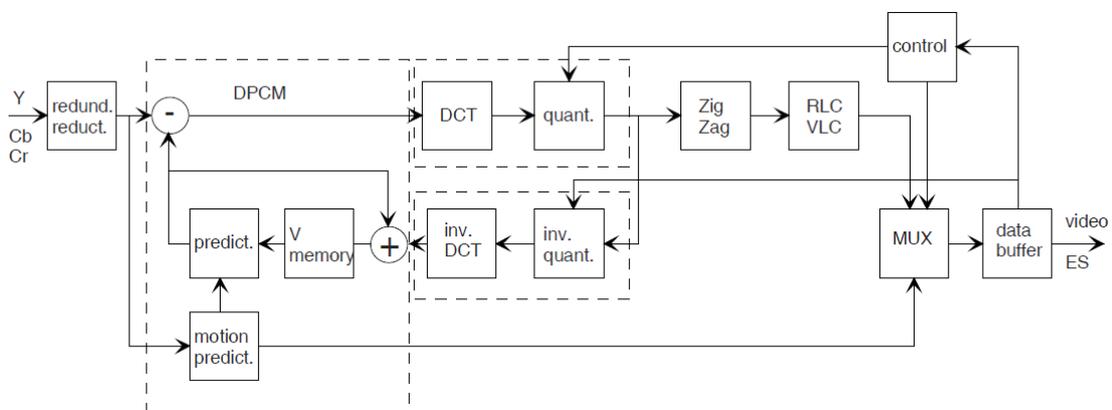


Figura 2.2. Diagrama de bloques del codificador MPEG.

- Reduccion de información redundante.

- Predicción por movimiento compensado (DPCM).
- Transformación a coeficientes (DCT) seguido de cuantización.
- Codificación Entrópica (RLC y VLC).

De todas las partes mencionadas, se analizará la codificación entrópica y la codificación diferencial seguida de cuantización, simulando el proceso que efectúan con MATLAB y midiendo las tasas de bits, entre otros parámetros, según sea el caso.

Las tasas de bits teóricas de entrada y salida del bloque codificador se resumen en la siguiente tabla^[1].

Tabla 2.1. Tasas de bits

	Tasa de bits (Entrada)	Tasa de bits (Salida)
SDTV	216 Mbps	2-7 Mbps
HDTV	> 1 Gbps	15 Mbps

En el siguiente capítulo, se demostrará experimentalmente la reducción de la tasa de bits en el bloque codificador del sistema.

2.2 Codificación de la fuente

La codificación de la fuente y la codificación de canal se basan en el **Teorema de Shannon** (1948), el cual señala (para la codificación de la fuente) que es posible una codificación con una pequeña probabilidad error (BER) si la tasa de bits generados R es mayor que la entropía de la fuente H. Si $R < H$ entonces el BER es más grande^[5].

La entropía de la fuente 'H' es una medida de la cantidad de información que genera dicha fuente y está en función de la probabilidad de cada símbolo generado. La expresión que calcula la entropía de la fuente es:

$$H(p) = - \sum_{i=1}^N p_i \log_2(p_i) \quad (1)$$

donde p_i es la probabilidad de símbolo o píxel.

2.2.1 Codificación entrópica

Aquí se usan códigos 'VLC' (variable length codes). Estos códigos tienen que tener 3

propiedades básicas. Deben ser **decodificables**, deben ser **instantáneos** y deben tener una **longitud mínima promedio de código**. Los códigos Huffman son los únicos que cumplen con las 3 propiedades anteriores, en este sentido, son óptimos.^[5]

Codificación Huffman

La codificación Huffman aplica un algoritmo de codificación que se basa en las probabilidades de cada símbolo generado por la fuente para crear códigos 'VLC'. El detalle del algoritmo está en la referencia bibliográfica^[5]. Aquí se muestran los resultados de su aplicación.

Por ejemplo, si se tiene 4 símbolos para transmitir con probabilidades {0.4, 0.3, 0.2, 0.1}, entonces la entropía de la fuente, la codificación de bit de longitud fija y la codificación Huffman (entrópica) se resumen en la siguiente tabla.

Tabla 2.2. Eficiencia de codificación.

Entropía H	Longitud media FLC	Longitud media VLC
1.85 bit/símbolo	2 bit/tx	1.9 bit/tx

Cuando se usa un código de longitud variable, se optimiza la transmisión, transmitiendo menos bits y ahorrando BW (Band Width). La longitud aproximada 'L' de un código codificado a través del algoritmo de Huffman es^[5]:

$$L = \log\left(\frac{1}{p_i}\right) \quad (2)$$

Esta longitud no es entera y hace que la codificación Huffman pierda eficiencia. Para solucionar ese inconveniente, se aplica la **codificación aritmética** en la norma MPEG-4^[2].

De la tabla anterior se puede determinar la eficiencia de codificación así:

$$\eta = \frac{\text{Entropía de la fuente}}{\text{longitud media FLC}} \leq \frac{\text{Entropía de la fuente}}{\text{longitud media VLC}} \quad (3)$$

donde 'FLC' es "Fixed length codes".

2.2.2 Codificación Predictiva

En la codificación predictiva se usa el DPCM (differential pulse code modulation) y se transmite la diferencia entre el pixel actual con el pixel predecido. Para predecir el valor de un pixel se usa generalmente un **predictor**. El predictor es lineal y tiene la forma:

$$X_n = \sum_{i=1}^K a_i X_{n-i} \quad (4)$$

Donde los coeficientes a_i óptimos están en función de la correlación entre los píxeles de una imagen y se ajustan de acuerdo a los valores de correlación^[6]. Cuando el predictor tiene la forma de la ecuación (4), también se dice que es causal^[4].

Un caso especial del predictor lineal es cuando $K = 1$ y $a = 1$. En este caso el predictor es unidimensional (1D) y el valor del pixel predicho es el anterior así:

$$X_n = \sum_{i=1}^A a_i X_{n-i} = \sum_{i=1}^1 1 \cdot X_{n-i}$$

$$X_n = X_{n-1}$$

De la ecuación anterior se tiene:

$$\hat{Y} = e = X_n - X_{n-1} \quad (5)$$

donde \hat{Y} es la imagen diferencial o imagen de error (e). En el caso de imágenes, la correlación entre píxeles es en ambas sentdos, horizontal y vertical. Por ello se emplea un predictor bidimensional (2D)^[4]. Este predictor se expresa mediante:

$$\hat{X}[m, n] = \sum_i^{K_1} \sum_j^{K_2} a_{i+j} X[m-i, n-j] \quad (6)$$

Si $K_1 = K_2 = 1$ y considerando 2 píxeles anteriores, el valor predecido es:

$$\hat{X}[m, n] = a_1 X[m, n-1] + a_2 X[m-1, n]$$

De aquí, se puede considerar un criterio de diseño para hallar los coeficientes a_1 y a_2

con el fin de minimizar el error cuadrático medio (MSE) dado en la ecuación (13). Si $X[m,n]$ es el pixel actual, entonces la imagen de error es:

$$e = X[m, n] - \hat{X}[m, n]$$

$$e = X[m, n] - a_1 X[m, n-1] - a_2 X[m-1, n]$$

Luego, el MSE es:

$$E[e^2] = E[(X[m, n] - a_1 X[m, n-1] - a_2 X[m-1, n])^2] \quad (7)$$

Donde $E[\cdot]$ es el operador valor esperado o valor medio. Para el propósito de esta tesis se toma en cuenta $a_1 = a_2 = 0.5$, debido a que se logra una reducción entrópica casi óptima con dichos valores y también, debido a la explicación dada en la sección 3.5^[2]. Para otras investigaciones donde se considere el diseño óptimo del predictor se tiene que resolver la ecuación (7) y determinar los coeficientes ' a_i ' que minimicen el MSE^[4].

Entonces la imagen diferencial $\hat{Y} = 'e'$, se transmite con menor entropía.

El objetivo de la codificación predictiva es reducir la entropía (cantidad de información) a transmitir. Si 2 imágenes consecutivas están altamente correlacionadas, la diferencia de las mismas será casi nula, lo que se traduce en menos tasa de bits.

En la codificación predictiva para imágenes fijas y móviles se distinguen 3 tipos, de acuerdo a la ubicación y uso de los pixeles en el predictor.^[2]

Codificación Intra – imagen

Usa los pixeles de la misma imagen para predecir el pixel actual. Puede tomar pixeles horizontales verticales o ambos para la predicción. Es usada generalmente para codificar imágenes fijas. Este tipo de codificación es empleado en la investigación en curso (ver sección 3.2).

Codificación Inter – imagen

Usa los pixeles de las imágenes anteriores (pasadas) para predecir el pixel actual. Es usada en los sistemas de codificación de DTV (digital TV), generalmente para imágenes en movimiento, pero combinada con otra técnica llamada **predicción por movimiento compensado**.

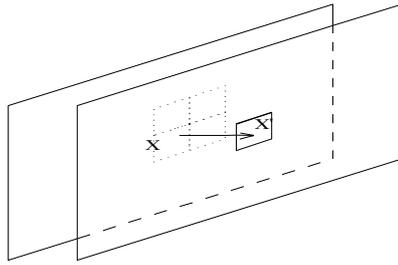


Figura 2.3. Predicción temporal de un píxel.

Codificación Mixta

Usa los píxeles de la misma imagen y de las imágenes anteriores para predecir el píxel actual. Es raramente usada en la práctica^[2].

2.2.3 Predicción por movimiento compensado

Es útil cuando las imágenes presentan objetos en movimiento, como el caso de muchas secuencias de video.

El proceso de codificación por éste método se ejecuta en dos partes. La **estimación de la imagen** la cual se realiza en el codificador y la **compensación de la imagen** que tiene lugar en el decodificador^[2].

La estimación consiste en obtener una imagen lo más parecida a la actual. Para ello, se toma como referencia las imágenes anteriores (pasadas).

Pasos para la estimación de movimiento

Se divide la imagen actual en bloques de tamaño $M \times M$. (generalmente 16×16). Cada bloque se compara con un bloque de similar tamaño del cuadro anterior. La comparación se efectúa dentro de un área de búsqueda de $\pm m$ píxeles (máximo ± 64), definidos a partir del borde de cada bloque^[8].

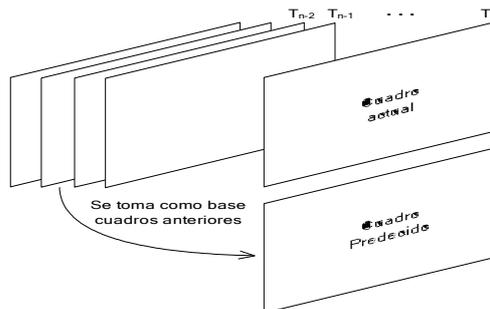


Figura 2.4. Predicción por movimiento compensado.

Para cuantificar la mejor coincidencia se usan varias medidas, las cuales se listan a continuación^[2].

- Promedio de la diferencia
- Promedio de la suma absoluta
- Promedio de la suma cuadrática

El tamaño de bloque y el tamaño del área de búsqueda es escogido tomando en cuenta el balance entre la **exactitud de la predicción** y el **tiempo de procesamiento**^[8].

Finalmente se calculan los **vectores de movimiento**, los cuales son codificados y enviados al decodificador para que se realice la compensación (reconstrucción) de la imagen^[2].

2.3 Correlación

Estima el grado de similitud entre imágenes o partes de ellas. Matemáticamente se usa el coeficiente de correlación 'r', el cual se calcula como sigue:

$$r = \frac{\sum_i \sum_j (A)(B)}{\sqrt{\sum_i \sum_j (A)^2 \sum_i \sum_j (B)^2}} \quad (8)$$

Donde $A=A_{ij} - m_A$ y $B=B_{ij} - m_B$. Además, A_{ij} y B_{ij} son pixeles de las imágenes comparadas y ' $m_{A,B}$ ' es la media de los pixeles de las imágenes^[2]. Esta fórmula se denomina producto-momento y se usará más adelante en la sección 3.2. La magnitud que representan los pixeles A_{ij} y B_{ij} en la ecuación (8) y en el resto de la tesina son los **niveles de intensidad luminosa** o **escala de grises** de una imagen. Como un pixel tiene 8 bits, entonces se tendrá $2^8 = 256$ niveles de intensidad.

Experimentalmente, cuando una imagen tiene poca cantidad de detalle, el coeficiente de correlación es mayor ' $r \rightarrow 1$ ' (ver figura 2.5a). Por otro lado cuando la imagen tiene altas variaciones, la correlación es menor ' $r \rightarrow 0$ ' (ver figura 2.5c).

2.4 Cuantización

La cuantización genera valores discretos menores que los valores de intensidad de los píxeles de una imagen de 256 niveles. Se usa luego de DPCM. Cuando se utiliza la

cuantización se genera pérdidas en la información. La función de cuantización se define como^[4]:

$$Q(x) = q \times [I/q] \quad (9)$$

Donde $[.]$ es la aproximación al entero más cercano y 'q' es un factor de conversión que generalmente es potencia de 2. ($q = 2^n$)^[4].

La fórmula (9) no se aplica en el sistema propuesto del capítulo 5 porque no produce datos uniformes, por ello se propone el siguiente diseño.

2.4.1 Diseño del cuantizador

Un cuantizador se diseña ajustando sus intervalos de decisión ' d_k ' y sus niveles de salida ' r_k ', que optimicen algún criterio como el MSE de la sección 2.5 ^[4].

Entonces considerando la PDF (función de densidad de probabilidad), $\rho(I(p))$ de los pixeles de la imagen ' $I(p)$ ' se tiene:

$$MSE = E[(I - \hat{I})^2] = \int_{d_1}^{d_{L+1}} (p - \hat{p})^2 \rho(p) dp$$

donde 'L' es el rango dinámico de la imagen. Expresado en función ' d_k ' y ' r_k ' se tiene:

$$MSE = \sum_{j=1}^L \int_{d_j}^{d_{j+1}} (p - r_j)^2 \rho(p) dp \quad (10)$$

Para minimizar el MSE se tiene que derivar la ecuación (10) respecto a ' d ' y a ' r '.

$$\frac{\partial MSE}{\partial d_j} = 0 \quad (11a)$$

$$\frac{\partial MSE}{\partial r_j} = 0 \quad (11b)$$

De la solución de las ecuaciones (11a) y (11b) resultan los valores de ' d_k ' y ' r_k ' óptimos^[4].

El MSE, que se explica en la sección 2.5, evalúa la calidad objetiva del sistema y es usado en el diseño de **cuantizadores no uniformes** (óptimos). Sin embargo, tal como se menciona en la última parte de la sección 1.1, esta tesis evalúa la calidad subjetiva, por lo tanto, el diseño considerando el MSE es no trascendente.

El cuantizador usado en el sistema propuesto del capítulo 5 es un **cuantizador uniforme**, el cual es subóptimo para los píxeles de una imagen que tienen una distribución uniforme, pero es preferido en la mayoría de sistemas de compresión prácticos debido a su simplicidad y rápida implementación^[8].

Para efectos de esta investigación se diseña un **cuantizador uniforme** como sigue:

$$r_j = \frac{d_j + d_{j+1}}{2}$$

Además:

$$d_{j+1} - d_j = d_j - d_{j-1} = \Delta, \quad 2 \leq j \leq 2^n$$

Y ' Δ ' se determina así:

$$\Delta = \frac{I_{max} - I_{min}}{2^n} = \frac{255 - 0}{2^n}$$

donde ' I ' son los niveles de intensidad luminosa, ' n ' son los bits de cuantización y ' Δ ' es el paso de cuantización.

Al cuantizar una imagen con pocos niveles de cuantización ocurre un efecto llamado "**falso contorno**" (ver figura 3.7) visible en las zonas de alta correlación de la imagen. (Generalmente cuando los bits de cuantización son menores a 6)^[10]. Este efecto se puede atenuar con la adición de ruido gaussiano antes de la cuantización y sustrayendo el mismo después de la cuantización. Esta técnica se denomina "Dithering" o "Agitación"^[4].

2.4.2 Resolución de la intensidad de una imagen

Esta sub sección es la más importante del capítulo en curso para efectos de **diseño**

del sistema con cuantización adaptativa propuesto en el capítulo 5.

Una imagen pierde calidad (subjetiva y objetiva) a medida que se reducen los niveles de intensidad de sus píxeles. Para cuantificar los efectos de la degradación cualitativa, Huang [1965] realizó una serie de pruebas subjetivas con imágenes con pequeñas, intermedias, grandes variaciones de intensidad (figura 2.5)^[10].



Figura 2.5. Imágenes: Luna (5a), leopardo (5b) y multitud (5c).

A dichas imágenes se les redujo sus niveles de intensidad 'n' (bits de cuantización) y luego, observadores fueron encuestados para clasificarlas de acuerdo a su calidad subjetiva. Los resultados se muestran en la figura 2.6 donde las curvas (llamadas curvas de isopreferencia), contienen puntos con la misma calidad subjetiva.

En los experimentos se determinó que a medida que 'n' decrece, las imágenes '5a' y '5b' (con mayor correlación) presentan una mayor degradación subjetiva, de manera que para conservar su calidad aparente, se tiene que aumentar su resolución espacial 'N'.

El punto clave del análisis de Huang es cuando la imagen '5c', (más baja correlación) pierde niveles de intensidad ('n' se reduce) su calidad subjetiva se conserva y no requiere un incremento de 'N'^[10].

De este experimento se concluye que las **imágenes con gran cantidad de detalle** (o

baja correlación) pueden ser representadas con **pocos niveles de intensidad** (o menos bits de información) sin **perder calidad subjetiva**.

Una probable causa de este resultado es que una reducción en 'n' incrementa **el contraste aparente** de una imagen, un efecto que la visión humana a menudo lo interpreta como una mejora en la calidad de la imagen^[10].

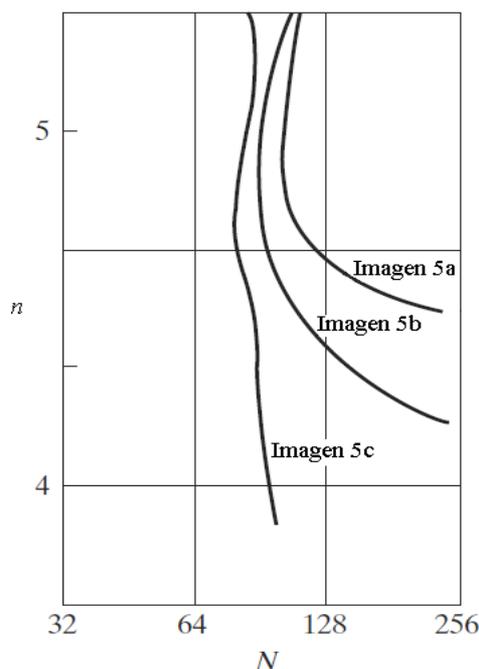


Figura 2.6. curvas de isopreferencia para las imágenes de la figura 2.5.

2.5 Sistemas de codificación adaptivos

Las técnicas adaptivas usadas en la compresión de imágenes actúan sobre ciertos parámetros del sistema de codificación y los ajustan dinámicamente en función de los datos de entrada del sistema, en el caso de la codificación de la fuente.

Los parámetros que usualmente se varían en la codificación de imágenes, cuando se usan métodos adaptivos, son: el paso de cuantización o los bits de cuantización 'n' (sección 2.4.1) y los pesos del predictor lineal 'a_i' (sección 2.2.2).

En la presente investigación se harán variar los **bits de cuantización 'n'** de acuerdo a las estadísticas de las imágenes de entrada para lograr un balance entre la tasa de bits y la calidad subjetiva de las imágenes a la salida del codificador. En tal sentido, existen investigaciones similares acerca de sistemas ADPCM, los cuales se describirán a continuación. (Bibliografía [4], Sección 6.3, 'Adaptive DPCM')^[4].

Cuando una imagen tiene varias regiones con finos detalles, las estadísticas de su

imagen diferencial cambiarán de acuerdo a los detalles en la imagen, por consiguiente, puede obtenerse un mejor desempeño ejecutando un cuantizador que, de algún modo, se adapte a las variaciones estadísticas de la imagen que es codificada.

En la bibliografía consultada ^[4], se mencionan 2 posibles aproximaciones para el diseño de un sistema ADPCM:

- DPCM con control de ganancia adaptivo.
- DPCM de clasificación adaptiva.

2.5.1 DPCM con control de ganancia adaptivo

En el primer caso se usa un cuantizador fijo cuya entrada (imagen diferencial) se ajusta para que tenga varianza unitaria. Para lograr una entrada con ganancia unitaria, se estima la varianza del pixel anterior de acuerdo a la expresión:

$$\sigma_e^2(n) = \gamma \sigma_e^2(n-1) + (1-\gamma) e_q^2(n-1)$$

Donde:

σ_e^2 : Varianza del pixel diferencial

e_q^2 : Pixel diferencial cuantizado

γ : Variable $0 \leq \gamma \leq 1$

De acuerdo al enfoque de la presente investigación, el siguiente método es el que más puntos en común tiene con el sistema adaptivo propuesto por el autor, por ello, se abordará seguidamente con mayor profundidad.

2.5.2 DPCM con clasificación adaptiva (Ref bibliográfica [4], [11], [12])

Esta alternativa usa diferentes cuantizadores para cuantizar los pixeles diferenciales de acuerdo a su clase o tipo. Cada cuantizador usa diferentes bits de cuantización y, por consiguiente, distintos niveles de entrada y salida. El elemento más importante en un sistema ADPCM es el **clasificador** (que en el diseño propuesto en la figura 5.1a. es el bloque “estadísticas de la imagen.”) El pixel actual se clasifica en una de las clases dadas y su correspondiente pixel diferencial se cuantiza usando el cuantizador perteneciente a cada clase (ver ecuación 25). Existen diferentes métricas posibles para seleccionar el cuantizador apropiado. Las métricas mencionadas en la referencia bibliográfica ^[4] son:

La **varianza** de una vecindad de píxeles cuando el píxel actual se sitúa en el centro.

El **contraste local**, definido como el cociente entre la desviación estandar y la media de los píxeles dentro de una vecindad local.

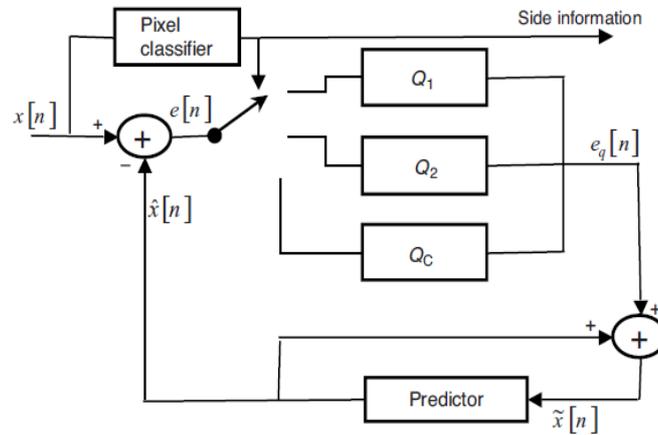


Figura 2.7. Codificador ADPCM con 3 cuantizadores.

El diseño del codificador adaptivo propuesto en el capítulo 5 usa como clasificador, **la correlación** entre los píxeles de una imagen y su predicción.

Para comparar la estructura entre el codificador ADPCM propuesto en esta investigación y el presentado en la bibliografía [4], se copia el diagrama de bloques de este último en la figura 2.7.

2.5.3 Diferencias entre el codificador propuesto y el codificador de la figura 2.7

Las principales diferencias entre el sistema de codificación propuesto en la figura 5.1a y el de la referencia bibliográfica, se señalan en la tabla 2.3.

Resultados obtenidos

En la tabla 2.4 se muestran los resultados de la implementación y simulación en software de los diversos métodos de codificación predictiva descritos en la sección 6.4 de la referencia bibliográfica [4].

Observaciones

La comparación realizada entre el método bibliográfico y el método propuesto por el autor se realizaría únicamente en términos objetivos ya que, en la bibliografía consultada,

no existen resultados de métodos subjetivos, sin embargo, **no es posible efectuar una comparación objetiva correcta** debido a las diferentes mediciones tomadas en los sistemas bibliográficos y los sistemas diseñados en esta investigación, como se aprecia en las 3 últimas diferencias de la tabla 2.3.

Tabla 2.3. Diferencias entre sistemas de codificación adaptiva.

ADPCM propuesto	ADPCM de la figura 2.7
Métrica o clasificador usado: Correlación entre todos los pixeles de la imagen.	Métrica o clasificador usado: Varianza entre una vecindad de pixeles.
Nº de cuantizadores implementados: Uno (cuyos bits de cuantización varían con la métrica.)	Nº de cuantizadores implementados: 3
No usa multiplexión.	Usa multiplexión de 3 a 1.
Uso de predicción 2D.	Uso de predicción 1D.
Medición de 'H' en las imágenes diferenciales cuantizadas (salida del codificador).	Medición de 'H' en las imágenes decodificadas.
Medición del 'PSNR' entre la imagen original y la decodificada.	Cálculo de 'SNR'

La medida de la carga computacional que se encuentra en la columna final de la tabla 2.4, se representa en MFLOPS (millones de operaciones punto flotante por segundo). Esta medida indica la complejidad del sistema adaptivo. Cuando un sistema adaptivo es más eficiente por lo general tiene un procesamiento más complejo y un MFLOPS elevado. Esta medida se usará más adelante para calcular la carga computacional de los sistemas de codificación propuestos en el capítulo 5.

Tabla 2.4. Resultados de los sistemas presentados en la bibliografía.

Tipo de DPCM	H (bits/pixel)	SNR (dB)	Procesamiento (MFLOPS)
2D No adaptivo	4	23.9	7.86
2D ADPCM de ganancia controlada	3	29.5	15.73
1D ADPCM de clasificación	5	20.1	51.12

De los métodos adaptivos hallados en el material bibliográfico [4], el mejor resulta ser: "2D ADPCM de ganancia controlada" de acuerdo a la tabla 2.4.

2.6 Evaluación de la calidad de las imágenes

Los criterios para evaluar la calidad del sistema de codificación pueden ser subjetivos o objetivos.

2.6.1 Evaluación subjetiva

El VQEG (video quality expert group) en la recomendación ITU-R 500, establece el método DSCQS (double stimulus continual quality scale) el cual sirve para evaluar subjetivamente una imagen^[1].

El método se basa en un procedimiento de comparación de dos imágenes (la original y la decodificada) hechas por un conjunto de personas. Las opiniones realizadas a cerca de la calidad de las imágenes son cuantificadas y luego se emite el resultado final. Este resultado a veces se conoce como MOS (Mean Opinion Score).

Esta evaluación será empleada para calificar subjetivamente las imágenes de prueba que se emplean en esta investigación.

Los datos de la evaluación subjetiva de esta investigación (dados en las tablas 5.1, 5.2, 5.4 y 5.6), fueron calculados sometiendo al autor de la misma a la comparación de las imágenes originales y decodificadas de los sistemas de codificación de la sección 5.2. Este hecho se considera porque:

a) Las diferencias de la comparación resultan evidentes. (Se distingue con claridad la pérdida o conservación de la calidad de una imagen).

b) En el anexo B, se presentan las imágenes originales y decodificadas aplicadas a los sistemas de la sección 5.2, con la finalidad de que los lectores puedan emitir sus propios juicios acerca de la calidad de las imágenes.

2.6.2 Evaluación objetiva

Las medidas objetivas se realizan a través de diversas mediciones tales como el S_v/N_q (Signal to Noise quantization)^[3], PSNR (Peak Signal to Noise Ratio)^[2] o las curvas R-D (Rate-Distortion curves)^[8].

El PSNR se calcula como:

$$PSNR = 10 \log \left(\frac{2^n - 1}{MSE} \right) \quad (12)$$

donde 'n' es el máximo número de bits usados para la codificación y el 'MSE' (Mean Square Error) se define como:

$$MSE = \frac{1}{NM} \sum_i^N \sum_j^M (I_{i,j} - \hat{I}_{i,j})^2 \quad (13)$$

y donde 'I' es la imagen original e ' \hat{I} ' es la imagen recuperada en el decodificador. Mientras mayor PSNR tenga una imagen decodificada, más similitud tendrá con la original.

CAPÍTULO 3

MEDICIONES Y SIMULACIONES DEL SISTEMA

3.1 Datos del codificador

El diagrama presentado en la figura 3.1 se usará para realizar las mediciones y representa un esquema simplificado del codificador MPEG de la figura 2.2. La figura 3.1 se diferencia del diagrama de la figura 2.2, principalmente porque no presenta la etapa de la codificación de transformación (DCT).

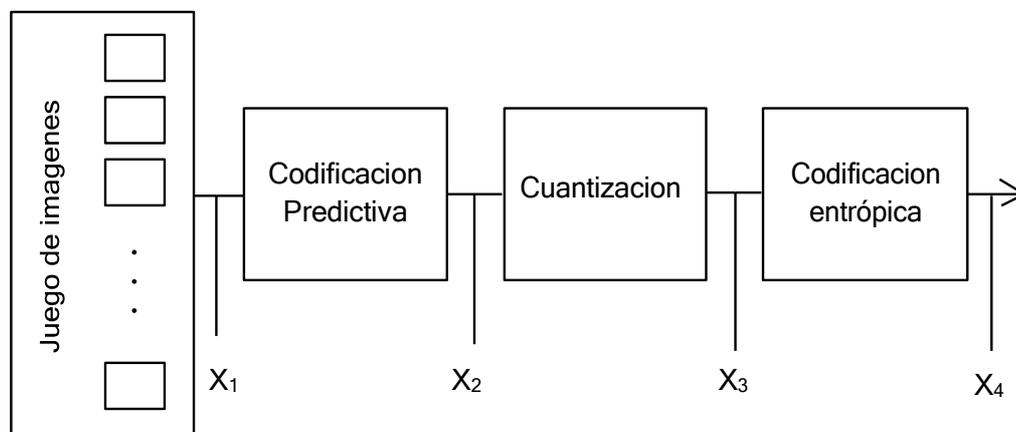


Figura 3.1. Datos del codificador.

La recopilación de datos del codificador se efectuará en 4 etapas ' X_i ', como lo indica la figura 3.1. La descripción de las etapas es la siguiente:

- X_1 : Mediciones estadísticas y entrópicas de las imágenes.
- X_2 : Medición de la tasa de bits luego de la codificación predictiva.
- X_3 : Medición de la tasa de bits y los efectos de cuantización.
- X_4 : Medición de la eficiencia de la codificación entrópica.

3.2 Mediciones cuantitativas de las imágenes

El proceso de simulación empieza con la selección de las imágenes. Aquí, distinguimos dos sub clases: **imágenes fijas** e **imágenes con movimiento**. Las mediciones a realizar en el punto X_1 de la figura 3.1 serán:

- Correlación espacial (r)
- Entropía (H)

Sólo se considera **imágenes fijas**, cuyo formato será el de la primera fila de la tabla 0.1, y la información de **luminancia (Y)** de la imagen analizada.

Las imágenes fijas escogidas fueron 15, tomadas de manera aleatoria en base a la programación de TV local.

Todos los cálculos se realizan en MATLAB ya que ofrece versatilidad para operar con imágenes y matrices. El código para calcular los parámetros de las imágenes es listado en **program_1.m** del anexo A.

Como ejemplo, se calculan los parámetros de la siguiente imagen:



Figura 3.2. Leopardo.

Al ejecutar el programa, obtenemos los siguientes datos:

$H=7.6582$

$r= 0.92159$

El coeficiente de correlación ' r ' indica la similitud de la imagen original de la figura 3.2 comparada con la misma, pero con un desplazamiento de un pixel. Si ' r ' se aproxima a 1

entonces habrá bastante correlación espacial. Si 'r' tiende a 0, entonces la imagen original y la desplazada estarán descorrelacionadas. Este coeficiente se calcula a través de la ecuación (6).

La entropía 'H' medida en bits/píxel, es una medida de la cantidad de información que contiene o que lleva la imagen de la figura 3.2 (el leopardo). Para calcular la entropía de la misma, se tiene que determinar la probabilidad de cada uno de los 256 píxeles o símbolos que contiene dicha imagen. Esto se hace clasificando los píxeles de la imagen a través de su histograma (figura 3.4)^[10].

El histograma muestra el número de píxeles que existen en una imagen en función de su nivel o intensidad. Así, en la figura 3.2 hay 0 píxel con intensidad 0, 0 píxel con intensidad 1, 2 píxeles con intensidad 2, y así sucesivamente. Entonces la probabilidad del símbolo 0 es:

$$p_0 = \frac{N^\circ \text{ de píxeles con intensidad '0'}}{N^\circ \text{ Total de píxeles de la imagen}}$$

Una vez que se tienen las probabilidades de todos los píxeles de la imagen se usa la fórmula (1). Como los píxeles tienen $2^8 = 256$ niveles, la máxima entropía de la fuente será 8 bits/píxel (lo cual ocurre cuando todos los píxeles son equiprobables).

$$H(p) = -\frac{1}{256} \sum_{i=1}^{256} \log_2\left(\frac{1}{256}\right) = 8 \text{ bit/píxel}$$

Para las 15 imágenes seleccionadas, los cálculos se resumen en la tabla 3.1.

Tabla 3.1. Datos cuantitativos de varias imágenes.

	Deporte 1	Deporte 2	Deporte 3	Dibujo 1	Dibujo 2	Dibujo 3	Noticia 1	Noticia 2	Noticia 3	Paisaje 1	Paisaje 2	Paisaje 3	Pelicula 1	Pelicula 2	Pelicula 3
Correlación (r)	0.9526	0.9754	0.9309	0.9792	0.9398	0.9716	0.9769	0.9883	0.9877	0.9790	0.9332	0.9034	0.966	0.9536	0.98
Entropía (H)	7.53	7.29	7.52	7.57	6.73	7.41	7.71	7.30	7.13	6.67	7.58	7.02	7.13	7.37	6.97

Un análisis adicional de la figura 3.2 es su respuesta en frecuencia, la cual se calcula a través de la transformada de Fourier discreta bidimensional (ecuación 14). En Matlab, el comando que lo ejecuta es **'fft2'**. Al efectuar el cálculo para la figura 3.2, el resultado para N=80 puntos, luego de escalar las amplitudes y centrar la gráfica se muestra en la figura 3.3:

$$F(u, v) = \frac{1}{N} \sum_a^{N-1} \sum_b^{N-1} f(a, b) \exp\left[\frac{-2\pi i}{N}(u \cdot a + v \cdot b)\right] \quad (14)$$

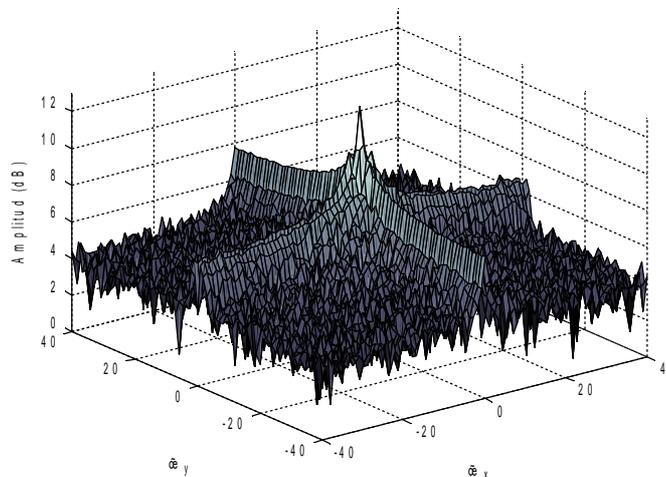


Figura 3.3. FFT bidimensional de la figura 3.2.

Finalmente, de la figura 3.3 se puede obtener una medida denominada potencia central 'P_c'. En la presente investigación no se profundiza el estudio de la 'P_c', sólo se menciona como una sugerencia para ser usada en futuras investigaciones (ver sección 6.3). En adelante no se tomará más en cuenta.

De acuerdo a la sección 2.4.2 'r', se podrá usar como métrica para variar o ajustar el paso de cuantización en el sistema propuesto del capítulo 5. La **entropía 'H'** junto a la **calidad de la imagen** son las variables que se desean controlar.

3.3 Análisis de la codificación entrópica

La eficiencia de codificación usando códigos de longitud fija (FLC) y códigos de longitud variable (VLC), se calcula a través de la fórmula (3). De este modo, se analizará el bloque X₃ – X₄ de la figura 3.1.

Como ejemplo, se calculan los datos, luego de la codificación Huffman, para la imagen de la figura 3.2, el leopardo.

La longitud de código promedio \bar{L} (de una imagen), se calcula con la siguiente fórmula:

$$\bar{L} = \sum_i p_i L_i \quad (15)$$

donde 'p' es la probabilidad de pixel y 'L' es la longitud del código para un pixel. Sin codificación, la longitud de un código para el pixel de una imagen de 256 niveles es $L = 8$ bits. Entonces aplicando la ecuación (15):

a) Sin codificación

$$\bar{L} = \sum_i p_i (8) = (8) \sum_i p_i = 8(1) = 8 \text{ bit/Tx}$$

Tomando el dato de la entropía de la figura 3.2 ($H = 7.66$), la eficiencia sin codificación será $H/L = 7.66 / 8 = 95.6 \%$

b) Con codificación

La codificación Huffman genera códigos de longitud variable, de este modo la ecuación (15) se convierte en:

$$\bar{L} = \sum_{i=1}^{256} p_i L_i$$

Se aplica el programa **program_2.m** de MATLAB para calcular \bar{L} . Los resultados son:

$L = 7.6831$

Eficiencia FLC = 95.7279%

Eficiencia VLC = 99.6758%

además, el programa muestra el histograma de la figura 3.2 (leopardo), en la figura 3.4.

El histograma muestra la cantidad de pixeles que tiene una imagen en función del valor o intensidad de cada pixel. En este caso, se usa para calcular las probabilidades "a

priori” de la imagen 'p_i' (suponiendo que es el primer fotograma clave de la secuencia).

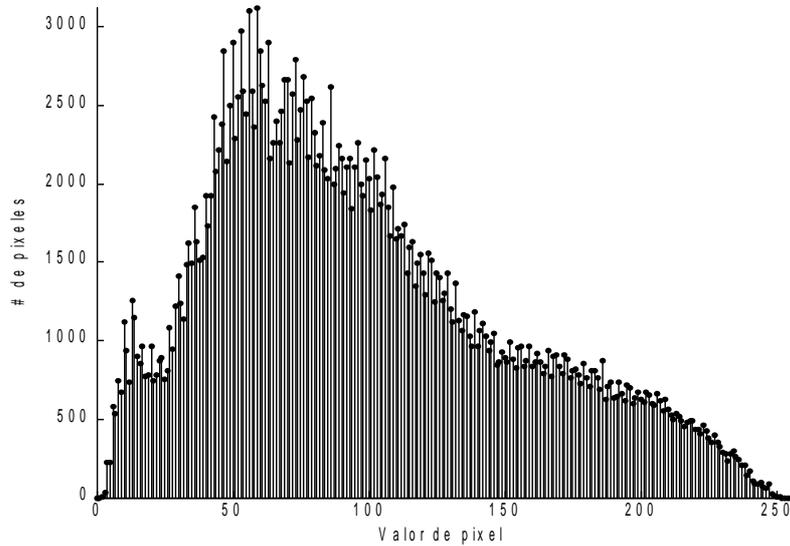


Figura 3.4. Histograma de la figura 3.2.

El objetivo de la codificación entrópica es aproximar tanto como sea posible la longitud de código promedio a la entropía de la imagen 'H'. Aplicando las imágenes de prueba se tienen los siguientes datos en la tabla 3.2.

Tabla 3.2. Eficiencia de codificación Huffman.

	Deporte 1	Deporte 2	Deporte 3	Dibujo 1	Dibujo 2	Dibujo 3	Noticia 1	Noticia 2	Noticia 3	Paisaje 1	Paisaje 2	Paisaje 3	Pelicula 1	Pelicula 2	Pelicula 3
<i>Eficiencia FLC</i>	94.1 %	91.1 %	94.0 %	94.7 %	84.1 %	92.6 %	96.3 %	91.3 %	89.1 %	83.4 %	94.8 %	87.8 %	89.1 %	92.1 %	87.1 %
<i>Eficiencia VLC</i>	99.6 %	99.5 %	99.7 %	99.6 %	99.6 %	99.6 %	99.6 %	99.6 %	99.5 %	99.6 %	99.6 %	99.6 %	99.6 %	99.6 %	99.6 %

3.4 Tasa de datos

Finalmente, para el caso de la figura 3.2, se calcula la tasa de datos a la entrada y a la salida del codificador entrópico.

3.4.1 Tasa de datos de entrada (R_i)

La tasa de datos teórica es 216 Mbps, pero esa tasa no la usamos en los cálculos. La tasa efectiva de datos (sólo para el caso de la luminancia 'Y'), se calcula como sigue^[7]:

$$Tasa = \left[(V \times H) \frac{\text{pixels}}{\text{cuadro}} \right] \left[(N) \frac{\text{cuadros}}{s} \right] \quad (16)$$

$$Tasa = (480 \times 640) \times 30 = 9.216 \times 10^6 \text{ pixel/s}$$

Como hay 8 bits/pixel, entonces la tasa de datos en la entrada (en Mbps) será:

$$R_o = L \times Tasa \quad (17)$$

$$R_i = 8 \times 9.216 \times 10^6 = 73.7 \text{ Mbps}$$

donde L es la longitud de código promedio calculada por la ecuación (15).

3.4.2 Tasa de datos de salida (R_o)

Cuando se usa VLC, la tasa de datos a la salida depende de las estadísticas de la imagen de entrada (longitud de código promedio). Por ejemplo, para la imagen de la figura 3.2 (leopardo), la tasa de salida de bits será:

$$R_o = 7.68 \times 9.216 \times 10^6 = 70.8 \text{ Mbps}$$

La reducción de la tasa usando sólo el bloque $X_3 - X_4$ de la figura 3.1 es:

$$1 - R_o/R_i = 1 - 7.68/8 \approx 5 \%$$

3.5 Análisis de la codificación predictiva

Para imágenes fijas se usa la codificación intra-imagen. La predicción de pixel actual se calcula en base a la ponderación de pixeles anteriores y/o vecinos, dentro de una misma imagen. Para la figura 3.2 (el leopardo), los pixeles p_i de la imagen predecida se calculan como:

$$\hat{p}[n, m] = 0.5 p[n, m-1] + 0.5 p[n-1, m] \quad (18)$$

Se demostró que el uso de más píxeles para predecir \hat{p} no produce una reducción significativa de la entropía (cerca del 0.1 % de reducción adicional usando más de 2 píxeles anteriores)^[2], por ello sólo se usan los píxeles $p[n, m-1]$ y $p[n-1, m]$.



Figura 3.5. Diferencia intra imagen.

Los datos de la codificación predictiva medidas en el punto X_2 de la figura 3.1, los calcula el programa **program_3.m** de MATLAB:

L=5.3455

H(Dif)=5.3184

Eficiencia DPCM = 0.59093

Eficiencia DPCM + VLC = 0.99493

La reducción de la entropía de 7.66 a 5.32 bit/píxel, permite reducir más la tasa de datos a la salida del codificador.

La imagen de la diferencia entre la imagen original y su predicción está el punto X_2 de la figura 3.1 y se muestra en la figura 3.5.

También se muestra el histograma de la diferencia en la figura 3.6, donde se observa, en este caso (y en todas las figuras propuestas), una mayor cantidad de píxeles con valores 0, lo que demuestra una alta correlación espacial de las imágenes diferenciales.

Usando la fórmula (5) y la codificación predictiva; la entropía normal y la entropía de la imagen diferencial para el conjunto de 15 imágenes de prueba se muestra en la tabla 3.3.

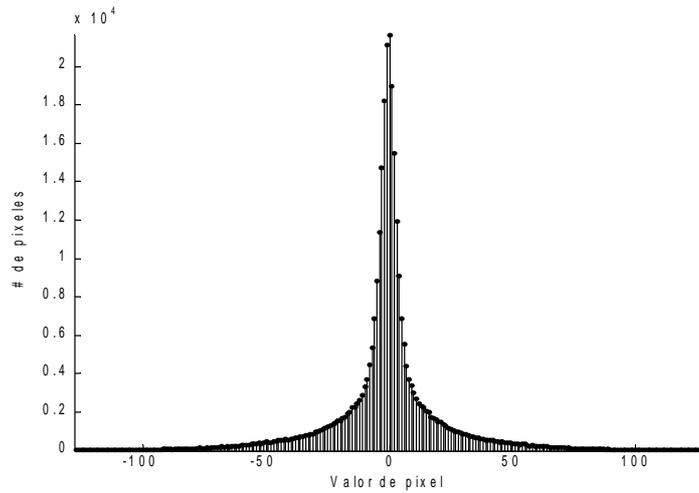


Figura 3.6. Histograma de la diferencia de la figura 3.5.

Tabla 3.3. Entropía diferencial de varias imágenes.

	Deporte 1	Deporte 2	Deporte 3	Dibujo 1	Dibujo 2	Dibujo 3	Noticia 1	Noticia 2	Noticia 3	Paisaje 1	Paisaje 2	Paisaje 3	Pelicula 1	Pelicula 2	Pelicula 3
<i>H</i>	7.53	7.29	7.52	7.57	6.73	7.41	7.71	7.30	7.13	6.67	7.58	7.02	7.13	7.37	6.97
<i>H diferencial</i>	5.10	4.53	4.99	4.33	4.82	4.43	4.07	3.67	3.31	4.10	5.18	5.54	4.27	4.67	3.63

Tasa de datos de salida

Considerando la misma tasa de datos de entrada, $Tasa = 9.216 \times 10^6 \text{ pixel/s}$ (calculada en la sección 3.4.1), la tasa de bits a la salida del codificador se obtiene de la siguiente forma:

Considerado $L = 5.35 \text{ bit/pixel}$:

$$R_o = L \times Tasa$$

$$R_o = 5.35 \times 9.216 \times 10^6 = 49.3 \text{ Mbps}$$

Si la tasa de bits de entrada es de 73.7 Mbps, la reducción de la tasa de bits usando la

codificación predictiva y la codificación entrópica se calcula del siguiente modo:

$$1 - R_o/R_i = 1 - 49.3/73.7 \approx 33 \%$$

3.6 Análisis de la codificación con cuantización

El proceso de cuantificar cualquier señal siempre ocasionará pérdidas en la información de la misma. Entonces, aparece el concepto de error o **ruido de cuantización**. En el caso del video, la diferencia entre la imagen original y la imagen recuperada en el decodificador se denomina ruido de cuantización. Para optimizar el sistema de codificación se trata de encontrar un balance entre la calidad perceptible de una imagen y la tasa de bits a la salida. Este balance se obtiene variando los parámetros del cuantizador.

El efecto de “falso contorno” o “contorneo” mencionado en la sección 2.4.1, se aprecia en las partes llanas de una imagen. Este efecto, es mucho menos notorio en imágenes descorrelacionadas o en zonas de la imagen con cambios bruscos en la intensidad de pixeles. En las figuras 14 y 15, se observa como aparece el efecto de contorneo cuando la imagen original se cuantiza con 3 bits (ver cielo).



Figura 3.7. Imagen original.

Para la figura 3.2 (el leopardo), los datos de la entropía sin cuantización y con ella, se calculan en el punto X_3 de la figura 3.1, con el programa **program_3.m** de Matlab. Los resultados para $n = 5$ bits de cuantización, son:

$$H(\text{Dif}) = 5.3184$$

$$H(\text{Dif}+\text{Quant}) = 2.6137$$



Figura 3.8. Imagen cuantizada con $n=3$ bits. ($2^3=8$ escalas de gris)
(Aparición de contornos en el cielo)

Además, se muestra la imagen diferencial cuantizada con $n = 5$ bit y el histograma de la misma en las figuras 15 y 16 respectivamente.

La reducción de la entropía en la figura 3.9 es evidente. Las contribuciones a la entropía en el caso de la codificación predictiva (figura 3.5) fueron 294 valores de probabilidad. En este último caso (figura 3.10), los valores de intensidad que contribuyen a la entropía son 37, de los cuales sólo 2 probabilidades (en -1 y en 0) contribuyen en mayor medida.

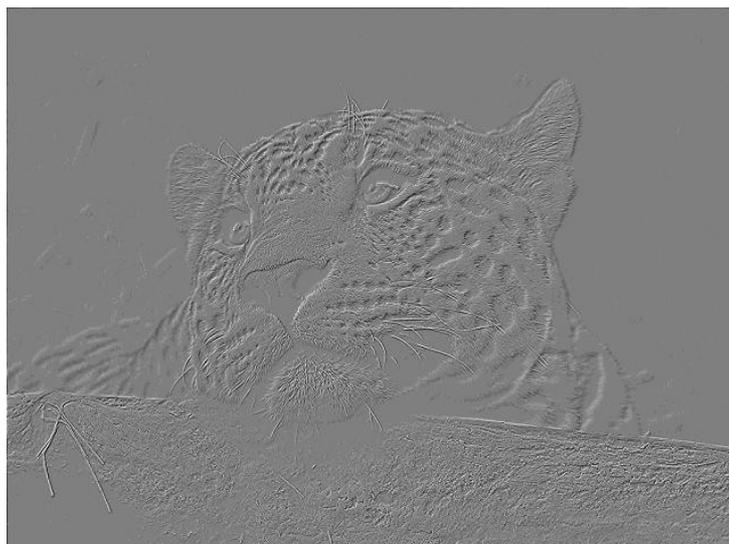


Figura 3.9. Imagen cuantizada de la figura 3.4 con $n=5$ bits

En la figura 3.9, se observa que desaparecen muchos detalles que se apreciaban en la figura 3.5 a causa de la cuantización.

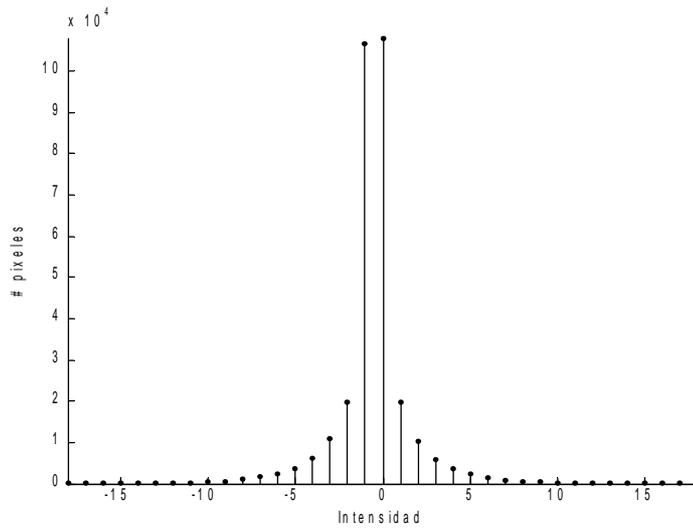


Figura 3.10. Histograma de la diferencia cuantizada.

Para las otras imágenes los datos se resumen en la tabla 3.4 (con $n = 5$ bits):

Tabla 3.4. Entropía de las imágenes cuantizadas

	Deporte 1	Deporte 2	Deporte 3	Dibujo 1	Dibujo 2	Dibujo 3	Noticia 1	Noticia 2	Noticia 3	Paisaje 1	Paisaje 2	Paisaje 3	Pelicula 1	Pelicula 2	Pelicula 3
<i>H diferencial</i>	5.10	4.53	4.99	4.33	4.82	4.43	4.07	3.67	3.31	4.10	5.18	5.54	4.27	4.67	3.63
<i>H diferencial + cuantizada</i>	2.40	1.87	2.13	1.75	2.63	1.89	1.79	1.42	1.45	1.42	2.54	2.78	1.73	2.28	1.50

Tasa de datos de salida

La tasa de entrada de bits se considera la misma que en los casos anteriores: 73.7 Mbps. La tasa de bits de salida se calcula con la fórmula $R_o = L \times Tasa$, donde 'L' es la longitud de código promedio luego de la codificación Huffman.

Con ayuda de Matlab, se calcula $L = 2.69$ bit/píxel, entonces:

$$R_o = 2.69 \times 9.216 \times 10^6 = 24.8 \text{ Mbps}$$

El porcentaje de reducción de datos usando codificación predictiva y cuantización (con

n=5 bits) para la figura 3.2 es:

$$1 - R_o/R_i = 1 - 24.8/73.7 = 66.4\%$$

CAPÍTULO 4

ANÁLISIS DE LOS RESULTADOS OBTENIDOS

Primero se presentan las mediciones cuantitativas de las imágenes a ser evaluadas. Con estos datos se ajustan los puntos de operación del sistema de cuantización adaptivo. La medida considerada como métrica es el coeficiente de correlación de las imágenes.

4.1 Correlación

Cuando la correlación se aproxima más a 1, los píxeles de la imagen están muy correlacionados. La correlación de las imágenes dibujo_2 y paisaje_3 son menores, entonces serán las que menos cambios subjetivos experimente cuando pierda niveles de intensidad.

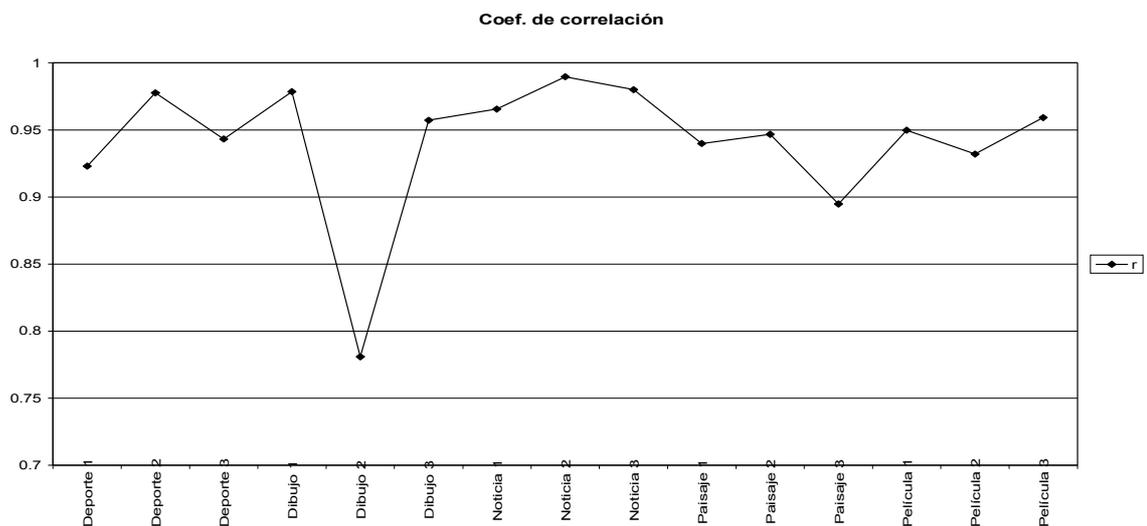


Figura 4.1. Correlación de las imágenes de prueba.

4.2 Codificación entrópica

La entropía máxima no supera los 8 bits/píxel. La codificación Huffman baja la tasa de

transmisión de 8 bits/píxel a un valor cercano a la entropía en todos los casos. La longitud de código variable no es menor que la entropía en ningún caso.

Cuando se aplica la codificación predictiva se reduce la entropía de las imágenes, parece existir una relación directa entre la reducción de la entropía y el coeficiente de correlación. Mientras mayor correlación tenga una imagen (noticia 2 y 3), mayor reducción de la entropía experimentará la diferencia y viceversa.

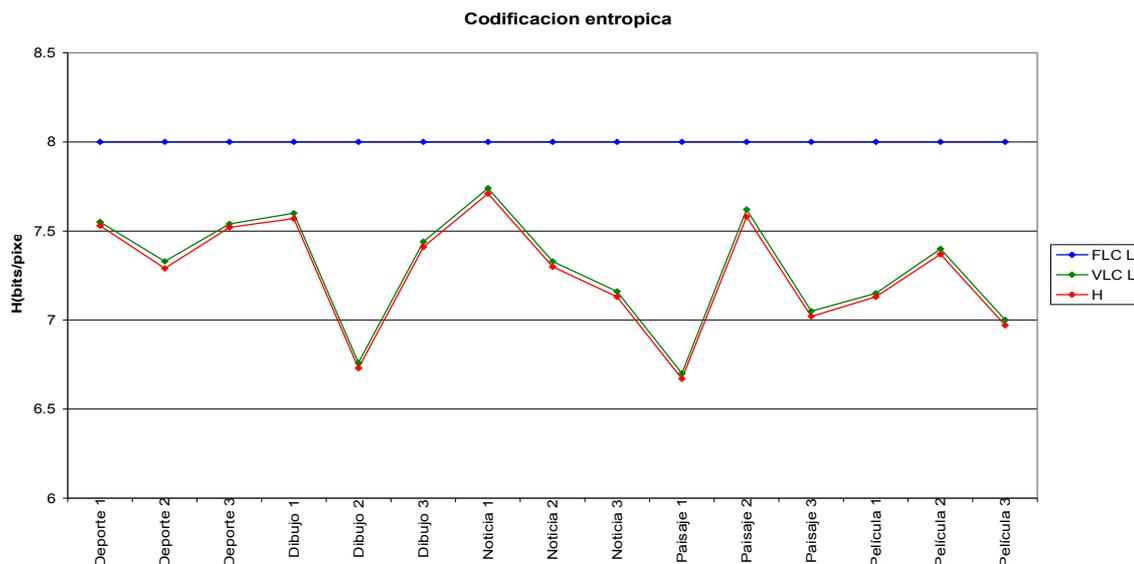


Figura 4.2. Eficiencia de la codificación entrópica.

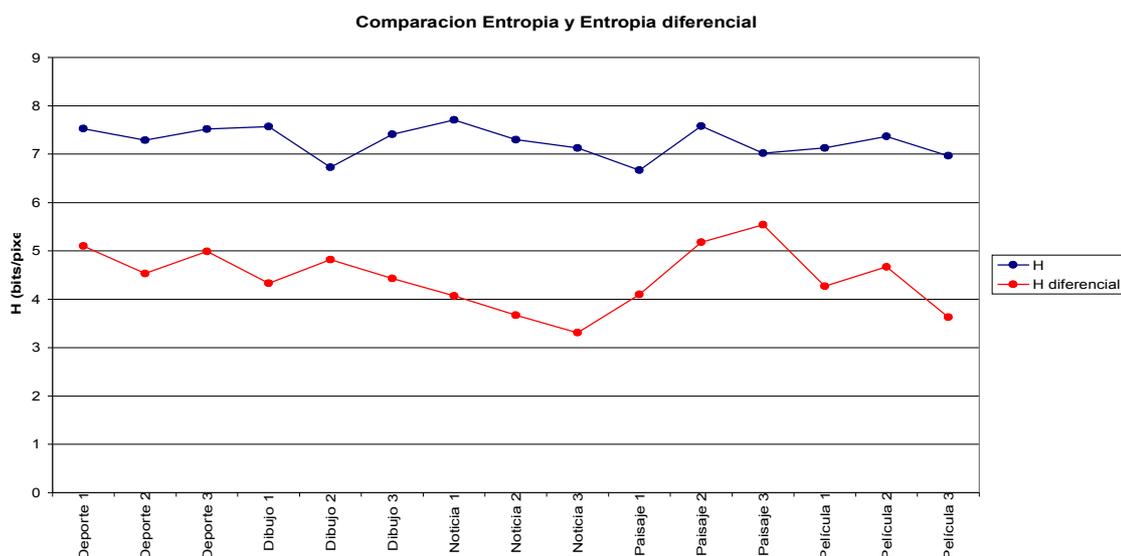


Figura 4.3. Comparación entre la entropía normal y predictiva.

4.3 Codificación predictiva y entrópica

La longitud de código es variable en el caso de la codificación predictiva, porque a

veces se tiene más de 256 símbolos. En un caso extremo, un pixel puede valer 0 y su predicción 255, entonces la resta resulta $0-255 = -255$. De este modo, el rango de valores que puede tener la diferencia es de -255 a 255.

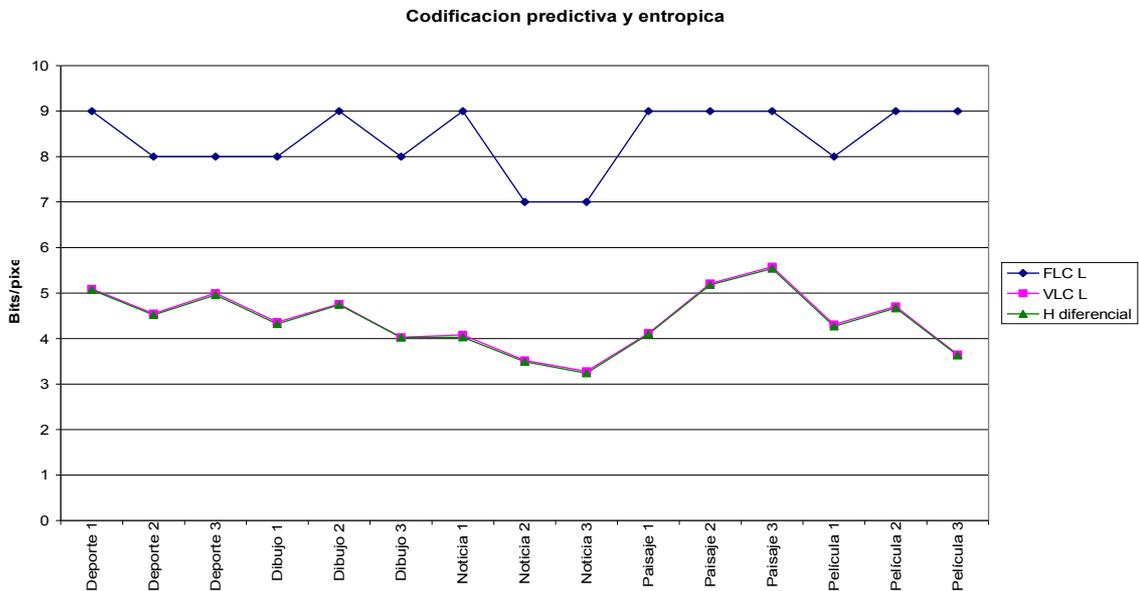


Figura 4.4. Eficencia de la codificación diferencial.

4.4 Codificación usando cuantización

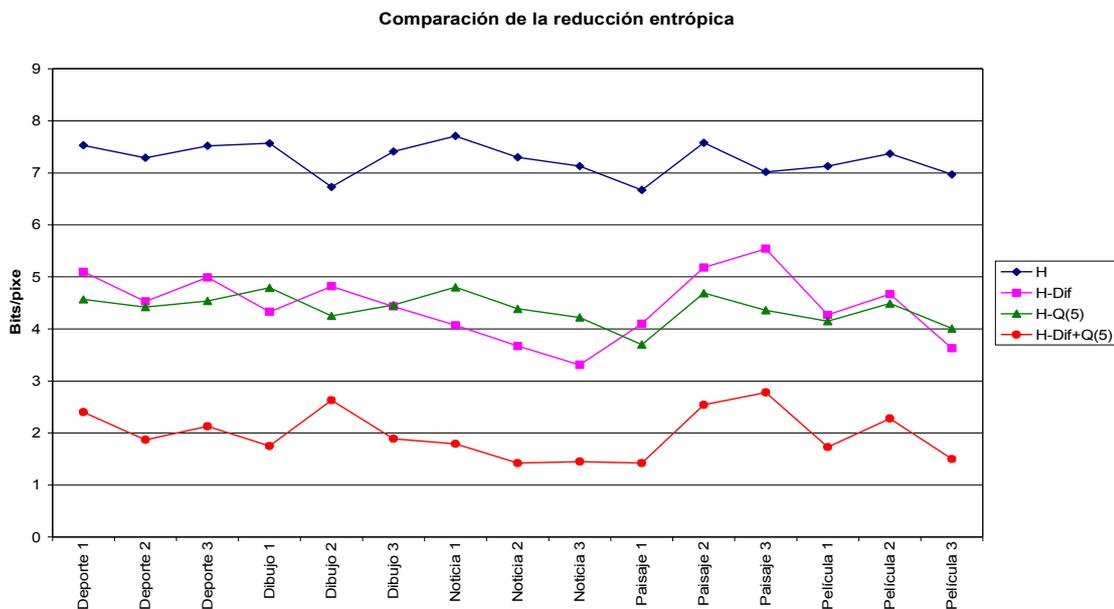


Figura 4.5. Reducción de la información entrópica en el sistema de la figura 3.1.

Usando sólo la codificación predictiva o sólo la cuantización como técnica de compresión, las tasas de bits a la salida del codificador resultan bastante similares. Se

logra un promedio de 2:1 en la reducción. Combinando ambas, el promedio de reducción se incrementa a 4:1 aproximadamente como se observa en la figura 4.5.

4.5 Evaluación de tasa de bits y distorsión

La evaluación objetiva se realiza a través de las curvas R-D. Para la imagen de la figura 3.2 se tiene:

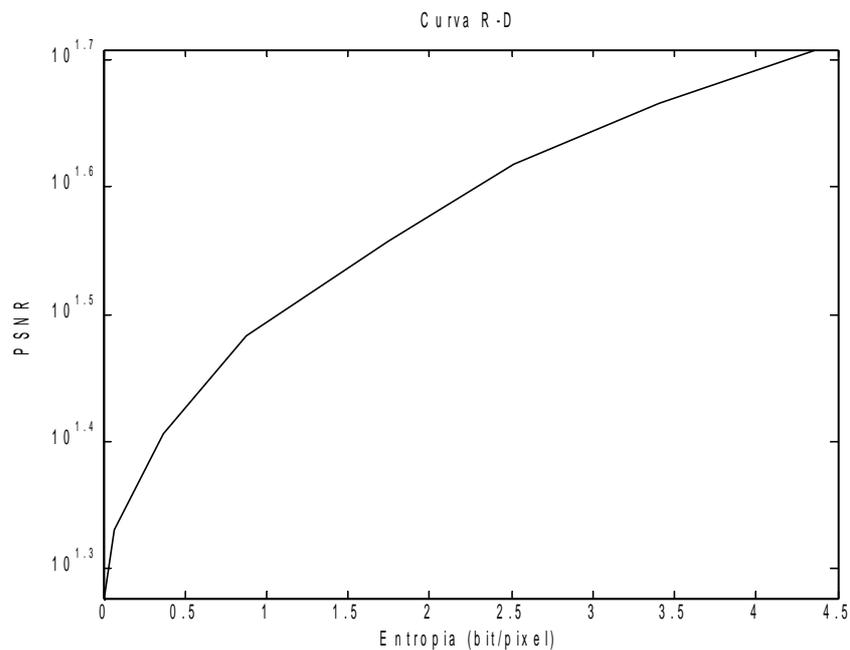


Figura 4.6. Curva R – D.

Esta curva sirve para determinar los parámetros (paso de cuantización 'q') para la operación deseada del sistema. Por ejemplo, si se desea trabajar a 1 bit/píxel como máximo, se calcula un valor de cuantización 'q₁' (diseño de los parámetros) y la máxima calidad esperada será alrededor de $10^{1.5} \approx 31$ dB.

El programa **program_4.m** calcula la PSNR y la 'H' para varias imágenes de entrada.

4.6 Reducción entrópica de la información

La reducción de la entropía aplicando las técnicas revisadas en el capítulo 3, se observa en la siguiente tabla considerando a la imagen de la figura 3.2 como referencia para los cálculos.

En la tabla 4.1, se observa que la máxima reducción de datos ocurre cuando se combina la DPCM y la cuantización.

Tabla 4.1. Reducción entrópica de la figura 3.2.

Entropía de la imagen 8. (El leopardo)			
H normal	H diferencial	H cuantizada (n=5)	H diferencial y cuantizada (n=5)
<i>7.66 bits/pixel</i>	<i>5.32 bits/pixel</i>	<i>4.67 bits/pixel</i>	<i>2.61 bits/pixel</i>

De la misma tabla, también se identifica la principal variable que se puede manipular para ajustar la tasa de bits a la salida del codificador. Ésta es el tamaño o paso de cuantización ' Δ '. En el siguiente capítulo se varía el paso de cuantización para lograr un balance entre la tasa de bits de salida y la calidad subjetiva de la imagen.

La técnica usada que adapta el paso de cuantización de acuerdo a las estadísticas de las imágenes de entrada se denomina ADPCM (Adaptive differential pulse code modulation).

CAPÍTULO 5

DISEÑO DEL SISTEMA PROPUESTO

El sistema propuesto adaptará el paso de cuantización de acuerdo a las **estadísticas de las imágenes** de entrada. En tal sentido el diagrama del sistema codificador es:

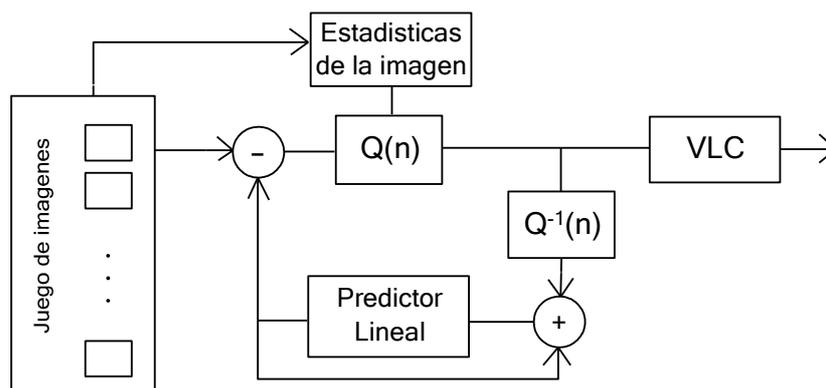


Figura 5.1a. Sistema codificador propuesto.

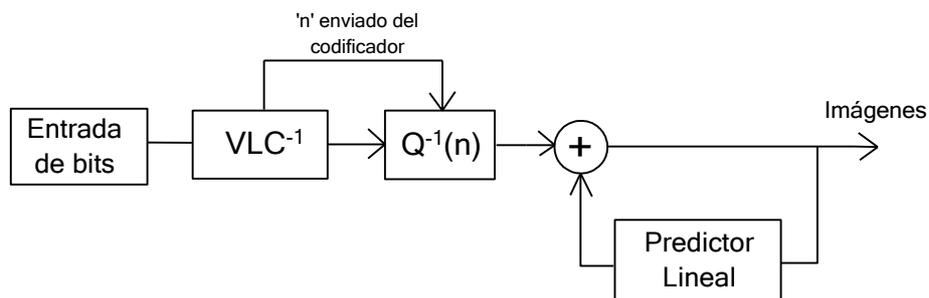


Figura 5.1b. Decodificador

El decodificador del sistema permanece invariable excepto por la recuperación de los bits de cuantización 'n' determinados en el codificador.

De acuerdo a las figuras 18 y 20, el **coeficiente de correlación** guarda relación con la reducción de la entropía, por eso se considera como medida estadística o métrica para ajustar el paso de cuantización.

Los principales sustentos en los que se basa el diseño del sistema son:

a) Las imágenes con correlación alta ' $r \approx 1$ ', presentan una distribución de píxeles uniforme, por tanto el paso de cuantización debería ser pequeño para no perder los finos detalles de la imagen y evitar el efecto de "falso contorno" (descrito en las secciones 2.4 y 3.6).

b) Las imágenes con baja correlación ' $r \rightarrow 0$ ', pueden ser cuantizadas por un cuantizador de paso ancho o con pocos niveles de intensidad porque no presentan finos detalles y el efecto de contorneo pasa desapercibido debido a los cambios bruscos entre píxeles. (Ver sección 2.4.2).

5.1 Diseño de los parámetros del sistema

En el sistema de la figura 5.1a se pueden diseñar o ajustar 2 parámetros: los pesos del predictor lineal 2D y el paso o intervalo de cuantización^[4]. Como se sugiere en la parte final de la sección 4.6, sólo se realizará el diseño de los pasos de cuantización. El predictor lineal queda establecido por la fórmula (18) de la sección 3.5.

El bloque VLC de la figura 5.1a no tiene parámetros de ajuste. De acuerdo a la sección 3.3, sólo sirve para aproximar la tasa de datos de salida a la entropía de la fuente. En la sección 5.6.1 se emplea la codificación VLC para estimar la tasa de bits óptimo del "sistema EC" (sección 5.4.2).

Experimentalmente las imágenes cuantizadas presentan degradación subjetiva (efecto de contorneo) cuando los niveles de cuantización son menores a 2^6 [10]. Por eso, en los sistemas estudiados se consideran cuantizadores con $n \leq 6$ bits (sección 2.4.2).

El paso de cuantización y los bits de cuantización estarán en función del coeficiente de correlación ' r ' de la imagen a codificar.

$$\begin{aligned}\Delta(r) &\propto r & 0 \leq r \leq 1 \\ n(r) &\propto r & 0 \leq r \leq 1\end{aligned}$$

De acuerdo a lo analizado en los capítulos 2 y 3, cuando ' r ' varía de menor a mayor, ' Δ ' debe disminuir y ' n ' debe aumentar; por tanto:

$$\begin{aligned}\Delta(r) &= k_1 \cdot r & 0 \leq r \leq 1 \\ n(r) &= k_2 / r & 0 \leq r \leq 1\end{aligned}$$

pero como ' r ' varía entre 0 y 1, entonces se cumple la relación inversa:

$$\begin{aligned}\Delta(r) &= \frac{k_1}{r} \\ n(r) &= k_2 \cdot r\end{aligned}\tag{19}$$

para todo k_1 y $k_2 \in \mathbb{R}^+$.

Las 2 fórmulas (19) son equivalentes. Por conveniencia se selecciona la segunda. Como existe la posibilidad que 'r' sea 0 ó negativo, se suma una constante 'c₂' a la ecuación (para evitar que n = 0) y luego se redondea con la función máximo entero ([.]) porque 'n' solo acepta valores enteros. De este modo la ecuación (19b) se convierte en:

$$n(r) = \lceil k_2 \cdot r + c_2 \rceil\tag{20}$$

La ecuación (20) constituye la fórmula base para el diseño de los cuantizadores uniformes y lineales de paso variable usados en el sistema de la figura 5.1a.

5.2 Descripción matemática de la codificación-decodificación

La operación efectuada en el codificador es:

$$Y_{TX} = Q\{Y - Y_{pre}\}\tag{21}$$

Donde 'Y' es la señal de luminancia de la imagen original y 'Y_{pre}' es su predicción (calculada por el bloque "predictor lineal" de la figuras 24a y 24b). La operación efectuada en el decodificador es:

$$Y_{RX} = Q^{-1}\{Y_{TX}\} + Y_{pre}\tag{22}$$

La entropía 'H', medida en todos los sistemas estudiados (A-E), es la de las imágenes diferenciales cuantizadas, es decir la entropía de 'Y_{TX}'.

Entonces, las imágenes que se comparan en la **evaluación subjetiva** son 'Y' e 'Y_{RX}'. Con éstas, también se calcula el MSE y el PSNR para tener una medida objetiva entre las imágenes original y decodificada.

En las siguientes 2 secciones del capítulo se diseñan y simulan varios sistemas de codificación con cuantización fija y adaptiva. Luego de recopilar la información respectiva, se compara el desempeño de los mismos y se obtienen las interpretaciones y conclusiones de la comparación.

Para una clara y precisa identificación de los sistemas usados, se los indexa con las letras mayúsculas del alfabeto y se anexa una breve descripción de los mismos.

- Sistema A: Sistema de codificación DPCM con $n=4$ bits de cuantización.
- Sistema B: Sistema de codificación DPCM con $n=5$ bits de cuantización.
- Sistema C: Sistema de codificación ADPCM con $n=4$ y $n=5$ bits de cuantización. Métrica usada: Correlación espacial 'r'.
- Sistema D: Sistema de codificación ADPCM con $n=3$, $n=4$ y $n=5$ bits de cuantización. Métrica usada: Correlación espacial 'r'.
- Sistema E: Sistema de codificación ADPCM con $n=3$, $n=4$, $n=5$ y $n=6$ bits de cuantización. Métrica usada: Correlación espacial 'r'.

*DPCM: Diferencial Pulse Code Modulation

*ADPCM: Adaptive Diferencial Pulse Code Modulation

5.3 Sistema de codificación sin cuantización adaptiva

La simulación de un sistema normal, no adaptivo, se ejecuta para comparar la eficiencia de ambos sistemas, con y sin cuantización adaptiva. El diagrama de bloques de un codificador convencional es el mismo que el mostrado en la figura 5.1a, pero sin la etapa “estadísticas de la imagen”.

Como se mencionó en el capítulo anterior, la entropía 'H' es la medida de la cantidad de información o tasa de bits de salida y el PSNR es la medida objetiva de la calidad de la imagen. La evaluación subjetiva se realiza tal como se describe en la sección 2.5.1.

Tabla 5.1. Sistema de codificación con $n=4$ bits de cuantización.

	Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5	Imagen 6	Imagen 7	Imagen 8	Imagen 9	Imagen 10	Imagen 11	Imagen 12	Imagen 13	Imagen 14	Imagen 15
Entropía (H)	1.61	0.77	1.1	1.16	1.0	0.81	0.17	1.66	0.92	0.14	0.22	0.54	0.37	0.31	0.8
PSNR	29.0	31.2	31.8	30.9	29.9	30.4	33.7	28.5	30.9	33.7	33.7	31.7	34.2	33.0	31.7
Deg. subjetiva	no	si	no	si	no	si	si	no	no	no	si	no	no	si	si

Los datos para 15 imágenes de prueba cuando se usa el “**Sistema A**”, se muestran

en la tabla 5.1 (para n = 4 bits). Estos bits de cuantización son típicos en sistemas DPCM^[4].

Para considerar la evaluación subjetiva de las imágenes, en el anexo 'B' se muestran las imágenes originales y las decodificadas con n=4 bits. La pérdida de la calidad subjetiva está en una proporción de 7:15 (casi el 50% del conjunto de imágenes se degenera).

Para mejorar la calidad subjetiva se plantea el “**Sistema B**” no adaptivo con n=5 bits. Los resultados se aprecian en la tabla 5.2.

Si bien la calidad subjetiva aumenta con n=5 bits, en una proporción de 3:15 (sólo el 20% de imágenes pierde calidad), esta mejora viene acompañada de un incremento en la entropía promedio del sistema, lo que se traduce en una mayor tasa de bits transmitida.

Tabla 5.2. Sistema de codificación con n=5 bits de cuantización.

	Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5	Imagen 6	Imagen 7	Imagen 8	Imagen 9	Imagen 10	Imagen 11	Imagen 12	Imagen 13	Imagen 14	Imagen 15
Entropía (H)	2.60	1.44	1.65	1.86	1.83	1.58	0.45	2.65	1.64	0.5	0.53	1.16	0.79	0.79	1.4
PSNR	35.0	36.6	37.6	36.6	35.5	36.0	38.5	34.9	36.5	36.9	38.4	36.6	37.0	37.3	37.2
Deg. subjetiva	no	no	no	no	no	no	si	no	no	no	no	no	si	no	si

5.3.1 Evaluación de los sistemas no adaptivos (Sistemas A y B)

La comparación se realiza en la tabla 5.3 para los sistemas A y B, con 4 y 5 bits de cuantización respectivamente, donde no se considera la correlación. Las variables de salida comparadas son la entropía promedio y el porcentaje de degradación subjetiva de las 15 imágenes. Opcionalmente también se compara el PSNR.

La entropía promedio se calcula a través de la fórmula (23) para 15 imágenes.

$$\bar{H} = \frac{1}{15} \sum_{i=1}^{15} H_i \quad (23)$$

Del mismo modo, el PSNR promedio se calcula como sigue:

$$PSNR = \frac{1}{15} \sum_{i=1}^{15} PSNR_i \quad (24)$$

Tabla 5.3. Comparación entre sistemas no adaptivos.

	Sistema B (n=5 bits)	Sistema A (n=4 bits)
\bar{H} bit/pixel	1.39	0.772
PSNR (dB)	36.7	31.6
Deg. subjetiva	20.0%	46.7%

La degradación subjetiva representa el porcentaje de imágenes que ha perdido calidad en términos subjetivos. Si bien la entropía se redujo al cambiar de n=5 a n=4 bits, la pérdida de la calidad se incrementa más del doble. El PSNR se reduce en 6 dB aproximadamente cuando se pasa del “sistema B” al “sistema A” (ver ecuación (28) de la sección 5.6).

5.3.2 Inconvenientes de los sistemas no adaptivos (Sistemas A y B)

Un sistema no adaptivo no es muy versátil en diseños con especificaciones que necesiten alta precisión. Por lo tanto, los resultados de los sistemas A y B (no adaptivos) servirán como base para diseñar nuevos sistemas de codificación cuyos datos de salida ('H', 'PSNR' y Deg. Subj.) sean más estrictos o requieran más exactitud.

Por ejemplo, se desea diseñar un sistema de compresión de video con las siguientes especificaciones de funcionamiento a su salida.

- 'H' máxima = 1.0 bit/pixel
- PSNR = (34 ± 0.5) dB
- Calidad subjetiva media a alta, (más del 60% de imágenes analizadas no debe presentar degradación subjetiva).

De acuerdo a la tabla 5.3, ninguno de los sistemas no adaptivos cumplen con las mencionadas especificaciones ya que, con el “sistema B” se satisface el PSNR propuesto, pero no se cumple con la tasa de bits porque en dicho sistema, $H = 1.39$ bit/pixel > ' H_{max} '. Por otra parte, el “sistema A” cumple con 'H' pero incumple la segunda condición de diseño. Para resolver este ejemplo, se usan los métodos de codificación adaptivos presentados en la siguiente sección.

5.4 Sistemas propuestos de codificación adaptiva

El proceso de diseño se realiza asociando ' n_k ' bits de cuantización a un rango de coeficientes de correlación, así:

$$\begin{aligned} r < r_1 & \leftarrow (n_1) \\ r_1 \leq r < r_2 & \leftarrow (n_2) \\ r_k \leq r < 1 & \leftarrow (n_k) \end{aligned} \quad (25)$$

De acuerdo a las ecuaciones (25), si una imagen o un grupo de imágenes tienen un coeficiente de correlación que se encuentra entre ' r_1 ' y ' r_2 ', se las cuantiza con ' n_2 ' bits; si el conjunto de imágenes tiene una correlación menor que ' r_1 ' entonces se las cuantiza con ' n_1 ' bits; y así sucesivamente se asigna bits de cuantización dependiendo del número de bits de cuantización ' n_k ' seleccionadas por el diseñador.

Los valores prácticos de los bits de cuantización suelen estar entre 3 y 7^[4]. Con $n=8$ se obtendría la imagen original y con $n<3$ (según pruebas experimentales)^[10], la mayoría de las imágenes se degrada notablemente (excepto el ruido gaussiano, ' $r \rightarrow 0$ ').

Para determinar los umbrales de cada región de correlación ' r_1 ', ' r_2 ', ..., ' r_k ' se tiene que considerar el tipo de relación que se desea para ' n ' y ' r '. Para un tipo de **relación lineal** entre ambas variables, las regiones se determinan a través de la fórmula de diseño (20).

Pueden existir otros tipos de relación entre ' n ' y ' r ' tales como:

- relación exponencial
- relación logarítmica
- relación con predominancia central
- relación con predominancia de extremos

Estos tipos de relación tienen fórmulas y ecuaciones complejas, y no serán vistas en esta investigación.

Como ejemplo, diseñaremos el “**Sistema C**” de la sección 5.2, con un cuantizador lineal adaptivo de $n=4$ y $n=5$ bits de cuantización para las 15 imágenes usadas en la sección 5.3.

Paso 1: Se determinan los coeficientes de correlación de cada una de las 15 imágenes usando la ecuación (8).

$$r = \begin{bmatrix} 0.7685 & 0.9171 & 0.7743 & 0.7260 & 0.8758 & 0.9349 & 0.9852 & 0.7182 & 0.8917 \\ 0.8972 & 0.9509 & 0.9552 & 0.9794 & 0.8549 & 0.8119 & & & \end{bmatrix}$$

Paso 2: Se selecciona el máximo y el mínimo valor de correlación de los datos anteriores.

$$r_{\min} = 0.7182; \quad r_{\max} = 0.9852;$$

Con estos valores y con los bits de cuantización $n_1=4$ y $n_2=5$, se resuelve la ecuación (20) y se determina ' k_2 ' y ' c_2 '.

$$\begin{aligned} n_1 &= [k_2 \cdot r_{\min} + c_2] \\ n_2 &= [k_2 \cdot r_{\max} + c_2] \end{aligned}$$

Luego, con ' k_2 ' y ' c_2 ' se evalúa la ecuación (20) para todos los coeficientes de correlación dados en el paso 1, resultando una lista de bits de cuantización óptimo para cada imagen.

$$n = [4 \quad 5 \quad 4 \quad 4 \quad 5 \quad 5 \quad 5 \quad 4 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 5 \quad 4]$$

Así, la imagen con $r = 0.7685$ se cuantizará con $n=4$ bits, la segunda con $n=5$ bits y así sucesivamente.

Paso 3: Se determina el umbral de correlación en mitad de la transición del máximo ' r ' cuantizado con $n=4$ y el mínimo ' r ' cuantizado con $n=5$ bits.

$$r_{\text{umbral}} = r_1 = \frac{r_{n=4} + r_{n=5}}{2} \quad (26)$$

Por lo tanto, para este ejemplo el umbral entre $n=4$ y $n=5$, será $r = 0.8334$; y la fórmula final para el sistema con cuantización adaptiva con 2 regiones de cuantización será:

$$\begin{aligned} r < 0.8334 & \rightarrow (n=4) \\ 0.8334 \leq r < 1 & \rightarrow (n=5) \end{aligned} \quad (27)$$

5.4.1 Resultados y evaluación del sistema adaptivo C

Cuando se aplican las 15 imágenes al “**Sistema C**” (ecuación (27)), las imágenes cuantizadas con $n=4$ y $n=5$ bits respectivamente se observan en las figuras 24.

Como se observa en la gráfica, la línea roja es el umbral de cuantización, el cual selecciona a 5 imágenes con baja correlación ($r < 0.8334$) y las cuantiza con $n=4$ bits. Las

demás imágenes que se encuentran en la zona superior del umbral ($r > 0.8334$) se cuantizan con $n=5$ bits.

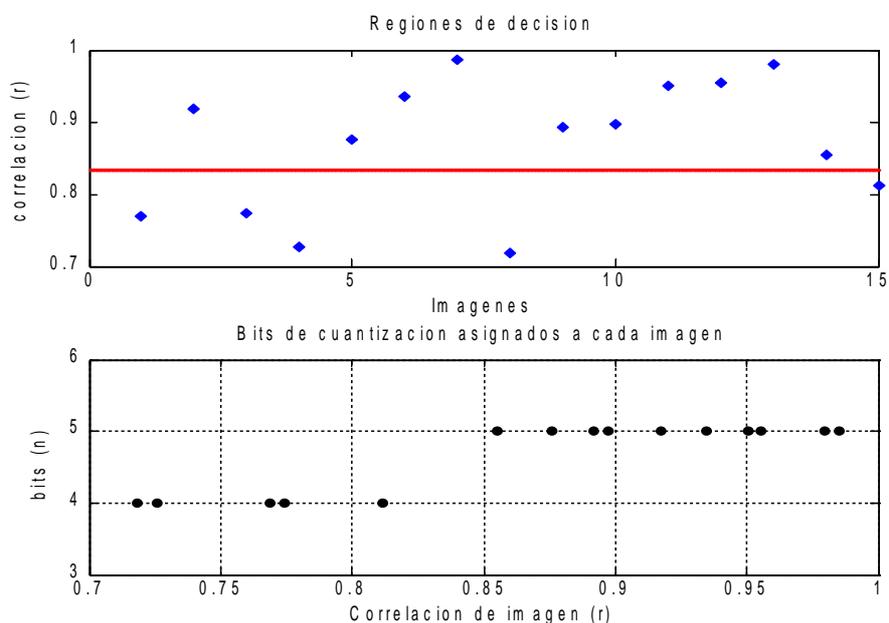


Figura 5.2. Datos de salida del sistema adaptivo C.

En la tabla 5.4 se recopilan los datos del sistema adaptivo con 2 regiones de cuantización para las 15 imágenes seleccionadas en la sección 5.3. La calidad subjetiva se prefiere como parámetro de evaluación de la calidad entre los sistemas. El PSNR se usa como un criterio de diseño establecido en el ejemplo de la sección 5.3.2.

Tabla 5.4. Datos del sistema adaptivo propuesto (C).

	Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5	Imagen 6	Imagen 7	Imagen 8	Imagen 9	Imagen 10	Imagen 11	Imagen 12	Imagen 13	Imagen 14	Imagen 15
(n) bits	4	5	4	4	5	5	5	4	5	5	5	5	5	5	4
Entropía (H)	1.61	1.44	1.1	1.16	1.83	1.58	0.45	1.66	1.64	0.5	0.53	1.16	0.79	0.79	0.8
PSNR	29.0	36.6	31.8	30.9	35.5	36.0	38.5	28.5	36.5	36.9	38.4	36.6	37.0	37.3	31.7
Deg. subjetiva	no	no	no	si	no	no	si	no	no	no	no	no	si	no	si

Como se ve en la tabla 5.5, este tipo de codificación adaptiva optimiza la tasa de bits y

la calidad subjetiva de las imágenes, ajustando las variables de salida a los requerimientos de diseño del ejemplo de la sección 5.3.2. Si bien es cierto el ajuste no es 100 % óptimo, se logra una mejor aproximación que los sistemas no adaptivos A y B.

En el caso de la entropía, el valor de diseño no debe superar a 1 bit/pixel, requisito que sólo lo cumple el sistema A, sin embargo este sistema no cumple con la condición de la calidad la cual debe estar entre 34 dB y menos aún cumple la condición de la calidad subjetiva, la cual indica que a lo mucho debe haber 40 % de degradación subjetiva.

Si analizamos el sistema B, este no cumple la condición de la entropía. Entonces el sistema adaptivo C, es el que mejor se aproxima a las condiciones de diseño del ejemplo de la sección 5.3.2, ya que el error de la entropía ' ΔH ' es 0.14 frente a 0.39 del sistema A.

EL sistema adaptivo C es, por consiguiente, el más óptimo de los 3 sistemas comparados en la tabla 5.5.

Tabla 5.5. Comparación entre sistemas no adaptivos (A, B) y adaptivo (C).

	Sistema B (n=5 bits)	Sistema A (n=4 bits)	Sistema C (n=4 y n=5)
\bar{H} bit/pixel	1.39	0.772	1.14
PSNR (dB)	36.7	31.6	34.7
Deg. subjetiva	20.0%	46.7%	26.7%

Para lograr un ajuste más óptimo a los parámetros de diseño, se considera la implementación de los sistemas adaptivos D y E, dados en la siguiente sección.

5.4.2 Optimización del sistema adaptivo C (Sistemas D y E).

Teniendo en cuenta las observaciones realizadas de la tabla 5.5, el sistema C no es 100% óptimo respecto a los parámetros de diseño establecidos en la sección 5.3.2. Para mejorar el sistema C se plantean los sistemas D y E. Las alternativas de optimización se plantean a continuación:

Criterio a considerar

Reducir la entropía del sistema C de 1.14 bit/pixel a menos de 1.0 bit/pixel.

Solución

Diseñar el sistema D con 3 regiones de cuantización fijando el bit inferior en n=3.

Diseñar el sistema E con 4 regiones de cuantización, 'n' se escoge entre 3 y 6. Este sistema será más selectivo.

Tomando en cuenta lo anterior, se diseña primero el sistema adaptivo D con 3 regiones

de cuantización. Siguiendo los pasos del ejemplo de la sección 5.4, se considera los siguientes parámetros de diseño:

- Bits de cuantización: $n=3$, $n=4$ y $n=5$.
- 2 umbrales de correlación determinados por la fórmula (26).
- Modo de operación lineal.

La información del sistema se resume en la figura 5.3 y los datos de salida del sistema se recopilan en la tabla 5.6, para las 15 imágenes seleccionadas.

De acuerdo a la figura 5.3, se tienen 3 regiones de cuantización. Para reducir la entropía al valor de diseño dado en la sección 5.3.2 (>1 bit/pixel), se selecciona la región inferior con $n=3$ bits de cuantización. Por lo tanto, este diseño cuantiza con 3 bits a las 4 imágenes más descorrelacionadas del conjunto. Los efectos en la entropía promedio y en la evaluación objetiva y subjetiva del grupo de imágenes se aprecian en la tabla 5.6.

El promedio de los datos de la tabla 5.6 se muestran a continuación, calculados con las ecuaciones (23) y (24).

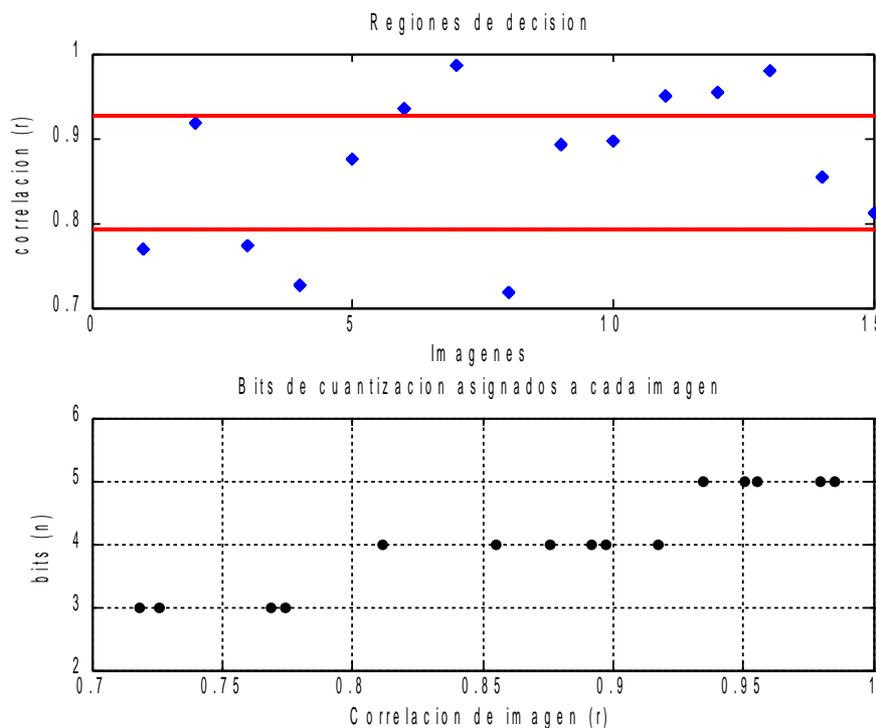


Figura 5.3. Datos de salida del sistema adaptivo D.

$$\bar{H} = 0.74$$

$$PSNR = 31.6$$

$$\text{Deg. Subjetiva} = 6/15 = 40 \%$$

Región 1	$r < 0.7931$	(n=3)
Región 2	$0.7931 \leq r < 0.926$	(n=4)
Región 3	$0.926 \leq r < 1$	(n=5)

Tabla 5.6. Datos del sistema adaptivo D.

	Imagen 1	Imagen 2	Imagen 3	Imagen 4	Imagen 5	Imagen 6	Imagen 7	Imagen 8	Imagen 9	Imagen 10	Imagen 11	Imagen 12	Imagen 13	Imagen 14	Imagen 15
(n) bits	3	4	3	3	4	5	5	3	4	4	5	5	5	4	4
Entropía (H)	0.76	0.77	0.62	0.6	1.0	1.58	0.45	0.73	0.92	0.14	0.53	1.16	0.79	0.31	0.8
PSNR	23.4	31.2	25.6	25.3	29.9	36.0	38.5	23.0	30.9	33.7	38.4	36.6	37.0	33.0	31.7
Deg. subjetiva	no	si	no	si	no	no	si	no	no	no	no	no	si	si	si

Comparando los datos con las restricciones de diseño, la entropía es menor a 1 bit / pixel (si cumple), el PSNR es menor a 33.5 dB (no cumple, mas no es relevante), y la calidad subjetiva si cumple, pero está en el límite. Cualquier diseño de ingeniería que opere al límite de las restricciones de diseño es errado pues la más ligera perturbación (ruido o variación brusca de las estadísticas de las imágenes), ocasionará que el sistema falle.

El sistema con n=4, n=5, y n=6 queda descartado porque indefectiblemente hará que se incremente la entropía promedio a una tasa mayor a 1.14 bits/pixel incumpliendo la restricción de diseño (>1 bit/pixel).

Finalmente se diseña un sistema lineal adaptivo de 4 regiones (sistema E). El desempeño de este último sistema debería ser óptimo pues discrimina mayores cambios en los pixeles de las imágenes.

Región 1	$r < 0.7472$	(n=3)
Región 2	$0.7472 \leq r < 0.8334$	(n=4)
Región 3	$0.8334 \leq r < 0.9429$	(n=5)
Región 4	$0.9429 \leq r < 1$	(n=6)

De acuerdo a los valores hallados por las fórmulas de diseño (20) y (25), las nuevas regiones de cuantización y sus respectivos umbrales de decisión se muestran en la figura 5.4.

Simulando nuevamente el sistema de codificación adaptiva con las nuevas regiones, se tiene los siguientes datos del “sistema E”:

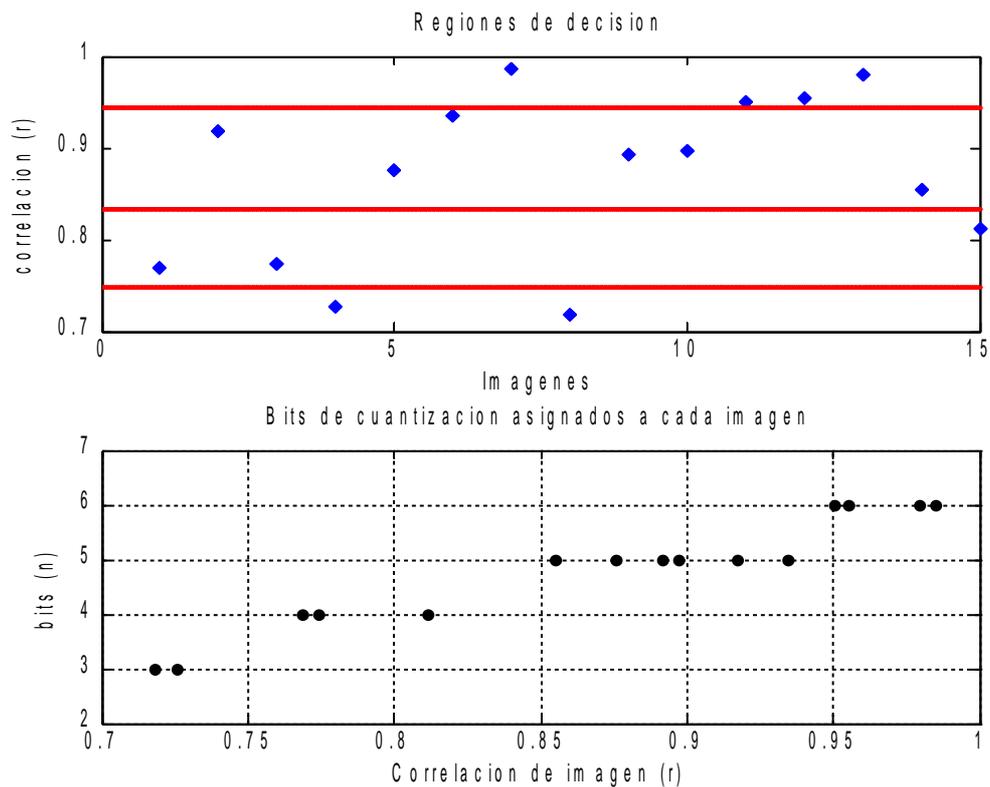


Figura 5.4. Datos de salida del sistema adaptivo y selectivo E.

$$\bar{H} = 1.19$$

$$PSNR = 35.3$$

$$\text{Deg. Subjetiva} = 2/15 = 13.3\%$$

La calidad subjetiva y objetiva mejoran respecto al “sistema D”. No obstante, La entropía promedio rebasa la entropía límite (1bit/pixel) impuesta por la condición de diseño. Con un mejor ajuste se logrará bajar la tasa de bits, manteniendo los índices de calidad de acuerdo a las condiciones de diseño.

La entropía del sistema E (1.19 bit/pixel) se puede bajar desplazando las regiones de cuantización de la figura 5.4 hacia arriba de tal manera que las imágenes 1 y 3 se cuanticen con 3 bits y las imágenes 11 y 12 se cuanticen con 5 bits (ver figura 5.5).

Para lograr ese desplazamiento se añade un “factor de compensación” |fc| a la

ecuación de diseño (20): $n(r) = \lceil [k_2 \cdot r + c_2] \rceil$. El $|fc|$ desplaza la gráfica de esta ecuación hacia la derecha, $n(r) = \lceil [k_2 \cdot r + c_2] \rceil - fc$ y hace que las imágenes se cuanticen con bits más bajos.

Este factor de compensación es definido por el usuario de acuerdo a los requerimientos de diseño. Entonces seleccionando $fc = 0.02$ para este caso, las nuevas regiones de cuantización serán:

Región 1	$r < 0.7931$	(n=3)
Región 2	$0.7931 \leq r < 0.8654$	(n=4)
Región 3	$0.8654 \leq r < 0.9673$	(n=5)
Región 4	$0.9673 \leq r < 1$	(n=6)

Estas regiones se aprecian gráficamente en la figura 5.5.

Este último sistema diseñado será llamado “**sistema EC**” (sistema E compensado) y los datos de salida del mismo son:

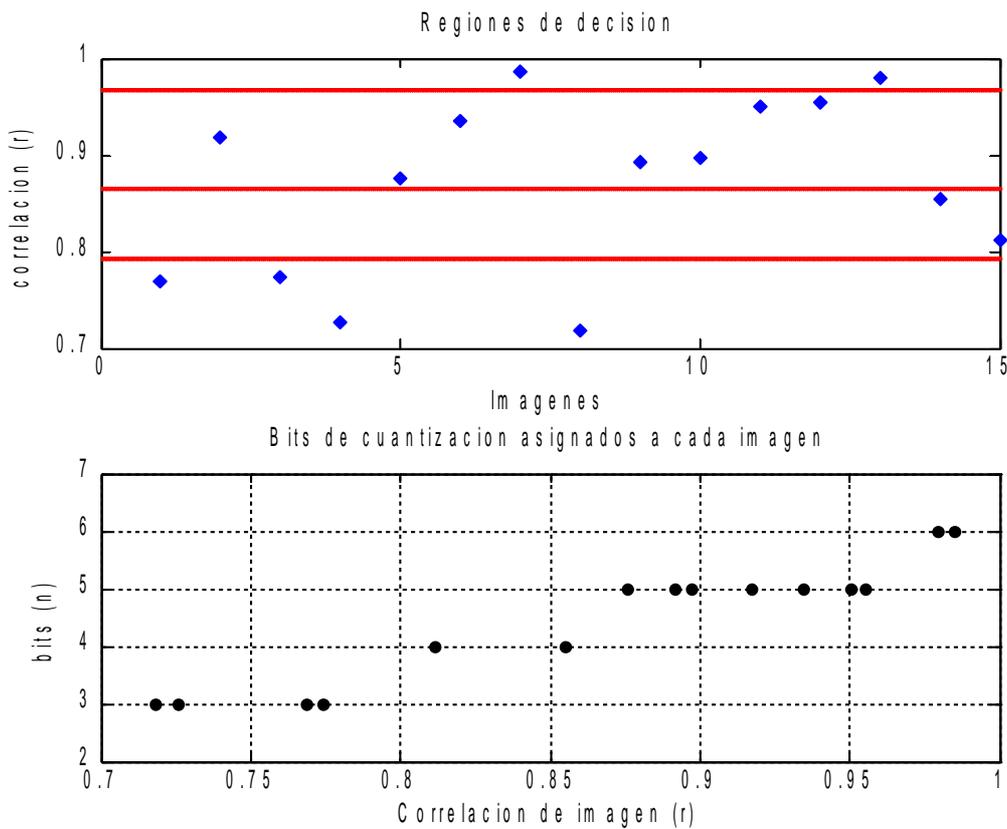


Figura 5.5. Datos de salida del sistema adaptivo EC.

$$\bar{H} = 0.98$$

$$PSNR = 33.6$$

Deg. Subjetiva = $3/15 = 20\%$

Logrando que el “sistema EC” se ajuste óptimamente a las condiciones de diseño de la sección 5.3.2.

5.5 Comparación entre sistemas de codificación

Para comparar los 5 sistemas estudiados se calcula la media de la entropía (tasa de bits) y la degradación subjetiva de las 15 imágenes de prueba para cada uno de ellos.

El PSNR no se considera como medida de comparación porque algunas imágenes, pese a tener bajo PSNR (imágenes 1, 3 y 8 de la tabla 5.6), no presentan degradación subjetiva considerable. Sin embargo, en la figura 5.6b. se muestra la comparación entre la entropía promedio y el PSNR promedio del sistema. Los resultados se muestran en los cuadros comparativos de las figuras 29a y 29b.

Para el conjunto de imágenes estudiadas, se logra mantener la calidad y reducir la tasa de bits usando el sistema de codificación con cuantización adaptiva propuesto “sistema EC”.

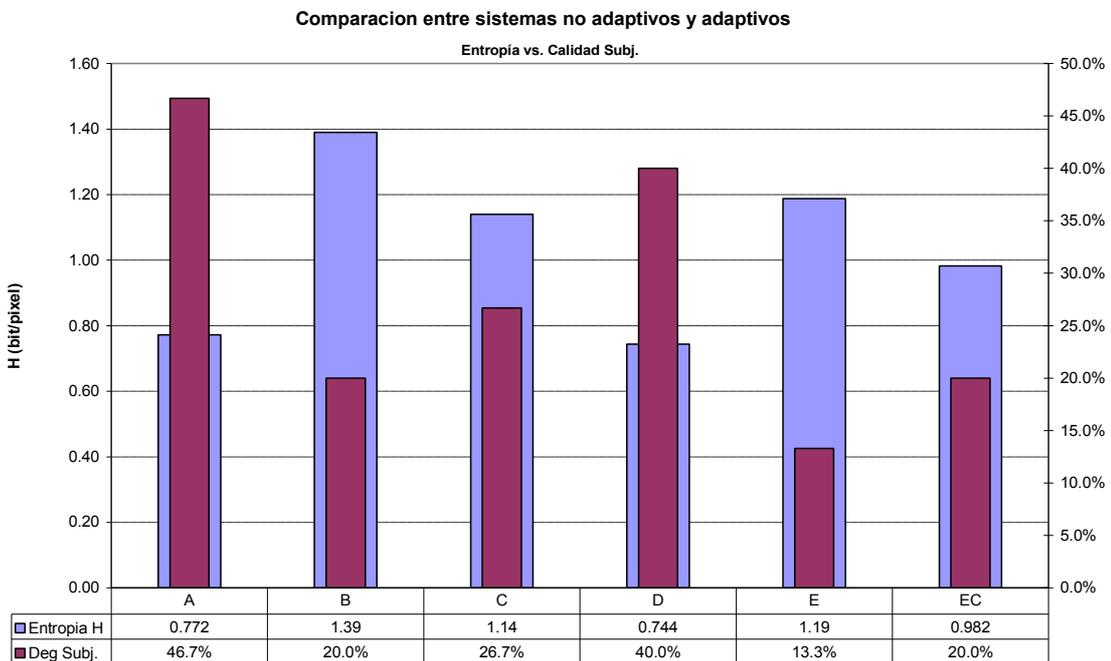


Figura 5.6a. Comparación entre sistemas de codificación no adaptivos (A y B) y adaptivos (C, D, E y EC).

5.6 Interpretación y conclusiones de los sistemas estudiados

Considerando la evaluación subjetiva como variable de comparación, se obtiene un

mejor análisis de los sistemas ADPCM propuestos debido a lo descrito en la sección 5.5 y a las comparaciones entre imágenes originales y decodificadas (Anexo B).

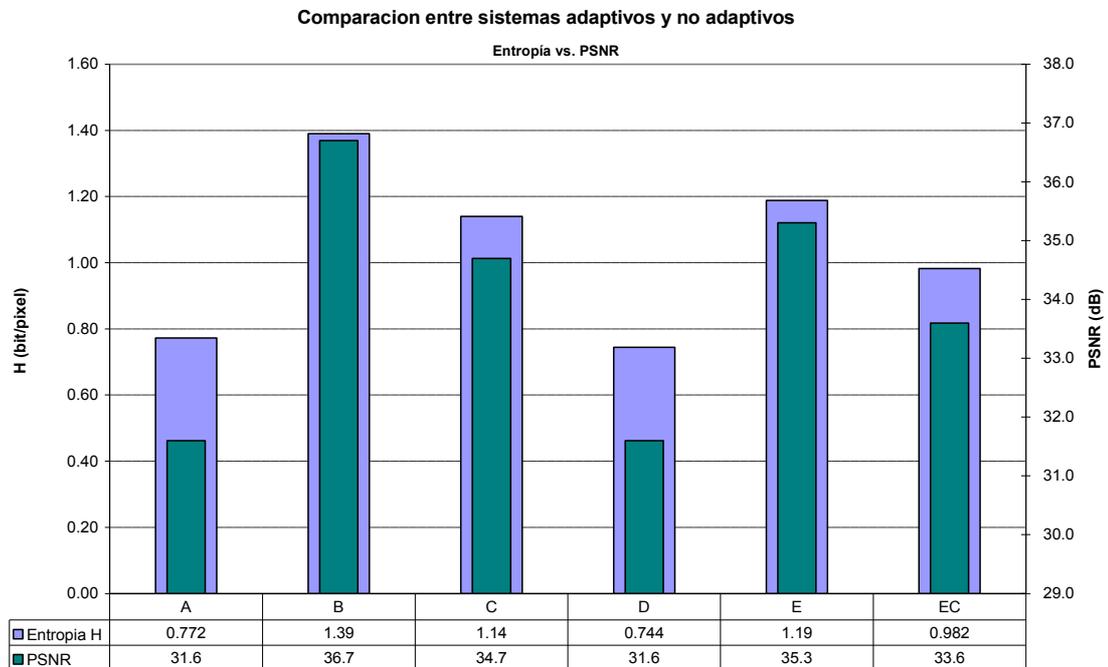


Figura 5.6b. Comparación entre sistemas de codificación no adaptivos (A y B) y adaptivos (C, D, E y EC).

La selección y diseño adecuados de las regiones de cuantización, permiten optimizar los sistemas de compresión ADPCM analizados que tienen como parámetro de diseño, la correlación entre los píxeles de una imagen.

En los sistemas A y B (no adaptivos) se observa que un incremento de 1 bit de cuantización, representa un incremento de 5.1 dB en el PSNR promedio (figura 5.6b). Esto se debe a la regla de los 6 dB que existe en los sistemas digitales PCM^[6].

$$PSNR = 6.02 \cdot n + \alpha \quad (28)$$

En la figura 5.6a. al pasar del sistema A al B se produce un incremento en la entropía y una minimización de la Deg. Subj., como era de esperarse.

Los datos de salida H, PSNR y Deg. subjetiva de los demás sistemas de codificación adaptivos son ponderaciones de los sistemas A y B. Esto ocurre porque los sistemas C, D, E y EC tratan de ajustar sus valores de salida a puntos de operación que se encuentran en el “medio aritmético” de los sistemas A y B, debido a las restricciones de diseño impuestas en el ejemplo de la sección 5.3.2.

$$H(\text{diseño}) = 1 \text{ bit/pixel}$$

$$H_A(0.772) < H(\text{diseño}) < H_B(1.39)$$

$$\text{Deg. Subj.}(\text{diseño}) < 40 \% \text{ de las imágenes.}$$

Como se describe en la sección 5.4.2 y se observa en la figura 5.6a. el único sistema cumple con todas las condiciones de diseño es el “sistema EC”. Una desventaja de este sistema es su complejidad: 4 veces más complejo que los sistemas A ó B y el doble que el “sistema C”.

No se debe realizar una comparación independiente entre 2 sistemas considerando sólo 'H' o sólo la Deg. subjetiva puesto que, como se observa en los sistemas A y B de la figura 5.6a., si optimizamos un factor (por ejemplo 'H') empeoramos el otro (la calidad subjetiva). Pero si podemos realizar comparaciones entre sistemas que tienen la misma 'H' o la misma Deg. Subj. Por Ejm.

Comparación entre el sistema A y el sistema D:

Ambos sistemas tienen similar entropía $H_A=0.772$ y $H_B=0.744$ pero el sistema D es mejor porque ofrece menor Deg. Subj.

Comparación entre el sistema B y el sistema EC:

Ambos sistemas tienen similar Deg. Subj. $D.S._B=D.S._{EC}=20\%$ (20% de las imágenes pierde calidad subjetiva), pero el sistema EC es superior porque puede operar a tasas de bit menores que el sistema B. (Sin embargo, el sistema EC es el más complejo de todos).

Otra consideración respecto a los sistemas adaptivos es que tienen que enviar información adicional hacia el decodificador acerca de cual cuantizador se usó en una determinada imagen, como se observa en la figura 5.1b.

Tomando en cuenta otros factores o variables como la complejidad del sistema propuesto o el costo de implementación del mismo, se puede diseñar un sistema de codificación de video más robusto y efectivo.

5.6.1 Tasa de bits de salida del sistema adaptivo EC

La tasa de bits efectiva (en Mbps) para el “sistema EC”, se calcula de la siguiente forma: Si nos referimos a la sección 3.4, la tasa de pixel/s de una imagen en SD con relación de aspecto 4:3 es $9.216 \times 10^6 \text{ pixel/s}$ y la entropía promedio del “sistema EC”

es 0.982 bit/pixel.

Entonces, si aplicamos codificación Huffman, la VLC tiende a aproximarse a la entropía del sistema, es decir $\bar{L} \rightarrow \bar{H}$. En la tabla 3.2 de la sección 3.3, se determina de forma experimental la eficiencia de la codificación Huffman; esta es 99.6 %. Luego aplicando la ecuación (3):

$$\eta = \frac{\bar{H}}{\bar{L}} \rightarrow \bar{L} = \frac{\bar{H}}{\eta} = \frac{0.983}{0.996} = 0.987 \text{ bit / pixel}$$

Finalmente, la tasa de bits real es:

$$R_0 = Tasa \times \bar{L} = 9.216 \times 0.987 \times 10^6 = 9.10 \text{ Mbps}$$

Si la tasa de entrada de datos es $R_i = 73.7$ Mbps, el porcentaje de reducción de tasa del “sistema EC” será:

$$1 - R_0/R_i = 1 - 9.10/73.7 = 87.7 \%$$

Entonces se concluye que, usando el “sistema EC”, se puede tener 87.7 % de compresión sin pérdidas relativas en la calidad.

5.7 Análisis de la carga computacional de los sistemas A – E.

Tal como se menciona en las observaciones de la sección 2.5.3, una medida objetiva de la carga computacional de un sistema de codificación es el FLOPS^[4] (Floating Points Operations per Second). Cabe señalar que el FLOPS considera operaciones punto flotante entre pixeles como **las multiplicaciones** y deja de lado a las sumas debido a que son operaciones más sencillas.

Como en un sistema de compresión de imágenes digitales existe una enorme cantidad de operaciones punto flotante, generalmente se consideran a los múltiplos del FLOPS como unidades de medida de la complejidad del sistema. Entonces se operará con el MFLOPS y el GFLOPS. Por convención se representará al FLOPS con una 'V'.

El procedimiento del cálculo será el siguiente:

- Se calculará los FLOPS de las etapas más relevantes de cada uno de los

sistemas de codificación no adaptivos (A y B) y adaptivos (C, D y E).

- Una vez hechos los cálculos, se realizará una comparación entre la **tasa de bits**, **calidad subjetiva** y **carga computacional** de los sistemas antes mencionados.

Entonces, primero se identifica las etapas del sistema de codificación que cuentan con la mayor cantidad de FLOPS. De la observación de la figura 3.1, se tienen 2 bloques funcionales que cuentan con FLOPS relevantes para los sistemas no adaptivos A y B: **el predictor 2D** y **el cuantizador lineal**. Adicionalmente, para el caso de los sistemas adaptivos C, D y E, se agregan los FLOPS del bloque: **estadísticas de la imagen** (ver figura 5.1a) y del promedio ponderado de los cuantizadores usados para tal sistema. De forma simbólica, representando los FLOPS por la variable 'V', se tiene:

Tabla 5.7: Representación simbólica de los FLOPS de cada sistema de codificación.

Sistemas de codificación	FLOPS por calcular
Sistema A	$V_{P2D} + V_{Q4}$
Sistema B	$V_{P2D} + V_{Q5}$
Sistema C	$V_{P2D} + V_r + \text{ponderación}(V_{Q4}, V_{Q5})$
Sistema D	$V_{P2D} + V_r + \text{ponderación}(V_{Q3}, V_{Q4}, V_{Q5})$
Sistema E	$V_{P2D} + V_r + \text{ponderación}(V_{Q3}, V_{Q5}, V_{Q4}, V_{Q6})$
Sistema EC	$V_{P2D} + V_r + \text{ponderación}(V_{Q3}, V_{Q5}, V_{Q4}, V_{Q6})$

Donde:

V_{P2D} : Flops del predictor lineal 2D

V_r : Flops del correlacionador

V_{Q3} : Flops del cuantizador de n = 3 bits

V_{Q4} : Flops del cuantizador de n = 4 bits

V_{Q5} : Flops del cuantizador de n = 5 bits

V_{Q6} : Flops del cuantizador de n = 6 bits

Entonces, se calculan los FLOPS arriba listados y luego, se los combina para cada sistema diseñado, de acuerdo a la expresión mostrada en la segunda columna de la tabla 5.7.

5.7.1 Cálculo de FLOPS en los sistemas ADPCM diseñados

Cálculo de los FLOPS del predictor lineal 2D

Para calcular los FLOPS se parte de la ecuación matemática que actúa en cada bloque. En el caso del predictor 2D la fórmula está dada por la ecuación (7):

$$\hat{X}[m,n]=a_1 X[m,n-1]+a_2 X[m-1,n]$$

Aquí se observa que hay dos multiplicaciones por cada pixel en una imagen SD de 640 x 480 pixeles. Además, la tasa de transferencia de las imágenes es de 30 cuadros/s. Entonces:

$$V_{P2D} = \frac{2 \text{ operations}}{\text{pixels}} \cdot \frac{640 \times 480 \text{ pixels}}{\text{frame}} \cdot \frac{30 \text{ frames}}{s} = 18.4 \text{ MFLOPS}$$

Cálculo del FLOPS del correlacionador

El coeficiente de correlación que es usado como métrica para determinar la variación de los bits de cuantización, se calcula a través de la ecuación (9):

$$r = \frac{\sum_i \sum_j (A)(B)}{\sqrt{\sum_i \sum_j (A)^2 \sum_i \sum_j (B)^2}}$$

Se calcula entonces la correlación entre la imagen de entrada y la misma recortada y desplazada $N=\pm 1$ pixel horizontal y verticalmente. En ese sentido se tiene $(2N+1)^2=9$ coeficientes de correlación, de los cuales se escoge el menor.

$$V_r = 9 \text{ coef.} \times \frac{3 \text{ operations}}{\text{pixels}} \cdot \frac{638 \times 478 \text{ pixels}}{\text{frame}} \cdot \frac{30 \text{ frames}}{s} = 247 \text{ MFLOPS}$$

Cálculo del FLOPS de los cuantizadores

Los cuantizadores usados están determinados por las fórmulas de diseño de la sección 2.4.1. (Pag. 13) así:

$$\Delta = \frac{I_{max} - I_{min}}{2^n}$$

$$\Delta = d_{j+1} - d_j = d_j - d_{j-1} \quad 2 \leq j \leq 2^n$$

$$r_j = \frac{d_j + d_{j+1}}{2}$$

Donde:

'I' son los niveles de intensidad luminosa, 'n' son los bits de cuantización, 'Δ' es el paso de cuantización, d_j son los niveles de decisión y r_j son los niveles de salida o cuantización.

La implementación matemática de cada cuantizador no representa una carga computacional relevante, como se aprecia en las fórmulas de diseño. Entonces, lo que determina la verdadera complejidad del sistema es el número de comparaciones lógicas en cada paso de cuantización 'Δ', así: (ver program_4 del anexo 1)

$$I_{logic} = \left(d_j + \frac{d_{j+1}}{2} \leq I \right) \wedge \left(d_{j+1} + \frac{d_{j+2}}{2} > I \right)$$

Adicionalmente se considera una operación punto flotante adicional para convertir la salida lógica en salida de intensidad luminosa.

$$I_Q = r_j \cdot I_{logic}$$

Dando como resultado 4 operaciones punto flotante por pixel.

Además, la diferencia entre la complejidad de los cuantizadores está determinada por el número de niveles de cuantización 'm', que en el caso del algoritmo ejecutado en MATLAB (Ver program_4 del anexo 1), esta dado por:

$$m = 2^n - 1 \quad 3 \leq n \leq 6$$

Por ejemplo, para n = 3 se tiene m = 2³ - 1 = 7 niveles y V_{Q3} resulta:

$$V_{Q3} = 7 \text{ levels} \times \frac{4 \text{ operations}}{\text{pixels}} \cdot \frac{640 \times 480 \text{ pixels}}{\text{frame}} \cdot \frac{30 \text{ frames}}{s} = 258 \text{ MFLOPS}$$

De similar manera se calcula los MFLOPS para Q4, Q5 y Q6.

$$V_{Q4} = 552 \text{ MFLOPS}$$

$$V_{Q5} = 1143 \text{ MFLOPS}$$

$$V_{Q6} = 2322 \text{ MFLOPS}$$

Tabla 5.8: MFLOPS de cada sistema de codificación diseñado

Sistemas de codificación	Carga computacional en MFLOPS
Sistema A	$18,4 + 552 = 570$
Sistema B	$18,4 + 1143 = 1161$
Sistema C	$18,4 + 247 + 1/3 \cdot 552 + 2/3 \cdot 1143 = 1211$
Sistema D	$18,4 + 247 + 4/15 \cdot 258 + 2/5 \cdot 552 + 1/3 \cdot 1143 = 936$
Sistema E	$18,4 + 247 + 2/15 \cdot 258 + 1/5 \cdot 552 + 2/5 \cdot 1143 + 4/15 \cdot 2322 = 1486$
Sistema EC	$18,4 + 247 + 4/15 \cdot 258 + 2/15 \cdot 552 + 7/15 \cdot 1143 + 2/15 \cdot 2322 = 1250$

Haciendo referencia a la tabla 5.7, con todos los MFLOPS calculados, se determina la complejidad o carga computacional de cada sistema en la tabla 5.8.

5.7.2 Comparación de los sistemas A – E considerando las variables de salida

Comparación entre sistemas adaptivos y no adaptivos

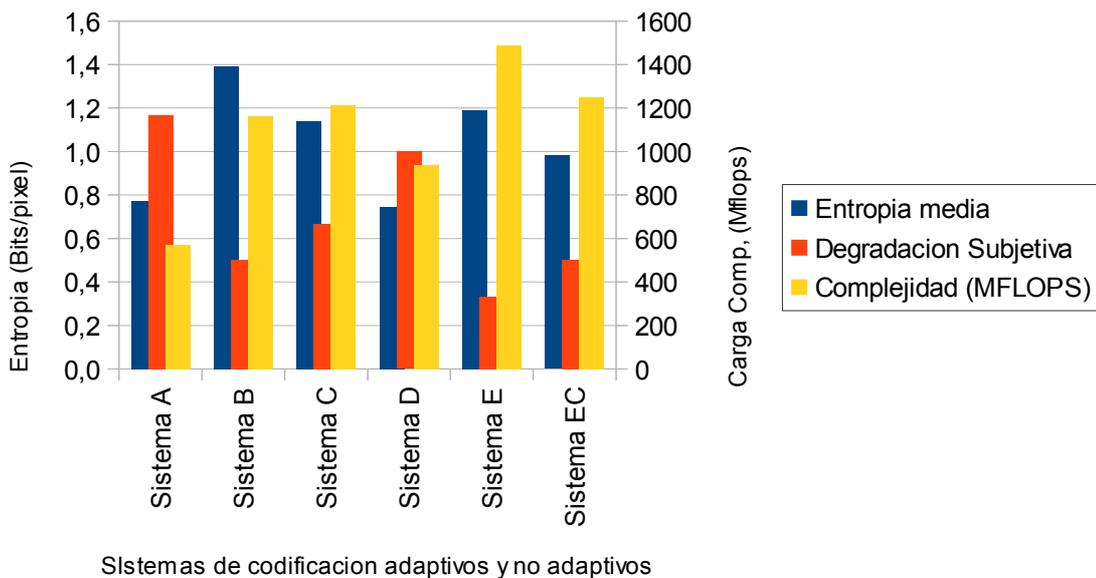


Figura 5.7: Comparación entre sistemas adaptivos y no adaptivos empleando tres parámetros de comparación.

De acuerdo a lo calculado en la tabla 5.8, en la figura 5.7 se muestra la relación entre las variables mencionadas: entropía, calidad subjetiva y carga computacional. De las mediciones y observaciones hechas, se concluye los siguientes puntos:

- La carga computacional determina la velocidad del sistema de codificación. Un sistema de codificación con una elevada carga computacional generará mayor retardo en la transmisión de las imágenes y/o el sobrecalentamiento del sistema, y viceversa.
- Una mayor carga computacional en un sistema de codificación no está determinado por la cualidad adaptiva que éste posea, como se observa cuando se compara el **sistema B** y el **sistema D**. En este caso, el sistema B es no adaptivo y tiene 1161 Mflops, lo cual indica que es más complejo que el sistema D que es adaptivo y presenta sólo 936 Mflops.
- Cuando se compara los sistemas E y EC se tiene que ambos sistemas son adaptivos e implementan 4 cuantizadores iguales para codificar al mismo grupo de imágenes, sin embargo, el sistema E es más complejo que el sistema EC. Esto se debe al punto de operación que determina la relación lineal entre el coeficiente de correlación de las imágenes y los bits de cuantización de las mismas. Dicho de otro modo, la complejidad del sistema de codificación diseñado es mayor cuando las imágenes son seleccionadas para cuantizarlas con mayores bits de cuantización.
- Para aplicaciones que requieran menor latencia (como videoconferencias), es posible diseñar un sistema lineal adaptivo que clasifique las imágenes para cuantizarlas con menos bits de cuantización, agilizando así, la transmisión de las mismas; no obstante, la calidad de las imágenes empobrece.

5.7.3 Relación entre las variables de entrada y salida del sistema

Hasta aquí, las variables de entrada y salida que se consideran son: El **coeficiente de correlación** y los **bits de cuantización** como variables de entrada y la **entropía**, la **calidad subjetiva** y la **carga computacional** como variables de salida. Observando los datos de salida obtenidos, se pueden formular las siguientes conclusiones:

- En un sistema **no adaptivo**, si 'r' aumenta o decrece, entonces: H se mantiene, la calidad se mantiene y la complejidad del sistema es constante.
- En un sistema **adaptivo**, si 'r' varía, entonces 'n' varía en proporción directa a la

variación de 'r'. Entonces 'n' es una variable dependiente de 'r'.

- Cuando 'n' varia, suele variar 'H' y la calidad subjetiva de las imágenes. Sin embargo esta variación de 'n' no determina un cambio en la complejidad del sistema.
- La complejidad o carga computacional del sistema es independiente del número de cuantizadores que se implementen.
- La carga computacional depende de la relación entre el coeficiente de correlación 'r' y los bits de cuantización 'n', existente en el sistema de codificación diseñado.

Si 'r' y 'n' tienen una relación lineal de acuerdo a la ecuación (20), entonces la complejidad del sistema está en función del factor de compensación 'fc' definido en la página 51 de la sección 5.4.2.

$$n(r) = \lceil [k_2 \cdot r + c_2] \rceil \pm |fc| \quad (29)$$

Una regla empírica muestra que si 'fc' es (+) entonces la complejidad se incrementa porque una mayor cantidad de imágenes se cuantizarán con bits mayores. Por otra parte, si 'fc' es (-) entonces la complejidad es menor debido a que la mayoría de las imágenes se cuantizarán con bits menores (para cualquier sistema lineal adaptivo diseñado con la ecuación 29).

CAPÍTULO 6

EVALUACIÓN DE RESULTADOS

6.1 Evaluación y conclusiones de las simulaciones

Correlación entre imágenes

- El coeficiente de correlación de las imágenes guarda relación directa con la reducción entrópica cuando se usa la codificación predictiva.
- Por lo general, las imágenes más decorrelacionadas incrementan más la entropía de las imágenes diferenciales en el sistema de transmisión.
- En imágenes con los mismos bits de cuantización, cuando la correlación tiende a 1, el PSNR es mayor, y cuando la correlación tiende a 0, el PSNR es menor.

Códigos de longitud variable VLC

- En imágenes con distribución de píxeles no uniforme, la codificación Huffman resulta efectiva.
- La longitud de código promedio, no es inferior en ningún caso, a la entropía de las imágenes (fuente). Se demostró experimentalmente el cumplimiento del **teorema de Shannon** para la codificación de la fuente.
- Para lograr una mayor eficiencia de codificación, es necesario **reducir la entropía** de la fuente (reducir la información de las imágenes).
- La codificación Huffman, sin importar la eficiencia sin codificación, logra un nivel estable de eficiencia promedio de 99.6 %.
- Imágenes muy claras o imágenes muy oscuras tendrán, por lo general, baja entropía. De acuerdo a la conclusión anterior, la ganancia de codificación VLC para ese tipo de imágenes es elevado.

Histograma de una imagen

- El histograma permite observar rápidamente la distribución de probabilidades de los píxeles en un cuadro o imagen. Gracias a ello, se puede deducir la entropía de una imagen (y por ende la tasa de bits) a la salida del codificador, observando la forma que tiene el histograma.
- Dependiendo de las características de las imágenes de entrada (su distribución de probabilidad de píxeles o su histograma), se obtendrán mayores o menores porcentajes de reducción de la tasa de bits a la salida.

Cuantización

- El intervalo o paso de cuantización determina la variación de la tasa de bits y la distorsión en las imágenes.
- Cuando se usa un cuantizador no uniforme en el sistema propuesto, la entropía diferencial de la imagen suele ser mayor (ver figura 3.10). Esto sucede porque los niveles de salida del cuantizador no coinciden con 0. Usando un cuantizador uniforme en el sistema propuesto, la entropía de la figura 3.2, (el leopardo) es 2.42 bit/píxel y utilizando un cuantizador no uniforme es 3.40 bit/píxel (para $n=6$ bit).

6.2 Evaluación subjetiva de las imágenes

Tal como se mencionó en la sección 2.6.1, existen varios métodos de evaluación subjetiva de la calidad de las imágenes, por lo tanto, en esta sección se describe la aplicación de dichos métodos para comparar los resultados objetivos y subjetivos, y tener un mejor criterio para validar la hipótesis de la investigación.

6.2.1 Fundamentación y descripción de la metodología

La recomendación ITU-R BT.500 sugiere 4 tipos generales de metodologías para evaluar la calidad subjetiva de las imágenes en un sistema^[17], no obstante, deja libre la posibilidad de crear metodologías específicas para sistemas diseñados que mejoran aspectos puntuales de la calidad.

En ese sentido se seleccionaron 2 metodologías de evaluación de acuerdo a la sección 6.2 de la recomendación, las cuales se refieren a los métodos de comparación de

estímulos (stimulus comparisson). Estas son:

- DSCS (3 escalas)
- DSCS (9 escalas)

Se seleccionaron estos métodos debido a su fácil implementación, a su corta duración de la evaluación (ya que las pruebas con otras metodologías duran cerca de 1 hora por sesión de evaluación) y finalmente, debido a la peligrosidad de obtener resultados erróneos cuando se usan otras metodologías (ya que estas tienen métodos de evaluación complejos y más susceptibles a cometer errores).

A continuación se describen las condiciones y los datos técnicos de los dos tipos de evaluación antes mencionados.

6.2.2 DSCS – Double Stimulus Comparative Scale (3 escalas)

Los métodos de comparación de estímulos se usan para evaluar nuevos sistemas de vídeo diseñados. Son considerados por su **rapidez al momento de la evaluación** y por la **sencilla manera de procesar la información recabada**. De las 3 maneras de efectuar la evaluación, se considera la siguiente^[17]:

Métodos de opinión usando adjetivos para la categorización: Usa adjetivos de comparación para la evaluación. Los adjetivos empleados pueden reportar la existencia de diferencias perceptibles (Igual que, diferente a, de mayor calidad que) o el grado de las diferencias percibidas (más o menos diferente, mucho peor a, casi igual que). Los datos analizados dependen del modo en el cual la información fue recolectada (cantidad de adjetivos empleados y/o valor cuantitativo que representan los adjetivos).

En la primera evaluación efectuada se emplean 3 adjetivos de comparación, los cuales tienen dos valores cuantitativos (0, 1), como se muestra en la tabla 6.1.

Tabla 6.1. Equivalencia entre adjetivos y rangos cuantitativos.

Rango o valor	Adjetivo empleado
1	Imagen A MEJOR que Imagen B
1	Imagen A PEOR que Imagen B
0	Imagen A IGUAL a Imagen B

Datos técnicos de la evaluación

Método de evaluación: DSCS (recomendación ITU BT-500 modificada).

Nº de imágenes evaluadas:	15 (con diferentes niveles de calidad)
Nº de encuestados:	24 personas
Formato de imágenes:	JPG
Resolución de las imágenes:	480 x 640 pixeles
Relación de aspecto:	4:3
Tiempo de exposición máximo:	8 seg.
Escalas de comparación:	3
Modo de exposición:	Aleatorio
Modo de evaluación:	Personal (asistido)
Sistema evaluado:	Sistema EC de la sección 5.4.2.

De acuerdo a la tabla 6.1 se confecciona el modulo de evaluación mostrado en la figura 6.1:

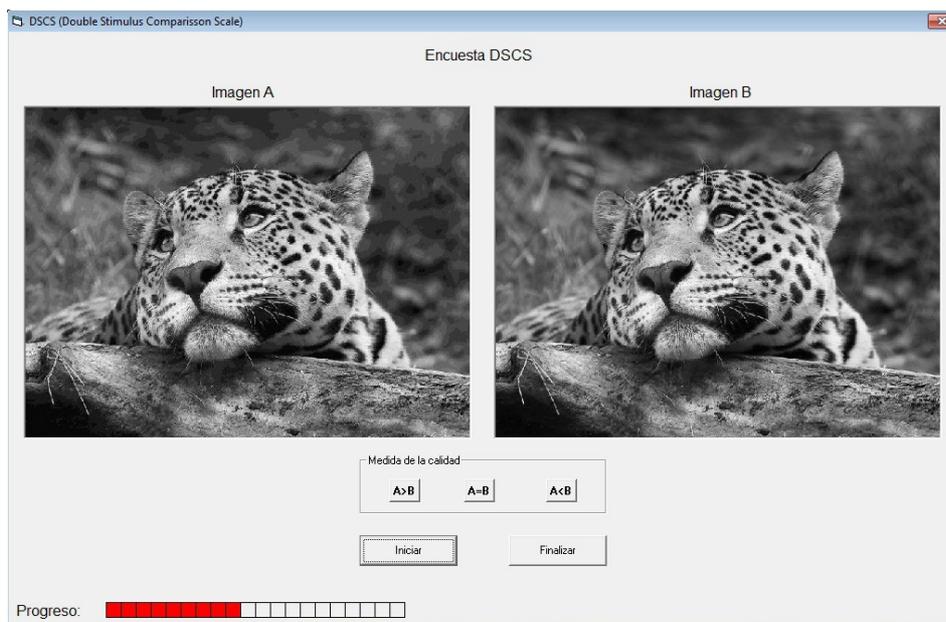


Figura 6.1. Módulo de evaluación DSCS (3 escalas).

Resultados de la evaluación

Los resultados porcentuales de las imágenes evaluadas se resumen a continuación:

$$Deg. Subj. = \frac{1}{24} \sum_{i=1}^{24} r_i = \frac{2060}{24} = 85.82 \%$$

La interpretación de este resultado dicta que un 85.8 % de las imágenes fueron

detectadas con pérdidas perceptibles (mínimas) de la calidad, por las personas evaluadas. De las 15 imágenes, 13 fueron observadas con pérdidas de la calidad y sólo 2 imágenes permanecieron invariables respecto a su calidad pese a que perdieron información en la compresión.

Comparación de los resultados de la evaluación DSCS con los resultados de la sección 5.6

En la sección 5.6, los resultados de la evaluación subjetiva hecha por el autor da 20 % de degradación subjetiva, la cual difiere del resultado obtenido en la encuesta (85.8 %). Esta diferencia se debe a la rigurosidad del método empleado en la evaluación. Los encuestados emitieron una opinión negativa al percibir una diferencia mínima, y muchas veces irrelevante, en la calidad de las imágenes comparadas. Si consideramos diferencias notables hasta un determinado umbral de degradación subjetiva, obtendremos resultados más aproximados a la evaluación hecha por el autor. Este proceso se realiza a continuación.

6.2.3 DSCS – Double Stimulus Comparative Scale (9 escalas)

Esta nueva metodología proporciona mayor flexibilidad y tolerancia al momento de recabar información de los encuestados. El módulo de evaluación presenta 9 escalas de diferenciación, las cuales permiten una evaluación cualitativa más detallada de las imágenes comprimidas. Esto quiere decir que el encuestado diferenciará cambios mínimos en la pérdida de la calidad de las distintas imágenes decodificadas y los resultados obtenidos serán más precisos.

Tabla 6.2. Escalas cuantitativas de la evaluación subjetiva.

Escala	-4	-3	-2	-1	0	1	2	3	4
Adjetivo	Muy diferente	Diferente	Ligeramente diferente	Similar	Igual	Similar	Ligeramente diferente	Diferente	Muy diferente
Umbral	Ptos. Negativos.			Ptos. Positivos.			Ptos. Negativos.		

De acuerdo a la tabla 6.2, si el encuestado califica la diferencia de 2 imágenes en el

rango central de la escala, entonces se obtendrán **puntos positivos** a favor del sistema evaluado (Sistema EC). Por otra parte, si el encuestado califica la diferencia entre dos imágenes en los extremos de la escala, habrá puntos en contra, es decir, se sumarán **puntos negativos** al sistema.

Resultados de la evaluación

Los datos técnicos, análisis cuantitativo y cálculos matemáticos empleados se muestran a continuación.

Datos técnicos de la evaluación

Método de evaluación:	DSCS (de acuerdo a la recomendación ITU BT-500).
Nº de imágenes evaluadas:	15 (con diferentes niveles de calidad)
Nº de encuestados:	22 personas
Tipo de imágenes:	JPG.
Resolución de las imágenes:	480 x 640 pixeles
Relación de aspecto:	4:3
Tiempo de exposición máximo:	9 seg.
Escalas de comparación:	3
Modo de exposición:	Aleatorio
Modo de evaluación:	Personal (asistido)
Sistema evaluado:	Sistema EC de la sección 5.4.2.

Módulo de evaluación empleado

En la figura 6.2, se observa el incremento de las escalas en las medidas de calidad, comparado con el módulo de la figura 6.1.

Resultados e interpretación de la evaluación

El promedio de los resultados porcentuales de las imágenes evaluadas se resumen a continuación:

$$Deg. Subj. = \frac{1}{22} \sum_{i=1}^{22} r_i = \frac{441.6}{22} = 20.07 \%$$

Haciendo un análisis más detallado, se presenta las siguientes observaciones:



Figura 6.2. Módulo de evaluación DSCS con 9 escalas de comparación.

a) Se observa que el promedio de imágenes que pierde calidad de manera significativa o relevante es $0.2 \times 15 = 3$ imágenes. Este resultado coincide con la evaluación hecha por el autor en la página 51 del cuerpo de la investigación.

b) La fórmula empleada por el programa de evaluación calcula el promedio ponderado de las escalas cuantitativas de la Tabla 6.2 por cada encuesta realizada, así:

$$P = \sum_{i=1}^N \frac{n_i \times Escala}{N \times \text{Max}(Escala)} \quad (30)$$

Donde:

- 'P' es el resultado porcentual que obtiene cada encuestado.
- 'N' es el número total de imágenes evaluadas.
- 'n_i' es la n-ésima imagen evaluada por el encuestado y
- la **Escala** se define numéricamente en la tabla 6.2 para cada adjetivo empleado en la evaluación.

Si el promedio porcentual de todos los participantes de la encuesta DSCS 2 es 20 %, entonces, reemplazando este resultado en la ecuación 30, da:

$$P = \sum_{i=1}^N \frac{n_i \times Escala}{15 \times 4} = 0.2$$

$$\sum_{i=1}^N n_i \times Escala = 12$$

$$Escala \times \sum_{i=1}^N n_i = 12$$

Si se supone que todos los encuestados calificaron todas las imágenes, el resultado de la sumatoria en la última ecuación es 15 y la escala promedio de todos los encuestados será:

$$Escala = \frac{12}{15} = 0,8 \approx 1$$

Finalmente, si se hace referencia tabla 6.2, el adjetivo que corresponde a la escala promedio de '1' es "Similar". Por lo tanto, bajo los términos presentados en el módulo de evaluación, se concluye que: Las personas evaluadas percibieron las imágenes decodificadas **similares** a las originales.

Esta última afirmación valida la hipótesis desde el punto de vista experimental.

6.2.4 Conclusiones de los métodos de evaluación subjetiva empleados

Se presentan los resultados de las tres formas de evaluación subjetiva empleadas para calificar al sistema de codificación ADPCM (Sistema EC). El sistema en mención cumple todos los requerimientos de diseño impuestos en la sección 5.3.2:

- $H < 1$ bit/pixel
- PSNR > 34 dB
- Deg. Sub. < 40 %

Las dos primeras condiciones fueron comprobadas analíticamente. La última condición fue comprobada con 2 de las 3 evaluaciones subjetivas presentadas. Los resultados se resumen en la figura 6.3.

En la gráfica se observa que la evaluación DSCS 9 y la evaluación realizada por el autor dan como resultado 20%. Este resultado satisface la última condición de diseño (Deg. Sub. < 40 %)

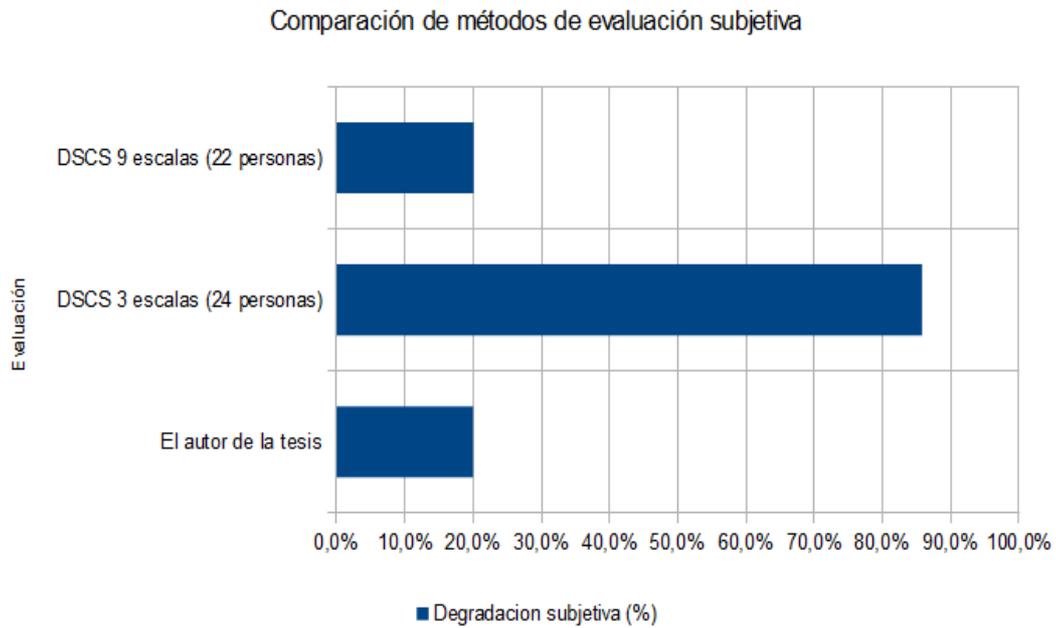


Figura 6.3. Comparación de los métodos de evaluación para el sistema EC.

Por lo tanto, queda demostrado que el sistema estudiado y diseñado en esta investigación (Sistema EC) logra el balance deseado entre la tasa de bits transmitidos y pérdida de la calidad subjetiva de las imágenes.

En la tabla 6.3 se resume las características del sistema EC comparadas con las restricciones de diseño. Los resultados de la tabla confirman lo concluido en la presente sección.

Tabla 6.3. Comparación entre los parámetros del sistema EC y las restricciones de diseño de la sección 5.4.

	Parámetros de diseño requeridos	Parámetros del sistema EC.
Ehtropía	< 1 bit/pixel	0,98 bit/pixel
PSNR	34 dB \pm 1	33,6 dB
Degradación subjetiva	< 40 %	20 %
MFLOPS	-	1250

6.3 Recomendaciones

- Es factible analizar variedad de imágenes con los programas de MATLAB propuestos en este trabajo.
- En los cálculos hechos se usa la correlación espacial. Cuando la imágenes tienen

movimiento es preferible usar la correlación temporal para ajustar el paso del cuantizador.

- En el sistema propuesto, es posible considerar varias métricas para variar el paso de cuantización. Una métrica propuesta puede ser **la covarianza** entre las imágenes^[4].
- Existen numerosas técnicas de codificación de la fuente aplicadas a la compresión de video. La elección o el desarrollo de alguna técnica en particular dependerá de las circunstancias o entornos que se presenten. Algunas de las nuevas técnicas de codificación se expondrán en la sección 6.3.
- Sólo se considera un reducido grupo de imágenes (15) para las simulaciones. Para obtener mejores resultados relacionados a la calidad subjetiva, el sistema debe ser ajustado con un mayor universo de imágenes.

6.4 Futuras investigaciones

Las futuras investigaciones están orientadas a la transmisión de video por internet y aplicaciones 4G, como videoconferencias entre 2 ó más usuarios móviles (con buena calidad y menos retardo)^[8].

Los nuevos algoritmos de compresión de video, que buscan cumplir los anteriores objetivos, requieren ejecutar millones de operaciones por segundo. Hoy, el avance tecnológico hace posible contar con equipos más veloces que permiten ejecutar operaciones complejas rápidamente.

Ejemplos de esas investigaciones son mencionadas a continuación:

- Eficiente asignación de bits para un conjunto de cuantizadores arbitrarios.
- Nuevo esquema de control de tasa usando modelos R-D cuadráticos.
- Algoritmo de optimización R-D rápido para codificación por movimiento compensado de video.
- Nuevo algoritmo de búsqueda diamante para rápida coincidencia en la estimación de video.
- Detalles técnicos de Wavelet y TCQ (Trellis Coded Quantizer).
- Cuantización asintóticamente eficiente.
- CELP (Code excited linear prediction), donde se usa algoritmos de cuantización avanzados para la codificación de la voz.

Estos y otros temas de investigación se mencionan en la referencia bibliográfica^[8] del

capítulo 7.

Del mismo modo, otros autores sugieren los siguientes temas:

- Uso de técnicas de segmentación para predicción por movimiento compensado^[2].
- Empleo de la codificación aritmética para mejorar la eficiencia de codificación VLC^[2].
- Estimación de coeficientes de transformación a través de las transformadas Wavelet, Hadamard, DST, KLT, Hartley y Slant^[4].

Por otra parte, el autor de esta investigación propone algunos temas que surgieron a partir de la misma:

- Codificación ADPCM usando otras métricas para determinar los bits de cuantización como la 'Pc' de la FFT2 de una imagen (ver sección 3.2).
- Codificación ADPCM aplicada a partes (bloques) de una imagen. Esto se debe a que se tienen imágenes compuestas por áreas de alta y baja correlación.
- Determinación de los umbrales de las regiones de cuantización usando relaciones no lineales (exponencial o logarítmica) entre 'n' y 'r' (ver sección 5.4).

CONCLUSIONES

Se verificó mediante las simulaciones y diseños efectuados que las técnicas de codificación MPEG reducen la tasa de bits a la salida del codificador. Una medida cuantitativa de esa reducción es la entropía de una imagen.

Bajo los términos de la presente investigación, se propuso y diseñó un nuevo sistema de codificación adaptivo ADPCM (sistema EC), con el cual se logró:

- 1) reducir la entropía de las imágenes,
- 2) mantener la calidad subjetiva de las imágenes,
- 3) minimizar la diferencia subjetiva entre las imágenes original y decodificada.

Estos 3 puntos están sujetos a condiciones de diseño preestablecidas como: La entropía máxima permitida y la calidad mínima aceptable.

La minimización de la diferencia subjetiva de las imágenes analizadas se comprobó con 2 de las 3 pruebas de evaluación subjetiva de comparación de estímulos propuestas por la ITU. En ellas, se observa que la degradación subjetiva de las imágenes decodificadas (20%), no supera el límite impuesto en las condiciones de diseño (40 %).

El uso del sistema de codificación adaptivo propuesto "Sistema EC" (y en general, todos los sistemas adaptivos del capítulo 5), permiten tener una mayor versatilidad para adecuar los parámetros de salida del sistema (tasa de bits y calidad subjetiva y objetiva) a las condiciones de diseño.

Finalmente, se demostró que es posible obtener un mejor balance entre la tasa de bits, la calidad de las imágenes decodificadas y la carga computacional de los sistemas, gracias a los codificadores adaptivos propuestos en esta investigación.

BIBLIOGRAFÍA

- [1] FISCHER W., "Digital Video and Audio Broadcasting Technology", 2da Ed., Springer Series, Alemania 2008.
- [2] ARNOLD J., FRATER M., P. M., "Digital Television Technology and Standards", Wiley & Sons, USA 2007.
- [3] BENOIT H., "Digital Television", 3ra Ed., Elsevier, USA 2008.
- [4] THYAGARAJAN K. S., "Still image and video compression with MATLAB" 1ra Ed., Wiley & Sons, Singapur 2011.
- [5] PROAKIS J., SALEHI, M., "Comunication Systems Engineering", 2da Ed., Prentice Hall, USA 2002.
- [6] COUCH L., "Sistemas de comunicación digitales y analógicos", 7ma Ed., Pearson-Prentice Hall, México 2008.
- [7] CARRION P. O., Tesis: "Análisis y diseño de un sistema de transmisión por RF y conversión de la señal de S-Video y Video Compuesto" UNSA, Arequipa 2009.
- [8] Editors: ANGELIDES M., AGIUS H., "THE HANDBOOK OF MPEG APPLICATIONS", 1ra Ed., Wiley & Sons, Reino Unido 2011.
- [9] RICHARDSON IAIN. E.G., "H.264 and MPEG-4 Video compression", 1ra Ed., Wiley & Sons, England 2003.
- [10] GONZALEZ R., WOODS R., "Digital Image Processing", 2da Ed., Prentice Hall, New Yersey 2002.
- [11] G. BOSTLEMANN, "A high quality DPCM system for video telephone signals using 8 Mb/s." Nach. Tech, 1974.
- [12] W. ZSCHUNKE, "DPCM picture coding with adaptive prediction." IEEE Trans. Commun., 1977.

Páginas de internet

- [13] <http://en.wikipedia.org/wiki/ISDB>

- [14]** http://en.wikipedia.org/wiki/H.262/MPEG-2_Part_2
- [15]** Image Processing Toolbox User's Guide, MATLAB 7.0
- [16]** RECOMMENDATION ITU-R BT.601-5, "Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios". 1995.
- [17]** RECOMMENDATION ITU-R BT.500-13, "Methodology for the subjective assessment of the quality of television pictures". 2012.
- [18]** M. PINSON, S. WOLF, "Comparing subjective video quality testing methodologies", ITS, NTIA, July 2003.

ANEXOS

ANEXO A

Código de programas en MATLAB

program_1.m

```
% Calculo de ancho de banda, correlacion doble y
% entropia de una imagen.
clear
close all;

% Adquisicion de la imagen
I=imread('D:\TESINA UNI\Prueba SD\luna.jpg');
Y=.3*I(:, :, 1)+.59*I(:, :, 2)+.11*I(:, :, 3);
figure(1);imshow(Y)

% Calculo de la entropia
for p=0:255          %Son 2^8= 256 pixeles
    S=(Y==p);
    y(p+1)=sum(sum(S));
end
% probabilidades a priori
N=size(Y,1)*size(Y,2);
pp=y./N;

H=0;
for i=1:256          %Son 2^8= 256 pixeles
    if pp(i)~=0      %evita el log(0)
        H=H-pp(i)*log2(pp(i));
    end
end
disp(['H=' num2str(H)])

% Calculo de la correlación
b=1;                %diferencia de 1 pixel
Yo=Y(1+b:end-b,1+b:end-b);
figure(3) imshow(uint8(Yo)) %vemos el recorte que queremos comparar
% datos estadísticos de la imagen a comparar
mu_a=round(mean(mean(Yo)));
A=Yo-mu_a;
cu_A=sum(sum(A.^2));
% calculo del coeficiente de correlacion de b*b figuras
k=1;
mat_r=[];
for y=0:2*b
```

```

    for x=0:2*b
        Yn=Y(y+1:end-2*b+y,x+1:end-2*b+x);
        mu_b=round(mean(mean(Yn)));
        B=Yn-mu_b;
        cu_B=sum(sum(B.^2));
        num=sum(sum(A.*B));
        den=sqrt(cu_A*cu_B);
        r(k)=num/den;
        k=k+1;
    end
    mat_r=[mat_r;r];
    k=1;
end
r=min(min(mat_r)); disp(['r=' num2str(r)])

```

program_2.m

```

% codificacion Huffman para una imagen
clear
close all

% lee una imagen de tamaño 640 x480
I=imread('D:\TESINA UNI\Prueba SD\luna.jpg');
Y=.3*I(:, :, 1)+.59*I(:, :, 2)+.11*I(:, :, 3);
figure(1); imshow(Y)

% calculo de las probabilidades a priori
for p=0:255 %Son 2^8= 256 pixeles
    S=(Y==p);
    y(p+1)=sum(sum(S));
end
figure(2); imhist(Y)
N=size(Y,1)*size(Y,2);
pp=y./N;

% calculo de los codigos VLC a partir de las prob.
pixels=[0:255];
D=huffmandict(pixels,pp);
avglen=0;
for l=1:256
    avglen=avglen+pp(l)*length(D{1,2});
end

H=0; %Calculo de la entropia
for i=1:256 %Son 2^8= 256 pixeles
    if pp(i)~=0 %evita el log(0)
        H=H-pp(i)*log2(pp(i));
    end
end
maxlen=8;
% comparacion de la eficiencia de codificacion
disp(['Eficiencia fixed = ' num2str(100*H/maxlen) '%'])
disp(['Eficiencia VLC = ' num2str(100*H/avglen) '%'])

```

program_3.m

```

% codificacion predictiva de imagenes fijas
clear
close all

```

```

% lee una imagen de tamaño 640x480 (SDTV)
I=imread('D:\TESINA UNI\Prueba SD\luna.jpg');
Y=.3*I(:, :, 1)+.59*I(:, :, 2)+.11*I(:, :, 3);
figure(1);imshow(Y)

% Calculo de la prediccion
w=1; pix_m=128;
Y_h=Y(:, 1:end-w);
Y_h=[pix_m*ones(size(Y,1),w) Y_h];
Y_v=Y(1:end-w, :);
Y_v=[pix_m*ones(w, size(Y,2));Y_v];
preY=.5*Y_h+.5*Y_v; %promedio de H y V
% diferencia
Dif=double(Y)-double(preY);
figure(2);imshow(uint8(Dif+pix_m-1)) %centrada en gris=127

% histograma de la diferencia (Dif)
q=min(min(Dif));
for p=min(min(Dif)):max(max(Dif))
    S=(Dif==p);
    y(p-q+1)=sum(sum(S));
end
x=min(min(Dif)):max(max(Dif));
figure(3);stem(x,y,'r. ');axis([-128 128 0 max(y)+1])
% Entropia diferencial
N=size(Dif,1)*size(Dif,2);
pp=y./N;
H_d=0;
for i=1:length(y) %varios niveles de pixeles
    if pp(i)~=0 %evita el log(0)
        H_d=H_d-pp(i)*log2(pp(i));
    end
end
disp(['H(Dif)= ' num2str(H_d)])

% Codificacion Huffman
pixels=[1:length(y)];
D=huffmandict(pixels,pp);
% para seleccionar elementos de celdas usar {}
avglen=0;
for l=1:length(y)
    avglen=avglen+pp(l)*length(D{l,2});
end
% Eficiencia DPCM
maxlen=ceil(log2(length(y)));
disp([maxlen avglen]);
disp(['Eficiencia DPCM = ' num2str(H_d/maxlen)])
% Eficiencia DPCM con Huffman
disp(['Eficiencia DPCM + VLC = ' num2str(H_d/avglen)])

```

program_4.m

```

% Sistema de codificacion con cuantizacion adaptiva
clear
close all

% lee una imagen de tamaño 640x480 (SDTV)
I=imread('D:\TESINA UNI\Prueba SD\leopardo.jpg');
Y=.3*I(:, :, 1)+.59*I(:, :, 2)+.11*I(:, :, 3);

```

```

figure(1);imshow(Y)

% calculo Coef de correlacion r
b=1; %diferencia de 1 pixel
Yo=Y(1+b:end-b,1+b:end-b);
% datos estadisticos de la imagen a comparar
mu_a=round(mean(mean(Yo)));
A=Yo-mu_a;
cu_A=sum(sum(A.^2));
% calculo del coef. de correlacion de b*b figuras
k=1;
mat_r=[];
for y=0:2*b
    for x=0:2*b
        Yn=Y(y+1:end-2*b+y,x+1:end-2*b+x);
        mu_b=round(mean(mean(Yn)));
        B=Yn-mu_b;
        cu_B=sum(sum(B.^2));
        num=sum(sum(A.*B));
        den=sqrt(cu_A*cu_B);
        r(k)=num/den;
        k=k+1;
    end
    mat_r=[mat_r;r];
    k=1;
end
r=min(min(mat_r)); %considerando el minimo valor
disp(['r=' num2str(r)])

% Ajuste del paso de cuantizacion con el coef. de
% correlacion (dos niveles n = 6 y n = 4)
r_umbral=0.9;
if r<r_umbral %imagenes con baja correlacion
    z=4; %cuantizador grueso
elseif r>=r_umbral %imagenes con alta correlacion
    z=5; %cuantizador fino
end
disp(['n=' num2str(z)])

% condiciones iniciales del sistema Y_pre=0 y n=8
% Datos de cuantizacion uniforme
n=z; %bits de cuantizacion < 8
m=2^n; %niveles de cuantizacion
d=round(255/m); %niveles de decision
dm=round(d/2);

% niveles de salida
Y1=zeros(size(Y,1),size(Y,2));
for u=1:m
    ri(u)=dm+(u-1)*d;
    umbral=(Y>=(u-1)*d)&(Y<u*d);
    Y1=Y1+double(ri(u)*umbral);
end

% Prediccion con diferencia de 1 pixel usando Y1
w=1; pix_m=128; %diferencia de w pixeles
Y_h=Y1(:,1:end-w);
Y_h=[pix_m*ones(size(Y,1),w) Y_h];
Y_v=Y1(1:end-w,:);
Y_v=[pix_m*ones(w,size(Y,2));Y_v];

```

```

Y_pre=.5*Y_h+.5*Y_v;      %promedio de H y V

% el segundo cuadro es el mismo que el primero
Y2=double(Y);
% diferencia
Y_dif=double(Y2)-double(Y_pre);
figure(2);imshow(uint8(Y_dif+pix_m-1))

% cuantizacion de la diferencia (uniforme)
% datos
n=z;      %bits de cuantizacion < 9
m=2^n-1;  %niveles de cuantizacion
fase=[-1 0 -1 0 7 3 1 0 0]; %offset para r=0
desp=-255+fase(n);
d=round(511/m); %niveles de decision
dm=round(d/2);
% cuantizacion
Y_dq=zeros(size(Y_dif,1),size(Y_dif,2));
for u=1:m
    rd(u)=(u-1)*d+dm+desp; %salidas del cuantizador
    umbral=(Y_dif>=(u-1)*d+desp) & (Y_dif<u*d+desp);
    Y_dq=Y_dq+double(rd(u)*umbral);
end

% Grafica de la diferencia cuantizada
figure(3)
imshow(uint8(Y_dq+pix_m-1))

% Calculo de la entropia de la dif cuantizada
q=min(min(Y_dq));
for p=min(min(Y_dq)):max(max(Y_dq))
    S=(Y_dq==p);
    y(p-q+1)=sum(sum(S));
end

N=size(Y_dif,1)*size(Y_dif,2);
pp=y/N;

Hdq=0;
for i=1:length(pp) %o y
    if pp(i)~=0 %evita el log(0)
        Hdq=Hdq-pp(i)*log2(pp(i));
    end
end
disp(['H = ' num2str(Hdq)])

% Recuperacion de la imagen en la decodificacion
% Cuantizacion inversa y suna con la prediccion
nn=z; %envio del valor del cuantizador
Y_r=double(Y_dq)+double(Y_pre);
figure(4);imshow(uint8(Y_r))
imwrite(uint8(Y_r),'decoded.jpg')

% Calculo del PSNR usando Y2 y Y_r
% MSE
Dif2=(round(Y2)-round(Y_r-1)).^2;
MSE=sum(sum(Dif2))/N;
PSNR=vpa(10*log10(255^2/MSE),3)

```

Programas efectuados para cada uno de los bloques del sistema propuesto de la figura 5.1.

Adquisición del conjunto o juego de imágenes (Y)

```
% lee una imagen de tamaño 640x480 (SDTV)
% y transforma la imagen en escala de grises
for N=1:15 %son 15 imagenes de prueba
I=imread(['D:\TESINA UNI\Prueba SD\' num2str(N) '.jpg']);
Y=.3*I(:, :, 1)+.59*I(:, :, 2)+.11*I(:, :, 3);
figure(1);imshow(Y)
Y=double(Y); ...
```

Bloque: estadísticas de la imagen (cálculo del coef. de correlación 'r') y selección de los bits de cuantización ('n')

```
% datos estadisticos de la imagen "original"
mu_a=round(mean(mean(Y)));
A=Y-mu_a;
cu_A=sum(sum(A.^2));

% Obtencion de la imagen con marco 0
Y_mar=[zeros(size(Y,1),1) Y zeros(size(Y,1),1)];
Y_mar=[zeros(1,size(Y,2)+2);Y_mar;zeros(1,size(Y,2)+2)];

b=1;
% calculo de la correlacion de la imagen

k=1;
mat_r=[];
for y=0:2*b
    for x=0:2*b
        Yn=Y_mar(y+1:end-2*b+y,x+1:end-2*b+x);
        mu_b=round(mean(mean(Yn)));
        B=Yn-mu_b;
        cu_B=sum(sum(B.^2));
        num=sum(sum(A.*B));
        den=sqrt(cu_A*cu_B);
        r(k)=num/den;
        k=k+1;
    end
    mat_r=[mat_r;r];
    k=1;
end
rr=[rr min(min(mat_r))]; %se escoge el minimo valor de r
N=N+1;
end
rr ...

% Salidad de matlab

rr=

    0.7685    0.9171    0.7743    0.7260    0.8758    0.9349    0.9852    0.7182    0.8917    0.8972    0.9509    0.9552    0.9794
    0.8549    0.8119

% Diseño del sistema con cuantizacion adaptiva
% mediante la formula 18 planteada en la tesis
% Selecccion de los maximos y minimos pasos de cuantizacion
```

```

n1=3;
n2=6;
comp=+0.00;
rmin=min(rr)+comp; rmax=max(rr)+comp;

% Calculo de las constantes c y d
c=(n2-n1)/(rmax-rmin);
d=n1-c*rmin; % o d1=n2-c*rmax;

%asignacion de las imagenes a los cuantizadores

for ri=1:length(rr)
    n(ri)=round(c*rr(ri)+d);
end
n

% Salidad de matlab

n =
    3  5  3  3  5  5  6  3  5  5  5  5  6  4  4

```

Bloque cuantizador y cuantizador inverso $Q(n)$ y $Q^{-1}(n)$

```

% cuantizacion de la diferencia (uniforme)
% datos
n=4; %bits de cuantizacion < 9
m=2^n-1; %niveles de cuantizacion
fase=[-1 0 -1 0 7 3 1 0 0]; %offset para r=0
desp=-255+fase(n);
d=round(511/m); %niveles de decision
dm=round(d/2);
% cuantizacion
Y_dq=zeros(size(Y_dif,1),size(Y_dif,2));
for u=1:m
    rd(u)=(u-1)*d+dm+desp; %salidas del cuantizador
    umbral=(Y_dif>=(u-1)*d+desp) & (Y_dif<u*d+desp);
    Y_dq=Y_dq+double(rd(u)*umbral);
end

% Grafica de la diferencia cuantizada
figure
imshow(uint8(Y_dq+pix_m-1))...

```

Bloque: predictor lineal (2D) y señal diferencial ('Y_dif')

```

% Prediccion con diferencia de 1 pixel usando Y1
w=1; pix_m=128; %diferencia de w pixeles
Y_h=Y1(:,1:end-w);
Y_h=[pix_m*ones(size(Y,1),w) Y_h];
Y_v=Y1(1:end-w,:);
Y_v=[pix_m*ones(w,size(Y,2));Y_v];
Y_pre=.5*Y_h+.5*Y_v; %promedio de H y V

% el segundo cuadro es el mismo que el primero
Y2=Y;
% diferencia
Y_dif=double(Y2)-double(Y_pre);

```

```
figure
imshow(uint8(Y_dif+pix_m-1))...
```

Bloque: codificación Huffman (VLC)

```
% Calculo del numero de pixels con determinada intensidad
% ver funcion imhist
for p=0:255          %Son 2^8= 256 pixeles
    S=(Y==p);
    y(p+1)=sum(sum(S));
end
x=0:255;
figure(2)
stem(x,y,'k. ');axis tight

% calculo de las probabilidades a priori
N=size(Y,1)*size(Y,2);
pp=y./N;

% calculo de los codigos VLC a partir de las prob.
% usando la funcion huffmandict
pixels=[0:255];
D=huffmandict(pixels,pp)...
```

Bloque del decodificador (figura 5.1b)

```
% Recuperacion de la imagen en la decodificacion
Y_r=double(Y_dq)+double(Y_pre);
figure
imshow(uint8(Y_r))
%guarda en un archivo la imagen decodificada para
%su comparacion con la original
imwrite(uint8(Y_r),['D:\imagen\dec_' num2str(N) '.jpg'])
end
```

ANEXO B

Código del módulo de evaluación

El siguiente programa esta redactado en Visual Basic 6.0 y consta de 2 formularios o ventanas. La ventana principal muestra las imágenes para ser evaluadas y las escalas de calificación que el encuestado usa para evaluar las diferencias subjetivas entre las dos imágenes presentadas, mientras que la segunda ventana muestra los resultados de la comparación hecha.

```
'=====
'Formulario Principal
'=====
'Programa de visuaizacion de imagenes para la evaluacion subjetiva
'de imagenes originales y decodificadas empleando el sistema ADPCM
'EC propuesto en la investigacion

Option Explicit
Dim ImageA As Picture, ImageB As Picture
Dim path1 As String, path2 As String, pathend As String, N As Single
Public k As Single

Private Sub Command1_Click()
'Cuando se inicia la encuesta se habilitan las opciones de seleccion
'y se muestran las imagenes

Command1.Enabled = False

'Luego se habilita el temporizador para configurar el tiempo de
exposicion
'de las imagenes
Timer1.Enabled = True
Timer1.Interval = 9000           'Temporizador a 9 seg.

'Variable para indexar las imagenes
k = 2

'Variables de recoleccion de datos de salida
Data1 = 0: Data2 = 0: Data3 = 0: Data4 = 0: Data5 = 0

'Se cargan las primeras imagenes debido al desfase del timer

Set ImageA = LoadPicture("C:\Bancos\Originales\ori_1.jpg")
Set ImageB = LoadPicture("C:\Bancos\Decod_EC\dec_1.jpg")

Image1.Picture = ImageA
Image2.Picture = ImageB

' Indicador de la primera imagen mostrada
Shape1(0).FillColor = vbRed
End Sub

Private Sub Command2_Click()
'Borra las imagenes
Image1.Picture = Nothing
Image2.Picture = Nothing
'resetea los controladores
Timer1.Enabled = False
```

```

Command1.Enabled = True

'Limpiar los indicadores de progreso
For N = 0 To 14
Shape1(N).FillColor = &H8000000F
Next N

Form2.Show vbModal

End Sub

Private Sub Form_Load()
'Carga los títulos de la presentación
Form1.Caption = "DSCS (Double Stimulus Comparisson Scale)"
Label1.Caption = "Encuesta DSCS VERSION 2"
Label5.Caption = "Comparar la Imagen A con la Imagen B y seleccionar la
alternativa que mejor se acomode a su juicio:"
Label5.Font.Size = 11

'Se deshabilita las opciones de seleccion

'Ajusta el tamaño de las imagenes al frame
Image1.Stretch = True
Image2.Stretch = True

End Sub

Private Sub Timer1_Timer()
'Lee la evaluacion hecha por el usuario
If Option1.Value = True Or Option9.Value = True Then
    Data1 = Data1 + 1
    Option1.Value = False: Option9.Value = False
ElseIf Option2.Value = True Or Option8.Value = True Then
    Data2 = Data2 + 1
    Option2.Value = False: Option8.Value = False
ElseIf Option3.Value = True Or Option7.Value = True Then
    Data3 = Data3 + 1
    Option3.Value = False: Option6.Value = False
ElseIf Option4.Value = True Or Option6.Value = True Then
    Data4 = Data4 + 1
    Option4.Value = False: Option6.Value = False
ElseIf Option5.Value = True Then
    Data5 = Data5 + 1
    Option5.Value = False
End If
'Muestra la siguiente imagen por un intervalo de tiempo de modo aleatorio

If Rnd() >= 0.5 Then
    path1 = "C:\Bancos\Originales\ori_" & k & ".jpg"
    Set ImageA = LoadPicture(path1)
    path2 = "C:\Bancos\Decod_EC\dec_" & k & ".jpg"
    Set ImageB = LoadPicture(path2)

    Image1.Picture = ImageA
    Image2.Picture = ImageB

Else
    path1 = "C:\Bancos\Originales\ori_" & k & ".jpg"

```

```

        Set ImageA = LoadPicture(path1)
        path2 = "C:\Bancos\Decod_EC\dec_" & k & ".jpg"
        Set ImageB = LoadPicture(path2)

        Image1.Picture = ImageB
        Image2.Picture = ImageA

End If

'Barra de progreso
Shape1(k - 1).FillColor = vbRed

'Limite del numero de imagenes para evaluar
If k >= 16 Then
    Image1.Picture = Nothing
    Image2.Picture = Nothing
    Shape1(k - 1).FillColor = &H80000004

    'cargamos la imagen que indica el final de la evaluacion
    pathend = "C:\Bancos\Originales\ori_16.jpg"
    Set ImageA = LoadPicture(pathend)
    Image1.Picture = ImageA
    Image2.Picture = ImageA
    Timer1.Enabled = False
    GoTo final
End If

'Contador de imagenes
k = k + 1
final:

End Sub

'=====
'Formulario de Resultados
'=====
Dim ratio As Single, percent As Single, pesos As Single, T As Integer

Private Sub Command1_Click()
Unload Me
End Sub

Private Sub Form_Load()

' Calcula los resultados finales
T = Data1 + Data2 + Data3 + Data4 + Data5

pesos = (Data2 + 2 * Data3 + 3 * Data4 + 4 * Data5) ' distribucion
ponderada

If T = 0 Then          ' Evita la division entre 0
    m = MsgBox("Debe calificar las imagenes", vbInformation, "Info")
Else
    ratio = 1 - pesos / (4 * T)
End If

percent = 100 * ratio

Text1 = FormatNumber(ratio, 2)

```

```
Text2 = FormatNumber(percent, 1) & "%"  
Text3 = T  
End Sub
```

```
'=====  
'Modulo de intercambio de variables  
'=====
```

```
Global Data1 As Single, Data2 As Single, Data3 As Single, Data4 As  
Single, Data5 As Single
```

ANEXO C

Hoja de datos de la evaluación subjetiva

Registro de participantes de la evaluación subjetiva de imágenes codificadas y decodificadas usando el algoritmo de compresión ADPCM (Sistema EC), Propuesto en el presente trabajo de investigación.

Nº	Nombre del participante	Edad	Sexo	Resultado
1	Luis Díaz Patiño	32	M	13.3%
2	Carola Soriano Paredes	52	F	20.0%
3	Yadira Tarazona Cuervo	28	F	13.3%
4	Flavio Montes Poma	35	M	20.0%
5	Herminio García G.	58	M	13.3%
6	Paulina Campos Delgadillo	33	F	13.3%
7	Érika Vidaurre García	30	F	33.3%
8	Juan López	30	M	6.7%
9	Percy San Miguel Sevincha	28	M	0.0%
10	Nery Pollo Martinez	29	F	0.0%
11	Pedro Villaroel Zúñiga	27	M	6.7%
12	José Aarón Villalonga	25	M	6.7%
13	Gabriela Torres Gárate	20	F	40.0%
14	Armando Espinoza Ocmin	21	M	40.0%
15	Víctor Córdoba Bemiy	22	M	6.7%
16	Luis Pinto	35	M	23.3%
17	Pedro Wischuquiyauri Moquillaza	30	M	10.0%
18	Maritza Milla Tarazona	27	F	41.1%
19	Sofía Gallarday Lucano	26	F	58.9%
20	Karen Montalván Rivera	26	F	38.6%
21	Jackeline Castillo Jayme	27	F	9.1%
22	Zuli Barrera Tarazona	22	F	27.3%

Anexo D

Imágenes originales y decodificadas del sistema EC

Las 15 imágenes mostradas, tienen una resolución de 480 x 640 píxeles y una relación de aspecto 4:3. Las mismas fueron usadas como imágenes de prueba en los sistemas de codificación A, B, C, D, E y EC del capítulo 5. Las imágenes de la izquierda son las originales y las de la derecha son las codificadas usando cuantización adaptiva y recuperadas en el decodificador.

1)



2)



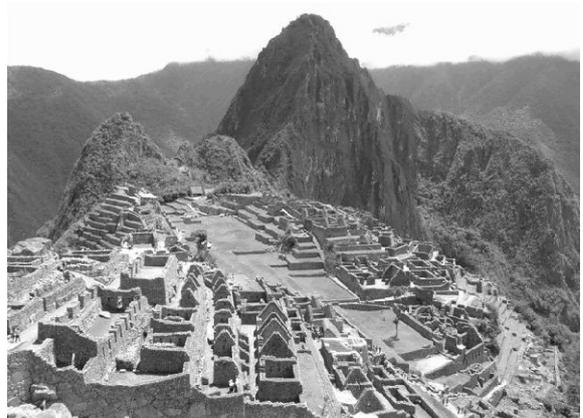
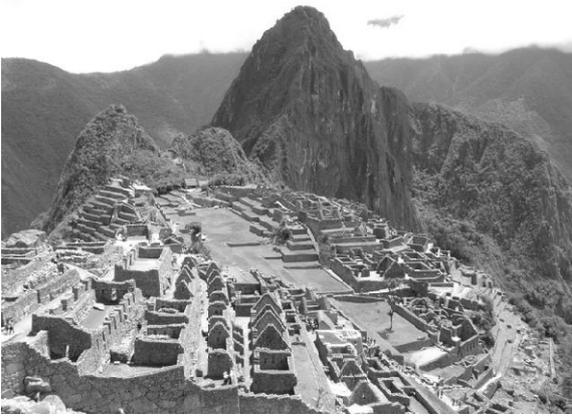
3)



4)



5)



6)



7)



8)



9)



10)



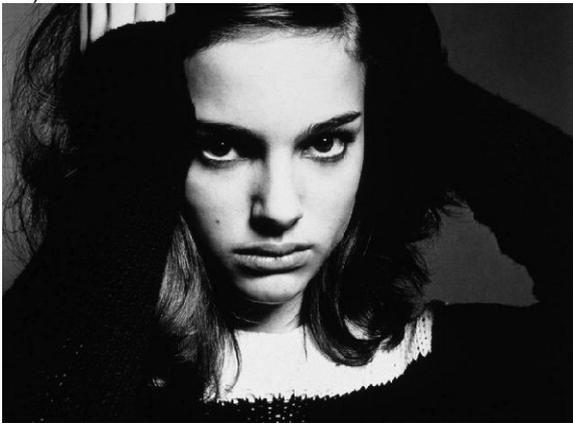
11)



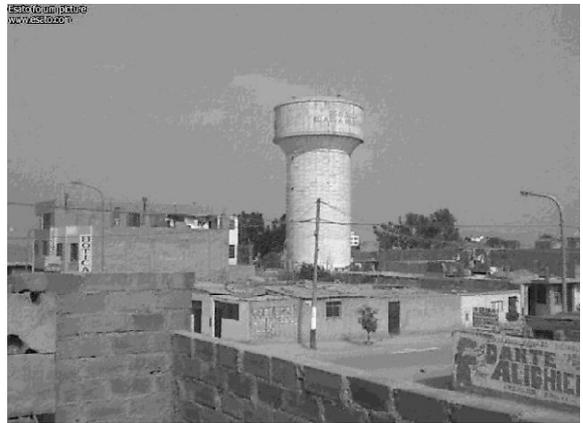
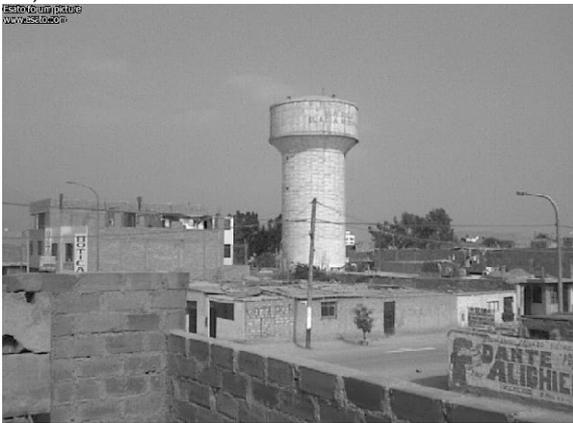
12)



13)



14)



15)

