



UNIVERSIDAD TÉCNICA DE AMBATO

CENTRO DE ESTUDIOS DE POSGRADO

**FACULTAD DE INGENIERÍA EN SISTEMAS,
ELECTRÓNICA E INDUSTRIAL**

**MAESTRÍA EN REDES Y TELECOMUNICACIONES
II EDICIÓN.**

TEMA:

**"COMPRESIÓN DE IMÁGENES EN MOVIMIENTO PARA
TRANSMISIÓN EN REDES IP EN BASE A LA APLICACIÓN DE LA
TRANSFORMADA DISCRETA COSENO CON DIFERENTES
RELACIONES DE COMPRESIÓN".**

Trabajo de Investigación como Requisito para optar por el Grado de Magíster en
Redes y Telecomunicaciones

Autor

Ing. José Gioberty Bucheli Andrade.

Tutor

Ing. Eddie Galarza Zambrano, M. Sc.

Ambato - Ecuador
Diciembre - 2010

APROBACIÓN DEL TUTOR

En mi calidad de Tutor del trabajo de investigación, nombrado por el H. Consejo Superior de Postgrado de la Universidad Técnica de Ambato:

CERTIFICO:

Que el trabajo de investigación: “COMPRESIÓN DE IMÁGENES EN MOVIMIENTO PARA TRANSMISIÓN EN REDES IP EN BASE A LA APLICACIÓN DE LA TRANSFORMADA DISCRETA COSENO CON DIFERENTES RELACIONES DE COMPRESIÓN”, presentada por el maestrante: José Bucheli Andrade, estudiante del programa de Maestría en Redes y Telecomunicaciones, de la Segunda Edición, reúne los requisitos y méritos suficientes para ser sometido a la evaluación del jurado examinador que el H. Consejo de Posgrado designe.

Ambato, 7 de Diciembre de 2010

TUTOR

Ing. Eddie Galarza Zambrano, M. Sc.
C. C. 1303128514

AUTORÍA DE TESIS

El abajo firmante, en calidad de estudiante de la Maestría en Redes y Telecomunicaciones, de la Segunda Edición, declaro que los contenidos de este trabajo de investigación científica “COMPRESIÓN DE IMÁGENES EN MOVIMIENTO PARA TRANSMISIÓN EN REDES IP EN BASE A LA APLICACIÓN DE LA TRANSFORMADA DISCRETA COSENO CON DIFERENTES RELACIONES DE COMPRESIÓN”, requisito previo a la obtención del Grado de Magíster en Redes y Telecomunicaciones, son absolutamente originales, auténticos, personales y de exclusiva responsabilidad legal y académica del autor.

Ambato, 7 de Diciembre de 2010

Ing. José Gioberty Bucheli Andrade
C. C. 0501513808

AL CONSEJO DE POSTGRADO DE LA UTA

El comité de defensa de la tesis “COMPRESIÓN DE IMÁGENES EN MOVIMIENTO PARA TRANSMISIÓN EN REDES IP EN BASE A LA APLICACIÓN DE LA TRANSFORMADA DISCRETA COSENO CON DIFERENTES RELACIONES DE COMPRESIÓN ” presentada por el Ing. José Gioberty Bucheli Andrade y conformada por: los Señores Miembros del Tribunal de Defensa y Director de Tesis Ing. Eddy Galarza M.Sc y presidido por: Ing....., Presidente del POSGRADO FISEI, Ing....., Director del CEPOS-UTA, una vez escuchada la defensa oral y revisada la Tesis de Grado escrita en la cual han constatado el cumplimiento de las observaciones realizadas por el tribunal de Defensa de la Tesis, remite la presente tesis para uso y custodia en la bibliotecas de la UTA.

Director del CEPOS – UTA

Director Académico Administrativo

Ing. Eddie Galarza M.Sc.
Director de Tesis

Miembro del Tribunal

Miembro del Tribunal

Miembro del Tribunal

DEDICATORIA

A mi Dios, a mi querida esposa Ana María por su amor, comprensión, apoyo y a mis hijos José Aníbal y Ana María, razón de mi esfuerzo y dedicación.

A mis padres (+)

José B.

AGRADECIMIENTO

Expreso mi agradecimiento al Ingeniero Eddie Galarza por su valiosa asesoría, colaboración, confianza, acertada recomendación en la ejecución del presente proyecto.

Un agradecimiento de manera especial, a la Facultad de Ingeniería en Sistemas de la Universidad Técnica de Ambato.

José B.

RESUMEN.

La reproducción de imágenes en movimiento (video) en un ordenador, pese a su aparente sencillez, constituye una de las tareas que más recursos consume. Para que una sucesión de imágenes produzca el efecto óptico del movimiento, es necesario que se sucedan a una determinada velocidad, que suele girar en torno a los 30 fps, es decir, 30 imágenes o fotogramas por cada segundo. Si se reduce este valor, el vídeo se reproduce a saltos, disminuyendo la sensación de continuidad. El problema es que el proceso de imágenes mediante una tarjeta gráfica es una de las tareas que más tiempo consume, de manera que, para transmitir tal cantidad de imágenes por segundo, es necesario disponer de un procesador y una tarjeta gráfica muy potentes, o bien reducir la calidad de la imagen con la que estamos trabajando.

Puesto que lo primero todavía no está al alcance de todos los usuarios, la mayor parte de los fabricantes han optado por la segunda opción, es decir, han ideado sistemas de compresión que reducen sensiblemente la calidad y el tamaño de la imagen.

Las imágenes en movimiento contienen grandes cantidades de información que requieren mucho espacio de almacenamiento, ancho de banda considerable y tiempos largos de la transmisión. Por lo tanto es ventajoso comprimir la imagen almacenando solamente la información esencial necesaria para reconstruir la imagen. Una imagen se puede pensar como matriz de los valores del píxel (o intensidad).

Para comprimir la imagen, las redundancias se deben explotar, por ejemplo, áreas donde hay poco o nada de cambio entre los valores del píxel. Por lo tanto imágenes que tienen grandes las áreas del color uniforme, tendrán redundancias grandes, e inversamente, imágenes que tengan cierta frecuencia y cambios grandes en color serán menos redundantes y más “duras” de comprimir.

En este proyecto se propone un método para realizar proceso de compresión y descompresión de imágenes en movimiento aplicando la técnica de la Transformada Discreta Coseno, como una alternativa para reducir la tasa de bits necesaria para la distribución o almacenamiento, sin detrimento de la calidad de la señal.

El sistema diseñado permite comprimir imágenes a color en movimiento con cuatro relaciones de compresión para mejorar o reducir la compresión y la calidad de la imagen. Una interfaz gráfica de usuario elaborada en MATLAB permite observar la señal original, la señal comprimida, además permite seleccionar el número de tramas que se van a procesar, almacenar y enviar a través de una red IP.

Siendo el proceso matemático muy demorado y complejo, la señal de entrada está limitada a tramas hasta de 300 x 300 pixeles y con un tamaño que no sobrepase los 25Mb (en formato sin compresión).

ÍNDICE GENERAL

	<i>Pág</i>
RESUMEN	VI
INDICE GENERAL.....	VIII
INDICE DE TABLAS.....	XI
INDICE DE FIGURAS.....	XII
INTRODUCCIÓN.....	1
CAPÍTULO I.....	3
1. EL PROBLEMA	3
1.1. Tema de investigación	3
1.2. Planteamiento del Problema	3
1.2.2. Análisis Crítico	5
1.2.3. Prognosis	6
1.2.4. Formulación del Problema.....	6
1.2.5. Delimitación	6
1.3. Justificación	7
1.4. Objetivos	7
1.4.1. Objetivos Generales	7
1.4.2. Objetivos Específicos	8
CAPÍTULO II	9
2. MARCO TEÓRICO	9
2.1. Antecedentes	9
2.2 Fundamentación Legal.....	10
2.2.1 Ley Especial de Telecomunicaciones y su reforma.....	10
2.2.2 Instituto Ecuatoriano de la Propiedad Intelectual.....	11

2.2.3 Dirección Nacional de Derecho de Autor y Derechos Conexos	11
2.2.4 LA UIT Unión Internacional de Telecomunicaciones	11
2.2.5 Estándares de codificación.....	13
2.3 Categorías fundamentales.....	17
2.3.1 Visión del ojo humano	17
2.3.2 Los tipos de visión	19
2.4 Definiciones Básicas.....	22
2.5 Imágenes estáticas	26
2.5.1 Imágen digital	27
2.5.2 Espacios de color	30
2.5.3 Compresión de imágenes estáticas	37
2.6 La transformada discreta del coseno (DCT).....	42
2.7 El Estándar JPEG.....	52
2.8 Compresión de imágenes en movimiento.....	54
2.8.1 Formatos de Video	56
2.9 Uso de la compresión para la transmisión.....	59
2.10 Redes basadas en IP.....	60
2.11 MATLAB	62
2.11.1 El espacio de trabajo de Matlab.....	63
2.11.2 Formatos de imágenes soportados por Matlab.....	64
2.11.3 Formatos de trabajo en Matlab.....	65
2.11.4 GUIDE (Graphical Users Interfase Development Environment).....	67
2.12 HIPÓTESIS	68
2.12.1 Señalamiento de Variables de la Hiótesis.....	68
CAPÍTULO III	69
3. METODOLOGÍA	69
3.1 Enfoque	69
3.2 Modalidad de Investigación	69
3.3 Nivel o Tipo de Investigación.....	69
3.4 Población y Muestra.....	70
3.5 Técnicas e Instrumentos de Recolección de Datos.....	71

3.5.1 MOS test.....	71
3.6 Procesamiento y Análisis.....	72
3.7 Guía para evaluar la calidad de la imagen.....	72
CAPÍTULO IV	73
4. PROPUESTA	73
4.1 Tema	73
4.2 Desarrollo de la propuesta	73
4.2.1 Introducción.....	73
4.2.2 Especificación de Requerimientos	74
4.2.3 Requisitos de Interfaces Externas.....	75
4.2.4 Complejidad computacional (Aritmética).....	78
4.3 Diseño de Software	79
4.3.1 Descripción del Sistema	80
4.3.1.1 Compresor de Imágenes en Movimiento.....	80
4.3.1.2 Descompresor de Imágenes en Movimiento	83
CAPÍTULO V	88
5. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS	88
5.1 Introducción.....	88
5.2 Procesamiento de la Información.....	88
5.3 Población y Muestra.....	89
5.4 Recopilación de datos	91
5.5 Pruebas de Hipótesis.....	91
5.5.1 Planteamiento de la Hipótesis	91
5.5.2 Selección del Nivel de Significancia.....	91
5.5.3 Determinación del Estadístico a ser Considerado en el Análisis.....	92
5.5.4 Protocolo de Pruebas.....	92

5.5.5 Resultados de la Prueba Subjetiva.....	93
5.5.6 Conclusiones de Tipo Estadístico.....	105
5.5.7 Resultados del proceso de transmisión de imágenes en movimiento sobre redes que utilizan el protocolo IP.....	107
5.5.8 Toma de Decisión de la Investigación.....	108

CAPÍTULO VI	109
--------------------------	-----

6. CONCLUSIONES Y RECOMENDACIONES	109
--	-----

6.1 Conclusiones	109
------------------------	-----

6.2 Recomendaciones	111
---------------------------	-----

ANEXO

BIBLIOGRAFÍA

ÍNDICE DE TABLAS

	<i>Pág</i>
Tabla 2.1 Estándar de codificación de video.....	14
Tabla 2.2 Equivalencia entre Bit y Byte.....	25
Tabla 2.3 Características de varios formatos de compresión.....	39
Tabla 3.1 Encuesta Mos Test.....	72
Tabla 4.1 Archivos para el compresor y descompresor de video.....	87
Tabla 5.1 Resultados de la prueba subjetiva mediante el Mos Test.....	93
Tabla 5.2 Resultado para relación de compresión de 75%	95
Tabla 5.3 Resultado para relación de compresión de 86%	97
Tabla 5.4 Resultado para relación de compresión de 93.75%	100
Tabla 5.5 Resultado para relación de compresión de 98.44%	103

Tabla 5.6 Resultado de aplicación del MOS Test a la encuesta realizada al personal de la ESPE Extensión Latacunga	105
Tabla 5.7 Tiempos de transmisión de imágenes en movimiento con diferentes relaciones de compresión utilizando el protocolo IP.....	107

ÍNDICE DE FIGURAS

Figura 2.1 Partes constitutivas del ojo.....	18
Figura 2.2 Curva de eficiencia luminosa para la visión fotópica.....	20
Figura 2.3 Fenómenos relativos al procesamiento de contraste.....	21
Figura 2.4 Vecindad entre píxeles (N _{4p}).....	23
Figura 2.5 Sistema de Coordenadas para una imagen.....	28
Figura 2.6 Modelo RGB.....	32
Figura 2.7 Imagen original y componente RGB.....	33
Figura 2.8 Modelo CMY.....	34
Figura 2.9 Modelo YIQ	35
Figura 2.10 Modelo HSI	36
Figura 2.11 Diagrama de bloque y rango de compresión en el caso del proceso con pérdida y sin pérdida	38
Figura 2.12 Fotografía guardada como GIF de 256 colores.	40
Figura 2.13 Fotografía guardada como JPEG con un preset de alta calidad.. .	40
Figura 2.14 Fotografía guardada como PNG8 en un archivo de 62.64K.....	40
Figura 2.15 Fotografía guardada como PNG24 en un archivo de 211K.....	41
Figura 2.16 Pasos del proceso de codificación basado en la DCT.....	44
Figura 2.17 Matriz original y matriz luego del proceso de la DCT.....	45
Figura 2.18 Aplicación de la DCT, IDCT, (a) imagen original (b) imagen luego de la descompresión con 9 de los 64 valores que corresponden a cada bloque en que es dividida una imagen.....	46
Figura 2.19 Proceso completo de compresión de imágenes	47

Figura 2.20 Barrido o exploración en zigzag.....	48
Figura 2.21 Secuencia en la compresión de imágenes en base a la DCT.....	48
Figura 2.22 Pasos del proceso de decodificación basado en la IDCT.....	49
Figura 2.23 Etapas que conforman el codificador JPEG.....	53
Figura 2.24 Etapas que conforman el decodificador JPEG.....	53
Figura 2.25 Ejemplo de una secuencia de imágenes.....	54
Figura 2.26 Esquema de la aplicación de un Códec de compresión.....	59
Figura 2.27 Red de conmutación de paquetes.....	60
Figura 2.28 Digitalización, compresión y transmisión de video	62
Figura 2.29 Pantalla de trabajo de Matlab. R2010a	64
Figura 2.30 Pantalla para acceder al Guide.....	67
Figura 2.31 Ventana principal del Guide.....	68
Figura 4.1 Imagen de MATLAB versión 7.10 R2010A utilizada	75
Figura 4.2 Panel virtual de interface de usuario para el compresor de imágenes en movimiento	81
Figura 4.3 Menú de propiedades de conexión de área local.....	83
Figura 4.4 Panel virtual de interface de usuario para el Descompresor de imágenes en movimiento	84
Figura 4.5 Ejemplo de una secuencia de imágenes obtenida luego del proceso de compresión y descompresión al 75% con el software diseñado ..	86
Figura 5.1 Resultado cuando se comprime el 75%.....	95
Figura 5.2 Resultado cuando se comprime el 86%	97
Figura 5.3 Resultado cuando se comprime el 93,75%.....	100
Figura 5.4 Resultado cuando se comprime el 98,44%.....	103

INTRODUCCIÓN

Debido al crecimiento de las aplicaciones multimedia muchos estándares de compresión de video han sido propuestos para diferentes aplicaciones, tales como H261, MPEG-1, MPEG-2 y MPEG-4. Cada uno de ellos buscando altos niveles de compresión, generalmente aprovechando la presencia de dos tipos de redundancia en las imágenes en movimiento: la espacial y la temporal

La compresión se basa en la eliminación de datos redundantes. Esta transformación (compresión) es aplicada a las imágenes antes de que sean almacenadas o antes de ser enviadas, por ejemplo vía red. La imagen una vez en su destino o almacenada es descomprimida.

En otras palabras, la compresión de imágenes trata de minimizar el número de bits necesarios para representar una imagen. Las aplicaciones de la compresión de imágenes son principalmente la transmisión y almacenamiento de información. En transmisión, sus principales aplicaciones son la televisión, radar y sonar, teleconferencias, comunicación por computadoras, transmisión por fax y otras así. En almacenamiento, la compresión de imágenes se utiliza sobre documentos, imágenes médicas, secuencia de imágenes, imágenes de satélite, mapas meteorológicos, etc.

Unas de las técnicas más comunes para lograr la compactación de una imagen es JPEG, que es un estándar para la compresión de imágenes. Así como JPEG representa un estándar para la compactación de imágenes el equivalente en transmisión de vídeo es el MPEG. MPEG puede considerarse como una extensión del JPEG.

El propósito de la presente tesis es diseñar implementar un compresor y un decodificador de imágenes en movimiento, el cual va a permitir seleccionar 4 niveles de compresión para ser almacenado o transmitidos a través de una red IP.

Este documento se lo ha dividido en cuatro partes fundamentales para el cumplimiento de sus objetivos y para facilitar la comprensión del trabajo.

En el capítulo I se realiza una descripción del problema de investigación incluyendo la justificación, la delimitación y los objetivos a alcanzar en el presente trabajo.

En el capítulo II se realiza una breve descripción de la información disponible sobre imágenes en movimiento, formatos de compresión, transmisión sobre redes IP y el software utilizado para la creación de una interface gráfica de usuario y que fundamentan el trabajo realizado.

El capítulo III se describe la metodología de la investigación, tipo de investigación y las características; el enfoque de la recolección de datos y su interpretación para el diseño del compresor y descompresor de imágenes en movimiento.

En el capítulo IV se describe la propuesta de la implementación del compresor, descompresor y la interface gráfica de usuario.

En el capítulo V se presenta el análisis de resultados, se presenta la validación del compresor, descompresor y la hipótesis de la investigación por medio de pruebas experimentales y en base a la opinión de un número de observadores.

Por último en el capítulo VI se indican las principales conclusiones y recomendaciones tras el análisis de los resultados del sistema.

Para finalizar en el presente trabajo se presentan la bibliografía y los anexos.

CAPÍTULO I

EL PROBLEMA

1.1. TEMA DE INVESTIGACIÓN

COMPRESIÓN DE IMÁGENES EN MOVIMIENTO PARA TRANSMISIÓN EN REDES IP EN BASE A LA APLICACIÓN DE LA TRANSFORMADA DISCRETA COSENO CON DIFERENTES RELACIONES DE COMPRESIÓN

1.2 PLANTEAMIENTO DEL PROBLEMA.

1.2.1. CONTEXTUALIZACIÓN

Uno de los tipos más importantes de información actualmente utilizados es la señal de vídeo, entendiendo como tal a la señal capaz de transmitir o almacenar una secuencia de imágenes sucesivas, es decir imagen en movimiento. Esta señal se utiliza en multitud de aplicaciones, que de entre las muchas clasificaciones que pueden hacerse, es posible dividir las en dos grandes grupos ¹:

1. Sistemas de propósito específico destinados a aplicaciones particulares, generalmente profesionales: enlaces de televisión o transmisión de imágenes.
2. Sistemas de aplicación masiva, destinados a la mayoría (o a una gran parte) de la población: televisión convencional, televisión DBS (difusión directa por satélite) o videoteléfono.

La señal convencional de vídeo es en principio analógica, ya que representa una secuencia de imágenes reales, por lo que es continua en el espacio (por lo menos en la dirección horizontal). No así en el tiempo, ya que lo que se transmite es un muestreo en el tiempo de la secuencia real (cada una de las imágenes es una

¹ http://sociedadinformacion.fundacion.telefonica.com/docs/repositorio/es_ES/TelefonicaySI/Publicaciones/imagenio.pdf La codificación y compresión digital de contenidos audiovisuales

muestra), que, al ser la frecuencia de muestreo suficientemente alta, es percibida por el ojo como una secuencia continua.

A través de la digitalización de la señal, se convierten las imágenes en matrices de puntos, cada uno de ellos de valor discreto, para permitir un manejo y transporte de la señal más sencillo gracias a las potentes técnicas existentes de procesamiento digital de señales.

Esta digitalización de la señal de vídeo tropieza con un gran problema característico del vídeo: su gran ancho de banda. El muestreo y cuantificación de la señal produce como resultado una secuencia digital con una tasa de bit o velocidad binaria muy alta y que al modularse por cualquiera de las técnicas existentes acaba presentando un ancho de banda bastante mayor que el que tiene la señal original analógica modulada.

Por ello, un objetivo básico de este proyecto es la reducción del número de bits de la señal de vídeo digital, de manera que disminuyan sus necesidades espectrales y se posibilite su uso cuando es reducido. Esta reducción se debe hacer sin disminución de la calidad subjetiva de la señal, o sea, la calidad de la secuencia de vídeo según la opinión del espectador que la está contemplando. Es decir, habrá que eliminar bits (o sea, información) de la secuencia sin afectar a su percepción.

Existen multitud de métodos para la compresión de una señal digital de vídeo, una de las seleccionadas en este proyecto es la codificación de transformadas, que consiste en transformar los valores de los puntos de una imagen en un conjunto de coeficientes con menor grado de correlación entre ellos. Para ello, en primer lugar se divide la imagen en bloques (por sencillez de cálculo) y se aplica a los valores de cada bloque una transformada generalmente lineal (Transformada Discreta Coseno), que a pesar de no ser óptimas, poseen buenas propiedades de compresión de energía y además existen algoritmos rápidos de cálculo. La

compresión se obtiene eliminando los coeficientes pequeños y cuantificando con menos precisión los poco importantes en cuanto a percepción visual.

El desarrollo de varios algoritmos rápidos para una implementación eficiente de la DCT, que emplean únicamente aritmética real, ha contribuido aún más a su popularidad.

Asimismo, la reducción en la complejidad de cálculo de estos algoritmos y la estructura recursiva contribuye a simplificar el hardware notablemente.

1.2.2. ANÁLISIS CRÍTICO.

Las Imágenes o el video digital sin compresión implican una gran cantidad de información. Si se desea guardar o transmitir a través de una red IP esta información debe requerir de una gran cantidad de espacio y ancho de banda.

Para resolver dicho inconveniente se analizarán técnicas de compresión de datos, con el fin de reducir el volumen de información que se transmite.

El proceso de compresión se basa en la definición de una gran cantidad de parámetros que determinan la calidad de la compresión, teniendo en cuenta que a mayor compresión, menor calidad, y viceversa.

Las imágenes constituyen gran parte del futuro de la información y de los servicios actuales multimedios, por ello, estándares de compresión de imágenes fijas y en movimiento han desarrollado técnicas de compresión de datos basadas en la transformada discreta de coseno que aprovecha algoritmos de computación veloces.

1.2.3. PROGNOSIS.

Los videos digitales ocupan mucho espacio, si no se realiza el proceso de compresión (propuesto en este proyecto), la transmisión sobre redes que utilicen el protocolo IP no se realizará de manera eficiente debido al limitado ancho de banda disponible, así como habrá un demora en el proceso de edición del video.

El no disponer además de un control de la relación de compresión conlleva a que no se disponga de un método que permita seleccionar la calidad de la imagen de acuerdo a la importancia que tenga ésta en un proceso de transmisión, lo que involucra una pérdida de tiempo para el usuario.

1.2.4. FORMULACIÓN DEL PROBLEMA.

¿Será posible diseñar un compresor de imágenes en movimiento en base a la transformada Discreta Coseno con el algoritmo mas adecuado, con diferentes relaciones de compresión que permita un uso óptimo de los recursos de memoria y un ancho de banda para la transmisión de la señal a través de la red vía el protocolo de internet (IP)?

1.2.5. DELIMITACIÓN.

El presente proyecto permitirá la compresión de imágenes digitales en movimiento, con diferentes relaciones de compresión, las cuales podrán ser transmitidas a través de la red vía el protocolo de internet (IP), utilizando un entorno de programación visual, para ser utilizado como material didáctico en las asignaturas de Procesamiento digital de señales, Comunicación para Instrumentación, Instrumentación Biomédica impartidas en la ESPE Sede Latacunga.

1.3. JUSTIFICACIÓN

Muchos procesos manejan cualquier cantidad de información, en este caso imágenes, que deben ser guardadas o transmitidas. La gran capacidad de estas imágenes dificultan el envío a través de la red.

Dado que el video sin comprimir demanda un gran ancho de banda, la necesidad de una compresión eficiente es de gran importancia en muchos tipos de aplicaciones.

El proyecto a desarrollar permitirá ofrecer un compresor de imágenes para la transmisión basados en estándares de codificación, con factores de calidad seleccionable es decir flexibilidad para poder seleccionar imágenes de alta calidad o baja calidad de imagen y mediante un programa que permita al operador un entorno visual de fácil accesibilidad y manejo para transmitirlo a la red vía el protocolo de internet (IP).

El proyecto propuesto permitirá además ser el sustento teórico y práctico de la asignatura de DSP, Comunicación para Instrumentación, Instrumentación Biomédica que se imparten a los estudiantes de ingeniería en Electrónica y además servirá como referencia para el desarrollo de nuevos proyectos de investigación en este campo.

1.4. OBJETIVOS.

1.4.1. GENERAL

Implementar un compresor de imágenes en movimiento para transmisión en redes IP en base a la aplicación de la Transformada Discreta Coseno con diferentes relaciones de compresión.

1.4.2. ESPECÍFICOS

- Estudiar de la transformada discreta coseno y su aplicación en el proceso de compresión de imágenes en movimiento.
- Analizar el proceso de transmisión de imágenes en movimiento sobre redes que utilicen el protocolo IP
- Desarrollar una propuesta para la compresión y la transmisión de imágenes en movimiento sobre la red vía protocolo de internet (IP).

CAPÍTULO II

MARCO TEÓRICO

2.1. ANTECEDENTES

No existe en la Universidad Técnica de Ambato proyecto relacionado con el tema propuesto.

El método de investigación que se utilizó es de tipo documental y experimental.

Hasta hace poco tiempo el tratamiento de las imágenes se basaba en procesos químicos, ópticos, mecánicos o artísticos. En definitiva, las modificaciones que podían aplicarse a una imagen fija o en movimiento no eran muchas y adicionalmente sumamente complicadas.

Por fortuna, la informática ha cambiado todo esto. La introducción de la informática ha revolucionado todos los procesos industriales, pero muy especialmente los del mundo de la imagen. La conversión de las imágenes y por lo tanto de las imágenes en movimiento a datos y su posterior proceso o transmisión digital han convertido esta época en la era de la imagen.

Otra característica muy importante de las imágenes en formato digital es su portabilidad y la facilidad que tienen los datos de ser almacenados en cualquier soporte informático, así como la posibilidad de ser transmitidos mediante cualquier método de transmisión digital de información: teléfono, radio, satélite, red, etc.

Las técnicas de compresión de imágenes se basan en la eliminación selectiva de sus componentes espectrales menos representativos, por lo que la pérdida de información es inevitable.

La función esencial de la compresión de imágenes es minimizar el tamaño de los datos y organizándolos en flujos escalables. Algunas aplicaciones hacen extensivo el uso de la información de la imagen digital, tales como internet, fotografía y televisión digital, video-telefonía móvil y sensores inteligentes. Dado que la compresión sin pérdidas es generalmente imposible de alcanzar a tasas de compresión mayores de 4¹, la compresión sin pérdidas sería la opción para algunos tipos de transmisión. Sin embargo, la codificación con pérdidas es frecuentemente inaceptable dado que inevitablemente introduce artefactos que pueden presentar errores en la validez de los datos médicos por ejemplo.

Bajo ciertas condiciones, los algoritmos de codificación recientes ² proveen calidad cercana a la percepción sin pérdidas. Todos los trabajos previos están relacionados con el desempeño de los algoritmos de compresión de imágenes

2.2 FUNDAMENTACIÓN LEGAL.

2.2.1 LEY ESPECIAL DE TELECOMUNICACIONES Y SU REFORMA

Disposiciones Fundamentales

Art. 1.- AMBITO DE LA LEY.- La Ley Especial de Telecomunicaciones tiene por objeto normar en el territorio nacional la instalación, operación, utilización y desarrollo de toda transmisión, emisión o recepción de signos, señales, imágenes, sonidos e información de cualquier naturaleza por hilo, radioelectricidad, medios ópticos y otros sistemas electromagnéticos.

Art. 3.- Servicios portadores son los servicios de telecomunicaciones que proporcionan la capacidad necesaria para la transmisión de señales entre puntos de terminación definidos de red. Los servicios portadores se pueden prestar en dos

¹ F. M. Giaime Ginesu and D. Giusto. A multi-factors approach for image quality assessment based on a human visual system model. *Elsevier Signal Processing: Image Communication*

² J.P.G Ortiz Transmisión Progresiva de imágenes jpeg 2000. Technical report Universidad de Almería. Departamento de Arquitectura de Computadores y Electrónica

modalidades: bajo redes conmutadas y bajo redes no conmutadas.

2.2.2 Instituto Ecuatoriano de la Propiedad Intelectual

El Instituto Ecuatoriano de la Propiedad Intelectual -IEPI- es una persona jurídica de derecho público, creada mediante la Ley de Propiedad Intelectual No. 83, publicada en el Registro Oficial 30, del 19 de mayo de 1998. El -IEPI- es el encargado de regular el control de la aplicación de las leyes de la propiedad intelectual.

2.2.3 Dirección Nacional de Derecho de Autor y Derechos Conexos

Es el área de la propiedad intelectual que protege las obras literarias, artísticas y científicas, garantiza los derechos que el autor tiene sobre sus obras y establece los mecanismos para que pueda, autorizar o prohibir su utilización a terceros. El derecho de autor protege por toda la vida y hasta 70 años después de la muerte del autor las obras del ingenio humano, con independencia del género, forma de expresión, mérito o finalidad, tales como: audiovisuales, esculturas, bases de datos, programas de ordenador, fotografías, mapas, adaptaciones, etc

2.2.4 LA UIT Unión Internacional de Telecomunicaciones³

La Unión Internacional de Telecomunicaciones es una organización progresista y dinámica que se ha dedicado a la tarea de ayudar a conectar el mundo por más de 140 años. En su carácter de organización flexible, la UIT se mantiene joven y lleva las riendas a la vanguardia de las comunicaciones mundiales.

La UIT es la organización más importante de las Naciones Unidas en lo que concierne a las tecnologías de la información y la comunicación. En su calidad de coordinador mundial de gobiernos y sector privado.

³ <http://www.itu.int/net/about/index-es.aspx>

Está compuesta por tres sectores:⁴

- UIT-T: Sector de Normalización de las Telecomunicaciones (antes CCITT).
- UIT-R: Sector de Normalización de las Radiocomunicaciones (antes CCIR).
- UIT-D: Sector de Desarrollo de las Telecomunicaciones de la UIT (nuevo).

En general, la normativa generada por la UIT está contenida en un amplio conjunto de documentos denominados *Recomendaciones*, agrupados por Series. Cada serie está compuesta por las Recomendaciones correspondientes a un mismo tema, por ejemplo Tarificación, Mantenimiento, etc. Aunque en las Recomendaciones nunca se "ordena", solo se "recomienda", su contenido, a nivel de relaciones internacionales, es considerado como mandatorio por las Administraciones y Empresas Operadoras.

El UIT-T coordina las iniciativas mundiales, propicia la excelencia técnica y la imparcialidad en la elaboración de normas y prepara el consenso necesario para garantizar la adopción de las nuevas tecnologías y los nuevos equipos en todo el mundo.

Desde Internet de banda ancha inalámbrica a la generación de tecnologías más recientes, desde la navegación aeronáutica y marítima a la radio astronomía y meteorología basada en satélites, de la convergencia de telefonía fija y móvil, acceso a Internet, datos, voz y televisión de las redes de próxima generación, la UIT se ha comprometido en conectar al mundo.

La UIT tiene su sede en Ginebra, Suiza, y como miembros se incluyen **192** Estados y **más de 700** Miembros de Sector y Asociados .

2.2.4.1 UIT-R BT.601 Recomendación, más conocido por las siglas **Rec. 601** o **BT.601** (o su nombre anterior, **CCIR 601**), es un estándar publicado en 1982 por la Unión Internacional de Telecomunicaciones - Sector de Radiocomunicaciones

⁴ http://es.wikipedia.org/wiki/Uni%C3%B3n_Internacional_de_Telecomunicaciones

(antes CCIR) para la codificación entrelazada de señales analógicas de vídeo en forma digital. Incluye los métodos de codificación de 525/ 60 y 625/ 50 , ambos con 720 luminancia y 360 muestras de crominancia por línea.

Rec. 601 de la señal: puede ser considerado como si se trata de una codificación digital de la señal analógica de vídeo por componentes, por lo que incluye datos para la sincronización horizontal y vertical y los intervalos de vaciado. Independientemente de la velocidad de los fotogramas, la frecuencia de muestreo de luminancia es de 13,5 MHz. La muestra de luminancia es de al menos 8 bits y las muestras de crominancia de por lo menos 4 bits cada uno.

2.2.5 Estándares de codificación

Los estándares son esenciales para la comunicación. Sin un lenguaje común que el transmisor y el receptor entiendan, la comunicación es imposible. Para comunicaciones de multimedios que involucran la transmisión de vídeo, los estándares juegan un papel aún más importante. La eficiencia tiene dos aspectos. Uno es que el estándar tiene que soportar un buen algoritmo de compresión que atenúe los requerimientos de ancho de banda para transmitir el vídeo. El otro es que el estándar tiene que permitir una implementación eficiente del codificador y decodificador y la complejidad del algoritmo de compresión debe ser tan baja como sea posible.

Los estándares para la codificación de vídeo son cruciales tanto para aplicaciones de almacenamiento como para la comunicación de multimedios. El éxito de las comunicaciones está principalmente determinado por el entendimiento de la fuente y el destinatario hacia el mismo lenguaje.

Para comunicaciones de multimedios existen dos grandes organizaciones de estandarización: *International Telecommunication Union - Telecommunication Sector* (ITU) e *International Organization for Standardization / International Electrotechnical Commission* (ISO/IEC). Los estándares recientes de codificación de vídeo están resumidos en la tabla 2.1. Estos estándares difieren principalmente

en el entorno de operación debido a las aplicaciones a las que fueron originalmente diseñados, sin embargo casi todos los estándares pueden esencialmente ser usados para un intervalo más amplio de operación. En términos de algoritmos de codificación, todos los estándares en la tabla 2.1 siguen un marco de trabajo similar, donde algunas características que los diferencian son el intervalo de parámetros y algunos modos específicos de codificación.⁵

Tabla 2.1 Estándar de codificación de video

Organización de estandarización	Estándar de codificación de vídeo	Valores típicos de tasa de bits	Aplicaciones típicas
ITU-T	H.261	$p \times 64$ kbits/s, $p=1 \dots 30$	ISDN Vídeo Teléfono
ISO	IS 11172-2 MPEG-1 Video	1.2 Mbits/s	CD-ROM
ISO	IS 13818-2 MPEG-2 Video	4-80 Mbits/s	SDTV, HDTV
ITU-T	H.263	64 kbits/s o inferior	PSTN Vídeo Teléfono
ISO	IS 14496-2 MPEG-4 Video	24-1024 kbits/s	Varias
ITU-T	H.263 Versión 2	< 64 kbits/s	PSTN Vídeo Teléfono
ITU-T	H.263L	< 64 kbits/s	Varias
ITU-T/ISO	H.264/MPEG-4 Parte 10	< 24 kbits/s	Varias

El estándar H.261 de codificación de vídeo fue definido por la ITU-T *Study Group XV* (SG15) para aplicaciones de vídeo telefonía y vídeo conferencia. Pone énfasis en tasas de bits pequeñas y bajos retardos de codificación⁶. Fue originado en 1984 y se previó ser usado para servicios audiovisuales a tasas de bits de cerca de $m \times 384$ kbits/s, donde m está entre 1 y 5. En 1988 se cambió su alcance y se decidió modificarlo para tasas de bits alrededor de $p \times 64$ kbits/s, donde p es de 1 a 30. Además, el estándar H.261 está diseñado para soportar dos formatos de imagen: el Formato Común Intermedio (CIF) y el Cuarto de CIF (QCIF). El algoritmo de codificación en el estándar H.261 es básicamente un híbrido de

⁵ T. Chen, Video Coding Standards for Multimedia Communications: H261, H263, and Beyond, Capítulo de visual Communications and Image Processing, 2002 pág. 25.

⁶ ITU-T Recommendation H.261: "Video Codec for audiovisual services at $p \times 64$ Kbits/s," Genova, 1990

compensación de movimiento para eliminar la redundancia temporal y la codificación por transformación para reducir la redundancia espacial.

El estándar MPEG-1 (de sus siglas en inglés *Motion Pictures Expert Group*) (ISO/IEC 11172-2) fue definido por el comité MPEG de ISO/IEC . El objetivo del estándar MPEG- 1 fue desarrollar un algoritmo que pudiera comprimir una señal de vídeo y ser capaz de reproducirla desde un CD-ROM (del inglés *Compact Disk - Read Only Memory*) o sobre líneas de teléfono a bajas tasas de bits (menos de 1.2 Mbits/s). El estándar MPEG-1 está principalmente diseñado para procesar vídeo con resolución QCIF, es decir, con dimensiones de 352 x 240 píxeles a 30 cuadros por segundo. El estándar MPEG-1 consta de tres partes: video, audio y sistemas. Las aplicaciones más comunes para MPEG-1 han sido el disco VideoCD y CDROMS para vídeo en computadoras.

La creación del estándar MPEG-2 (ISO/IEC 13818-2) por el comité de la ISO fue para mejorar el estándar MPEG-1, ya que no servía para los requerimientos de la industria de radiodifusión⁷. Así que MPEG desarrolló un algoritmo de compresión que procesaba vídeo a resoluciones completas equiparables a vídeo CCIR (del francés *Comité Consultatif International de la Radio*) 601 [704 x 480 NTSC (del inglés *National Television System committee*), 704 x 576 PAL (del inglés *Phase Alternation Line*)]. El estándar MPEG-2 tomó ventaja del ancho de banda disponible para la distribución de imágenes de alta resolución.

Sus objetivos son un incremento en la calidad de imagen, escalabilidad multi-resolución. Permite la compresión a resoluciones y tasas de bits mayores a los del estándar MPEG-1. Por su parte el estándar MPEG-2 está diseñado para manejar vídeo de calidad de radiodifusión que proporciona una mejor calidad a tasas de transferencia de datos más altas. Como su predecesor, consta de tres partes: vídeo, audio y sistemas. Las aplicaciones más comunes para el estándar MPEG-2 son la televisión digital y DVD (del inglés *Digital Versatil Disk*). Este estándar fue

⁷ J. Watkinson, MPEG-2 primera edición, Focal Press, 1999, pág. 244

creado por la ISO y el ITU-T, y tienen exactamente el mismo contenido en la parte de vídeo.

El estándar MPEG-4 de la ISO/IEC fue desarrollado por el grupo de trabajo MPEG⁸.

Luego de MPEG-2, finalizado en 1995, el grupo Moving Picture Experts Group (MPEG) continúa su trabajo de estandarización. En 1993 (antes de terminado MPEG-2) comienza el trabajo en el estándar MPEG-4. MPEG-4 abarca muchos más aspectos de los sistemas de video digital que sus predecesores MPEG-1 y MPEG-2.

La versión 1 del estándar MPEG-4 fue terminada en octubre de 1998 y se convirtió en un estándar internacional en los primeros meses de 1999. A principios de 2000 se terminaron los trabajos sobre las extensiones del estándar MPEG-4 Versión 1, llamadas MPEG-4 Versión 2, cuyas características son la compatibilidad que tiene hacia la versión anterior. El estándar MPEG-4 Visual (ISO/IEC 14496-2) permite la codificación híbrida de imágenes naturales (basadas en píxeles) así como de vídeo con escenas sintéticas (generadas por computadora). El estándar MPEG-4 provee elementos tecnológicos estandarizados que habilitan la integración de la producción, distribución y paradigmas de acceso de contenidos de los tres campos: televisión digital, aplicaciones gráficas interactivas (contenido sintético) y multimedia interactivos.

El estándar H.263⁹ fue definido por el grupo ITU-T SG15, que es el mismo grupo que definió H.261. Las actividades del estándar H.263 iniciaron alrededor de Noviembre de 1993, y fue adoptado en marzo de 1996.

Después de que terminó la estandarización de H.263, el continuo interés en la codificación de vídeo con tasas de bits muy bajas hizo claro que mejoras

⁸ F. Pereira, T. Ebrahimi, The MPEG-4 book, primera edición, Prentice Hall, 2002 pag 849

⁹ ITU-T Recommendation H.263: "Video coding for low bitrate communication" 1996

posteriores eran posibles al estándar H.263. Por lo tanto el grupo de trabajo ITU-T SG16 estableció un esfuerzo, informalmente conocido como H.263+, para satisfacer los requerimientos de estandarización de tales mejoras sobre el estándar H.263. El resultado fue una nueva versión conocida como H.263 Versión 2 (H.263+). Esta versión está orientada hacia aplicaciones de telecomunicaciones en tiempo real y servicios no-conversacionales. Posteriormente se realizaron extensiones al estándar H.263 Versión 2, informalmente llamadas H.263++, que crearon una tercera generación de sintaxis H.263. Las principales consideraciones para estas extensiones son: tipo de transformación variable, codificación aritmética adaptativa, tablas VLC (del inglés *Variable Length Code*) de resistencia al error, y filtrado de desanillado.

Estándar H.264/AVC

El estándar H.264 fue comenzado por el grupo de trabajo Video Coding Experts Group (VCEG) de la International Telecommunication Union (ITU-T). Las últimas etapas del trabajo se realizaron por la Joint Video Team (JVT) un grupo conformado por los grupos VCEG y el MPEG. El estándar final fue publicado en conjunto en 2003, por MPEG como la parte 10 del estándar MPEG-4 (ISO/IEC 14496-10) y por ITU-T como H.264.

2.3 CATEGORÍAS FUNDAMENTALES

2.3.1 Visión del ojo humano

Aunque el ojo es denominado a menudo el órgano de la visión, en realidad, el órgano que efectúa el proceso de la visión es el cerebro; la función del ojo es traducir las vibraciones electromagnéticas de la luz en un determinado tipo de impulsos nerviosos que se transmiten al cerebro a través del nervio óptico.¹⁰

¹⁰ <http://www.juntadeandalucia.es/averroes/~29701428/salud/ojo.htm>

El globo ocular es una estructura esférica de aproximadamente 2,5 cm de diámetro con un marcado abombamiento sobre su superficie anterior. La parte exterior, o la cubierta, se compone de tres capas de tejido: la capa más externa o esclerótica tiene una función protectora, cubre unos cinco sextos de la superficie ocular y se prolonga en la parte anterior con la córnea transparente; la capa media o úvea tiene a su vez tres partes diferenciadas: la coroides muy vascularizada, reviste las tres quintas partes posteriores del globo ocular, continúa con el cuerpo ciliar, formado por los procesos ciliares, y a continuación el iris, que se extiende por la parte frontal del ojo. La capa más interna es la retina, sensible a la luz. La figura 2.1 muestra los componentes principales del ojo humano.

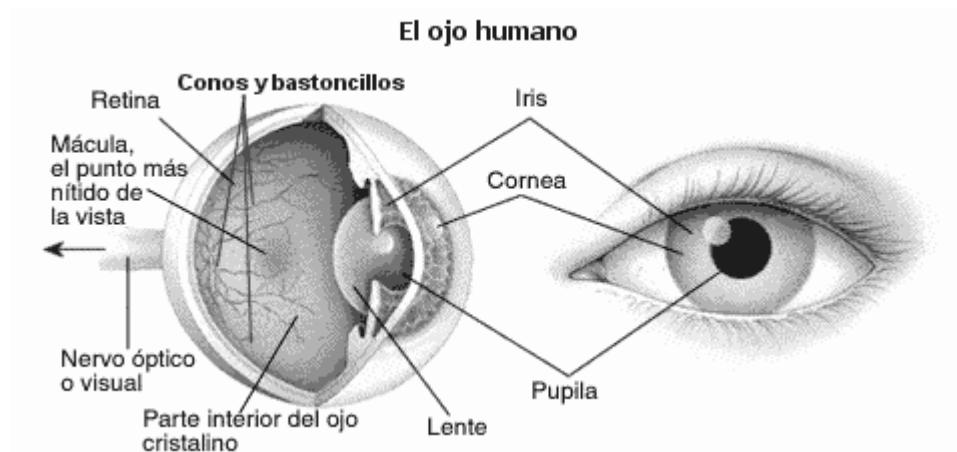


Figura 2.1 Partes constitutivas del ojo

- **La córnea** es una membrana resistente, compuesta por cinco capas, a través de la cual la luz penetra en el interior del ojo. Por detrás, hay una cámara llena de un fluido claro y húmedo que separa la córnea de la lente del cristalino. En sí misma, la lente es una esfera aplanada constituida por un gran número de fibras transparentes dispuestas en capas. Está conectada con el músculo ciliar, que tiene forma de anillo y la rodea mediante unos ligamentos. El músculo ciliar y los tejidos circundantes forman el cuerpo ciliar y esta estructura aplanada o redondea la lente, cambiando su longitud focal.
- **El iris** es una estructura pigmentada suspendida entre la córnea y el cristalino y tiene una abertura circular en el centro, la pupila. El tamaño de la pupila

depende de un músculo que rodea sus bordes, aumentando o disminuyendo cuando se contrae o se relaja, controlando la cantidad de luz que entra en el ojo.

- **La retina** es una capa compleja compuesta sobre todo por células nerviosas. Las células receptoras sensibles a la luz se encuentran en su superficie exterior detrás de una capa de tejido pigmentado. Estas células tienen la forma de conos y bastones y están ordenadas como los fósforos de una caja. Situada detrás de la pupila, la retina tiene una pequeña mancha de color amarillo, llamada mácula lútea; en su centro se encuentra la fovea central, la zona del ojo con mayor agudeza visual. La capa sensorial de la fovea se compone sólo de células con forma de conos, mientras que en torno a ella también se encuentran células con forma de bastones. Según nos alejamos del área sensible, las células con forma de cono se vuelven más escasas y en los bordes exteriores de la retina sólo existen las células con forma de bastones.
- **El nervio óptico** entra en el globo ocular por debajo y algo inclinado hacia el lado interno de la fovea central, originando en la retina una pequeña mancha redondeada llamada disco óptico. Esta estructura forma el punto ciego del ojo, ya que carece de células sensibles a la luz.

En el ojo existen millones de células especializadas en detectar las longitudes de onda procedentes de nuestro entorno. Estas maravillosas células, principalmente los conos y los bastoncillos, recogen las diferentes partes del espectro de luz solar y las transforman en impulsos eléctricos, que son enviados luego al cerebro a través de los nervios ópticos, siendo éste el encargado de crear la sensación del color.

2.3.2 Los tipos de visión¹¹

Al igual que en la fotografía, la cantidad de luz juega un papel importante en la visión. Así, en condiciones de buena iluminación (más de 3 cd/m²) como ocurre

¹¹ HUBEL, D.H. "Ojo, cerebro y visión." Servicio de Publicaciones, Universidad de Murcia, 2000.

de día, la visión es nítida, detallada y se distinguen muy bien los colores; es la visión fotópica. Para niveles inferiores a 0.25 cd/m^2 desaparece la sensación de color y la visión es más sensible a los tonos azules y a la intensidad de la luz. Es la llamada visión escotópica. En situaciones intermedias, la capacidad para distinguir los colores disminuye a medida que baja la cantidad de luz pasando de una gran sensibilidad hacia el amarillo a una hacia el azul. Es la visión mesiópica.

2.3.2.1 Sensibilidad a la intensidad luminosa

Según los tipos de visión, se definen unas curvas de sensibilidad del ojo a la luz visible para un determinado observador patrón que tiene un máximo de longitud de onda de 555 nm (amarillo verdoso) para la visión fotópica y otro de 480 nm (azul verdoso) para la visión escotópica. Al desplazamiento del máximo de la curva al disminuir la cantidad de luz recibida se llama efecto Purkinje, esta curva se indica en la figura 2.2



Figura 2.2 Curva de eficiencia luminosa para la visión fotópica

Toda fuente de luz que emita en valores cercanos al máximo de la visión diurna (555 nm) tendrá un rendimiento energético óptimo porque producirá la máxima sensación luminosa en el ojo con el mínimo consumo de energía. No obstante, si la fuente no ofrece una buena reproducción cromática puede provocar resultados contraproducentes.

2.3.2.2 Sensibilidad al contraste

Los mecanismos básicos de la visión trabajan con el contraste. Experimentos psicofísicos que evidencian los fenómenos visuales relativos a la percepción de

contraste pueden verse en la figura 2.3. En la figura de las bandas de Mach, se muestra el fenómeno de inhibición lateral que se produce en los campos receptores y a pesar de que cada barra representa un nivel constante, se aprecian diferencias subjetivas en los bordes de separación de las bandas. Se puede observar también un ejemplo de contraste simultáneo, en el que se muestra un rectángulo interior del mismo nivel de intensidad pero que se percibe más claro u oscuro según el nivel de luminosidad del rectángulo exterior.

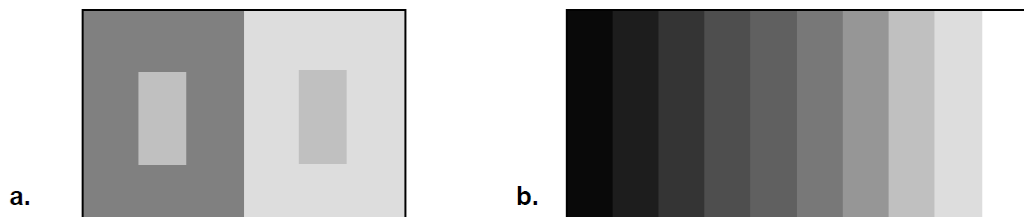


Figura 2.3 Fenómenos relativos al procesamiento de contraste. **a.** Contraste simultáneo **b.** Bandas de Mach

2.3.2.3. Sensibilidad en frecuencia

El estudio de la anatomía y fisiología del SVH ha permitido que se conozca que la percepción del estímulo depende de la frecuencia del mismo. Tradicionalmente una de las caracterizaciones más importantes que se hacen del sistema de visión humano consiste en la máxima frecuencia que es capaz de detectar o diferenciar y que se denomina agudeza visual.

2.3.2.4 La acomodación

Se llama acomodación a la capacidad del ojo para enfocar automáticamente objetos situados a diferentes distancias. Esta función se lleva a cabo en el cristalino que varía su forma al efecto. Pero esta capacidad se va perdiendo con los años debido a la pérdida de elasticidad que sufre; es lo que se conoce como

presbicia o vista cansada y hace que aumente la distancia focal y la cantidad de luz mínima necesaria para que se forme una imagen nítida.

2.3.2.5 La adaptación

La adaptación es la facultad del ojo para ajustarse automáticamente a cambios en los niveles de iluminación. Se debe a la capacidad del iris para regular la abertura de la pupila y a cambios fotoquímicos en la retina. Para pasar de ambientes oscuros a luminosos el proceso es muy rápido pero en caso contrario es mucho más lento. Al cabo de un minuto se tiene una adaptación aceptable. A medida que pasa el tiempo, vemos mejor en la oscuridad y a la media hora ya vemos bastante bien. La adaptación completa se produce pasada una hora.

2.4 Definiciones Básicas¹²

- **Imagen**

El término imagen monocroma o simplemente imagen, se refiere a una función bidimensional $f(x,y)$, representando intensidad de luz, donde x e y son las coordenadas espaciales y el valor de f en cualquier punto (x,y) es proporcional al brillo (o nivel de gris) de la imagen en ese punto.

- **Píxel**

Acrónimo formado con las palabras inglesas 'Picture' 'element'. Es el elemento simple de una imagen digital. Los monitores gráficos muestran imágenes dividiendo la pantalla en miles (o millones) de píxeles, dispuestos en filas y columnas. Los píxeles están tan juntos que parecen estar conectados.

¹² <http://www.innovanet.com.ar/gis/TELEDETE/TELEDETE/tradiimg.htm>

El número de bits usados para representar cada pixel determina cuántos colores o gamas de gris pueden ser mostrados. Por ejemplo, en modo color de 8-bits, el monitor en color utiliza 8 bits para cada pixel, permitiendo mostrar 2 elevado a 8 (256) colores diferentes o gamas de gris. En monitores de color, cada pixel se compone realmente de tres puntos: uno rojo, uno azul, y uno verde. Idealmente, los tres puntos convergen en el mismo punto, pero todos los monitores tienen cierto error de convergencia que puede hacer que el color los pixeles aparezca borroso.

La calidad de un sistema de visualización depende en gran medida de su resolución, es decir, cuántos bits utilizan para representar cada pixel, por ejemplo, una cámara que puede tomar fotografías con una resolución de 2048×1536 píxeles se dice que tiene 3,1 megapíxeles ($2048 \times 1536 = 3.145.728$).

- **Relaciones entre píxeles**

Un píxel p con coordenadas (x,y) tiene cuatro vecinos horizontales y verticales, cuyas coordenadas son: $(x+1,y)$, $(x-1,y)$, $(x,y-1)$, $(x,y+1)$. A este conjunto de píxeles se llama vecindad 4 o 4 vecinos de p y se denota por $N4(p)$, ver la figura 2.4.

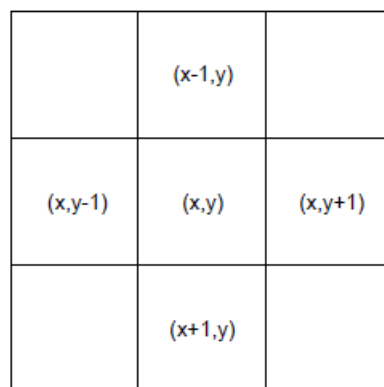


Figura 2.4 Vecindad entre pixeles (N4p)

- **Profundidad del Píxel**

La profundidad del píxel determina el número de bits que un píxel puede almacenar, también se la conoce como resolución de bits o profundidad del color.

Se encuentra estrechamente relacionada con la nitidez de la imagen digital, ya que a una mayor profundidad se tendrá un mayor nivel de tonos, lo que permite una representación más real de la imagen.

Así un píxel puede tener profundidad de 1, 2, 4, 8, 16, 24 y 32 bits, aunque los valores normalmente utilizados son de 8, 16 o 24.

Un píxel con profundidad de 1 bit puede representar dos tonos (blanco y negro), un píxel con profundidad 8 tiene 256 valores posibles (2^8), uno con profundidad de 24 tiene 16'777.216 valores posibles (2^{24}) en este caso se habla de *color verdadero* debido a que se asemeja a lo que el ojo humano puede distinguir en el mundo real. Para el caso de una profundidad de 32 bits en realidad no se añaden más niveles de tonos que los definidos con 24 bits, lo que se hace con los últimos 8 bits es indicar el grado de transparencia (desde 0 para totalmente transparente hasta 255 para totalmente visible).

Una profundidad de 1, 2 o 4 bits resulta limitada para imágenes de video debido a la baja calidad que ofrecería. En el caso de las imágenes de video monocromáticas generalmente se trabaja con profundidad de 8, es decir 256 tonalidades de grises que varían gradualmente.

- **Bit**

Es el acrónimo de Binary digit. (dígito binario). Un bit es un dígito del sistema de numeración binario. Mientras que en el sistema de numeración decimal se usan diez dígitos, en el binario se usan sólo dos dígitos, el 0 y el 1. Un bit o dígito binario puede representar uno de esos dos valores, 0 ó 1.

- **Byte**

Es una palabra inglesa, la Real Academia Española ha aceptado como equivalente a octeto (es decir a ocho bits), para fines correctos, un byte es considerado como una secuencia de bits contiguos, cuyo tamaño depende del código de información o código de caracteres en que sea definido. La ISO y la IEC en la norma 80000-

13:2008 recomiendan restringir el empleo de esta unidad a los octetos (bytes de 8 bit).

Las velocidades de transmisión de datos se miden normalmente en kilobits por segundo (kbps), mientras que el almacenamiento o tamaño de los archivos se especifica en bytes (Kilobytes, Megabytes, etc.). Sin embargo, muchos equipos de comunicación (módems, 'routers', etc.) y los navegadores o los programas de descarga de archivos especifican la velocidad en kilobytes por segundo, creando bastante confusión entre los usuarios.¹³

Para convertir kilobits a kilobytes se debe dividir para 8, por ejemplo en velocidades de transmisión: 256 kbits por segundo = 32 kbytes por segundo. Con esa velocidad de transmisión de datos, para descargar un Megabyte hay que emplear 32 segundos, considerando un Megabyte 1.024 Kilobytes.

Tabla 2.2 Equivalencia entre Bit y Byte

1 byte		8 bits
1 kilobyte	1.024 bytes	8,192 kilobits
1 megabyte	1.024 kilobytes	8,3886 megabits
1 gigabyte	1.024 megabytes	8,5899 gigabits

- **Color.**- El color se forma mediante la combinación de los tres colores básicos rojo, azul y verde (en inglés RGB).
- **Brillo.**- Indica si un área está más o menos iluminada.
- **Tono.**- Indica si un área parece similar al rojo, amarillo, verde o azul o a una proporción de ellos.
- **Luminosidad.**- Brillo de una zona respecto a otra zona blanca en la imagen.
- **Croma.**- Indica la coloración de un área respecto al brillo de un blanco de referencia.

¹³ <http://www.consumer.es/web/es/tecnologia/internet/2004/12/22/114135.php>

- **La Saturación.-** Se refiere al grado en el que un color puro se diluye con luz blanca, esto significa que indica la concentración de color. Una saturación nula corresponde a una ausencia de color, que es el caso de la escala de grises
- **Mapa de color**
El mapa de color es una matriz de $n \times 3$, donde cada renglón es una tripleta de colores. El primer renglón corresponde al valor mínimo del eje de color y el último renglón al máximo. Al definir diferentes distribuciones de intensidad de los tres colores básicos, se crean diferentes mapas de color.
- **Histograma de una imagen.**

El histograma de una imagen es una representación del número de píxeles de cierto nivel de gris en función de los niveles de gris.

2.5 Imágenes Estáticas

La imagen digital estática es aquella que carece de movimiento, es decir, es una representación de un objeto u objetos, reales o ficticios, en un instante en el tiempo. Este tipo de imagen a su vez se clasifica en dos categorías: imágenes vectoriales e imágenes de mapa de bits o bitmaps

- **Imágenes vectoriales:** Se componen de objetos definidos mediante operaciones matemáticas. A estos objetos (líneas, círculos, rectángulos, etc.) también se les conoce como vectores. Cada objeto es una entidad independiente, con propiedades como forma, color, tamaño y posición en el plano; puede tratárselo separado del resto de objetos sin afectarlos, pero si se requiere transformar la imagen completa se asocia a todos los objetos y se aplican los cambios, incluso se pueden hacer rellenos de formas.

Son imágenes que no requieren gran cantidad de espacio para su almacenamiento ya que sólo guardan la información necesaria para generar cada uno de los objetos. También se puede reducir o aumentar su tamaño sin

disminuir su calidad, un cambio en tamaño implica el cálculo de la posición y longitud de los objetos pero a diferente escala.

- **Imágenes de mapa de bits, bitmaps o imágenes ráster:** Están formadas por matrices de píxeles, cada píxel tiene una posición y color determinado. Este tipo de imágenes permite recoger una amplia gama de colores o tonos, por lo que son adecuadas para representar eficientemente el mundo real. Pueden ser creadas manualmente mediante un programa de pintura o retoque; obtenidas desde documentos con información analógica por medio de escáneres, desde el mundo real por medio de cámaras digitales u otro tipo de dispositivos con similar funcionamiento y desde cualquier aplicación de software por medio de programas para captura de pantallas.

2.5.1 Imagen digital

Es una imagen que ha sido discretizada tanto en coordenadas espaciales como en brillo. Podemos considerar una imagen digital como una matriz cuyos índices de filas y columnas identifican un punto en la imagen y el correspondiente elemento de matriz identifica el valor de gris en ese punto.

En principio las imágenes se representan con un patrón bidimensional de intensidad de luz y color. Este patrón está definido en un sistema de coordenadas en donde el origen corresponde a la esquina superior izquierda de la imagen y se describe como una función $f(x, y)$ como se ilustra en la figura 2.5.



Figura 2.5 Sistema de Coordenadas para una imagen

Una imagen digital puede también indicarse como un arreglo rectangular de puntos, o elementos de imagen, arreglados en m filas por n columnas. La expresión $m \times n$ es llamada la resolución de imagen. Una imagen puede ser de uno de los siguientes tipos: ¹⁴

- **Una imagen bi-nivel (o monocromática).** Esta es una imagen donde los pixeles pueden tener uno o dos niveles, normalmente referidos como blanco y negro. Cada pixel es en este caso representa un pixel para este tipo de imagen simple.
- **Una imagen en escala de grises.** Un pixel en este caso puede tener uno de los n valores de 0 a $n-1$, indicando uno de los 2^n matices de los grises (o matices de algún otro color). El valor de n es normalmente compatible con un tamaño de un byte, es decir 4, 8, 12, 16, 24 o algún otro múltiplo conveniente de 4 o de 8.

¹⁴ <http://gemini.udistrital.edu.co/comunidad/profesores/jruiz/jairocd/texto/uit/compimage.pdf>

- **Una imagen de tonos continuos.** Este tipo de imagen puede tener similares colores (o escalas de grises). Cuando pixeles adyacentes difieren por una unidad exactamente, es difícil para el ojo humano poder percibir esta diferencia de color. Como consecuencia, tal imagen puede contener áreas con colores que parecen variar continuamente cuando el ojo se mueve sobre el área. Un pixel en tal imagen es representado tanto por un número largo simple (en el caso de muchas escalas de grises) o por tres componentes (en el caso de una imagen de color). Una imagen de tonos continuos es normalmente una imagen natural y es obtenida de tomar una fotografía con una cámara digital o al escanear una fotografía o una pintura.
- **Una imagen en tono discreto (también llamada una imagen grafica o imagen sintética).** Esta es normalmente una imagen artificial. Esta puede tener pocos colores o muchos colores, pero ésta no tiene el ruido y manchas de una imagen natural. Un ejemplo de este tipo de imágenes son las fotografías de un objeto artificial o máquina, una página de texto, una carta, una caricatura y el contenido de una pantalla de computador.

Los pixeles adyacentes en una imagen de tono discreto frecuentemente son idénticos o varían significativamente en valor. Tal imagen no comprime bien con los métodos de pérdidas, ya que las pérdidas de algunos pixeles puede generar una carta ilegible. Los métodos de compresión para imágenes de tono continuo frecuentemente no manejan muy bien los límites abruptos de una imagen de tono discreto, además métodos especiales son necesarios para lograr una compresión eficiente de estas imágenes. Nótese que una imagen de tono discreto puede ser altamente redundante, ya que el mismo patrón o carácter puede aparecer muchas veces en una imagen.

- **Una imagen parecida a una caricatura.** Esta es una imagen de color que consiste de áreas uniformes. Cada área tiene un color uniforme no obstante áreas adyacentes tienen colores diferentes. Esta característica puede ser explotada para obtener una mejor compresión.

Intuitivamente es claro que cada tipo de imagen puede ofrecer redundancia, pero ellas son redundantes en diferentes formas. Por lo tanto, cualquier método de compresión dado, no se ejecuta bien con todas las clases de imágenes y por lo tanto existen diferentes métodos de acuerdo al tipo de imagen.

Hay métodos de compresión para imágenes binivel, para imágenes de tono continuo y para imágenes de todo discreto. Hay métodos que intentan desintegrar una imagen en partes de tono continuo y tono discreto y así comprimirlas cada una por separado.

2.5.2 Espacios de color ¹⁵

El uso del color en el procesamiento de imágenes está principalmente motivado por dos factores:

El color es un poderoso descriptor que, en la mayoría de los casos simplifica la identificación y extracción de los objetos de una escena. Los humanos podemos distinguir miles de colores y sin embargo sólo dos docenas de niveles de gris.

Debido a las características del ojo humano y a la teoría tricromática ¹⁶, todos los colores que podemos reconocer en una imagen son una combinación de los llamados colores primarios: R(Red/Rojo), G(Green/Verde) y B(Blue/Azul).

En esencia, un modelo de color es la especificación de un sistema de coordenadas tridimensional y de un subespacio de este sistema en el que cada color queda representado por un único punto. Entre los espacios de color utilizados más frecuentemente para el procesamiento de imágenes se encuentran el RGB, YIQ, CMY, YCbCr y HSI.

Generalmente las características para poder distinguir un color de otro son: brillo, tono y saturación. El brillo es la luminosidad u oscuridad relativa del color y

¹⁵ http://es.wikipedia.org/wiki/Teor%C3%ADa_del_color

¹⁶ http://www.gusgsm.com/teoria_tricromica

normalmente se expresa como un porcentaje comprendido entre 0% (negro) y 100% (blanco). El tono es el color reflejado o transmitido a través de un objeto. Se mide como un ángulo en grados, entre 0° y 360°. Normalmente, el tono se identifica por el nombre del color, como rojo, naranja o verde.

Por último la saturación, a veces llamada cromatismo, se refiere a la pureza relativa de la cantidad de luz blanca mezclada con el tono, es decir, es la fuerza o pureza del color. La saturación representa la cantidad de blanco que existe en proporción al tono y se mide como porcentaje entre 0%(gris) y 100%(saturación completa).

- **Modelo RGB¹⁷**

En el modelo RGB utilizado en el presente proyecto, cada color aparece en sus componentes espectrales primarias: rojo, verde y azul. Este modelo está basado en el sistema de coordenadas cartesianas. El subespacio de color de interés es el tetraedro mostrado en la Figura 2.6, en el cual los valores RGB están en tres vértices: cian, magenta y amarillo y se sitúan en otros tres vértices, el negro corresponde al origen y el blanco se sitúa en el vértice más alejado del origen. En este modelo, la escala de grises se extiende desde el negro al blanco a lo largo de la diagonal que une esos dos puntos, y los colores son puntos dentro del tetraedro definidos por los vectores desde el origen. Por conveniencia, se asume que todos los vectores han sido normalizados, de modo que el tetraedro de la figura es el tetraedro unitario, es decir, todos los valores de R, G y B están en el rango [0,1]. Las imágenes en este modelo se forman por la combinación en diferentes proporciones de cada uno de los colores primarios RGB.

Las imágenes del modelo de color RGB consisten en tres planos de imagen independientes, uno por cada color primario. Cuando llegan a un monitor RGB, estas tres imágenes se combinan en la pantalla fosforescente para producir una

¹⁷ <http://bibing.us.es/proyectos/abreproy/11875/fichero/Proyecto+Fin+de+Carrera%252F3.Espacios+de+color.pdf>

imagen en color compuesta. De forma alternativa, la mayoría de las cámaras de color empleadas para la adquisición de imágenes digitales utilizan el formato RGB, lo que por sí solo hace de éste un modelo importante en el procesamiento de imágenes.

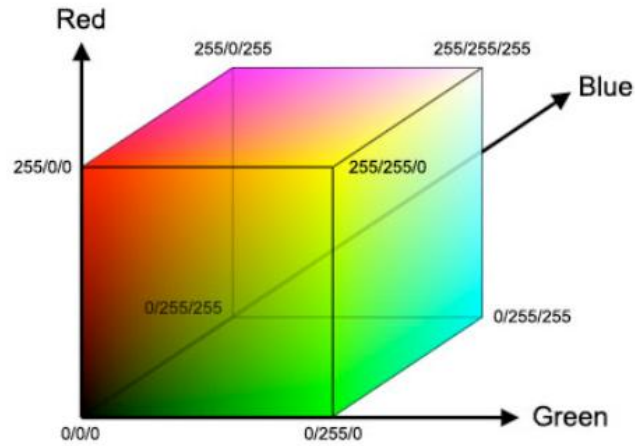


Figura 2.6 Modelo RGB

Una imagen a color está representada por una profundidad de 8 y 24 bits; para el caso de 24 bits por lo general se los divide en tres bytes: 8 bits para el color rojo, 8 para el verde y 8 para el azul, para representar otros colores se realizan combinaciones de éstos (RGB).

En la figura 2,7 se muestra la descomposición de una imagen de color en formato RGB



Figura 2.7 Imagen original y componentes RGB

- **Modelo CMY**¹⁸

Cian (C), magenta (M) y amarillo (Y) son los colores secundarios de la luz, o bien los colores primarios de los pigmentos. Se denominan sustractivos ya que se

¹⁸ <http://bibing.us.es/proyectos/abreproy/11875/fichero/Proyecto+Fin+de+Carrera%252F3.Espacios+de+color.pdf>

utilizan como filtros para sustraer colores de la luz blanca. La mayoría de los dispositivos que depositan pigmentos coloreados sobre papel, tales como impresoras y fotocopiadoras en color, necesitan una entrada CMY o bien una conversión interna de RGB a CMY. El sistema coordinado es el mismo que en modelo RGB pero donde había negro ahora existe blanco y viceversa.

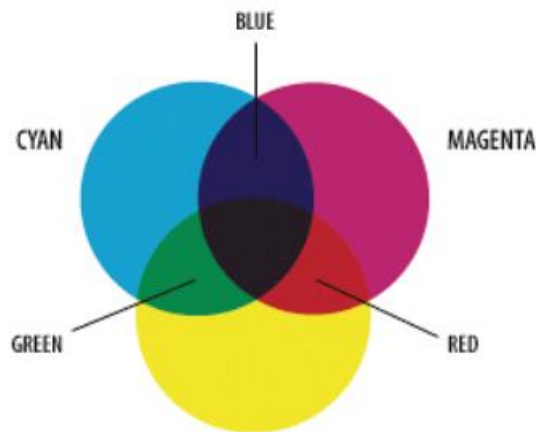


Figura 2.8 Modelo CMY

- **Modelo YIQ¹⁹**

El modelo YIQ se utiliza en las emisiones comerciales de televisión. Básicamente el YIQ es una recodificación del RGB utilizada por su eficacia en la transmisión y para mantener la compatibilidad con los estándares de televisión en blanco y negro. La componente Y del sistema YIQ, proporciona toda la información de vídeo que necesita un sistema de televisión monocromático.

El modelo YIQ fue diseñado para aprovechar la mayor sensibilidad del sistema visual humano a los cambios de la saturación. Así los estándares YIQ emplean más bits para representar la Y y menos para representar la I o Q.

Además de tratarse de un estándar ampliamente difundido, la ventaja principal del modelo YIQ en el procesamiento de imágenes es que la luminancia (Y) y la

¹⁹ <http://bibing.us.es/proyectos/abreproy/11875/fichero/Proyecto+Fin+de+Carrera%252F3.Espacios+de+color.pdf>

información del color (I y Q) están desacopladas, así la importancia de este desacoplamiento radica en que la componente de luminancia de una imagen puede procesarse sin afectar a su contenido cromático.

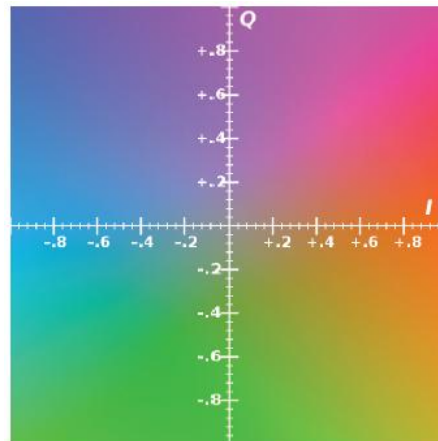


Figura 2.9 Modelo YIQ

- **Modelo YCbCr²⁰**

Se trata de una codificación no lineal del espacio RGB. El parámetro Y indica la luminancia, los parámetros Cb y Cr indican el tono del color: Cb ubica el color en una escala entre el azul y el amarillo, Cr indica la ubicación del color entre el rojo y el verde. El espacio YCbCr es una versión escalada y desplazada del espacio de color YUV. El parámetro Y representa la luminancia (es decir, información en blanco y negro), mientras que U y V representan la cromaticidad (es decir, información con respecto al color). Cb (Cr respectivamente) es la diferencia entre la componente azul (rojo, respectivamente) y un valor de referencia. La transformación de RGB a YCbCr puede ser ejecutada utilizando la ecuación de matrices 2.1, donde R, G y B oscilan en el rango de [0, 1]. Y en el de [16,235], Cb y Cr en el de [16,240].

²⁰ <http://bibing.us.es/proyectos/abreproy/11875/fichero/Proyecto+Fin+de+Carrera%252F3.Espacios+de+color.pdf>

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{pmatrix} 65.481 & 128.553 & 24.966 \\ -39.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{pmatrix} * \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (2.1)$$

En la práctica, generalmente se usa en los estudios de televisión europeos y en la compresión de imágenes.

- Modelo HSI

Sus siglas corresponden a H (Hue/Tonalidad), S (Saturation/Saturación), I (Intensity/Intensidad). El modelo HSI debe su utilidad a dos hechos básicos. Primero, la componente de intensidad, I, está desacoplada de la información cromática contenida en la imagen. Segundo, las componentes de tono y saturación están íntimamente relacionadas con la forma en que los seres humanos percibimos el color. Estas características hacen que el modelo HSI sea una herramienta ideal para desarrollar algoritmos de procesamiento de imágenes basados en alguna de las propiedades de la percepción del color del sistema visual humano.

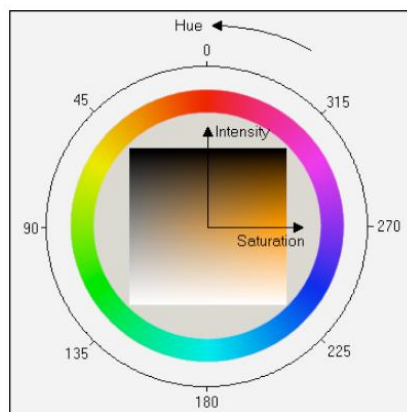


Figura 2.10 Modelo HSI

2.5.3 Compresión de imágenes estáticas

La mayoría de las imágenes contienen grandes cantidades de información redundante. La eliminación de esta información permite reducir el tiempo de transporte, así como la cantidad de espacio necesario para su almacenamiento.

Las técnicas de compresión se dividen en dos grupos:

2.5.3.1 Compresión sin pérdida: Cuando se desea que los datos digitales, al volver a descomprimirlos sean exactamente igual a los originales, sin degradaciones ni pérdidas, se utiliza la compresión sin pérdidas. Esto ocurre en muchas aplicaciones en las cuales detalles muy sutiles de brillo y espaciales son relevantes para la calidad de la imagen. A menudo, un observador no puede ni siquiera ver estos detalles. Sin embargo, la calidad de la imagen ha de ser mantenida intacta, debido a que la imagen puede tener que someterse a operaciones de procesado.

La compresión sin pérdidas²¹ se hace necesaria siempre que se desconozca el uso analítico que se dará a la imagen. Por ejemplo en la exploración del espacio, las imágenes se estudian durante años, después de que se originan. Al no saber cómo será procesada la imagen en el futuro, es imposible saber qué degradaciones o pérdidas podría tolerar. Es mejor no introducir entonces ninguna degradación. La relación de compresión puede ser hasta de 3:1

²¹ H. Lee, Y. Kim, A. Rowberg, and E. Riskin, "Statistical distributions of dct coefficients and their application to an interframe compression algorithm for 3-d medical images," IEEE Trans. Medical Imaging, vol. 12, no. 3, pp. 478–485, 1993.

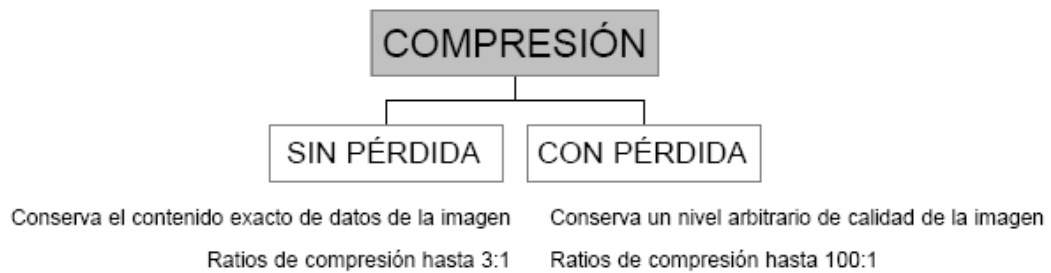


Figura 2.11 Diagrama de bloque y rango de compresión en el caso del proceso con pérdida y sin pérdida.

Algunos ejemplos de técnicas de compresión sin pérdidas son:

- Run-Length Coding(Código de longitud de cadenas)
- Codificación de Huffman
- Codificación en planos de bit
- Codificación del contorno
- Algoritmo de Lempel-Ziv-Welch

2.5.3.2 Compresión con pérdida: Conserva la calidad de la imagen hasta un nivel dado, pero no conserva el contenido total de datos de la imagen original (Existe pérdida de la información), es un compresión irreversible.

Permiten obtener un alto nivel de compresión, por lo que son las más utilizadas para el tratamiento digital de las imágenes. Sin embargo su éxito depende en gran medida de lograr que la distorsión causada por la pérdida de cierta información no pueda ser distinguida por el ojo humano.

Dentro de la compresión con pérdidas existen diferentes técnicas tales como: compresión mediante transformadas (PCA, Transformada Discreta Coseno, Wavelets, etc.), cuantificación umbral, codificación predictiva con pérdidas, LOT (*Lapped Orthogonal Transform*), codificación en subbandas, cuantificación

vectorial, quadrees, compresión mediante fractales, compresión mediante redes neuronales, entre otros. La relación de compresión puede ser de hasta 100:1

La primera forma de compresión a que se somete una imagen (analógica) es la cuantización del brillo de los pixels a valores digitales, llamada PCM (modulación código pulsación). El número de bits asignados a cada valor de brillo de los pixels, determina el tamaño (en datos) y la calidad de la imagen digital resultante. La imagen es codificada en una secuencia de valores de brillo de cada pixel.

Para la compresión con pérdidas existen diferentes estándares de compresión para las imágenes estáticas, tanto para las imágenes bitmap, como para las imágenes vectoriales y para los metaficheros (híbridos) estos son:

- Formatos bitmap: TIFF, JPEG, GIF, PNG, PCX, BMP, RAW, PICT, PSD, TGA
- Formatos vectoriales: DXF, CDR, entre otros.
- Formatos híbridos: CGM, WMF, WPG.

Tabla 2.3 Características de varios formatos de compresión

Formato	Compresión / Tipo	Profundidad de color	Uso típico
TIFF	Opcional / Sin pérdidas	1 a 64 bits	Imágenes de alta calidad, cámaras digitales, escáneres, impresión
RAW	Sin pérdidas	48 bits	Cámaras digitales
JPEG	Con pérdidas	8 o 24 bits	Cámaras digitales, Internet, impresión, intercambio de imágenes
GIF	Sin pérdidas	1 a 8 bits	Internet, imágenes de reducido tamaño, logos...
PNG	Sin pérdidas	1 a 48 bits	Internet, gráficos, iconografía software
PSD	Sin pérdidas	1 a 64 bits	Edición y manipulación



Figura 2.12 Fotografía guardada como GIF de 256 colores. El tamaño de archivo es 71.77K



Figura 2.13. Fotografía guardada como JPEG con un preset de alta calidad. El tamaño es 24.28K .



Figura 2.14 Fotografía guardada como PNG8 en un archivo de 62.64K



Figura 2.15. Fotografía guardada como PNG24 en un archivo de 211K.

- **Redundancia en las imágenes.**

Las técnicas de compresión son posibles porque todo conjunto de datos normalmente contiene redundancias.

En el caso de una imagen, dicha redundancia existe en forma de pautas de repetición y otras formas de información de brillo común entre varios píxeles de la imagen.

El objetivo de la compresión de la imagen reside en caracterizar estas redundancias y codificarlas de una forma distinta que requiera menos datos que la original.

Tipos de redundancia en las imágenes:

- **Interpíxel:** Los píxeles adyacentes tienden a estar altamente correlacionados. En muchas imágenes los niveles de gris no cambian rápidamente, sino gradualmente.
- **Psicovisual:** Cierta tipo de información es más importante para el sistema de visión humano que otro tipo de información.

Por ejemplo, el sistema de visión humano sólo puede percibir frecuencias espaciales por debajo de 50 ciclos (cualquier información de mayor frecuencia carece de interés).

Normalmente, cuando una operación de compresión se aplica iterativamente a una imagen, el tamaño de la imagen comprimida alcanza un límite a partir del cual ya no se reduce, o incluso puede aumentar.

La reducción del tamaño de una imagen tiene un límite, marcado por las características intrínsecas de la redundancia y por la calidad requerida en la imagen cuando se descomprima posteriormente.

2.6 La transformada discreta del coseno (DCT)

También denominada transformada del coseno es la más ampliamente utilizada en compresión de imágenes. Esta transformada cuenta con una buena propiedad de compactación de energía.

La DCT está bastante relacionada con la DFT, con la diferencia de que es una transformada real, debido a que los vectores base se componen exclusivamente de funciones coseno muestreadas

La DCT es, sin ninguna duda, una de las transformadas más populares en el mundo de la compresión de imágenes y video. La DCT es utilizada en normas de codificación como JPEG, H.261, H.263, MPEG-1, MPEG-2 Y MPEG-4, entre otras, lo que proporciona una prueba de sus elevadas prestaciones.

Básicamente, la DCT cambia los valores originales que representan intensidades de color en valores que representan frecuencias del coseno.

El proceso de descomponer un conjunto de muestras en un conjunto ponderado de funciones base cosinusoidales se denomina Transformada Discreta del Coseno

Directa (DCT), y al método de reconstruir el conjunto de muestras a partir del conjunto ponderado de funciones base cosinusoidales se le conoce como Transformada Discreta del Coseno Inversa (IDCT).

- **Concepto Matemático de la DCT y de la IDCT**

A la entrada del codificador, la imagen original se agrupa en bloques de 8 x 8, se traslada de enteros sin signo en el rango a enteros con signo con rango y se le da como la entrada de la transformada DCT. En la salida del codificador, la transformada DCT inversa (IDCT) devuelve los bloques 8 x 8 para formar la imagen reconstruida.

Las siguientes ecuaciones son las definiciones matemáticas idealizadas de la DCT 8x8 (2.2) y de la IDCT 8x8 (2.3)

$$F(u, v) = \frac{C(u)}{2} \frac{C(v)}{2} \left[\sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \right] \quad (2.2)$$

$$f(x, y) = \sum_{u=0}^7 \frac{C(u)}{2} \sum_{v=0}^7 \frac{C(v)}{2} F(u, v) \cos \frac{(2x+1)u\pi}{16} \cos \frac{(2y+1)v\pi}{16} \quad (2.3)$$

- 1: F es el valor del coeficiente de la línea “u” y columna “v” (dominio de la frecuencia), u y v varían de 0 a 7.
- 2: Estas constantes valen 0.707 ($\sqrt{1/2}$) para cuando “u” y/o “v” valen 0, y valen 1 en cualquier otro lugar del tablero.
- 3: Sumatoria de todos los términos de “x” entre 0 y 7.
- 4: Sumatoria de todos los términos de “y” entre 0 y 7.
- 5: Valor del píxel en una columna dada y línea (dominio del tiempo).

- 6: Generador de funciones vertical, produce la forma de onda cosenoidal vertical:
 x es el número de la columna de los píxeles y u es la frecuencia. La unidad se expresa en radianes.
- 7: Generador de funciones horizontal, produce la forma de onda cosenoidal horizontal: “ y ” es el número de la línea de los píxeles y “ v ” es la frecuencia.

La DCT toma cada bloque de 8x8 y lo descompone en 64 señales de una base ortogonal. Cada una de estas señales contiene una de las 64 frecuencias espaciales únicas que comprenden al espectro de la señal de entrada. La salida de la DCT es un conjunto de 64 amplitudes de señal de la base o coeficientes cuyos valores son únicamente determinados por una señal de 64 puntos de entrada en particular. Al coeficiente con frecuencia cero en las dos dimensiones se le llama coeficiente DC y a los otros 63 coeficientes, coeficientes AC. Estos nombres vienen del uso habitual de la DCT en electrónica para analizar circuitos donde hay simultáneamente términos de corriente continua (DC) y de corriente alterna (AC). En la figura 2.16 se muestra los pasos típicos para el proceso de compresión basado en la DCT.

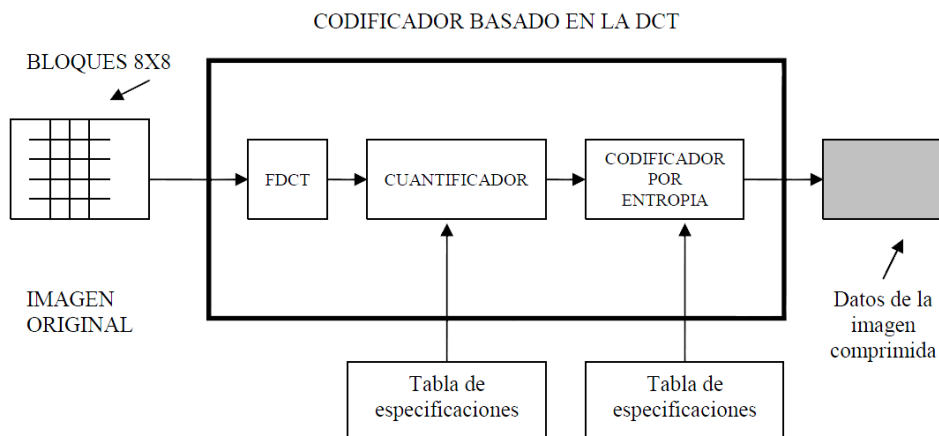


Figura 2.16 Pasos del proceso de codificación basado en la DCT

Por ejemplo, si se tiene la matriz de 8x8 y se aplica la DCT, se tendrá otra matriz con los coeficientes de frecuencia dados, como se observa en la figura 2.17.

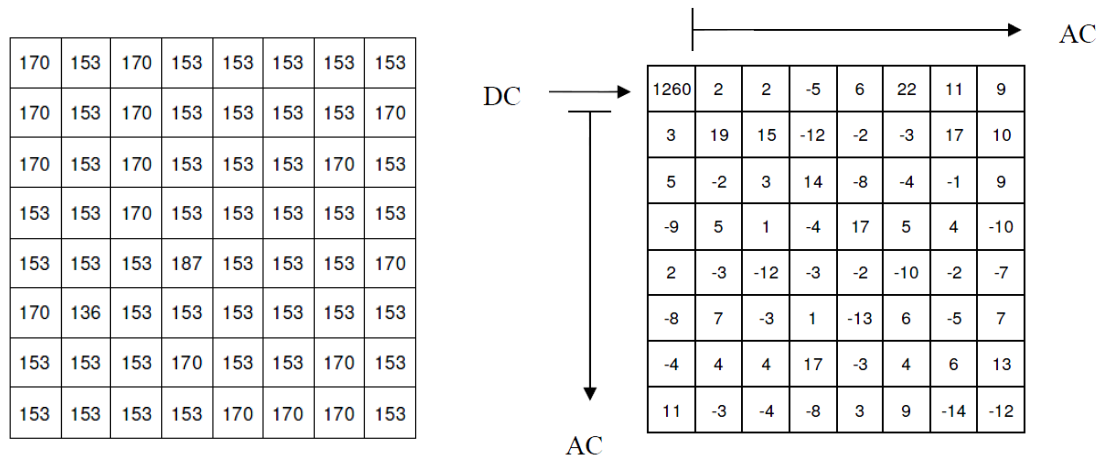


Figura 2.17 Matriz original y matriz luego del proceso de la DCT

En la figura 2.17 se observa que el primer coeficiente tiene el valor más alto, es la posición en donde está concentrada la mayor parte de la energía, a este valor se le llama valor DC.

Ya que los valores de la imagen varían lentamente entre píxeles consecutivos, el proceso de la DCT realiza la compresión de los datos concentrados la mayor parte de la señal en las frecuencias espaciales bajas. Para un bloque típico de 8 x 8 de una imagen típica, la mayoría de las frecuencias espaciales tienen una amplitud cero o casi cero y no serán codificadas.

La Transformada Discreta del Coseno (DCT) es utilizada para pasar la información temporal a información en la frecuencia. La DCT permite agrupar las componentes de baja y alta frecuencia en un patrón ordenado facilitando la codificación. La figura 2.18 muestra un ejemplo de una imagen en color a la cual se le aplica el proceso de compresión y descompresión mediante el método empleado en este proyecto.



(a)



(b)

Figura 2.18 Aplicación de la DCT, IDCT, (a) imagen original (b) imagen luego de la descompresión con 9 de los 64 valores que corresponden a cada bloque en que es dividida una imagen

El algoritmo de la Transformada del Coseno es el proceso que mayor carga computacional demanda en el proceso de codificación de una imagen o una secuencia de video y, por lo tanto, la implementación de un algoritmo rápido es importante para la optimización de tiempos en procesos en tiempo real.

- **Cuantificación**

Después de salir de la DCT, cada uno de los 64 coeficientes se cuantifica uniformemente en conjunción con una tabla de cuantificación de 64 elementos, que debe ser especificado por la aplicación o el usuario como entrada al codificador. Cada elemento de la tabla de cuantificación puede ser cualquier valor entre 1 y 255, que especifica el tamaño del paso del cuantificador para su coeficiente DCT correspondiente. El propósito de la cuantificación es conseguir aun más compresión representando los coeficientes DCT sin más precisión que la que es necesaria para conseguir la calidad de la imagen deseada.

Visto de otra forma, lo que se pretende con este proceso es descartar la información que no es visualmente relevante. La cuantificación es una aplicación

muchos a uno y, por tanto, es fundamentalmente con pérdidas. De hecho es la principal fuente de pérdidas en los codificadores basados en la DCT.

La cuantificación se define como una división de cada coeficiente DCT por su coeficiente de cuantificación, es decir, el tamaño del paso de la cuantificación correspondiente redondeado al entero más cercano.

$$F^*(u, v) = \text{Redondeo} \left(\frac{F(u, v)}{Q(u, v)} \right) \quad (2.4)$$

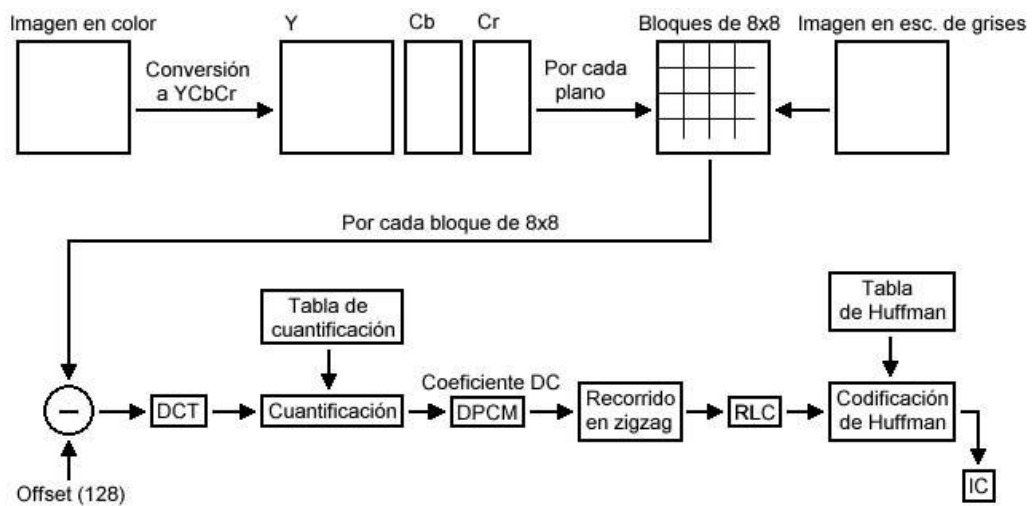


Figura 2.19 Proceso completo de compresión de imágenes

- **Algoritmo de Zigzag:**

Para aprovechar el patrón ordenado de componentes de frecuencia que se obtienen a partir de la DCT, se utiliza el algoritmo de Zigzag para hacer un mejor ordenamiento de las componentes de frecuencia. Con dicho algoritmo se logra obtener un vector en el cual las primeras componentes son las de baja frecuencia mientras que las últimas representan las componentes de alta frecuencia. Con ello

se efectúa un paso previo a los algoritmos de codificación que aprovechan tal circunstancia. Esta matriz se indica en la figura 2.10.

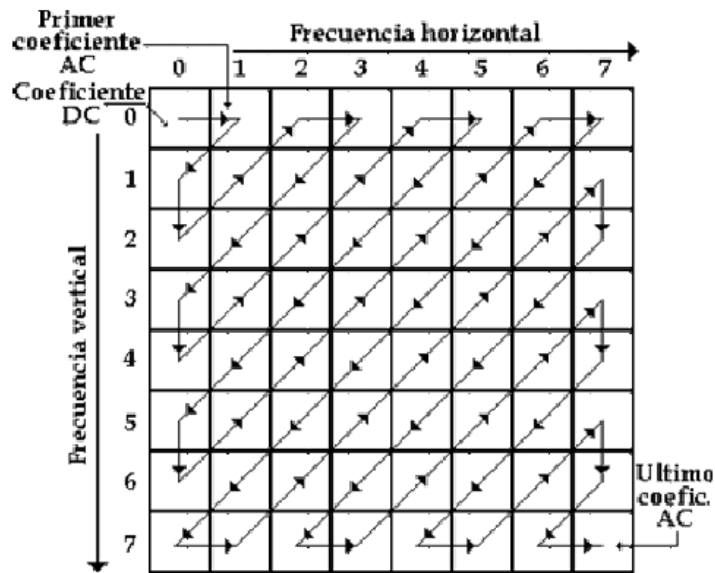


Figura 2.20 Barrido o exploración en zigzag

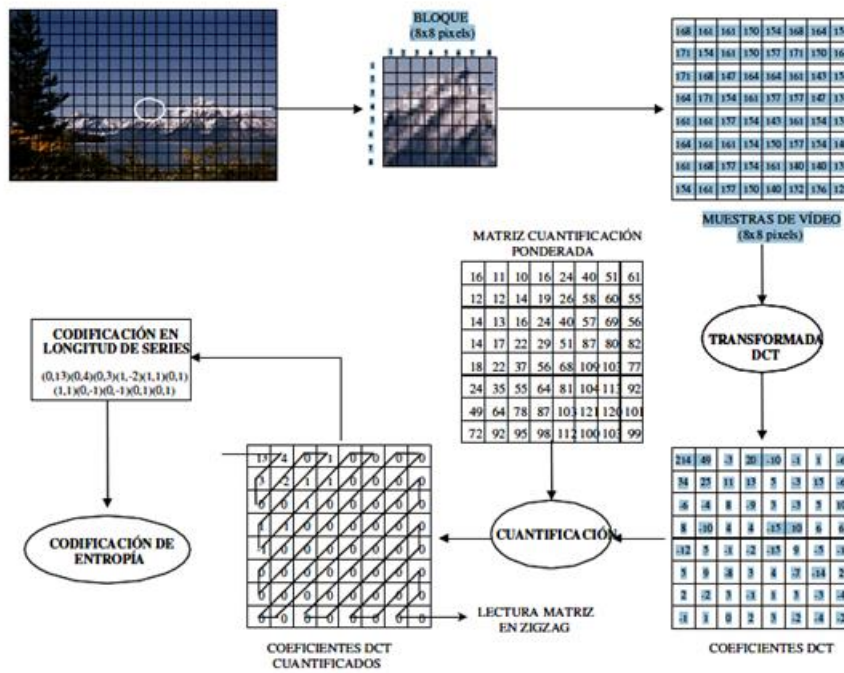


Figura 2.21 Secuencia en la compresión de imágenes en base a la DTC

En el decodificador, la IDTC realiza el proceso inverso. Toma 64 coeficientes DCT (que para entonces ya han sido cuantificados) y reconstruye una imagen de 64 puntos de salida sumando las señales base. Si la DCT y la IDCT pudieran ser calculadas con precisión absoluta y los coeficientes de la DCT no se cuantificaran se podría reconstruir perfectamente la imagen original, simplemente la transforma a un dominio en el que puede ser codificada más eficientemente.

La decuantificación es la función inversa que, en este caso, significa simplemente que en la normalización se eliminan multiplicando por el tamaño del paso, lo que devuelve el resultado a la representación apropiada para la entrada de la IDCT.

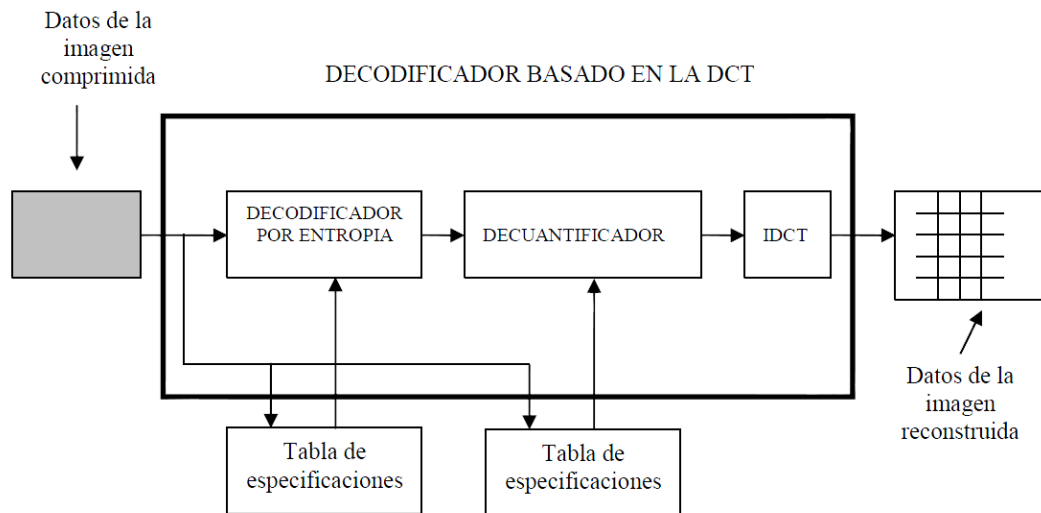


Figura 2.22 Pasos del proceso de decodificación basado en la IDCT

Para evaluar el rendimiento de un compresor de datos, hay que definir algunas mediciones.

- **Relación de compresión:** es la relación entre el tamaño de los datos de salida y el tamaño de los datos de entrada. Esta relación se indica en la ecuación 2.5

$$\text{Relación de compresión} = \frac{\text{Tamaño de los datos de salida}}{\text{Tamaño de los datos de entrada}} \quad (2.5)$$

Un valor de 1 significa que los datos de salida son idénticos a los datos de entrada, y no se ha encontrado redundancia en los datos de entrada. Un valor de 0.7 significa que el archivo comprimido ocupa el 70% del tamaño que ocupaba el archivo original.²²

- **Factor de compresión:** es el inverso de la relación de compresión como se indica en la ecuación 2.6.

$$\text{Factor de compresión} = \frac{\text{Tamaño de los datos de entrada}}{\text{Tamaño de los datos de salida}} \quad (2.6)$$

En este caso, un valor superior a 1 significa que el archivo ha sido comprimido. Este valor es más familiar para los usuarios debido a que mientras el factor de compresión es mayor, significa que los datos tienen mejor compresión.

- **Explosión de datos.**

Existen muchas técnicas de compresión de imagen. Cada una de ellas es eficiente en la compresión de ciertos tipos de imágenes. Pero cuando una técnica no es adecuada para un tipo particular de imagen, el resultado puede ser una pobre compresión, o lo que es peor, puede ocurrir una “explosión de datos”, es decir, que la imagen comprimida sea mayor en tamaño que la original.

Este fenómeno suele ocurrir cuando mediante una técnica de compresión se intenta eliminar un tipo de redundancia que no está presente en la imagen. Sucede cuando una imagen contiene pautas con características desconocidas por el algoritmo de compresión.

²² D. Salomon, Data compression: The Complete Reference, Springer, pp. 1-13, 719- 734, 2007.

Normalmente, un esquema de compresión suele tener una válvula de escape, de manera que si detecta que sus resultados están expandiendo el tamaño de la imagen, hace que el esquema procese la imagen con otro algoritmo o la deje intacta.

- **Entropía**

La entropía es una medida de su contenido en información. Si la entropía es elevada, la información contenida en la imagen tiende a ser muy impredecible. Es decir, una imagen con alta entropía contiene mucha aleatoriedad y poca redundancia. Si la entropía es baja, la información de la imagen es más predecible, contiene poca aleatoriedad y mucha redundancia.

Se puede medir la entropía de una imagen como la probabilidad de que ocurra dicha imagen. Se mide por un número que representa el número de bits necesarios para representar esa probabilidad. Para una imagen aleatoria cualquiera, sería:

$$\text{Entropía} = (\text{N}^\circ \text{ pixels}) \times (\text{N}^\circ \text{ líneas}) \times (\text{N}^\circ \text{ bits por pixel}) \quad (2.7)$$

- **Codificación de Huffman**

Es una técnica de codificación de entropía. Convierte los valores de brillo de los píxeles en nuevos códigos de longitud variable, basados en su frecuencia de aparición en la imagen. Así, a los brillos que aparecen frecuentemente se les asigna códigos más cortos, y a los menos frecuentes, códigos más largos.

Los códigos de Huffman se asignan creando un árbol de pares de valores de brillo, basado en sus frecuencias combinadas de ocurrencia. Este árbol asegura que los códigos más largos se asignen a los brillos menos frecuentes y viceversa.

- **Compresión de imágenes en color**

La compresión de imágenes en color se realiza comprimiendo cada imagen componente del color individualmente. En el caso de una imagen RGB (rojo,

verde, azul) se aplicaría el mismo esquema de compresión a las tres imágenes componentes del color. En la descompresión se aplicaría la misma a las tres imágenes componentes comprimidas, recuperándose la imagen original RGB.

Cada componente (píxel) de una imagen consta de 3 componentes R (Rojo), G (Verde) y B (Azul). Las combinaciones de dichas componentes generan la variedad de colores que percibe el ser humano. El estándar JPEG propone una representación distinta que permite una codificación más eficiente de las componentes de color. Mediante una transformación lineal se obtiene el formato Y (Luminancia), Cb (Crominancia Azul), Cr (Crominancia Roja). Luego de obtener el formato Y Cb Cr (YUV) se aprovecha el hecho de que el sistema visual humano es mucho más sensible a la información contenida en la Luminancia(Y) y menos sensitiva a la guardada en la Crominancia. Por lo tanto, se submuestran los componentes Cb y Cr. Con esto se reduce la cantidad de información sin afectar significativamente la calidad visual.

2.7 El estándar JPEG

Durante los últimos años de la década de los ochenta y los primeros de la de los noventa, un comité mixto ISO y CCITT conocido como JPEG (*Joint Photographic Experts Group*, Grupo Mixto de Expertos Fotográficos) trabajó para establecer el primer estándar internacional para compresión de imágenes estáticas de tono continuo tanto en escala de grises como en color.

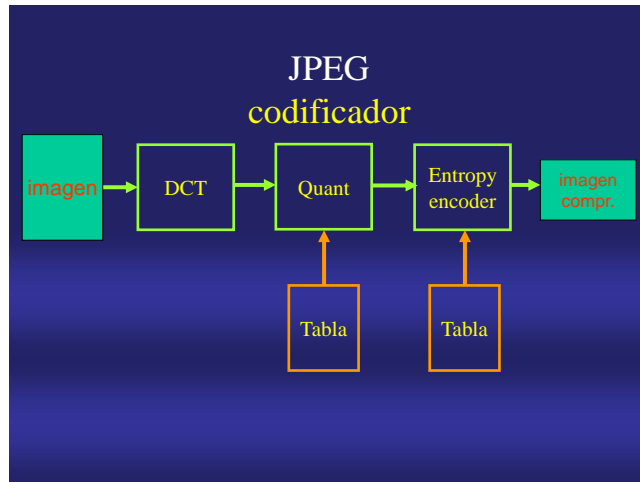


Figura 2.23 Etapas que conforman el codificador JPEG

El denominado *estándar JPEG* incluye dos métodos de compresión básicos, cada uno con varios modos de operación. El estándar especifica un método basado en la *transformada discreta del coseno* (TDC) para compresión con pérdidas y un método predictivo para compresión sin pérdidas.

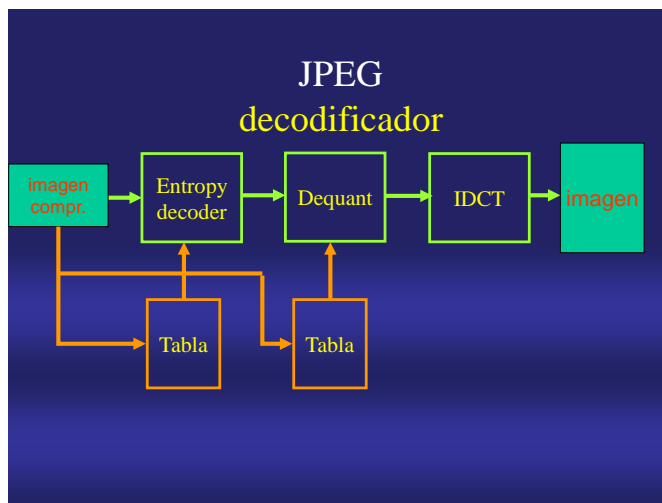


Figura 2.24 Etapas que conforman el decodificador JPEG

2.8 COMPRESIÓN DE IMÁGENES EN MOVIMIENTO

Las imágenes de vídeo están compuestas de información en el dominio del espacio y el tiempo. La información en el dominio del espacio es provista por los *píxeles*, y la información en el dominio del tiempo es provista por imágenes que cambian en el tiempo. Puesto que los cambios entre cuadros colindantes son diminutos, los objetos aparentan moverse suavemente.

Cuando se digitaliza una secuencia de video analógico cualquiera de acuerdo al estándar ITU-R BT.601 (CCIR 601), se requiere un ancho de banda de 116 Mbit/segundo ó de 116 millones de bits cada segundo. Dado que la mayoría de las redes son sólo de 100 Mbit/segundo, no es posible ni deseable transmitir las secuencias de vídeo sin alguna modificación. Para solucionar este problema se han desarrollado una serie de técnicas denominadas técnicas de compresión de vídeo e imágenes, que reducen el alto nivel de bits precisos para transmisión y almacenamiento.

En una imagen dinámica el contenido de ésta varía de acuerdo al movimiento, por lo que es necesario aprovechar la redundancia espacial y redundancia temporal para obtener niveles significativos de compresión.

Por ejemplo, para el caso de cuadros contiguos en los que no existe movimiento, la secuencia de éstos puede codificarse como una sola imagen e incluir solamente algunos bits para indicar que esa imagen se debe repetir. Si por el contrario, existe movimiento, se codifica la información que varía entre un cuadro y otro.



Figura 2.25 Ejemplo de una secuencia de imágenes

En muchos casos, la información de un mismo cuadro puede ser reducida para así evitar redundancia.

Algunos de los estándares de compresión de video que se pueden mencionar son: Norma H.261, MPEG, Norma H.263 y Norma H.264.

La compresión de movimiento es un caso particular de compresión de imagen. Se trata de eliminar la redundancia entre las imágenes (*frames*) de una secuencia de imágenes.

Es aplicable sólo para una sucesión de imágenes conexas, como ocurre en producciones de vídeo y películas cinematográficas, secuencias de imágenes médicas, y aplicaciones de teleconferencia.

La señal de video tiene 4 dimensiones: magnitud, tiempo, espacio horizontal X y espacio vertical Y. La eliminación de bits se aplica a las 4 características. Si bien los métodos prácticos se basan en compresión espacial (plano X, Y) y temporal, la magnitud está implícitamente afectada al cambiar la cuantificación de las muestras en el plano espacial.

- **Spatial Compression** (*compresión espacial*): Consiste en eliminar la redundancia en una imagen fija (still frame), basada en la similitud entre píxeles vecinos. Se conoce como codificación interna del frame (intraframe coding). Una de las técnicas más eficientes es en base a transformadas, como por ejemplo la transformada de coseno discreto (**DCT**).
- **Temporal Compression** (*compresión temporal*): La señal de video es una secuencia de frames en el tiempo, donde frames vecinos normalmente son similares y cambian solo algunos píxeles. La técnica para reducción de bits en función del tiempo se conoce como codificación inter-frame (interframe coding) y aprovecha la similitud de frames vecinos para transmitir solo la diferencia entre ellos.

En la codificación *interframe* normalmente, las diferencias entre las imágenes sucesivas de una secuencia son pequeñas y pueden deberse principalmente a objetos individuales que se mueven o a movimientos de la cámara. En estos casos, la mayor parte de una imagen es idéntica, entre dos imágenes sucesivas. Esto implica la existencia de mucha redundancia entre las imágenes. Las técnicas de codificación *interframe* eliminan esta redundancia basándose en la información de una imagen previa para obtener la siguiente, en lugar de recodificar la misma información varias veces.

Sin embargo, cuando cambia una escena entera, se debe tratar la nueva imagen como inconexa con ninguna previa imagen y recurrir al uso de la codificación típica de imágenes simples. Hay varias técnicas de codificación. Todas se basan en codificar, en una forma más concisa, la información que está duplicada entre dos imágenes. Así se reducen los tamaños resultantes de imágenes individuales. Para determinar qué información está entre dos imágenes, se mide el “movimiento” entre dos imágenes (frames). Este proceso se llama estimación de movimiento.

Las técnicas de compresión de movimiento generalmente parten de la compresión por bloques de imágenes estáticas. La primera imagen en una sucesión se codifica como si fuera una imagen individual. Las imágenes siguientes se codifican de la misma manera, pero con la adición de códigos especiales de bloque. Estos códigos relativos se usan para indicar que un bloque es idéntico al bloque correspondiente de la imagen de referencia, o que es una versión movida de un bloque de la imagen de referencia. Los códigos especiales de bloque se usan cuando un bloque es posible referirlo a un bloque similar en la imagen, en lugar de ser grabado.

2.8.1 Formatos de Video²³

Un archivo de video es una mezcla de imagen y audio en un sólo archivo digital. Los formatos de los archivos de vídeo han ido evolucionando con el paso del

²³ http://es.wikipedia.org/wiki/Formato_de_archivo_inform%C3%A1tico

tiempo, debido al enorme tamaño que ocupaban y al lento proceso de las descargas de las páginas Web que los contenían, se han desarrollado nuevos formatos de compresión para mejorar la velocidad en las conexiones en la transmisión de secuencias de video. A continuación se muestran los formatos más comunes que podemos encontrar

- **MPEG** (Moving Pictures Experts Groups). El estándar MPEG comprime secuencias de vídeo, además de audio, para poder ser almacenadas en los 650 ó 700 MB que puede contener un CD-ROM o para ser transferidos por la red. Los archivos MPEG corresponden a archivos con extensiones: .mpg, .mpeg, .m1v, .mp2, .mpa, .mpe, .mpm, .mpv, .m1 .a la mayoría de los navegadores de la WWW ya traen MPEG integrado.
- **Mov / Movie** Es el formato de vídeo desarrollado por Apple. Es el formato común para películas en QuickTime, la plataforma nativa para películas en Macintosh, pero puede utilizarse con diferentes reproductores en Windows.
- **Avi y Avi 2.0** (Audio Video Interleaved o Video for Windows) El formato AVI es el formato propietario desarrollado por Microsoft. La calidad de vídeo es moderada, sin embargo, se usan diferentes codecs (Codificador y Decodificador) para obtener resultados diferentes.

Su funcionamiento es muy simple pues almacena la información por capas, guardando una capa de video seguida por una de audio. Cuando capturamos video hacia nuestra computadora llega en formato *AVI*. Se generan archivos muy grandes y de difícil manejo.

- **Divx** El DivX es muy popular y está bastante relacionado con los DVDs y su piratería, ya que con dicho formato muchas personas pueden pasar sus DVDs a CDs con una calidad aceptable de Video.

Otros formatos²⁴

- **Microsoft Windows Media Video** : Funciona con el Windows Media player viene integrado en el sistema operativo. Las extensiones de este tipo de contenidos son: .asf y .wmv para el video y .wma para el audio.
- **Apple Quicktime .mov** Sus archivos .mov requieren de un player especial que es el Quicktime player para visualizarlos.
- **Real Video** Requiere de su propio player que es el Real Player, para hacer streaming requiere del Real Server.

La arquitectura de un vídeo define la forma en que se manejan y sincronizan los datos de Vídeo. Cada arquitectura está optimizada para una plataforma específica. Por ejemplo, Windows Media está optimizado para el PC, QuickTime para Apple y Real Systems para la Web.

Todas las arquitecturas de vídeo son similares en funciones, pero difieren ampliamente en los detalles. Todas definen los formatos de los archivos y sus extensiones, pero no necesariamente especifican el formato actual de los datos de vídeo almacenado en el archivo. La arquitectura puede restringir el formato de vídeo, pero no siempre se da el caso. Por ejemplo, *Windows Media* y *QuickTime* manejan una amplia variedad de formatos de vídeo. Esto es porque la arquitectura de vídeo sólo define el entorno.

Este entorno maneja los datos de vídeo con la información necesaria para reproducirse. Por consiguiente, cuando se abre un archivo con las extensiones **.avi** o **.mov**, lo que realmente se abre es el entorno. El computador lee la información del entorno y llama al software apropiado para que se reproduzca ese archivo correctamente.

²⁴ <http://www.slideshare.net/chulasan/formatos-de-video>

No todas las arquitecturas son flexibles. *Real Systems* sólo soporta video comprimido con su propio *codec propietario*, así la información del entorno no tiene el codec específico para poder descodificar el video comprimido. Las extensiones **.avi** y **.mov** identifican la arquitectura de vídeo. El resto se encuentra dentro del entorno **.avi o .mov**

2.9 Uso de la compresión para la transmisión

Ya que el ancho de banda disponible en Internet (a un coste razonable, claro está) no aumenta del mismo modo que aumenta en la red local, hay que buscar productos de vídeo que optimicen la calidad de la imagen con el ancho de banda disponible. Eso es lo que hace el uso de estándares como el H.261 y MPEG4.

La optimización de la calidad de la imagen y de la frecuencia de imágenes por segundo para el ancho de banda de que se dispone, permite que éste ya no sea una limitación para la transmisión de imágenes en tiempo real por Internet con una calidad mucho más que aceptable. Y no se trata solamente de la transmisión de imágenes utilizando líneas ADSL.

En el mercado actual coexisten diversas técnicas de codificación, unas basadas en la compresión espacial de las imágenes, y otras en la compresión temporal de las secuencias de vídeo analizando las variaciones de movimiento entre una imagen y la siguiente. El resultado de las primeras es el envío de imágenes comprimidas que reproducidas secuencialmente componen el vídeo.

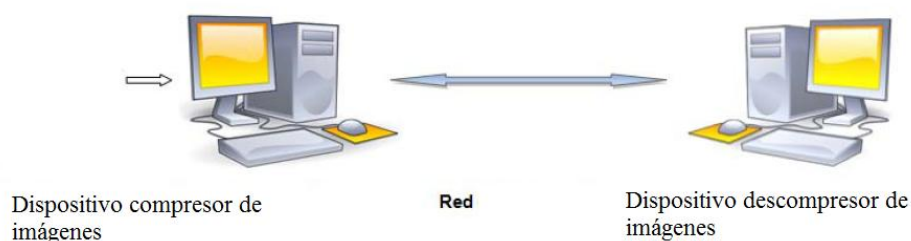


Figura 2.26 Esquema de la aplicación de un Códec de compresión de imágenes

La técnica más extendida es el Motion-JPEG o MJPEG. En el segundo caso, el resultado es la transmisión de vídeo comprimido tanto en el espacio como en el tiempo. Con esto se obtiene un uso más eficiente del ancho de banda, así como del espacio de almacenamiento en disco.

2.10 Redes basadas en IP

Las redes basadas en IP utilizan la tecnología de conmutación de paquetes, que usa la capacidad disponible de una forma mucho más eficiente y que minimiza el riesgo de posibles problemas como la desconexión. Los mensajes enviados a través de una red de conmutación de paquetes se dividen primero en paquetes que contienen la dirección de destino. Entonces, cada paquete se envía a través de la red y cada nodo intermedio o router de la red determina a donde va el paquete. Un paquete no necesita ser enrutado sobre los mismos nodos que los otros paquetes relacionados. De esta forma, los paquetes enviados entre dos dispositivos de red pueden ser transmitidos por diferentes rutas en el caso de que se caiga un nodo o no funcione adecuadamente

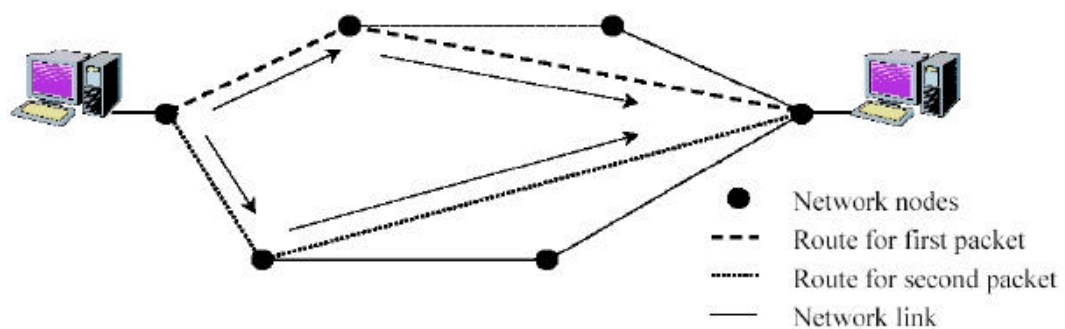


Figura 2.27 Red de conmutación de paquetes

Las redes basadas en IP solucionan muchos de los problemas a los que se enfrenta en un entorno complejo, a la vez que proporciona una solución elegante que cubre las necesidades actuales y las venideras. Últimamente, todas las formas de

comunicación que incluyan datos, voz imágenes en movimiento y entretenimiento convergirán en una red de transporte común.

La ventaja fundamental de las redes con IP es que dado que todos los elementos de la red son identificables por su dirección IP, la gestión del sistema se hace más sencilla simplificando las labores de operación y maximizando la utilidad de la red de comunicación. Las labores de conmutación son realizadas por software dirigiendo una dirección IP-fuente de vídeo, a una dirección IP-equipo de destino; de esta forma, los ordenadores de gestión del sistema de video se convierten en las matrices de conmutación de los sistemas de vídeo tradicionales.

En un sistema de video IP hay múltiples procesos ejecutándose simultáneamente. A continuación se describen algunos de los más importantes relacionados con la compresión:

- **Codificación:**

El proceso que se realiza en la cámara de red o el servidor de vídeo que codifica (digitaliza y **comprime**) la señal de vídeo analógico de manera que pueda transmitirse a través de la red.

- **Transmisión IP:**

Transmisión sobre una red de datos basada en el protocolo IP, inalámbrica o con cableado, desde una fuente a hardware variado de grabación o visualización (por ejemplo un servidor de PC's).

- **Grabación:**

Datos transferidos a discos duros estándar conectados a un dispositivo de almacenamiento como puede ser un servidor, NAS (Network Attached Server) o SAN (Storage Area Network).

- **Decodificación:**

El vídeo codificado debe ser traducido, o decodificado, con el fin de ser visualizado/monitorizado. Este proceso se realiza en un PC o en otro sistema decodificador que se emplee para visualizar el vídeo.

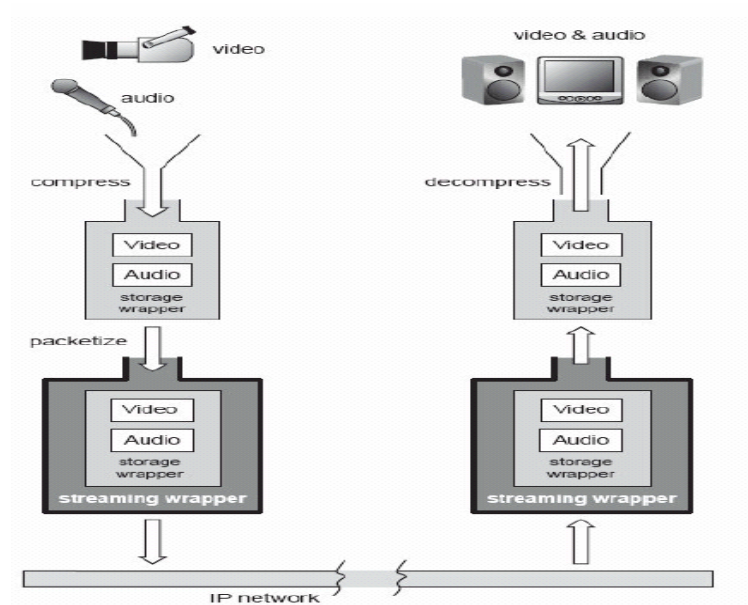


Figura 2.28 Digitalización, compresión y transmisión de video

2.11 MATLAB

MATLAB es el nombre abreviado de “MATrix LABoratory”. Es un programa para realizar cálculos numéricos con vectores y matrices, y por tanto se puede trabajar también con números escalares (tanto reales como complejos), con cadenas de caracteres y con otras estructuras de información más complejas.

Matlab es un lenguaje de alto rendimiento para cálculos técnicos, es al mismo tiempo un entorno y un lenguaje de programación. Uno de sus puntos fuertes es que permite construir herramientas propias reutilizables. Se puede crear

fácilmente funciones propias del usuario y programas especiales (conocidos como M-archivos) en código Matlab, se pueden agrupar en Toolbox (también llamadas librerías): colección especializada de M-archivos para trabajar en clases particulares de problemas.

Matlab, aparte del cálculo matricial y álgebra lineal, también puede manejar polinomios, funciones, ecuaciones diferenciales ordinarias, gráficos, etc....

Matlab puede llamar a una gran variedad de funciones, tanto propias como programadas por los usuarios. A veces, puede haber funciones distintas que tengan el mismo nombre. Por tanto, es interesante saber cómo Matlab busca cualquier función que se le pida que ejecute. La clave es el camino de búsqueda (search path) que el programa utiliza cuando encuentra el nombre de una función. El search path es una lista de directorios que se puede ver y modificar mediante la orden path, o utilizando el Path Browser²⁵

2.11.1 El espacio de trabajo de Matlab

Nada más abrir Matlab (podemos hacerlo pinchando en el icono que aparece en el escritorio o en su defecto en Inicio->Todos los programas) aparecerá una pantalla como la siguiente:

²⁵ <http://personal.us.es/contreras/practica1.pdf>

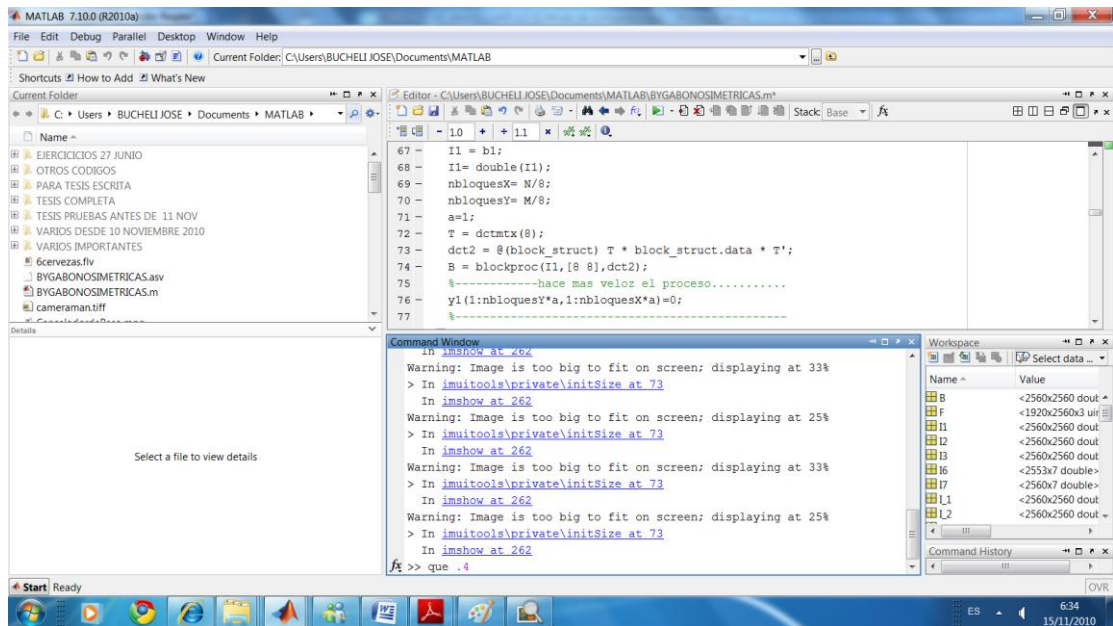


Figura 2.29 Pantalla de trabajo de MatLab. R2010a

2.11.2 Formatos de imágenes soportados por Matlab²⁶

MatLab soporta los siguientes formatos de imágenes: .BMP, .HDF, .JPEG, .PCX, .TIFF, .XWB. La mayoría de las imágenes disponibles hoy en día, se encuentran en el formato JPEG, el cual es el nombre para el estándar de compresión de imágenes que más se utiliza.

Una imagen puede ser “truecolor” o indexada. Una imagen indexada, de altura h y ancho w , se almacena mediante dos matrices: una matriz de índices tamaño $h \times m$ y un mapa de colores. Cada celda de la matriz de índices corresponde a un pixel de la imagen y contiene un índice a un color concreto del mapa de colores. Por ejemplo, para una imagen en color de 8 bits hay 256 posibles colores; cada celda de la matriz de índices tiene un valor entero entre 0 y 255 que es un índice a uno de los 256 colores descritos en el mapa de colores.

²⁶ <http://www.udb.edu.sv/Academia/Laboratorios/biomedica/Imagenes%20Medicas/guiaIIM.pdf>

Al trabajar en MatLab, se deben considerar múltiples factores como la manera en la que se carga una imagen, el formato correcto, guardar la información de diferentes maneras, como desplegar una imagen, conversiones entre diferentes formatos de imágenes, etc.

2.11.3 Formatos de trabajo en Matlab²⁷

- **Imagen de Intensidad (Imagen en escala de grises)**

Representa una imagen como una matriz, donde cada elemento tiene un valor que se corresponde a cuan brillante u oscuro debería ser coloreado el píxel en la posición correspondiente. Existen dos formas de representar el número que a su vez representa la brillantez del píxel. La clase **double**, asigna un número flotante (número con decimales), entre 0 y 1 a cada píxel. El valor 0 corresponde al Negro y el 1 al Blanco. La otra clase es la **uint8** la cual asigna un entero entre 0 y 255 para representar la brillantez de un píxel. El valor 0 corresponde al Negro y 255 al Blanco.

La clase **uint8** requiere aproximadamente un octavo de almacenamiento comparado con lo que requiere la clase **double**. Sin embargo, debe considerarse el hecho que muchas funciones matemáticas pueden aplicarse únicamente con la clase **double**.

- **Imagen Binaria**

Este formato de imagen también guarda una imagen como una matriz pero un píxel de la misma únicamente tiene dos posibles colores Blanco o Negro (y nada entre ellos). Este formato de imagen asigna 0 para el Negro y 1 para el Blanco.

²⁷ <http://www.udb.edu.sv/Academia/Laboratorios/biomedica/Imagenes%20Medicas/guiaIIM.pdf>

- **Imagen Indexada**

Es una forma práctica de representar las imágenes a color. Una imagen indexada guarda una imagen como dos matrices. La primera matriz tiene el mismo tamaño que la imagen y un número para cada píxel. La segunda matriz se conoce como Mapa de Color (Colormap) y su tamaño puede diferir del de la imagen. Los números en la primera matriz indican el número a usar en la matriz Mapa de Color.

El mapa de colores es una matriz que tiene tres columnas, una por cada color básico (rojo, verde y azul o RGB) y una fila por cada posible color. Para el ejemplo de 256 colores, el mapa de colores sería una matriz de tamaño 256 x 3. Cada fila del mapa de colores describe un color concreto y contiene tres números en punto flotante, con un valor entre 0 y 1, que dan las cantidades de rojo, verde y azul que tiene el color

- **Imagen RGB (Truecolor)²⁸**

Este es otro formato para imágenes a color. Se presenta en la forma de un arreglo de $m \times n \times 3$ de clases **uint8**, **uint16**, **single**, o **double** cuyos valores de píxel especifican valores de intensidad. Este representa una imagen con tres matrices de tamaños que concuerdan con el de la imagen. Cada matriz corresponde a uno de los colores básicos Rojo, Verde o Azul y establecen cuanto de cada uno de estos colores debe ser usado en un determinado píxel.

- **Imagen Múltiples-Frames**

Cuando se requiere hacer estudios de secuencias de imágenes, este formato es el indicado. Por ejemplo en imágenes médicas donde se requiere el estudio de secuencias de cortes transversales, longitudinales, etc.

²⁸ <http://www.udb.edu.sv/Academia/Laboratorios/biomedica/Imagenes%20Medicas/guiaIIM.pdf>

2.11.4 GUIDE (Graphical User Interfase Development Environment) es un juego de herramientas que extiende por completo el soporte de MATLAB, diseñadas para crear GUIs (Graphical User Interfaces) fácil y rápidamente dando auxilio en el diseño y presentación de los controles de la interfaz, reduciendo la labor al grado de seleccionar, tirar, arrastrar y personalizar propiedades.

Una vez que los controles están en posición se editan las funciones de llamada (Callback) de cada uno de ellos, escribiendo el código de MATLAB que se ejecutará cuando el control sea utilizado.

A la herramienta GUIDE se accede de varias maneras, la primera de ellas es tecleando guide en la ventana de comando.

```
>> guide
```

Otra manera de entrar a GUIDE es través del File opción New y por último el GUI, (como se muestra en la figura).

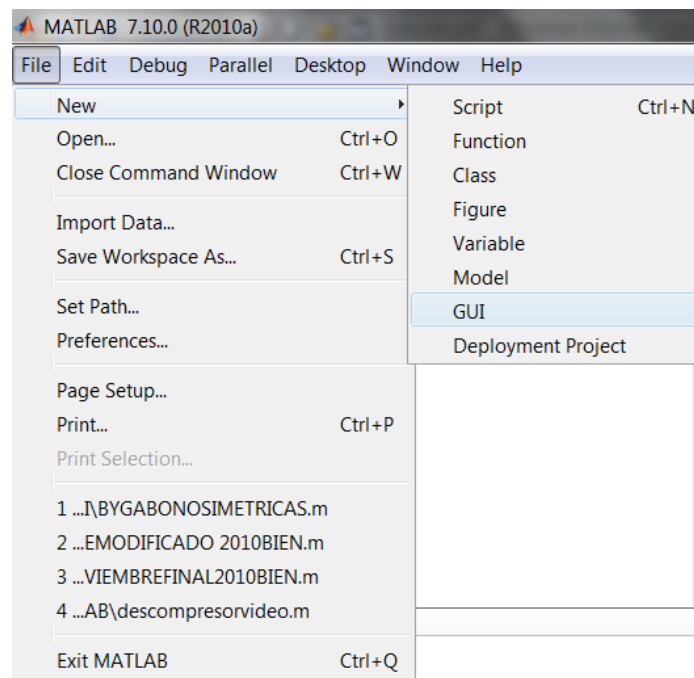


Figura 2.30 Pantalla para acceder al Guide

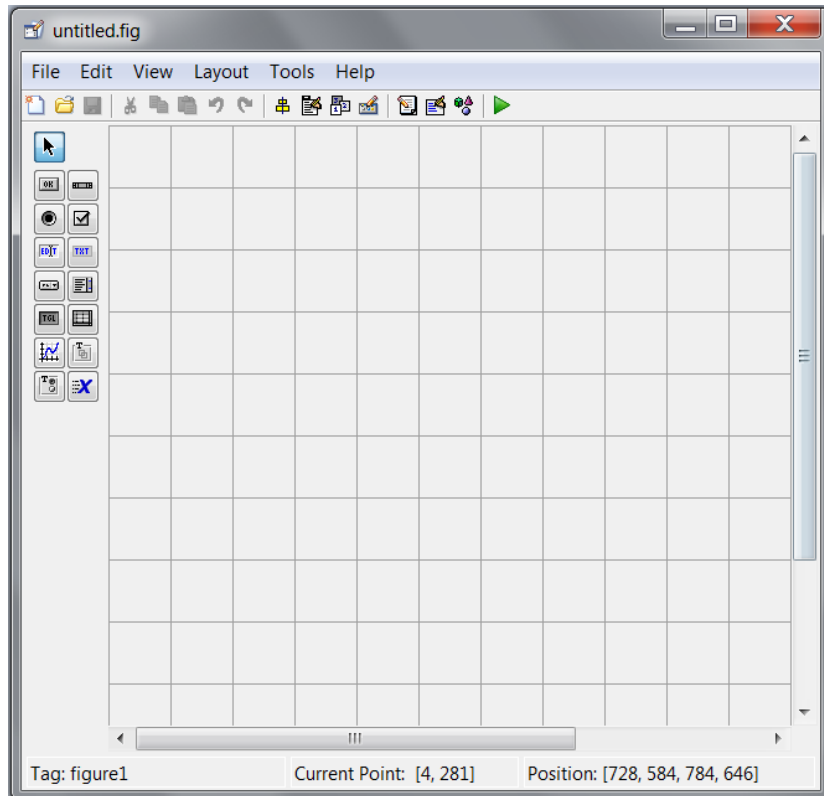


Figura 2.31 Ventana principal del Guide

2.12. HIPÓTESIS

La Transformada Discreta Coseno para la compresión de imágenes en movimiento permitirá optimizar la transmisión en redes que utilicen el protocolo IP con una menor cantidad de bits y aumentado la velocidad de transmisión.

2.12.1. SEÑALAMIENTO DE VARIABLES DE LA HIPÓTESIS

2.12.1.1. VARIABLE INDEPENDIENTE

Transformada discreta coseno para la compresión de imágenes en movimiento

2.12.1.2 VARIABLE DEPENDIENTE

Transmisiones de imágenes en movimiento con una menor cantidad de bits y mayor velocidad en redes que utilizan el protocolo IP.

CAPÍTULO III

METODOLOGÍA

3.1. ENFOQUE

El presente proyecto de investigación tuvo un enfoque analítico, deductivo comparativo, entre las diferentes técnicas de compresión y se experimentará el compresor y descompresor diseñado con sus características y técnicas de codificación sobre una red que utilice el protocolo IP, que determine su desempeño.

3.2. MODALIDAD DE INVESTIGACIÓN

Este proyecto se llevó a cabo mediante una investigación bibliográfica – documental, así como también mediante la experimentación en laboratorio.

La investigación bibliográfica - documental, permitió la recolección de información, datos y fórmulas que se encuentran en los textos, papers, y en la Web como aporte científico y como soporte técnico para fundamentar el empleo de estándares, algoritmos, y programas computacionales.

La experimentación en laboratorio permitió determinar el algoritmo más adecuado para el funcionamiento óptimo de todo el proceso de compresión y descompresión de imágenes digitales en movimiento y su interface con el usuario.

3.3. NIVEL O TIPO DE INVESTIGACIÓN

- La metodología de la presente investigación fue descriptiva, explicativa y experimental.
- Se realizó un estudio sobre la representación de imágenes en movimiento

- Se realizó un estudio sobre las etapas que abarca el proceso de compresión y descompresión utilizando como método de codificación de la Transformada Discreta Coseno.
- Se realizaron pruebas de utilización del algoritmo para el desarrollo de la DCT que determine el mejor proceso de compresión de imágenes fijas y luego de imágenes en movimiento.
- Mediante el manejo y observación por parte de varias personas encuestadas del proceso de compresión y descompresión, se validará el proyecto

3.4. POBLACIÓN Y MUESTRA

El personal de la Escuela Politécnica del Ejército Sede Latacunga entre estudiantes y docentes es de 3100, se determina el tamaño de la muestra y se realizó la encuesta a un grupo de personas con diferentes grados de conocimiento del tema, pertenecientes a la población (Muestreo Aleatorio Simple (MAS)); a los cuales se les encuestó sobre: *“la calidad de la imagen transmitida y recuperada”*.

Por lo tanto, el tamaño muestral para estimar el parámetro en el muestreo aleatorio simple, fijado el error de muestreo se utiliza¹:

$$n = \frac{Nz^2 pq}{e^2 N + z^2 pq} \quad \text{Ec. 3.1}$$

Donde:

n es el tamaño de la muestra;

z es el nivel de confianza;

p.- Probabilidad a favor

q.- $q = 1 - p$

¹ Alba Fernández, María Virtudes. Muestreo estadístico. España: Septem Ediciones, 2005. p 42.

N es el tamaño de la población;
e es la precisión o el error.

3.5 TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

La recolección de la información se la realizó mediante la aplicación de una prueba estándar para la determinación de la calidad de las imágenes en movimiento recuperada, esta prueba se la realizó en la Escuela Politécnica del Ejército Extensión Latacunga.

El instrumento preparado está apoyado completamente en el test internacional de procesamiento de voz e imágenes conocido como MOS TEST.

3.5.1. MOS TEST

MOS (Mean Opinion Score). Nota media de opinión. Se refiere a una valoración cuantitativa de la calidad subjetiva de una señal (por ejemplo, voz, imágenes o vídeo).²

El sistema de medida subjetivo MOS constituye la medida de mayor difusión para la determinación de la calidad de imagen, pero como todo método subjetivo su ejecución acarrea una serie de desventajas obvias: el tiempo y los recursos necesarios, el coste y la imposibilidad de automatizar el proceso³

Según la recomendación de **UIT-R BT.500-11**⁴, una prueba MOS realiza comparaciones entre la imagen a transmitir, que es la referencia de la prueba, con la imagen recibida y descomprimida y produce una valoración numérica de 1 (pobre) a 5 (excelente), esta medida relaciona la calidad de la imagen descomprimida que visualiza el observador.

² <http://www.zonagratis.com/curiosidades/DicTelefonica/M.htm>

³ <http://www.televisiondigital.es/TecnologiasRelacionadas/AltaDefinicion/ForoTVAD>

⁴ http://www.dii.unisi.it/~menegaz/DoctoralSchool2004/papers/ITU-R_BT.500-11.pdf

Este test permitió obtener respuestas concretas sobre la calidad de la reproducción de imágenes en movimiento proporcionado por el compresor y descompresor desarrollado.

3.6 PROCESAMIENTO Y ANÁLISIS

Los diferentes imágenes en movimiento pasaran a través del compresor y se harán pruebas con relaciones de compresión seleccionable y se lo enviará a otra PC a través de una red IP. En el decodificador se selecciona el archivo comprimido y se lo descomprime. Unas interface gráfica (realizada en MATLAB) permiten al operador un entorno visual de fácil accesibilidad y manejo.

Se presentará cuadros del grado de aceptación visual, resultado de la muestra obtenida al personal de la ESPE Extensión Latacunga.

3.7 GUÍA PARA EVALUAR LA CALIDAD DE LA IMAGEN

A la hora de evaluar las diferencias observadas entre la imagen de referencia y la imagen a evaluar, se utiliza la escala discreta de apreciación de cinco notas que se indica en la tabla 3.1:

El formato de encuesta aplicado se muestra en el ANEXO A

Tabla No. 3.1 Encuesta Mos Test

ESCALA DE CINCO NOTAS DE CALIDAD	
CALIFICACIÓN	DEFINICIÓN
5	Imperceptible
4	Perceptible pero no molesta
3	Ligeramente Molesta
2	Molesta
1	Muy molesta

CAPÍTULO IV

PROPUESTA

4.1 TEMA: DISEÑO DE UN COMPRESOR Y DESCOMPRESOR PARA LA TRANSMISIÓN DE IMÁGENES EN MOVIMIENTO CON DIFERENTES RELACIONES DE COMPRESIÓN SOBRE UNA RED VÍA PROTOCOLO DE INTERNET (IP).

4.2 DESARROLLO DE LA PROPUESTA

La siguiente etapa del estudio tiene el propósito de describir los requerimientos, las características técnicas del diseño para la implantación del sistema de compresión, descompresión y transmisión de imágenes en movimiento sobre una red en base al protocolo IP.

4.2.1. INTRODUCCIÓN

Aunque las tasas de bits de las redes, tanto de las LAN como de las WAN, se incrementen dramáticamente, las conexiones de banda ancha se hagan comunes en la mayoría de los hogares y la capacidad de almacenamiento de los discos duros, memorias flash y medios ópticos sea mucho más grande que antes, la compresión de imágenes en movimiento sigue siendo necesaria puesto que esta trae consigo dos importantes beneficios. Primero, la posibilidad de utilizar un video digital en ambientes de transmisión y almacenamiento que no soportan video sin comprimir y segundo, la compresión permite un uso más eficiente de los recursos de transmisión y almacenamiento, ya que es preferible enviar por el mismo medio un video de alta resolución comprimido o varios canales de video comprimidos que enviar una cadena de video de baja resolución sin comprimir por el mismo canal.

Las necesidades de comunicación están cambiando y el resultado es la creciente demanda de servicios de contenido Multimedia sobre las redes actuales y su rápida evolución. Es también indiscutible la implantación definitiva del Protocolo IP desde los ámbitos empresariales a los domésticos, por lo tanto la intención de esta proyecto es ubicar al lector en el entorno de la compresión, transmisión sobre redes que usan e protocolo IP y la respectiva descompresión.

4.2.2 ESPECIFICACIÓN DE REQUERIMIENTOS

En este apartado se describe las características de la solución computacional que se desarrolló. Se definen los requerimientos funcionales, requerimiento de desempeño y operacionales, requerimiento de usuario y del entorno.

Esta especificación de requisitos se ha estructurado tomando como referencia las directrices dadas por el standard “IEEE Recommended Practice for Software Requerements Specification ANSI/IEEE 830 1998”¹

4.2.2.1 Funciones del Sistema

La solución computacional que se va a desarrollar no constituye un elemento o componente de una solución mayor. Sin embargo debe ser soportada por la herramienta de desarrollo utilizada (Matlab Versión 7.10).

¹ Especiación de Requisitos según el estándar de IEEE 830 -IEEE Std. 830-1998

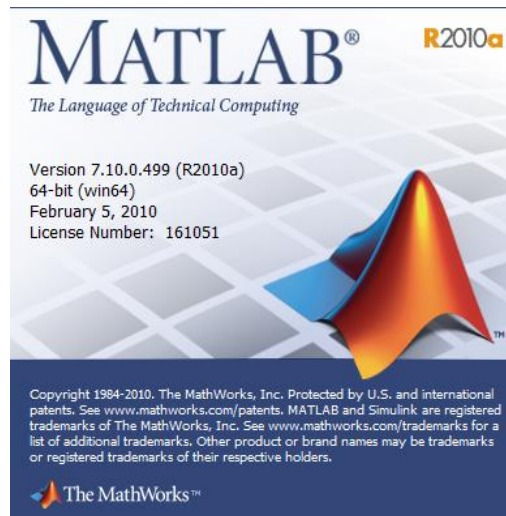


Figura 4.1 Imagen de MATLAB versión 7.10 R2010 utilizada en el proyecto

En términos generales, el sistema deberá proporcionar soporte a las siguientes tareas de gestión.

- Compresión de la información que representa las imágenes en movimiento
- Transmisión de la información a la red vía protocolo IP
- Descompresión de la información que representa las imágenes en movimiento.

4.2.2.2 Fuera del Alcance

Existen ciertas funciones que se encuentran fuera del alcance de la solución computacional.

- El formato del video de entrada es RGB 24 bits

- Siendo el proceso matemático muy demorado y complejo, la señal de entrada está limitada a tramas hasta de 300 x 300 pixeles y con un tamaño que no sobrepase 1MB.
- El número de tramas no debe ser elevado, se requiere mucho tiempo para el proceso de descompresión

4.2.2.3 Requisitos Específicos

El sistema deberá permitir:

- Abrir un archivo de imágenes en movimiento con o sin compresión RGB de 24 bits
- Reproducir el archivo en formato original
- Selección del número de tramas a comprimir
- Selección de la relación de compresión
- Compresión de la información de video
- Envío a otra PC a través de la red Ethernet vía protocolo IP
- Selección del Frame Rate en el descompresor
- Descompresión de la información que representa las imágenes en movimiento
- Visualización de la imagen comprimida

4.2.3 REQUISITOS DE INTERFACES EXTERNAS

4.2.3.1 Interfaces de Usuario

Utilizamos una herramienta en Matlab la cual nos permite mostrar en forma didáctica y de manera sencilla, la compresión y descompresión. Para mostrar lo anterior utilizamos una interface gráfica de usuario (GUIs).

4.2.3.2 Interfaces Software.

- Cada dispositivo del sistema de video sobre IP, poseerá una dirección única y exclusiva.

4.2.3.3. Interfaces de Comunicación.

- La tarjeta de Red o (NIC)
- La red LAN deberá emplear un cableado de red estándar, incluyendo: CAT5 o CAT6 (Par Trenzado no Apantallado – UTP).
- Equipo de conectividad (Routers, switch)

4.2.3.4 Requisitos Tecnológicos

La aplicación cliente se ejecutará sobre un PC con la siguiente configuración mínima:

- Procesador Corel Duo de 2.4 GHz o superior.
- Memoria 1 GB

- Espacio libre en disco 500 Mb
- Tarjeta de video

El sistema operativo en la que se debe ejecutar la aplicación será el windows XP, Windows 7.

4.2.4 Complejidad computacional (Aritmética)

La complejidad computacional aritmética se define como el procesamiento realizado sobre los datos usados. Los tipos de datos de punto flotante comprometen mucha mayor complejidad que la que necesitan los datos de tipo enteros o boléanos. Del mismo modo algunas de las operaciones aritméticas como la sustracción o la adicción utilizan mucha menos área de implementación y tiempo de ejecución que otras, como por ejemplo divisiones, raíces cuadradas y funciones trascendentales.

4.3 DISEÑO DEL SOFTWARE

Las características del video natural se pueden separar en dos importantes grupos. Las características espaciales (la variación de la textura al interior de la escena, el número y la forma de los objetos, el color, etc.) y las características temporales (el movimiento de los objetos, cambios en la iluminación, movimiento del punto de vista de la cámara, etc.)

Para el proceso de compresión, en este proyecto se optó por el aprovechamiento de las características espaciales de las imágenes.

En este proyecto se aprovecha las características del sistema visual humano en el proceso de compresión. Se ha demostrado que el sistema visual humano es más insensible a las variaciones de alta frecuencia. Aprovechando esta cualidad, se eliminará información de las imágenes a codificar (es por ello que el sistema tiene pérdidas) y se aprovechará este hecho para lograr la compresión.

Puesto que un video se compone de imágenes en movimiento, el proceso de compresión elaborado consiste en la compresión de cada imagen a color con el procedimiento en forma general a continuación indicado.

Después de que se selecciona un archivo de imágenes en movimiento a color, se selecciona cada una de los frames (tramas), la cual es representada por tres matrices, una para cada color (RGB). Esta matriz es dividida en submatrices (bloques) de 8x8 con la finalidad de que se efectúe con rapidez la DCT en cada uno de estos bloques.

La decodificación es esencialmente el proceso inverso al de la codificación. Se llevan a cabo los mismos procesos, pero en orden inverso.

4.3.1 DESCRIPCIÓN DEL SISTEMA

La ventana del sistema está conformada por diferentes opciones que permitirán al usuario escoger el archivo de imágenes en movimiento almacenadas en su PC. La ventana mostrará la imagen original. Otra opción que el usuario podrá escoger es el número de coeficientes de la imagen desea eliminar, cuantas tramas, así como la posibilidad de enviar el archivo vía protocolo IP a otra máquina conectada a la red.

4.3.1.1 COMPRESOR DE IMÁGENES EN MOVIMIENTO

Procesamiento

- Archivo de imágenes en movimiento: Recibe un archivo en formato MPG y lo almacena en matrices para cada uno de los tres colores (RGB), donde cada coeficiente requiere de 24 bits para ser representado.
- División de la imagen en bloques de 8x8. Aquí se divide cada una de las tres matrices de la imagen en submatrices (bloques) de dimensiones 8x8, obteniéndose 64 valores para cada una.
- En caso de que las matrices no sean cuadradas se las rellena con valores “Cero” para conseguirlo.
- Transformación DCT. Se aplica la DCT a cada uno de los bloques 8x8. Los resultados obtenidos luego de aplicada la DCT, determinan que el primer coeficiente de cada bloque es siempre positivo y mayor que cualquiera de los otros elementos del bloque, este término se le denomina DC , a los restantes se les llama AC.
- Eliminación de coeficientes. Elimina el número de coeficientes de cada uno de los bloques de acuerdo con la selección del usuario.

Interface Gráfica de usuario: La figura 4.2 muestra la interface gráfica de usuario elaborada en MATLAB para el proceso de compresión.

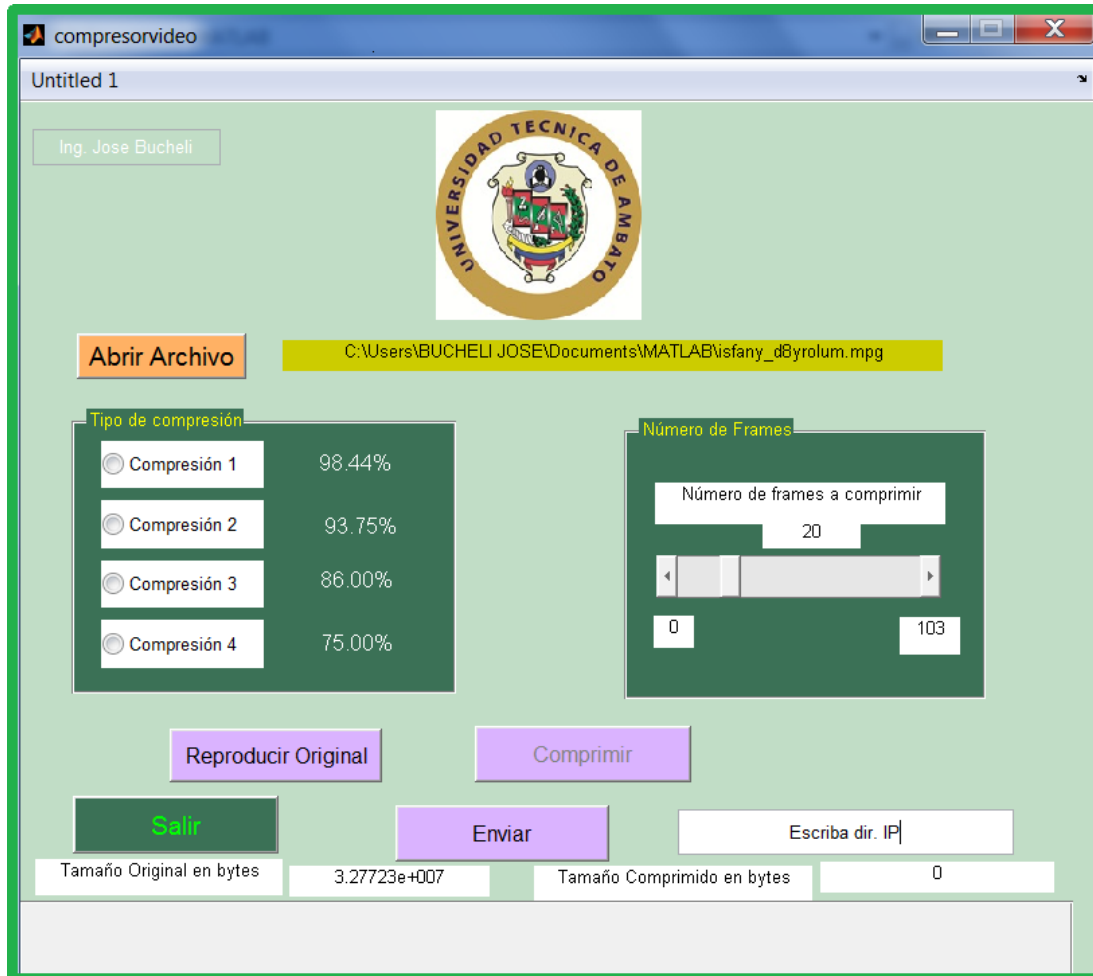


Figura 4.2 Panel virtual de interface de usuario para el compresor de imágenes en movimiento.

- Abrir archivo: Seleccione un archivo en formato MPG y espere hasta que se cargue el archivo.
- Si desea reproducir el video pulse sobre botón Reproducir Original.

- Antes de comprimir seleccione “Número de frames a comprimir”, si no selecciona, por defecto toma el valor del número de tramas originales.
- El botón comprimir, realiza el proceso de compresión. El tiempo de ejecución es de aproximadamente 4 segundos por trama.

Así al seleccionar compresión 1, solo la componente DC de cada bloque de 8×8 servirá para almacenar la información, los restantes coeficientes son eliminados. Es decir la imagen se forma con solo 1 de los 64 valores de cada bloque de 8×8 .

Al seleccionar la opción compresión 2, se selecciona una matriz de 2×2 de cada bloque de 8×8 , los restantes coeficientes son eliminados. Es decir la imagen se forma con 4 de los 64 valores de cada bloque de 8×8 .

Al seleccionar la opción compresión 3, se selecciona una matriz de 3×3 de cada bloque de 8×8 , los restantes coeficientes son eliminados. Es decir la imagen se forma con 9 de los 64 valores de cada bloque de 8×8 .

Al seleccionar la opción compresión 4, se selecciona una matriz de 4×4 de cada bloque de 8×8 , los cuales son almacenados, los restantes coeficientes son eliminados. Es decir la imagen se forma con 16 de los 64 valores de cada bloque de 8×8 .

- El botón enviar abre primeramente el menú de propiedades de conexión de área local como se indica en la figura 4.3, en la cual se puede seleccionar los parámetros de la red de acuerdo a los requerimientos.

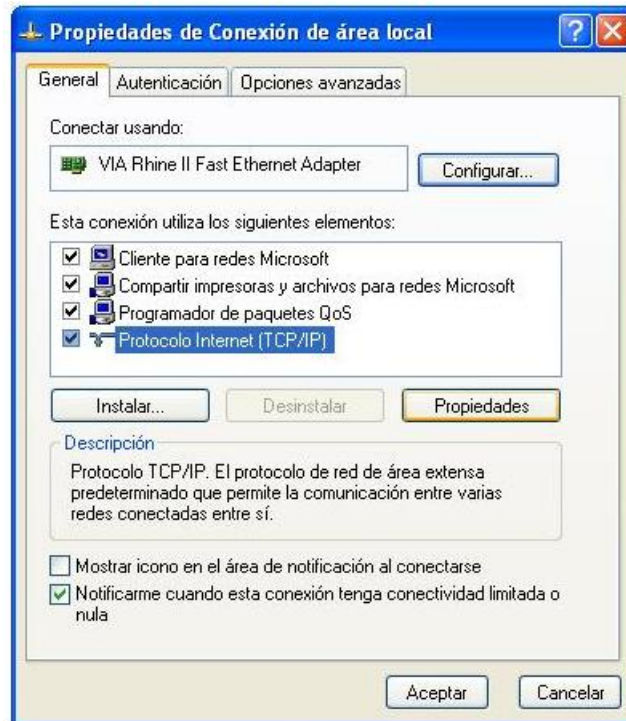


Figura 4.3 Menú de propiedades de conexión de área local

4.3.1.2 DESCOMPRESOR DE IMÁGENES EN MOVIMIENTO

Procesamiento

- Leer archivo comprimido: Lee un archivo comprimido en formato de MATLAB. extensión .mat en matrices para cada uno de los tres colores (RGB), donde cada coeficiente requiere de 24 bits para ser representado.
- Decodificación: Realiza el proceso inverso para leer los valores de DC y AC, los almacena en una matriz. Se rellena con ceros los valores que fueron eliminados según el grado de compresión seleccionado por el usuario en el proceso de compresión.

- Transformada inversa IDCT. Se aplica la DCT inversa a cada bloque de 8x8 y se forma la imagen aproximada a la imagen original antes del proceso de compresión.

En la figura 4.4 se indica la interface gráfica con todos los botones para el proceso de cargar archivo, descomprimir, visualizar el video.



Figura 4.4 Panel virtual de interface de usuario para el Descompresor de imágenes en movimiento

Interface Gráfica de usuario: La figura 4.4 muestra la interface de usuario elaborada en MATLAB para el proceso de descompresión.

- Abrir: Permite seleccionar un archivo comprimido previamente enviado
- Al pulsar sobre el botón descomprimir, se realiza el proceso de compresión, el tiempo de duración depende del número de tramas seleccionadas en el proceso de compresión
- Antes de reproducir, seleccione la velocidad de reproducción.
- El botón reproducir permite visualizar el video, su calidad dependerá del grado de compresión seleccionado.

En la figura 4.5 se muestra un ejemplo de imágenes en movimiento del compresor y descompresor para una relación de compresión del 75 %.



Figura 4.5 Ejemplo de una secuencia de imágenes obtenida luego del proceso de compresión y descompresión al 75 % con el software diseñado en el que se ve una posición diferente en cada campo.

El software diseñado tanto para el proceso de compresión y descompresión utilizan los archivos ejecutados en MATLAB que se presentan en la tabla 4.1

Tabla 4.1 Archivos para el compresor y descompresor de video

Función en MATLAB	Nombre de archivo	Descripción
Compresión	Compresorvideo.fig	Programa de interfaz de usuario del compresor
Compresión	Compresorvideo.m	Programa para el proceso de compresión de imágenes en movimiento
Descompresión	Descompresorvideo.fig	Programa de interfaz de usuario del compresor
Descompresión	Descompresorvideo.m	Programa para el proceso de compresión de imágenes en movimiento

CAPÍTULO V.

ANÁLISIS E INTERPRETACIÓN DE RESULTADOS.

5.1. INTRODUCCIÓN.

Conforme la informática avanza, las imágenes se han convertido en un área muy importante de esta. Hoy en día surgen más entornos gráficos orientados a múltiples aplicaciones. Las imágenes sean hecho con la informática, de ahí la necesidad de compactarlas, hay que reducir la cantidad de datos necesarios para representar una imagen digital. En esencia, la compresión es un proceso que intenta alcanzar una representación más compacta de la señal digital, mediante la eliminación de la redundancia existente en dicha señal¹

El sistema descrito en este documento está implementado como un prototipo, desarrollado en MatLab. Concretamente, este proyecto está destinado al análisis, implementación de un algoritmo para el proceso de compresión, transmisión y descompresión de videos con limitado número de tramas.

5.2 PROCESAMIENTO DE LA INFORMACIÓN

Para evaluar el desempeño del sistema y los algoritmos utilizados, fue necesario determinar la calidad de las imágenes que componen el video a partir del descompresor diseñado e implementado

La calidad de los resultados obtenido es medida a menudo mediante tests subjetivos controlados cuidadosamente en los que se reproducen muestras de video a espectadores, a quienes se les pide que las puntúen en una escala. Las calificaciones

¹ Compresion de datos , Multimedia<http://www.dsi.uclm.es/assignaturas/300204/>

asignadas por cada espectador son promediadas para así obtener la puntuación de opinión media (MOS, Mean Opinion Score)²

En función del Test MOS TEST, el cual constituye un estándar internacional, ya que es utilizada como una norma para la valoración de la calidad imagen subjetivas. En esta prueba se procede a analizar la frecuencia de los criterios que respondieron los encuestados, obteniéndose un criterio numérico sobre la muestra, la cual permitirá tomar decisiones de aceptar o rechazar la hipótesis nula.

5.3. POBLACIÓN Y MUESTRA

Determinar el método de selección de la muestra y, por tanto, su tamaño, ha de establecer los procedimientos que habrán de aplicarse a los resultados para sacar conclusiones válidas. Por lo general, resulta virtualmente imposible estudiar a todos los individuos que integran una población (o sea, realizar un censo)³. El cálculo del tamaño de la muestra se encuentra sustentado en la Teoría de la Probabilidad y en las pruebas de hipótesis, de manera que la estimación de la realidad a partir de la medición de una muestra se encuentra sujeta a cierta imprecisión.

El error muestral es inversamente proporcional al tamaño de la muestra, es decir, a un tamaño muestral mayor, el error muestral será menor. De lo anterior podemos deducir que si un estudio se realiza involucrando a todo el universo, entonces el error muestral será del cero por ciento; sin embargo, no siempre es posible realizar la recolección de datos atendiendo a la totalidad de las unidades de análisis, ya sea por lo numeroso del universo, la falta de recursos destinados a la investigación o algún otro factor⁴.

² http://es.wikipedia.org/wiki/Calidad_de_experiencia

³ Silva Ayçaguer, Luis Carlos. Muestreo para la investigación en ciencias de la salud. España: Ediciones Díaz de Santos, 2007.

⁴ Pazmiño Cruzatti, Iván. Tiempo de investigar, investigación científica 2: cómo hacer una tesis de grado. Ecuador: EDITEKA Ediciones, 2008. p 10.

Por lo tanto, el tamaño muestral para estimar el parámetro en el muestreo aleatorio simple, fijado el error de muestreo se utiliza la ecuación 3.1 ya indicada en el capítulo III

Para una población de 3100 personas de la ESPE Extensión Latacunga, un nivel de confianza del 95% ($z = 1.96$)

Probabilidad a favor del 50% $p = 0.5$

$$q = 1 - p \quad q = 0.5$$

Un error del $e = 10\%$, el tamaño de la muestra sería de:

$$n = \frac{3100 * 1.96^2 * 0.5 * 0.5}{0.08^2 * 3100 + 1.96^2 * 0.5 * 0.5} = 143$$

La muestra a la que se aplicará la encuesta para determinar la calidad de la imagen recuperada será de 143 personas de la población total de la ESPE Extensión Latacunga.

Se debe indicar que los datos son del tipo cualitativo, y que, para ser analizados se empleó el Test MOS TEST, el cual permite realizar un análisis estadístico sobre proporciones.

Los resultados en función del MOS TEST, permiten visualizar los elementos de la muestra como aquellos que cumplen con la característica y aquellos que no la cumplen.

5.4. RECOPIACIÓN DE DATOS.

La recopilación de la información, se la realizó en las instalaciones de la Escuela Politécnica del Ejército Extensión Latacunga, la encuesta fue aplicada a los docentes, y estudiantes.

5.5 PRUEBAS DE HIPÓTESIS

5.5.1 PLANTEAMIENTO DE LA HIPÓTESIS

Hipótesis de la investigación.

La Transformada Discreta Coseno para la compresión de imágenes en movimiento permitirá optimizar la transmisión en redes que utilicen el protocolo IP con una menor cantidad de bits y aumentado la velocidad de transmisión.

Hipótesis nula.

La Transformada Discreta Coseno para la compresión de imágenes en movimiento no permitirá optimizar la transmisión en redes que utilicen el protocolo IP debido a que no se tiene una menor cantidad de bits y por lo tanto no se podrá aumentar la velocidad de transmisión.

5.5.2 SELECCIÓN DEL NIVEL DE SIGNIFICANCIA

Considerando los resultados que se obtienen del análisis de los datos, de la encuesta realizada a la muestra determinada, se espera que la hipótesis nula sea rechazada por la naturaleza de la misma. Pero para llegar a esta determinación se debe procesar los resultados obtenidos.

El nivel de significancia es la probabilidad máxima que se espera para cometer un error, el nivel de significancia o de error esperado es del cinco por ciento para un número de observadores muy elevado, es decir que expresa un nivel de confianza del 95% de haber tomado la decisión correcta.

5.5.3 DETERMINACIÓN DEL ESTADÍSTICO A SER CONSIDERADO EN EL ANÁLISIS

La etapa de experimentación consiste en permitir al personal encuestado observar la secuencia de imágenes en movimiento original, seleccionar el número de tramas, para luego someterle a las pruebas de compresión y descompresión, seleccionando 4 relaciones de compresión y observar el efecto en el tamaño del video comprimido y en la calidad de la secuencia de las imágenes descomprimidas.

5.5.4 PROTOCOLO DE PRUEBAS

Para realizar las pruebas se llevó a cabo el siguiente protocolo:

1. Informar a la persona a encuestar en lo que consiste el proceso de compresión y descompresión de la secuencia de imágenes en movimiento.
2. Explicar el formato y la metodología de la encuesta.
3. Ubicar al usuario, frente a la pantalla del computador y explicarle como funciona la interface de usuario que le permita observar las imágenes en movimiento de video de corto tamaño.
4. Ejecutar el módulo de prueba.

5. Después de observar el video el usuario debe llenar el formato de encuesta respectivo, respondiendo las preguntas relacionadas con las características del video

5.5.5 RESULTADOS DE LA PRUEBA SUBJETIVA

Una vez obtenido el conjunto de votaciones de los 143 observadores válidos de la prueba, se calculan los valores medios, en función del MOST Test que se codifica como se indicó en el capítulo III con valores de 1 a 5 (desde muy molesto a imperceptible) las diferencias entre la imagen original y la imagen luego del proceso de compresión, transmisión y descompresión.

Entonces para un intervalo de confianza del 95%, según lo anterior, los resultados obtenidos tras el procesamiento de los datos de la prueba subjetiva con las 4 relaciones de compresión se resumen en la siguiente tabla:

Tabla No. 5.1 Resultados de la prueba subjetiva mediante el Mos Test

Relación de Compresión	Imperceptible (5)	Perceptible pero no molesta (4)	Ligeramente molesta (3)	Molesta (2)	Muy Molesta (1)
4 (75%)	131	12			
3 (86%)	85	41	17		
2 (93,75%)		100	40	3	
1 (98,44%)				41	102

Se debe aclarar que el valor de la relación de compresión corresponde, a que en proceso de compresión la imagen que está dividida previamente en bloques de 8x8, es decir, la compresión de las imágenes en movimiento consiste principalmente en reducir el número de píxeles de cada imagen, esto se lo hace escogiendo matrices de

4x4, 3x3, 2x2 y 1x1 de cada bloque de 8x8 en que previamente se dividió la imagen (es decir cada trama del video original se la divide en bloque de 8x8 antes de aplicar la DCT).

Por la Teoría del Muestreo, se determina el Intervalo de Confianza para la proporción (p) de unidades en la muestra que poseen la característica a estudiar.

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n-1}} \quad (5.1)$$

$$\text{con } \hat{p} = \frac{a}{n} = \bar{x} \quad (5.2)$$

Donde:

a .- Número de unidades de la muestra que poseen la característica a observar

n .- Número de unidades de la muestra

\hat{p} .- Proporción de unidades en la muestra que poseen la característica estudiar.

\hat{q} .- Proporción de unidades en la muestra que no poseen la característica a ser estudiada, es decir: $\hat{q} = (1 - \hat{p})$

$z_{\alpha/2}$.- Área bajo la curva de Distribución de Probabilidad Normal Estándar, entre la media y valores positivos de z.

$$Z_{\alpha/2} = \frac{95\%}{2} = 47.5\% = 0.475 = 0,5.$$

Según la Tabla de Distribución de Probabilidad Normal Estándar, para una proporción corresponde a:

$$Z_{\alpha/2} = 1.96$$

En base a las ecuaciones 5.1 y 5.2 y los resultados de la encuesta descrita en la tabla 5.1, se determinaron los siguientes intervalos de confianza

5.5.5.1 INTERVALO DE CONFIANZA PARA LA RELACIÓN DE COMPRESIÓN (4) 75%.

Tabla No. 5.2 Resultado para relación de compresión de 75%.

Calificación	Frecuencia	Porcentaje %
a. Imperceptible	131	91,6
b. Perceptible pero no molesta	12	8,4
	143	100,0

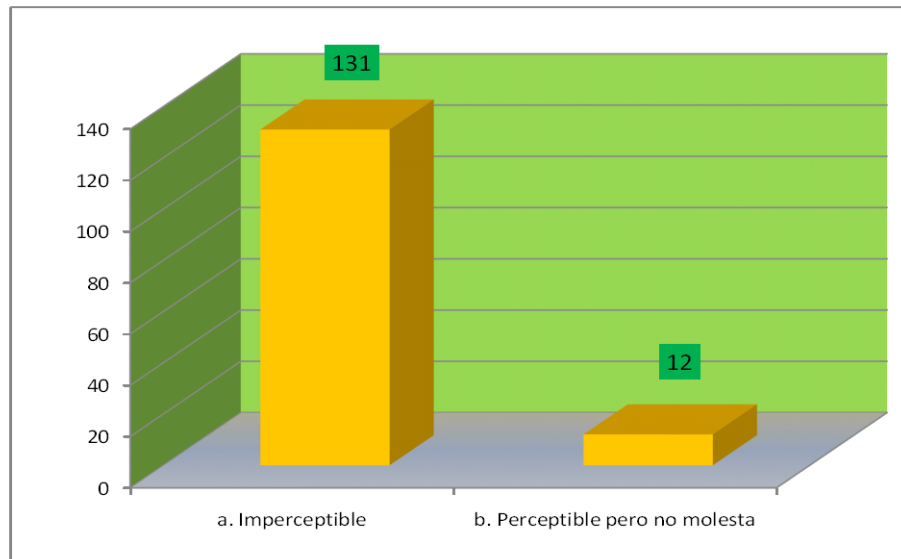


Figura No. 5.1 Resultado cuando se comprime el 75%

Evaluación de la Calidad de las imágenes en movimiento con compresión del 75% y respuesta diferencia Imperceptible que responden 131 observadores

$$\hat{p} = \frac{a}{n} = \frac{131}{143} = 0.916 = 91.60 \% . \text{ Entonces:}$$

$$\begin{aligned} \hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n-1}} &= 0.916 \pm 1.96 \sqrt{\frac{(0.916)(0.084)}{143-1}} \\ &= 0.916 \pm 1.96 (0.02327) \\ &= 0.916 \pm 0.0456 \\ &= 0.9616, \quad 0.8704 \end{aligned}$$

Por lo tanto, calculando el rango total entre valor máximo y valor mínimo se tiene un intervalo de probabilidad de 87,04% al 96.16% que representa el intervalo de confianza para los observadores que contestaron con 5 valorado como diferencia entre imagen original y procesadas como imperceptible

Evaluación de la Calidad de las imágenes en movimiento con compresión del 75% y respuesta diferencia Imperceptible que responden 12 observadores

$$\hat{p} = \frac{a}{n} = \frac{12}{143} = 0.0839 = 8,391 \% . \text{ Entonces:}$$

$$\begin{aligned} \hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n-1}} &= 0.0839 \pm 1.96 \sqrt{\frac{(0.0839)(0,9161)}{143-1}} \\ &= 0.0839 \pm 1.96 (0.02326) \\ &= 0.0839 \pm 0,0456 \\ &= 0,1295, \quad 0.0383 \end{aligned}$$

Por lo tanto, calculando el rango total entre valor máximo y valor mínimo se tiene un intervalo de probabilidad de 3.83al 12,95 % que representa el intervalo de confianza para los observadores que contestaron con 4 valorado como diferencia entre imagen original y procesadas como perceptible pero no molesta.

5.5.5.2 INTERVALO DE CONFIANZA PARA LA RELACIÓN DE COMPRESIÓN (3) 86%.

Tabla No. 5.3 Resultado para relación de compresión de 86%.

Calificación	Frecuencia	Porcentaje %
a. Imperceptible	85	59,4
b. Perceptible pero no molesta	41	28,7
b. Ligeramente molesta	17	11,9
	143	100,0

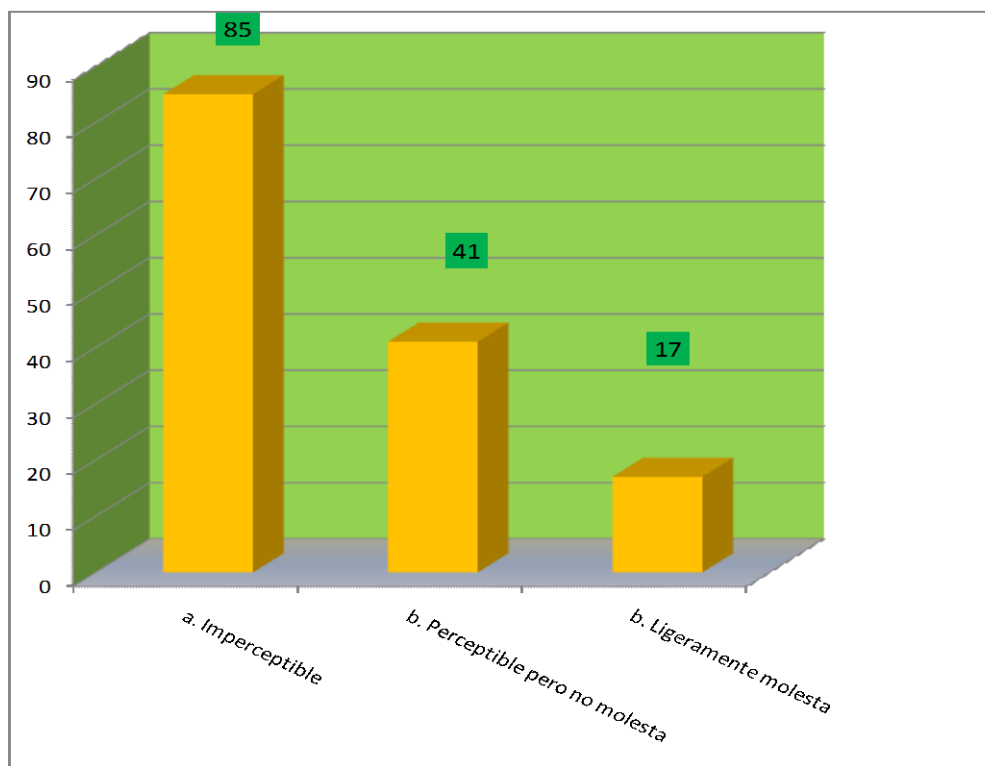


Figura No. 5.2 Resultado cuando se comprime el 86%

Evaluación de la Calidad de las imágenes en movimiento con compresión del 86% y respuesta diferencia Imperceptible que responden 85 observadores

$$\hat{p} = \frac{a}{n} = \frac{85}{143} = 0.5944 = 59.44 \% . \text{ Entonces:}$$

$$\begin{aligned} \hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n-1}} &= 0.5944 \pm 1.96 \sqrt{\frac{(0.5944)(0.4056)}{143-1}} \\ &= 0.5944 \pm 1.96 (0.04120) \\ &= 0.5944 \pm 0.08076 \\ &= 0.67516, \quad 0.51364 \end{aligned}$$

Por lo tanto, calculando el rango total entre valor máximo y valor mínimo se tiene un intervalo de probabilidad de 51,36% al 67,51% que representa el intervalo de confianza para los observadores que contestaron con 5 valorado como diferencia entre imagen original y procesadas como imperceptible

Evaluación de la Calidad de las imágenes en movimiento con compresión del 86% y respuesta diferencia Perceptible pero no molesta y que responden 41 observadores

$$\hat{p} = \frac{a}{n} = \frac{41}{143} = 0.2867 = 28.67 \% . \text{ Entonces:}$$

$$\begin{aligned} \hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n-1}} &= 0.2867 \pm 1.96 \sqrt{\frac{(0.2867)(0.7133)}{143-1}} \\ &= 0.2867 \pm 1.96 (0.03794) \\ &= 0.2867 \pm 0.07438 \\ &= 0.36108, \quad 0.2132 \end{aligned}$$

Por lo tanto, calculando el rango total entre valor máximo y valor mínimo se tiene un intervalo de probabilidad de 21.32% al 36.108% que representa el intervalo de

confianza para los observadores que contestaron con 4 valorado como diferencia entre imagen original y procesadas como perceptible pero no molesta.

Evaluación de la Calidad de las imágenes en movimiento con compresión del 86% y respuesta diferencia Ligeramente molesta que responden 17 observadores

$$\hat{p} = \frac{a}{n} = \frac{17}{143} = 0.1188 = 11.88 \% . \text{ Entonces:}$$

$$\begin{aligned} \hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n-1}} &= 0.1188 \pm 1.96 \sqrt{\frac{(0.1188)(0.8812)}{143-1}} \\ &= 0.1188 \pm 1.96 (0.02715) \\ &= 0.1188 \pm 0.05321 \\ &= 0.1720, \quad 0.06559 \end{aligned}$$

Por lo tanto, calculando el rango total entre valor máximo y valor mínimo se tiene un intervalo de probabilidad de 6.55% al 17.20% que representa el intervalo de confianza para los observadores que contestaron con 3 valorado como diferencia entre imagen original y procesadas como Ligeramente molesta.

5.5.5.3 INTERVALO DE CONFIANZA PARA LA RELACIÓN DE COMPRESIÓN (2) 93,75%.

Tabla No. 5.4 Resultado para relación de compresión de 93.75%.

Calificación	Frecuencia	Porcentaje %
a. Perceptible pero no molesta	100	69,9
b. Ligeramente molesta	40	28,0
c. Molesta	3	2,1
	143	100,0

Figura No. 5.3 Resultado cuando se comprime el 93,7586%

Evaluación de la Calidad de las imágenes en movimiento con compresión del 93.75% y respuesta diferencia Perceptible pero no molesta que responden 100 observadores

$$\hat{p} = \frac{a}{n} = \frac{100}{143} = 0.6993 = 69.93 \% . \text{ Entonces:}$$

$$\begin{aligned} \hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p} \hat{q}}{n-1}} &= 0.6993 \pm 1.96 \sqrt{\frac{(0.6993)(0.3007)}{143-1}} \\ &= 0.6993 \pm 1.96 (0.03848) \\ &= 0.6993 \pm 0.07542 \\ &= 0.7747, \quad 0.6238 \end{aligned}$$

Por lo tanto, calculando el rango total entre valor máximo y valor mínimo se tiene un intervalo de probabilidad de 62.38% al 77.47% que representa el intervalo de confianza para los observadores que contestaron con 4 valorado como diferencia entre imagen original y procesadas como **Perceptible pero no molesta**.

Evaluación de la Calidad de las imágenes en movimiento con compresión del 93.75% y respuesta diferencia Ligeramente molesta que responden 40 observadores

$$\hat{p} = \frac{a}{n} = \frac{40}{143} = 0.2797 = 27.97 \% . \text{ Entonces:}$$

$$\begin{aligned} \hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p} \hat{q}}{n-1}} &= 0.2797 \pm 1.96 \sqrt{\frac{(0.2797)(0.7203)}{143-1}} \\ &= 0.2797 \pm 1.96 (0.03766) \\ &= 0.2797 \pm 0.07382 \\ &= 0.3535, \quad 0.2058 \end{aligned}$$

Por lo tanto, calculando el rango total entre valor máximo y valor mínimo se tiene un intervalo de probabilidad de 20,58% al 35.35% que representa el intervalo de confianza para los observadores que contestaron con 3 valorado como diferencia entre imagen original y procesadas como **Ligeramente molesta**.

Evaluación de la Calidad de las imágenes en movimiento con compresión del 93.75% y respuesta diferencia Molesta que responden 3 observadores

$$\hat{p} = \frac{a}{n} = \frac{3}{143} = 0.02097 = 2.097 \% . \text{ Entonces:}$$

$$\begin{aligned} \hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p} \hat{q}}{n-1}} &= 0.02097 \pm 1.96 \sqrt{\frac{(0.02097)(0.9790)}{143-1}} \\ &= 0.02097 \pm 1.96 (0.01202) \\ &= 0.02097 \pm 0.02356 \\ &= 0.04453, \quad -0.00259 \end{aligned}$$

Por lo tanto, calculando el rango total entre valor máximo y valor mínimo se tiene un intervalo de probabilidad de -0.259% al 4.45% que representa el intervalo de confianza para los observadores que contestaron con 2 valorado como diferencia entre imagen original y procesadas como **Molesta**.

5.5.5.4 INTERVALO DE CONFIANZA PARA LA RELACIÓN DE COMPRESIÓN (1) 98.44%.

Tabla No. 5.5 Resultado para relación de compresión de 98.44%.

Calificación	Frecuencia	Porcentaje %
a. Molesta	41	28,7
b. Muy Molesta	102	71,3
	143	100,0

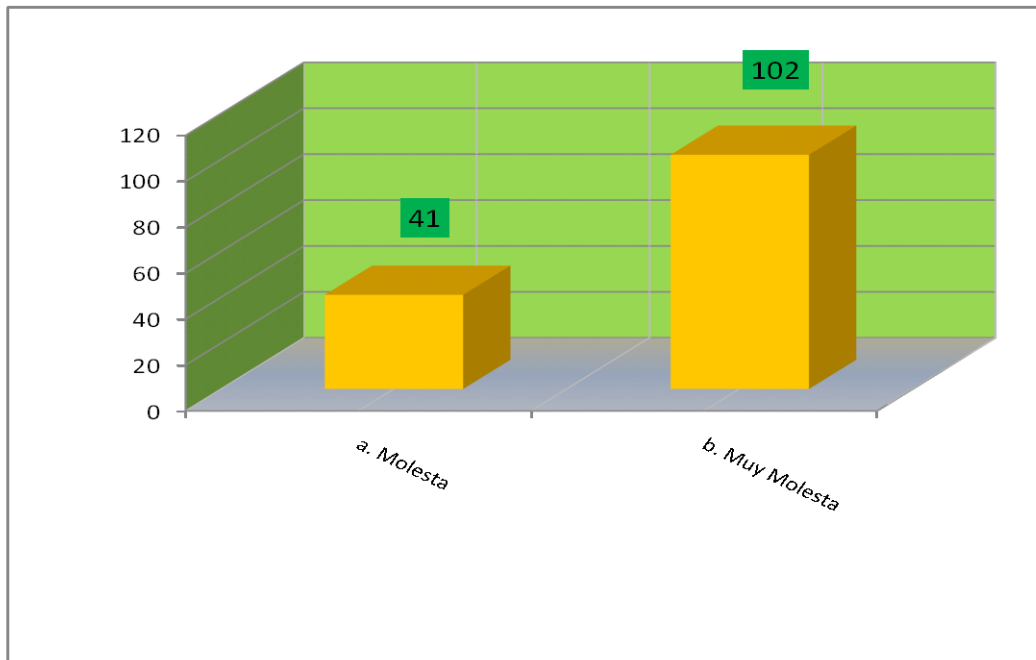


Figura No. 5.4 Resultado cuando se comprime el 98.44%

Evaluación de la Calidad de las imágenes en movimiento con compresión del 98.44% y respuesta diferencia Molesta que responden 41 observadores

$$\hat{p} = \frac{a}{n} = \frac{41}{143} = 0.2867 = 28.67 \% . \text{ Entonces:}$$

$$\begin{aligned}
\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n-1}} &= 0.2867 \pm 1.96 \sqrt{\frac{(0.2867)(0.7133)}{143-1}} \\
&= 0.2867 \pm 1.96 (0.0379) \\
&= 0.2867 \pm 0.07428 \\
&= 0.3609, \quad 0.2124
\end{aligned}$$

Por lo tanto, calculando el rango total entre valor máximo y valor mínimo se tiene un intervalo de probabilidad de 21.24% al 36.09% que representa el intervalo de confianza para los observadores que contestaron con 2 valorado como diferencia entre imagen original y procesadas como **Molesta**.

Evaluación de la Calidad de las imágenes en movimiento con compresión del 98.44% y respuesta diferencia Muy Molesta que responden 102 observadores

$$\hat{p} = \frac{a}{n} = \frac{102}{143} = 0.7132 = 71.32 \% . \text{ Entonces:}$$

$$\begin{aligned}
\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}\hat{q}}{n-1}} &= 0.7132 \pm 1.96 \sqrt{\frac{(0.7132)(0.2868)}{143-1}} \\
&= 0.7132 \pm 1.96 (0.03795) \\
&= 0.7132 \pm 0.07438 \\
&= 0.7875, \quad 0.6388
\end{aligned}$$

Por lo tanto, calculando el rango total entre valor máximo y valor mínimo se tiene un intervalo de probabilidad de 63.88% al 78.75% que representa el intervalo de confianza para los observadores que contestaron con 2 valorado como diferencia entre imagen original y procesadas como **Muy Molesta**.

5.5.6 CONCLUSIONES DE TIPO ESTADÍSTICO

Según lo anterior, los resultados obtenidos tras el procesamiento de los datos de la prueba subjetiva se resumen en la tabla 5.6:

Tabla 5.6 Resultado de aplicación del MOS Test. A la encuesta realizada al personal de la ESPE Extensión Latacunga.

Relación de Compresión	Imperceptible (5)	Perceptible pero no molesta (4)	Ligeramente molesta (3)	Molesta (2)	Muy Molesta (1)
4 (75%)	87,04% al 96,16%	3,83al 12,95 %			
3 (86%)	51,36% al 67,51%	21,32% al 36,108%	6,55% al 17,20%		
2 (93,75%)		62,38% al 77,47%	20,58% al 35,35%	de - 0,259% al 4,45%	
1 (98,44%)				21,24% al 36,09%	63,88% al 78,75%

De acuerdo al análisis estadístico realizado en la investigación, al determinarse los intervalos de confianza para valores de relación de compresión del 75%, 86 %, 93.75% y 98.44% de las imágenes en movimiento, se obtuvieron las siguientes conclusiones de orden estadístico:

- Para una relación de compresión del 75%, se evidencia que este evento tiene mayor probabilidad de obtener una diferencia entre la imagen original y la tratada de Imperceptible.

- Para una relación de compresión del 86%, se evidencia que este evento tiene mayor probabilidad de obtener una diferencia entre la imagen original y la tratada de Imperceptible.
- Para una relación de compresión del 93.75%, se evidencia que este evento tiene mayor probabilidad de obtener una diferencia entre la imagen original y la tratada de Perceptible pero no molesta.
- Para una relación de compresión del 98.44%, se evidencia que este evento tiene mayor probabilidad de obtener una diferencia entre la imagen original y la tratada de Muy Molesta.

5.5.7 RESULTADOS DEL PROCESO DE TRANSMISIÓN DE IMÁGENES EN MOVIMIENTO SOBRE REDES QUE UTILIZAN EL PROTOCOLO IP

Las pruebas se la realizaron entre dos PC, una con sistema operativo Windows XP y la otra con Windows 7, vía inalámbrica con protocolo TCP/IP. El ancho de banda disponible durante las pruebas era de alrededor de 48Mbs.

Tabla 5.7 Tiempos de transmisión de imágenes en movimiento con diferentes relaciones de compresión utilizando el protocolo IP.

RELACIÓN DE COMPRESIÓN	Tiempo1	Tiempo3	Tiempo de Transmisión	Bytes Transmitidos
Sin compression	19,23,38	19,24,19	48 Seg	26,60MByte
	19,30,15	19,30,51	46 Seg	26,70MByte
	19,37,17	19,37,50	44 Seg	26,45Mbyte
	19,56,07	19,56,45	45 Seg	26,44Mbyte
Compresión 4 (menor relación de compresion)	20,25,27	20,26,06	39 Seg	23,60MByte
	20,27,43	20,28,20	37 Seg	23,70MByte
	20,30,07	20,30,43	36 Seg	23,63MByte
Compresión 3	20,37,47	20,38,19	32 Seg	13,83MByte
	20,40,19	20,40,40	21 Seg	13,84MByte
	20,41,53	20,42,12	19 Seg	13,82MByte
	20,43,12	20,43,34	22 Seg	13,89MByte
Compresión 2				
	20,52,54	20,53,05	11 Seg	6,51MByte
	20,54,10	20,54,20	10 Seg	6,47MByte
	20,55,29	20,55,40	11 Seg	6,50MByte
Compresión 1 (mayor relación de compresión)				
	21,05,16	21,05,19	3 Seg	1,70Mbyte
	21,06,06	21,06,09	3 Seg	1,68Mbyte
	21,07,54	21,07,57	3 Seg	1,69Mbyte

Tiempo 1 y Tiempo 3 corresponde a la hora de llegada de 2 archivos simples que son enviados desde el software realizado en MATLAB, antes y después del archivo de interés, esto es el archivo comprimido con el fin de determinar en base a la diferencia horaria el tiempo que dura la transmisión.

Esta tabla demuestra que para una mayor relación de compresión el tiempo utilizado en la transmisión se reduce notablemente.

5.5.8 TOMA DE DECISIÓN EN LA INVESTIGACIÓN

5.5.8.1 PRUEBA DE HIPÓTESIS

Para la validación de la Hipótesis de la Investigación, se analizó el nivel de incidencia correspondiente a Imperceptible y Perceptible pero no molesta, que se obtuvo de los diferentes intervalos de confianza, cuando el observador experimentó al variar el nivel de compresión entre los 4 valores seleccionados en este proyecto.

Por lo tanto, la Hipótesis de la Investigación es aceptada para los eventos en los que se comprime al 75%, 86% y 93,75%, y no es aceptada cuando se realiza con un mayor grado de compresión aunque en este caso se reduce en mayor cantidad el número de bits, y por lo tanto se aumenta la velocidad de transmisión comparado con el que se emplea en el video original, pero se pierde el nivel de calidad de las imágenes en movimiento.

CAPÍTULO VI

CONCLUSIONES Y RECOMENDACIONES

6.1 CONCLUSIONES

- Se ha cumplido el objetivo principal de este proyecto el cual consiste en implementar un compresor de imágenes en movimiento para transmisión en redes IP en base a la aplicación de la Transformada Discreta Coseno con diferentes relaciones de compresión.
- El proyecto desarrollado permitió ofrecer un compresor de imágenes tomando como referencia estándares de codificación existentes, con factores de calidad seleccionable es decir flexibilidad de poder seleccionar imágenes de alta calidad o baja calidad de imagen y mediante una interface gráfica realizada en MATLAB, permitir al operador un entorno visual de fácil accesibilidad y manejo para transmitirlo a la red vía el protocolo de internet (IP).
- El video digital ha tomado una importancia mayor en las últimas décadas, debido a las nuevas facilidades que provee la tecnología actual y el desarrollo y perfeccionamiento de los estándares de codificación que hacen que cada vez más se incremente la calidad del video digital utilizando menos recursos de almacenamiento o menos datos para transmitir una secuencia de video
- El Vídeo no es nada más que la reproducción en forma secuencial de imágenes (cuadros, tramas o fotogramas por segundo), que al verse con una determinada velocidad y continuidad dan la sensación al ojo humano de apreciar el movimiento natural.

- Una imagen a color está compuesta de tres fotogramas, uno para cada componente de color (rojo, verde y azul) para ser interpretados por el ojo humano y que son cuantificados con una profundidad de pixel de 24 bits en total, 8 para cada color. El sistema de codificación usado en este proyecto es el RGB (Red, Green, Blue).
- Las imágenes de vídeo están compuestas de información en el dominio del espacio y el tiempo. La información en el dominio del espacio es provista por los píxeles, y la información en el dominio del tiempo es provista por imágenes que cambian en el tiempo. En el presente proyecto se aprovecha el efecto de la información redundante entre píxeles cercanos, es decir el análisis se lo hace en el dominio espacial.
- La compresión basada en el dominio espacial es llamada codificación intracuadro, y puede ser realizada mediante el método de predicción o por codificación por transformada. En el proyecto desarrollado se utilizó la Transformada Discreta Coseno o DCT donde la imagen es transformada de su representación espacial a su frecuencia espacial equivalente, cuyos valores son colocados en una matriz de M por N.
- Para obtener mejores resultados, a la matriz original de la imagen se la divide en bloques de 8x8 píxeles y entonces se aplica la DCT a cada bloque. Los resultados mostraron que al aplicar DCT a una imagen, las zonas con colores similares se representan con coeficientes de baja frecuencia y las imágenes con mucho detalle, con coeficientes de alta frecuencia.
- Los 64 valores resultantes de aplicar la DCT se colocan en una nueva matriz del mismo orden de la original, donde la componente de mayor valor llamada DC se coloca en la parte superior izquierda y los otros componentes que son de menor valor son llamados AC y ocupan las otras posiciones de la matriz.

- Una vez aplicada la DCT , el sistema de compresión se obtiene principalmente mediante la selección de los coeficientes de mayor valor que están colocados desde el extremo superior izquierdo de la matriz resultante, creando por lo tanto con los valores seleccionados una nueva matriz de menor número de datos.
- El proceso de descompresión de imágenes en movimiento utilizado en este proyecto es básicamente el inverso del proceso de codificación.
- Los resultados experimentales y las pruebas del MOS test demuestran que la calidad de la imagen es considerada como buena hasta un valor de relación de compresión de 93.75% (este valor relativo se lo determinó considerando cuantos valores de la matriz de 8x8 se toman para formar y guardar la imagen luego de aplicar el proceso de descompresión.
- Al seleccionar solo uno de los 64 coeficientes, se observa que en el video reconstruido no se puede distinguir la imagen.
- Las pruebas realizadas (MOS Test) que validan la hipótesis, es subjetiva por cuanto no todos tenemos igual apreciación ante un mismo evento. Sin embargo es un buen referente para determinar el desempeño del sistema desarrollado.
- El mayor problema que se encontró en este proyecto es el proceso matemático muy demorado y complejo, lo que determina un alto tiempo de ejecución de las pruebas.

6.2 RECOMENDACIONES

- Incentivar la investigación a través del desarrollo de proyecto que involucren el procesamiento de imágenes tanto para aplicaciones industriales como médicas.

- Para el presente trabajo, seleccionar archivos cuya resolución no sea superior al especificado (en este caso 300x300) y tamaño no sobrepase los 25Mb (en formato sin compresión) para su correcto funcionamiento. En el caso de imágenes de mayor tamaño, se recomienda utilizar la tecnología cluster que permitirá aumentar su capacidad de procesamiento.
- Se recomienda tener instalado MATLAB versión 7.10 (R2010), en vista de varias funciones utilizadas que no están presente en versiones anteriores y dificultarían la ejecución del programa.
- Se recomienda el desarrollo de trabajos posteriores que utilicen otras técnicas y otros programas como C++ que permiten optimizar el tiempo de ejecución de la parte matemática.
- Generar nuevos proyectos en el campo del procesamiento de imágenes que estén relacionadas con las líneas de investigación relacionadas al presente trabajo, que permitan generar soluciones del entorno y de esta manera vincular la universidad con la colectividad.

BIBLIOGRAFÍA

- GONZALEZ, RAFAEL C.; WOODS, RICHARD E. “Digital Image Processing”, 1992, Addison-Wesley Publishing Company, Inc
- HERNANDEZ E. WEISS G. “A First Course on Wavelets” CRC press, New York, 1996.
- NOHAMMAD NURUZZAMAN: “Digital Image Fundamentals In Matlab”, author house , indiana 2005.
- HARALICK, ROBERT M., “Recognition Methodology: Algorithms & Architecture”, “Image Pattern Recognition: Algorithm Implementations, Techniques, and Technology”. Corbett, Francis J. Editor, Proceedings of SPIE 755, 1987.
- HOLLINGUM, JACK, “Machine Vision – The Eyes of Automation”, 1984, IFS Publications) Ltd, UK.
- J. P. G. Ortiz. Transmisión progresiva de imágenes jpeg2000. Technical report, Universidad de Almería, Departamento de Arquitectura de Computadores y Electrónica, 2003.
- F. M. GIAIME GINESU AND D. GIUSTO. A multi-factors approach for image quality assessment based on a human visual system model. Elsevier Signal Processing: Image Communication
- Chen, Video Coding Standards for Multimedia Communications: H261, H263, and Beyond, Capítulo de visual Communications and Image Processing, 2002 pág. 25.
- J. WATKINSON, MPEG-2 primera edición, Focal Press, 1999, pág. 244

- F. PEREIRA, T. EBRAHIMI, The MPEG-4 book, primera edición, Prentice Hall, 2002 pag 849
- Alba Fernández, María Virtudes. Muestreo estadístico. España: Septem Ediciones, 2005. p 42.
- Chen, Video Coding Standards for Multimedia Communications: H261, H263, and Beyond, Capítulo de visual Communications and Image Processing, 2002 pág. 25.
- H. Lee, Y. Kim, A. Rowberg, and E. Riskin, “Statistical distributions of dct coefficients and their application to an interframe compression algorithm for 3-d medical images,” IEEE Trans. Medical Imaging, vol. 12, no. 3, pp. 478–485, 1993.

ENLACES

- División de servicios de Documentación de Telefónica I + D (Septiembre de 2004) La codificación y compresión digital de Contenidos Audiovisuales. Recuperado el 15 de Febrero del 2010, de <http://info.telefonica.es/sociedadde/informacion/pdf/publicaciones/imagenio/capitulos/imageniocap7.pdf>.
- Acerca de la UIT (enero 2009). Recuperado julio de 2009, de <http://www.itu.int/net/about/index-es.aspx>
- Unión Internacional de Telecomunicaciones (marzo 2009). Recuperado agosto de 2009, de http://es.wikipedia.org/wiki/Uni%C3%B3n_Internacional_de_Telecomunicaciones.

- Tratamiento Digital de imágenes (julio 2008). Recuperado enero 2010, de <http://www.innovanet.com.ar/gis/TELEDETE/TELEDETE/pant3.htm>
- Eroski Consumer (28 enero de 2005). Diferencia entre byte y bit. Las diferencia entre las medidas para la velocidad de transmisión y el almacenamiento confunde a los usuarios. Recuperado febrero 2010, de <http://www.consumer.es/web/es/tecnologia/internet/2004/12/22/114135.php>
- Compresión de Imagen (2009). Recuperado mayo 2010, de <http://gemini.udistrital.edu.co/comunidad/profesores/jruiz/jairocd/texto/uit/compimage.pdf>.
- Wikipedia (2008). Teoría del Color. Recuperado julio de 2010, de http://es.wikipedia.org/wiki/Teor%C3%ADa_del_color
- Stephen Westland, 2001.Imagen Digital. Qué es la teoría tricrómica. Recuperado Agosto de 2010, de http://www.gusgsm.com/teoria_tricromica
- Subgrupo de Normalización Técnica de la presentación. Grupo Técnico del Foro de la Televisión de Alta Definición en España Coordinado por UNIVERSIDAD POLITÉCNICA DE MADRID Abril de 2008. Recuperado agosto 2010, de <http://www.televisiondigital.es/TecnologiasRelacionadas/AltaDefinicion/ForoTVAD>
- Rec. ITU-R BT.500-11 2002. RECOMMENDATION ITU-R BT.500-11 Methodology for the subjective assessment of the quality of television pictures. Recuperado mayo 2010, de http://www.dii.unisi.it/~menegaz/DoctoralSchool2004/papers/ITU-R_BT.500-11.pdf

ANEXO A

FORMATO DE ENCUESTA



UNIVERSIDAD TÉCNICA DE AMBATO
FACULTAD DE INGENIERÍA EN SISTEMAS, ELECTRÓNICA E
INDUSTRIAL

MAESTRÍA EN REDES Y TELECOMUNICACIONES

II EDICIÓN

ENCUESTA SOBRE LA CALIDAD DE LAS IMÁGENES EN MOVIMIENTO
LUEGO DEL PROCESO DE DESCOMPRESIÓN PARA CUATRO
DIFERENTES RELACIONES DE COMPRESIÓN

El presente instrumento tiene como objeto evaluar la calidad de un archivo de video luego del proceso de descompresión, archivo al que previamente se le aplicó la compresión en base a la Transformada Discreta Coseno y un algoritmo para reducir datos redundantes. Las preguntas deben ser contestadas en una escala de Imperceptible a Muy molesta en base a la diferencia entre el video de referencia y la imagen de prueba.

- Observe el video original las veces que sea necesario
- En la interfase gráfica de usuario para el descompresor, pulse sobre el botón abrir y seleccione un archivo de la lista, una vez cargado el archivo, pulse descomprimir y luego reproducir, que mostrará el video descomprimido.

De ante mano agradezco por su amable y subjetiva respuesta a las preguntas aquí planteadas.

Pregunta	Relación de compresión	Imperceptible	Perceptible pero no molesta	Ligeramente Molesta	Molesta	Muy molesta
Cuál es su criterio sobre la diferencia entre el archivo descomprimido con el archivo original	Compresión 4 75%					
	Compresión 3 86%					
	Compresión 2 93.75%					
	Compresión 1 98.44%					