

## Índice de Contenidos

I - OBJETIVOS .....	11
II - ESTRUCTURA DE LA TESIS .....	13
III - INTRODUCCIÓN.....	15
IV - MODELO BIOLÓGICO DE LA VISIÓN HUMANA.....	17
FENÓMENOS VISUALES .....	18
<i>Sensibilidad al Contraste</i> .....	19
<i>Ruido</i> .....	19
<i>Contraste de Luminosidad</i> .....	20
<i>Percepción de Profundidad</i> .....	21
Profundidad Óculmotora .....	21
Profundidad Binocular.....	22
Profundidad Monoocular .....	23
ILUSIONES .....	23
V - PROCESAMIENTO DE IMÁGENES.....	27
ESTRUCTURA DE UNA IMAGEN .....	27
REPRESENTACIÓN DE IMÁGENES.....	29
ETAPAS FUNDAMENTALES DE PROCESAMIENTO DE IMÁGENES .....	29
VI - COMPRESIÓN DE DATOS .....	33
TERMINOLOGÍA.....	33
TÉCNICAS BÁSICAS DE COMPRESIÓN DE DATOS.....	34
CANTIDAD DE SÍMBOLOS DEL CONJUNTO .....	34
FRECUENCIA DE APARICIÓN DE LOS SÍMBOLOS EN EL TEXTO.....	35
CONTEXTO EN EL QUE APARECEN LOS SÍMBOLOS.....	35
COMPRESIÓN CON PÉRDIDA VS. COMPRESIÓN SIN PÉRDIDA.....	36
PRESERVACIÓN DE LA INFORMACIÓN .....	36
COMPRESIÓN DE IMÁGENES .....	37
REDUNDANCIA DE CÓDIGO .....	38
REDUNDANCIA INTERPIXELS .....	39
REDUNDANCIA PSICOVISUAL .....	41
VII - JOINT PHOTOGRAPHICS EXPERT GROUP (JPEG) .....	43
¿POR QUÉ USAR JPEG?.....	44
DESCRIPCIÓN DEL ALGORITMO JPEG BASELINE .....	45
¿CÓMO FUNCIONA EL JPEG? .....	46
TRANSFORMACIÓN.....	47
CUANTIFICACIÓN .....	48
CODIFICACIÓN .....	51

---

VIII - CLASES DE MEDICIONES .....	55
CRITERIO DE FIDELIDAD .....	55
MEDIDAS OBJETIVAS DE LA CALIDAD .....	55
MEDICIONES SUBJETIVAS DE LA CALIDAD .....	57
IX - EXPERIENCIA DESARROLLADA.....	61
RESUMEN.....	61
OBJETIVOS .....	61
IMÁGENES UTILIZADAS PARA LA EXPERIENCIA .....	62
EVALUACIÓN OBJETIVA.....	63
EVALUACIÓN OBJETIVA.....	64
EVALUACIÓN SUBJETIVA .....	64
DESARROLLO DE LA EVALUACIÓN .....	65
X - SISTEMA DE EVALUACIÓN SUBJETIVA DE IMÁGENES.....	67
SOFTWARE DESARROLLADO .....	67
<i>Valorización</i> .....	67
<i>Evaluación</i> .....	69
<i>Aplicaciones de soporte</i> .....	71
APLICACIÓN ESTADÍSTICA SYSTAT .....	72
XI - PRESENTACIÓN DE LOS RESULTADOS .....	73
RESULTADOS OBJETIVOS .....	73
<i>Clase Caras</i> .....	73
<i>Clase Formulas</i> .....	74
<i>Clase Paisajes</i> .....	75
RESULTADOS SUBJETIVOS.....	75
<i>Modelo 1</i> .....	78
<i>Modelo 2</i> .....	79
<i>Modelo 3</i> .....	79
<i>Modelo 4</i> .....	80
Análisis de los Inexpertos.....	82
Análisis de los Expertos .....	85
XII - CONCLUSIONES .....	88
OBJETIVO 1 .....	92
OBJETIVO 2 .....	92
OBJETIVO 3 .....	92
XIII - TRABAJOS FUTUROS.....	95
XIV - BIBLIOGRAFÍA.....	96

## Índice de Imágenes

IMAGEN 1: FÓSFOROS 1 .....	39
IMAGEN 2: FÓSFOROS 2 .....	40
IMAGEN 3: LENA.BMP .....	47
IMAGEN 4: CARA1 .....	63
IMAGEN 5: CARA2 .....	63
IMAGEN 6: CARA3 .....	63
IMAGEN 7: FÓRMULA1 .....	63
IMAGEN 8: FÓRMULA2 .....	63
IMAGEN 9: FÓRMULA3 .....	63
IMAGEN 10: PAISAJE1 .....	63
IMAGEN 11 PAISAJE3.....	63
IMAGEN 12: PAISAJE2 .....	63

## Índice de Fórmulas

FÓRMULA 1: SEÑAL DE LUZ.....	27
FÓRMULA 2: COMPONENTES DE ILUMINACIÓN Y REFLEJO .....	27
FÓRMULA 3: COMPONENTE DE REFLEXIÓN.....	27
FÓRMULA 4: SUPERPOSICIÓN DE IMÁGENES IDEALES .....	28
FÓRMULA 5: FUNCIÓN BRILLO.....	28
FÓRMULA 6: DEFINICIÓN DE IMAGEN.....	28
FÓRMULA 7: REDUNDANCIA DE DATOS RELATIVA .....	37
FÓRMULA 8: RADIO DE COMPRESIÓN .....	38
FÓRMULA 9: PROBABILIDAD DEL NIVEL DE GRIS EN UNA IMAGEN.....	38
FÓRMULA 10: PROMEDIO DE BITS.....	39
FÓRMULA 11: CORRELACIÓN DISCRETA .....	40
FÓRMULA 12: CORRELACIÓN DISCRETA NORMALIZADA .....	41
FÓRMULA 13: ERROR ENTRE F Y EN (X,Y).....	55
FÓRMULA 14: ERROR TOTAL .....	56
FÓRMULA 15: ERROR CUADRÁTICO MEDIO .....	56
FÓRMULA 16: $SNR_{MS}$ .....	56
FÓRMULA 17: PROBABILIDAD DEL NIVEL K.....	56
FÓRMULA 18: INFORMACIÓN DEL NIVEL K .....	57
FÓRMULA 19: ENTROPÍA .....	57
FÓRMULA 20: HIPOTESIS DE ANOVA.....	76
FÓRMULA 21: MODELO ESTADÍSTICO .....	76
FÓRMULA 22: HIPÓTESIS MODIFICADAS .....	76

## Índice de Tablas

TABLA 1: BLOQUE DE 8X8 DE LENA .....	45
TABLA 2: COEFICIENTES DCT DEL BLOQUE DE PÍXELES DE LENA .....	48
TABLA 3: MATRIZ DE CUANTIFICACIÓN POR DEFECTO DE JPEG.....	49
TABLA 4: CÁLCULO DE ETIQUETAS .....	50
TABLA 5: CÓDIGO POR DEFECTO DE JPEG .....	51
TABLA 6: CATEGORÍAS DE COEFICIENTES Y SUS RESPECTIVOS RANGOS .....	52
TABLA 7: ALGUNAS PALABRAS DE CÓDIGO DE LA TABLA DE CODIFICACIÓN POR DEFECTO.....	54
TABLA 8: ESCALA DE CALIDAD .....	58
TABLA 9: OTRA POSIBLE ESCALA .....	59
TABLA 10: ESCALA PARA IMÁGENES TELEVISIVAS.....	59
TABLA 11: TAMAÑOS Y RADIOS DE LAS IMÁGENES .....	66
<b>TABLA 12: VALORES OBJETIVOS CARAS.....</b>	<b>73</b>
TABLA 13: VALORES OBJETIVOS FÓRMULAS .....	74
TABLA 14: VALORES OBJETIVOS PAISAJES .....	75
TABLA 15: ANOVA.....	77
TABLA 16: ANOVA MODELO 1.....	78
TABLA 17: ANOVA MODELO 2.....	79
TABLA 18: ANOVA MODELO 3.....	80
TABLA 19: ANOVA MODELO 4 PARA EXPERTOS .....	80
TABLA 20: ANOVA MODELO 4 PARA INEXPERTOS .....	81

## Índice de Funciones

FUNCIÓN 1: AUTOCORRELACIÓN DE IMAGEN 2 .....	40
FUNCIÓN 2: AUTOCORRELACIÓN DE IMAGEN 3 .....	40
FUNCIÓN 3: RESIDUOS - DISTRIBUCIÓN NORMAL INEXPERTOS.....	83
FUNCIÓN 4: RESIDUOS – ESTIMADOS INEXPERTOS.....	84
FUNCIÓN 5: RESIDUOS - DISTRIBUCIÓN NORMAL INEXPERTOS.....	87
FUNCIÓN 6: RESIDUOS – ESTIMADOS EXPERTOS.....	87

## Índice de Figuras

FIGURA 1: INTENSIDADES.....	19
FIGURA 2: FRACCIÓN DE WEBER.....	19
FIGURA 3: CONTRASTE DE LUMINOSIDAD .....	20
FIGURA 4: CONTORNOS SUBJETIVOS.....	24

FIGURA 5: ILUSIÓN DE MULLER-LYER .....	25
FIGURA 6: ILUSIÓN POR TAMAÑO .....	25
FIGURA 7: A) TRIPLE IMPOSIBLE B) ESCALERA INTERMINABLE.....	25
FIGURA 8: ETAPAS FUNDAMENTALES DEL PROCESAMIENTO DE IMÁGENES .....	30
FIGURA 9: SECUENCIA EN ZIGZAG PARA UN BLOQUE DE 8X8 DE COEFICIENTES .....	53
FIGURA 10: ELIMINACIÓN DE LA VARIABLE FIGURA.....	77
FIGURA 11: CORRELACIÓN CALIDAD-RADIO PARA EXPERTOS-ÍNEPERTOS.....	88
FIGURA 12: CORRELACIÓN SNR - RADIO .....	89
FIGURA 13: SNR - RADIO PARA CARAS.....	90
FIGURA 14: CALIDAD - RADIO PARA CARAS .....	90
FIGURA 15: SNR - RADIO PARA LOS PAISAJES .....	91
FIGURA 16: .CALIDAD - RADIO PARA PAISAJES.....	91
FIGURA 17: CALIDAD SUBJETIVA EXPERTOS-ÍNEPERTOS.....	93

## Índice de Pantallas

PANTALLA 1: VALORIZACIÓN .....	68
PANTALLA 2: FIN DE LA EVALUACIÓN .....	69
PANTALLA 3: PANTALLA DE EVALUACIÓN.....	70
PANTALLA 4: REPORTE DE DATOS COMPLEMENTARIOS.....	70
PANTALLA 5: REPORTE DE CALIFICACIÓN.....	71

## Índice de LSD

LSD 1: CALIDAD-CLASE PARA ÍNEPERTOS .....	82
LSD 2: CALIDAD-COMPRESIÓN PARA ÍNEPERTOS.....	83
LSD 3: CALIDAD-OBSERVADOR PARA ÍNEPERTOS.....	83
LSD 4: CALIDAD-CLASE PARA EXPERTOS .....	85
LSD 5: CALIDAD-COMPRESIÓN PARA EXPERTOS.....	86
LSD 6: CALIDAD-OBSERVADOR PARA EXPERTOS.....	86

### **III - Introducción**

La compresión de una imagen digital puede facilitar su procesamiento, almacenamiento y transmisión. A medida que los Departamentos de las grandes Organizaciones se vuelven cada vez más digitales y distribuidos, la cantidad de datos multimediales (en particular imágenes) que tienen, obligan a considerar su compresión para su almacenamiento y transmisión.

El objetivo general de la compresión es representar una imagen con la menor cantidad posible de bits con la menor pérdida de calidad posible, acelerando así la transmisión y minimizando los requerimientos de almacenamiento. De manera alternativa, el objetivo es lograr la mejor fidelidad posible para una capacidad de almacenamiento disponible limitada.

En el campo de las comunicaciones se ha destinado mucho trabajo a la codificación de imágenes digitales, dirigido a reducir los requerimientos de la tasa de bits para la transmisión de las imágenes. La experiencia ha mostrado que cada esquema de codificación está sujeto a su propio y único conjunto de causas de pérdida que a menudo son difíciles de caracterizar [González 1996]. Esto se debe a la forma en la que se diseñan los esquemas de codificación para variar selectivamente la precisión de la representación; el observador puede ser insensible a los errores en algunas partes de la imagen pero no en otras. Se puede lograr una compresión considerable sólo mediante algoritmos con pérdida, los que no permiten la recuperación exacta de la imagen original. Esta pérdida de información hace que la compresión y otros algoritmos de procesamiento de imágenes con pérdida tengan problemas de confiabilidad debido a la potencial pérdida de calidad [Cosman 1994].

Hay una necesidad de contar con mediciones precisas de las pérdidas subjetivas que puedan usarse para predecir la calidad de una imagen. El objetivo de este trabajo es determinar estas mediciones de la distorsión y probar que se pueden usar diferentes medidas objetivas para predecir la calidad subjetiva de una imagen.

La necesidad de este tipo de mediciones está particularmente reconocida en el área de la codificación de imágenes digitales. La existencia de mediciones de la distorsión subjetivamente relevantes que reflejen las evaluaciones de la calidad de una imagen que hace un espectador, harían considerablemente más fácil la tarea de diseñar y optimizar los esquemas de codificación.

Estas distorsiones se comparan y contrastan con un conjunto de imágenes representativas de varios dominios de aplicación, y se examina cuán buenas son las mediciones de distorsiones que se

pueden obtener fácilmente para predecir las evaluaciones subjetivas, que son más caras en tiempo. Los ejemplos son de imágenes tradicionales que se utilizan en el área de procesamiento y compresión de imágenes comprimidas utilizando JPEG estándar [Wallace 1991] [Russo 1998].

## IV - Modelo Biológico de la Visión Humana

El objetivo de esta sección es dar una descripción elemental sobre el funcionamiento del sistema de visión humano. Las razones para incluir esta descripción en este trabajo se fundamenta en la necesidad de mostrar algunos fenómenos visuales característicos de nuestro sistema visual y que son significativos para explicar las diferencias que pueden existir entre una medición objetiva de pérdida de calidad y la apreciación subjetiva de un observador humano.

Una vez enfocada la luz en la retina, muchas capas del sistema fisiológico procesan la información, algunas eliminan o rechazan cierta información, y otras hacen énfasis en otro tipo de información, dándole a la señal una forma que se pueda interpretar representado estructuras físicas.

Las células fotosensibles (bastoncillos y conos) del ojo no responden en forma uniforme a todas las longitudes de onda en el rango visible y el procesamiento que viene después sirve para refinar más la importancia final de diversas regiones del espectro para una interpretación de la imagen.

El primer paso para procesar la información de la luz es la recepción de la señal luminosa por las células fotosensibles en la retina. Aunque la mayoría de estas células tienen una estructura larga y delgada, en la retina no están agrupadas unas paralelas a las otras. Más bien están inclinadas hacia el centro de la pupila. El resultado es una sensibilidad direccional, conocida como *efecto de Stiles-Crawford*, en el que los conos responden más a la luz que llega en línea recta que a la que llega con un ángulo de inclinación a través del borde de la pupila.

Una vez que la luz ha logrado alcanzar el material fotosensible en un bastoncillo o en un cono, causa una acción química que resulta en una señal nerviosa. El químico tiene el nombre general de *fotopigmento*.

La señal transportada por el cambio en el potencial de la membrana forma el mensaje completo enviado por el fotorreceptor al resto del sistema visual. De esta manera, el único mensaje enviado por un bastoncillo o por un cono es que la luz ha llegado y ha estimulado el fotopigmento; no se transmite información que describa la longitud de onda del fotón. Este efecto se llama *principio de univarianza*.

Hasta el momento sólo se ha discutido la respuesta del ojo a un único fotón. De hecho, los procesos químicos que ocurren dentro de un fotorreceptor duran varios milisegundos, y los fotones



adicionales que llegan al receptor durante ese tiempo se agregan a la respuesta general. De esta manera la salida de un receptor es realmente una respuesta promediada en el tiempo, un efecto llamado *suavización temporal*. En efecto, los sensores imponen un filtro de paso bajo sobre su respuesta temporal, aunque la frecuencia crítica de ese filtro cambie con respecto al nivel de luz de fondo: cuando llega poca luz, hay poca suavización.

Se ha visto que un sensor reacciona a un fotón entrante con un cambio químico y que luego se comunica al circuito nervioso en el ojo por un cambio en el potencial eléctrico en la membrana de la célula. Sin embargo hay una complicación adicional que le posibilita al ojo responder a variaciones enormes en los niveles de luz entrante. El fenómeno de *adaptación* le da al sistema una gran sensibilidad cuando la iluminación general es baja, y algo (aunque menos) de sensibilidad cuando la iluminación general es alta. Aunque sería mejor una sensibilidad máxima sobre todos los rangos de iluminación, éste parece ser un problema difícil para cualquier sistema receptor. Dada la necesidad de resignar ciertos beneficios a favor de otros, parece muy deseable tener la mayor sensibilidad a niveles de luz bajos, en donde variaciones pequeñas transportan una gran cantidad de información. El rango de adaptación es extremadamente grande. La luminancia se mide en *candelas* por metro cuadrado; que puede considerarse como la luz generada por una vela típica.

Como los bastoncillos son unas diez veces más sensibles que los conos, son más útiles para visión nocturna, cuando los niveles de luz ambiental son bajos. En niveles de iluminación diurna, son los conos los detectores más útiles de la información de la luz. Cuando un cono se ha adaptado a un nivel particular de intensidad de luz, funciona igual que los bastoncillos: las intensidades de luz más allá de un nivel particular harán que el cono se sature y cese de enviar señales nerviosas.

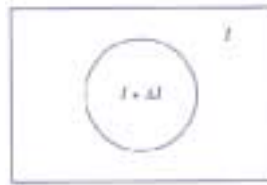
Los bastoncillos y los conos sólo responden a la luz que los alcanza. La luz debe pasar a través de las capas internas de la retina, la que puede absorber algo de luz. La luz también debe pasar a través del ojo mismo, atravesando la lente y los otros componentes del ojo. Por ejemplo, la lente en el ojo humano cambia de color con el tiempo, volviéndose cada vez más amarilla a medida que la persona envejece. De esta manera, la lente actúa como un filtro amarillo, lo que obviamente afecta la distribución espectral de la luz que incide sobre la retina.

## **Fenómenos Visuales**

La complejidad del sistema visual humano hace que sea necesario recurrir a experimentos físicos para comprender su funcionamiento.

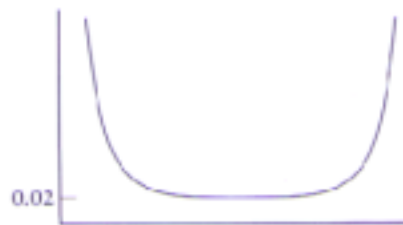
## ***Sensibilidad al Contraste***

A un observador se le muestra una hoja de papel con una intensidad reflejada  $I$ , dentro de la cual hay una hoja más pequeña con una intensidad levemente diferente  $I + \Delta I$ , como en la Figura 1.



**Figura 1: Intensidades**

Se debe encontrar la menor *diferencia notable*  $\Delta I$  tal que el observador informe que la región interna es de una intensidad diferente que la región externa. Sobre un amplio rango de intensidades, la relación  $\Delta I/I$  (llamada *fracción de Weber*) es casi constante con un valor de alrededor de 0,02, según se muestra en la Figura 2. Esta curva se conoce como *función de sensibilidad de contraste*, y sugiere que el sistema visual humano responde a relaciones de intensidades y no a valores absolutos.



**Figura 2: Fracción de Weber**

## ***Ruido***

Muchos sentidos humanos son tolerantes al *ruido*. Por ahora, simplemente se considerará al ruido como una señal que parece tener un componente aleatorio fuerte que se agrega a la señal que nos interesa. Un ejemplo del dominio del audio es el siseo de los cassettes, que es el sonido producido por una cinta de audio en blanco. Un ejemplo visual es la estática en una señal de televisión, donde los colores están ocasionalmente mal y hay puntos blancos o negros.

Siempre y cuando este ruido no sea demasiado extremo, el sistema visual humano tiende a ser muy bueno en ignorarlo. Esto es probablemente el resultado de cómo se distribuyen los fotorreceptores dentro de la retina.

## ***Contraste de Luminosidad***

El fenómeno de *contraste de luminosidad* (también llamado *contraste simultáneo*) se ilustra en la Figura 3



**Figura 3: Contraste de Luminosidad**

Aquí se tiene un recuadro de un valor dado de gris rodeado por una cantidad de otros recuadros con diferentes valores de gris. La luminosidad aparente del recuadro parece depender del valor de gris que lo rodea; cuanto más oscuro es el valor de gris del recuadro externo, más claro parece el cuadrado interno.

Este fenómeno hace difícil escoger dos intensidades al azar (o dos colores) y esperar que se comporten en forma predecible en toda una imagen. Por ejemplo, una forma rápida de representar una escena nocturna es con un degradé horizontal de color, claro en la parte de abajo (para representar la luz del sol poniente) y oscuro en la parte de arriba (para mostrar el cielo de la noche).

El fenómeno de *contraste de luminosidad* nos permite aceptar una misma escena válida tanto para representar el día como la noche, cuando el nivel de iluminación es muy diferente. Por ejemplo, si se supone que una noche usted está leyendo en su escritorio. El libro frente a usted está impreso en papel blanco que refleja un 40% de la luz incidente y la tinta negra de la impresión refleja sólo el 5%. Ahora usted enciende otra lámpara que duplica la iluminación de la habitación. La impresión negra ahora refleja el doble de energía luminosa hacia usted, pero no parece el doble de brillante. Esto es porque la página blanca también está el doble de brillante, por lo que la *relación* se ha mantenido igual.

La constancia en la iluminación es una característica poderosa del sistema visual humano y es uno de los fenómenos que hace posible que se mantenga una imagen mental constante del mundo, a pesar de cambios bruscos en el nivel de iluminación.

## ***Percepción de Profundidad***

El sistema visual humano es capaz de construir una visión del mundo en 3D. Esta capacidad, llamada *percepción de profundidad*, proviene de muchas clases diferentes de información visual, algunas de las cuales pueden reunirse de un solo ojo, y otras que requieren los dos ojos.

Para ver cómo los dos ojos funcionan juntos, se verá la ubicación de los ojos en la cabeza de un conejo y en la de una persona. El conejo tiene casi 300° de campo visual, aunque sólo una pequeña parte de este campo visual se ve con los dos ojos. El ojo humano tiene un campo de vista más chico, pero los dos ojos se superponen en una región mucho mayor. Otros ejemplos de ubicación del ojo son los caracoles, que tienen los ojos en el extremo de dos antenas flexibles de manera que las regiones de visibilidad y superposición pueden cambiarse a voluntad, y las ballenas, que tienen los ojos tan separados en los costados de la cabeza que son completamente ciegas en línea recta. Las arañas y los escorpiones tienen clusters de por lo menos 6 ojos, y algunos 8; hay una cantidad importante de superposición de campos.

En general, los animales depredadores tienen los ojos cerca del frente de la cabeza con mucha superposición de campo para una mejor estimación de la profundidad cuando están persiguiendo a una presa. En forma inversa, los animales que son cazados tienen los ojos muy separados, lo mejor para ver más del entorno y responder a ataques potenciales. Por ejemplo, los ojos del búho tienen una porción muy grande de superposición.

## **Profundidad Óculmotora**

Los efectos *óculomotores* provienen de los ajustes musculares en los ojos. Cuando se mira algo, se usan los músculos alrededor del ojo para hacer que los ojos *converjan*, o para rotarlos físicamente para hacer que el punto de atención, o *punto de fijación*, caiga dentro de la fovea. También *se acomoda* cambiando el foco, tensando o relajando el cuerpo ciliar para ajustar el espesor de la lente cristalina dentro del ojo.

Ninguno de estos efectos es un indicador particularmente robusto o preciso de información de profundidad, ya que sólo brindan información útil acerca de objetos cerca del observador. Cuando se observa un objeto ubicado a más de unos 6 metros de distancia, el cuerpo ciliar está en su estado más relajado, y los ojos convergen en el infinito (mirando derecho hacia adelante). Así, para 6 metros y más básicamente no hay datos óculomotores que contribuyan a la percepción de la profundidad.

Las otras categorías principales de datos de profundidad son los datos *visuales*, que se dividen en las clases *binocular* y *monocular*, dependiendo de si involucran los dos ojos o uno solo.

### **Profundidad Binocular**

La actividad ocular se denomina binocular cuando están involucrados dos ojos. La capacidad de hacer juicios de profundidad basados en información proveniente de una visión binocular se llama *estereopsis* (visión estereoscópica). La estereopsis puede proporcionar información muy precisa sobre la profundidad de los objetos en una escena.

Por ejemplo, sostiene 2 lápices verticalmente a una distancia de 1 metro de sus ojos. La estereopsis hace posible que se distinga una disparidad de 1 mm en las distancias de estos dos lápices; ésta es una precisión bastante notable de 1 unidad en 1000.

Para percibir profundidades basadas en información binocular, el sistema visual necesita realizar dos tareas que son (por lo menos conceptualmente) distintas. La primera es *combinar las características* de las dos imágenes; a continuación se realiza un cálculo de su *disparidad retinal*, o desplazamiento relativo en las imágenes retinales.

Se puede imaginar que la combinación de las características comienza con la *extracción de las características*, o búsqueda de objetos significantes en ambas imágenes, seguida de una *correspondencia de características* que identifica características similares en las dos imágenes. Se sugiere un ejemplo de esto mirando una habitación llena de libros; la primera etapa del procesamiento identificaría cada bulto con forma de libro en cada imagen como un *libro*. Dos libros rojos pueden luego ponerse en correspondencia.

Por el momento no se conoce la explicación completa de percepción de profundidad y de combinación de imagen binocular, pero parecería ser que la fisiología del sistema visual y del cerebro jugarían un papel muy importante en la resolución de la disparidad retinal para crear una imagen unificada del mundo. Parece haber células que están específicamente diseñadas para buscar equivalencias entre partes particulares de cada retina. Cuando estas células encuentran una equivalencia, la profundidad del punto de fijación puede usarse para determinar si el objeto bajo escrutinio está más cerca o más lejos que el punto de enfoque.

Al igual que con el resto del sistema visual (y todo el cuerpo humano), la estereopsis es a la vez robusta y frágil. Si cualquiera de los pasos de la estereopsis no se satisface, entonces se dice que la persona es *estereociega*. En vez de tolerar dos imágenes competitivas o no resueltas, el sistema visual

parece seleccionar una imagen para la presentación al resto del cerebro, y suprime la información que viene del otro ojo. La selección de qué imagen elige para procesar puede ser fija, o puede cambiar, dependiendo del individuo.

### **Profundidad Monoocular**

Pueden extraerse varios datos en cuanto a la profundidad de una sola imagen; éstos se conocen como datos de profundidad *monoculares*. Hay dos categorías generales de tales casos: datos *estáticos*, que pueden extraerse a partir de una sola escena, y datos *dinámicos*, que requieren varias imágenes en un período de tiempo.

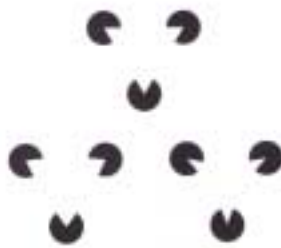
- *Interposición*: El dato de interposición es cómo se entiende la escena: si un objeto A tapa a un objeto B, se supone que A está más cerca que B. La interposición es muy poderosa; si en un experimento se le muestra a un sujeto una escena en la que los datos de disparidad retinal y de interposición se contradicen, los datos de interposición son los que triunfan.
- *Tamaño*: El dato llamado *tamaño* resume la experiencia de que los objetos más grandes parecen estar más cerca que los objetos más chicos. Incluso en un conjunto abstracto de objetos, el objeto más grande parece estar más cerca que los otros. También parece existir una noción de *tamaño familiar*; si se ve la cara de un amigo, rápidamente se estima la distancia a la que está esa persona porque se conoce aproximadamente el tamaño real de su cara.
- *Perspectiva*: Las pistas de profundidad clasificadas como fenómenos de *perspectiva* están relacionadas con los cambios percibidos en las estructuras físicas con la distancia. La perspectiva es un resultado natural de la pupila que actúa como la puerta de entrada a nuestro sistema visual. Se podría pensar en la pupila como un punto a través del cual debe pasar toda la luz, creando una *proyección de perspectiva*. La perspectiva no es la única manera de proyectar un mundo 3D en una superficie 2D, pero es aquella con la cual se está más familiarizado en la vida cotidiana.

## **Ilusiones**

Las ilusiones ópticas han contribuido mucho con nuestra comprensión del sistema visual. Sirven para aislar y demostrar efectos y fenómenos que normalmente o se toman por seguros o se ignoran. [Luckiesh 1965] presenta un compendio clásico de ilusiones visuales.

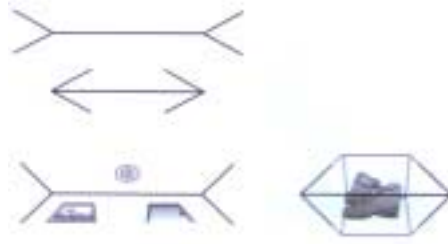
Aquí se presentan algunas ilusiones con el objeto de demostrar que puede haber una gran disparidad entre la descripción mecánica de una imagen y su percepción. En estos casos la definición de *precisión* no incluye al observador, ya que la gente a menudo *ve* objetos, contornos y relaciones en imágenes que no son explícitamente parte de la imagen.

Si se considera por ejemplo la Figura 4. En estas imágenes la mayoría de los observadores percibe un triángulo cuyas esquinas están sugeridas por los puntos negros eliminados. El sistema visual completa el resto de los contornos del triángulo, aun cuando los bordes resultantes no sean rectos; estos contornos se denominan *contornos subjetivos*.



**Figura 4: Contornos subjetivos**

Muchas ilusiones famosas ubican objetos de igual tamaño en contextos diferentes, con el resultado de que parecen de diferente tamaño. La ilusión de Muller-Lyer en la Figura 5 muestra dos líneas horizontales de igual longitud, una con dos puntas de flecha orientadas hacia adentro y la otra con dos puntas de flecha orientadas hacia afuera. Normalmente la primera parece más larga. Una explicación de este efecto es que las flechas parecen sugerir dos planos que se interceptan. Las flechas hacia adentro sugieren que el ángulo es cóncavo desde nuestro punto de vista; por ejemplo, si se está mirando la unión entre la pared y el techo dentro de una habitación. Las flechas hacia afuera sugieren un ángulo convexo, el lado de afuera de una caja por ejemplo. Los observadores tal vez suponen que la intersección cóncava está más lejos que la convexa, ya que parece estar en retroceso. Como ambos segmentos horizontales tienen la misma longitud, *se sabe* a partir de la perspectiva que la línea que está más lejos debe ser por lo tanto más grande. Este argumento no es seguro, pero sugiere el tipo de fenómenos de alto y bajo nivel que probablemente se combinan para crear algunas ilusiones.



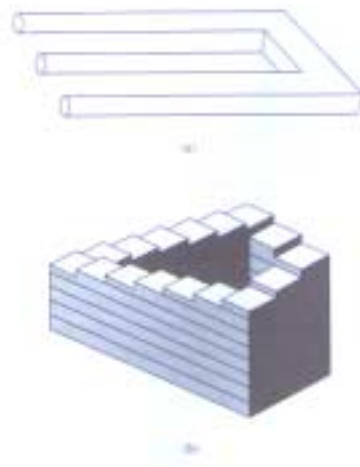
**Figura 5: Ilusión de Muller-Lyer**

En la Figura 6 se muestra una ilusión similar; los dos círculos internos son del mismo tamaño, aunque normalmente no lo parecen. La explicación para esta ilusión es todavía más tenue.



**Figura 6: Ilusión por tamaño**

Otras ilusiones clásicas incluyen *figuras imposibles*, en donde se nos presenta una proyección plana de una forma en 3D que es lógica localmente pero globalmente inconsistente. Ejemplos famosos incluyen la barra triple imposible de Penrose y la escalera interminable de la Figura 7.



**Figura 7: a) Triple imposible b) Escalera interminable**



## V - Procesamiento de Imágenes

En esta sección se define qué es una imagen digital y las etapas de su procesamiento.

El término *imagen* se refiere a una función bidimensional de intensidad de luz  $f(x,y)$ , donde  $x,y$  representan las coordenadas espaciales. El valor de  $f$  en un punto cualquiera  $(x,y)$  es proporcional al brillo o nivel de gris de la imagen en ese punto. Una imagen digital es una imagen  $f(x,y)$  que se ha discretizado tanto en las coordenadas espaciales como en el brillo. Una imagen digital puede considerarse como una matriz cuyos índices de fila y columna identifican un punto de la imagen, y el valor correspondiente del elemento de la matriz indica el nivel de gris en ese punto. Los elementos de una distribución digital se denominan *pixels* o *pels* [González 1996].

### Estructura de una Imagen

La señal de la luz es positiva y distinta de cero, como lo indica la Fórmula 1:

$$\infty > I_{x,y} > 0$$

#### Fórmula 1: Señal de luz

$I$  representa la energía o intensidad y  $x$  e  $y$  representan el dominio espacial de la imagen. Además, como la imagen también se forma con la luz reflejada de los objetos, la estructura física de una imagen se divide en dos partes, una es la cantidad de luz disponible para iluminar los objetos, y otra es la habilidad de esos objetos de reflejar luz. Estas partes se llaman componente de iluminación y componente de reflejo respectivamente, Fórmula 2.

$$\infty > i_{x,y} > 0$$

$$1 > r_{x,y} > r_{min} \approx 0.005$$

#### Fórmula 2: Componentes de iluminación y reflejo

Combinando éstas de acuerdo a la ley de la reflexión para formar una imagen  $I_{x,y}$ , se obtiene la Fórmula 3

$$\infty > I_{x,y} = i_{x,y} \cdot r_{x,y} > 0$$

#### Fórmula 3: Componente de reflexión

Con esta fórmula se ve que existen en una imagen dos tipos diferentes de información. La primera información es  $i_{x,y}$ , está relacionada con la iluminación de la escena. La segunda es  $r_{x,y}$ , y está relacionada con la naturaleza de los objetos de la escena. Estas componentes están separadas en términos de la naturaleza del mensaje que posee cada una.

Hasta ahora se ha asumido que existe un método ideal para generar una imagen. Sin embargo, tal método no existe y sólo se puede obtener una aproximación a la fórmula anterior. Dado que muchos de los métodos para formar una imagen utilizan o involucran mecanismos lineales, se puede ver una imagen como una superposición de imágenes ideales y expresarla como en la Fórmula 4:

$$\infty > \hat{I}_{x,y} = \int_{-\infty}^{\infty} I_{x,y} h_{x,X;y,Y} dXdY > 0$$

#### Fórmula 4: Superposición de Imágenes Ideales

Donde  $\hat{I}_{x,y}$  representa la imagen práctica y  $h_{x,X;y,Y}$  representa la función lineal del mecanismo que forma una imagen, o función brillo.  $h_{x,X;y,Y}$  es la imagen práctica de una imagen ideal que consiste de un punto de luz localizado en  $x=X$  e  $y=Y$ .

Si la función brillo tiene la misma forma para todos los puntos de luz de la imagen ideal, se puede representar con la integral de convolución de la Fórmula 5:

$$\infty > \hat{I}_{x,y} = \int_{-\infty}^{\infty} I_{x,y} h_{x-X;y-Y} dXdY = I_{x,y} * h_{x,y} > 0$$

#### Fórmula 5: Función Brillo

Combinando se obtiene la Fórmula 6:

$$\infty > \hat{I}_{x,y} = (i_{x,y} \cdot r_{x,y}) * h_{x,y} > 0$$

#### Fórmula 6: Definición de Imagen

Esta última expresión pone en evidencia las tres componentes esenciales de una imagen práctica. Si  $h_{x,y}$  es lo suficientemente pequeña en el dominio extendido, la imagen práctica puede tomarse como una aproximación adecuada a la imagen ideal. Si  $h_{x,y}$  falla, se puede procesar la imagen práctica con varios métodos para intentar remediar esta situación.

## Representación de Imágenes

Para ser utilizable en el procesamiento de imágenes, una función de imagen  $f(x,y)$  debe ser digitalizada tanto espacialmente como en su amplitud. La digitalización de las coordenadas espaciales  $(x,y)$  se denomina muestreo de la imagen y la digitalización de la amplitud se conoce como cuantificación del nivel de gris.

Una imagen continua se describe en forma aproximada por una serie de muestras igualmente espaciadas en forma de una matriz de  $N \times M$ , donde cada elemento de la matriz es una cantidad discreta (nivel de gris). Cada elemento de la matriz se denomina pixel.

Tanto el muestreo como la cuantificación se pueden expresar en términos matemáticos formales. Si  $Z$  y  $R$  representan a los enteros y reales respectivamente, el proceso de muestreo puede entenderse como una partición de la cuadrícula del plano  $xy$ , siendo las coordenadas del centro de cada elemento de la cuadrícula un par de elementos de producto cartesiano  $Z \times Z$ . Se puede decir que  $f(x,y)$  representa una imagen digital si  $(x,y)$  son enteros de  $Z \times Z$  y  $f$  es una función que asigna un nivel de gris a cada par de coordenadas  $(x,y)$  distintas. Este es el proceso de cuantificación descrito anteriormente. Puede ser que los niveles de grises también sean enteros, con lo cual se reemplaza  $R$  por  $Z$ .

La digitalización necesita decisiones sobre los valores para  $N$  y  $M$  y el número de niveles de grises válidos para cada pixel. En la práctica estos valores normalmente son potencias de 2.

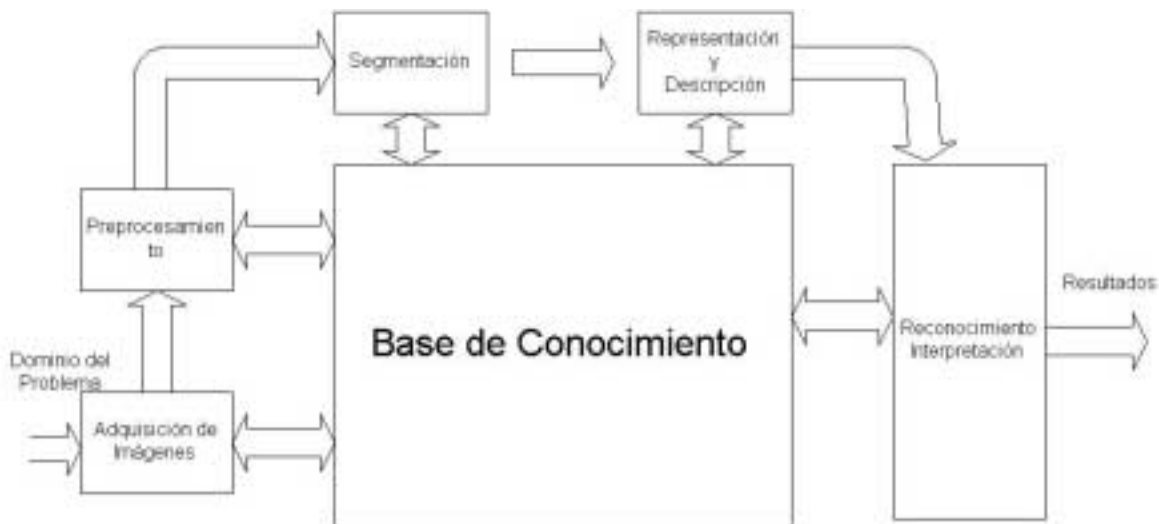
Puesto que se tiene una aproximación a una imagen continua, es razonable preguntarse ¿qué muestreo y cuántos niveles de grises son necesarios para una buena aproximación?. La resolución (grado de detalle) de una imagen depende estrechamente de estos dos parámetros, es obvio que cuanto más se incrementan estos, más cerca se está de la imagen continua. Sin embargo, existe un problema en cuanto al almacenamiento y las necesidades de procesamiento, ya que éstos crecen en función de  $N$ ,  $M$  y  $m$  (niveles de grises).

## Etapas Fundamentales de procesamiento de Imágenes

El tratamiento de imágenes digitales incluye áreas como hardware, software y aspectos teóricos.

La Figura 8 representa la obtención de un objetivo global a partir de algún problema determinado por medio del procesamiento de imágenes. Estos problemas corresponden a diferentes

dominios como por ejemplo leer los destinatarios de las cartas de correo o buscar patrones parecidos en huellas dactilares.



**Figura 8: Etapas Fundamentales del Procesamiento de Imágenes**

La primera etapa es la *adquisición de la imagen*. Para esto se necesita un sensor de imágenes y la posibilidad de digitalizar la señal producida, el sensor puede ser una cámara de televisión que produce una imagen del dominio del problema cada 1/30 segundos.

Una vez obtenida la imagen, la siguiente etapa consiste en el *preprocesamiento* de esa imagen. La función básica es mejorar la imagen de forma que se aumenten las posibilidades de éxito en los procesos posteriores. Típicamente trata sobre técnicas de mejora del contraste, eliminación de ruido y aislamiento de regiones que sirvan para el objetivo determinado por el dominio del problema.

La *segmentación* consiste en separar la imagen de entrada en sus partes constituyentes. Es una de las etapas más difíciles del procesamiento de imágenes.

A la salida de la segmentación habitualmente se tienen los datos que constituyen el contorno de una región o todos los puntos de ellas. Luego los datos se deben convertir a una forma adecuada para el procesamiento, esta es la etapa de *representación y descripción*. Se debe decidir si es importante el contorno, donde solo se tiene información del exterior del objeto, como esquinas, puntos de inflexión, etc. La información regional sirve cuando se necesita información sobre la textura o estructuración. En esta etapa debe especificarse un método para describir los datos de forma que resalten los datos de interés; la descripción o selección de rasgos consiste en extraer alguna información cuantitativa de interés que sirva para diferenciar una clase de objeto de otra.

La última etapa incluye el *reconocimiento*, que es el proceso que asigna una etiqueta a un objeto basándose en la información de los descriptores. La *interpretación* implica asignar significado a un conjunto de objetos reconocidos.

El conocimiento sobre un dominio de problema está codificado en una base de datos. Ésta puede ser tan simple como detallar los lugares de las imágenes en donde se sabe que se ubica la información de interés, limitando así la búsqueda para hallar información o puede también puede ser muy compleja, como una lista de todos los posibles defectos en un problema de inspección de materiales o una base de datos de imágenes de huellas digitales.

Durante la etapa de procesamiento, según el dominio de la aplicación, puede ser útil almacenar una imagen  $f$  en la base de datos. Esto implica optimizar el espacio físico para lo que es necesario comprimir con pérdida la imagen  $f$ , obteniendo  $f'$ . Visualmente  $f$  y  $f'$  deben ser parecidas.

Para evaluar esta diferencia es más eficaz la evaluación subjetiva dadas las características del sistema visual humano, pero es más costosa en su ejecución. Por lo tanto, sería deseable evaluar las imágenes de acuerdo a alguna medida objetiva modificada de rápida ejecución.

El material de este capítulo es a título informativo, y sirve como base para las presentaciones posteriores. Si el lector estuviere interesado, existe mucha bibliografía sobre procesamiento de imágenes que está fuera del alcance de este trabajo [González 1996] [Jain 1989] [Stockham 1972].

## VI - Compresión de Datos

Se analizan los conceptos básicos relacionados con la redundancia de la información y cómo se puede utilizar esto para crear modelos de compresión. En particular, se estudiarán los esquemas de compresión con pérdida.

### Terminología

La compresión de un bloque de datos original se llama proceso de codificación, y el código resultante se llama bloque de datos comprimido. Inversamente a este proceso, el bloque de datos comprimido se descomprime para reproducir el bloque de datos original.

La codificación puede ser con o sin pérdida. En la codificación sin pérdida los datos descomprimidos son exactamente iguales a los originales. Este tipo de codificación se da en información textual como planillas, documentos, etc. La codificación con pérdida realmente elimina información del archivo original y el usuario al descomprimir tiene menos información que la que originalmente tenía.

El grado de reducción obtenido como resultado del proceso de compresión se conoce como *Radio de Compresión*. Este grado mide la cantidad de datos comprimidos en comparación con la cantidad de datos originales, de manera que:

$$\text{Radio de Compresión} = \text{Long. del bloque original} / \text{Long. bloque comprimido}$$

Es obvio que mientras mayor sea el radio de compresión, más efectiva será la técnica empleada. De manera alternativa, la técnica empleada no será efectiva si el radio de compresión es menor o igual a 1.

Cuando se comprimen los datos, el radio de compresión varía en proporción a la sensibilidad de los datos a el/los algoritmo/s usado/s. Así, se debe centrar la atención sobre un radio de compresión promedio y no sobre un radio alcanzado con un archivo en particular. En general, buenos algoritmos pueden alcanzar un radio de compresión promedio de 1,5 mientras que excelentes algoritmos basados en técnicas de procesamiento sofisticadas alcanzarán un promedio de radio de compresión mayor a 2.0 para archivos de texto y compresión sin pérdida [Nelson, 1991]. Si se tiene una compresión con pérdida se puede llegar a radios 1:30 [Wallace 1991] [Russo 1998].

## Técnicas básicas de compresión de datos

Las técnicas de compresión de datos tienen por objetivo aplicar una transformación  $m'=f(m)$ , de manera de aprovechar la redundancia del dato  $m$ , y minimizar así el mensaje  $m'$ .

En general, cuanta más información del objeto a comprimir se tenga, más eficiente será el método de compresión, dado que la codificación puede estar basada en caracteres simples, bloques de dos, tres o más caracteres, estructuras completas de un texto (palabras o párrafos), bits, bloques de bits, etc. [Held 1991].

Por ejemplo, si el algoritmo tiene en cuenta la sintaxis de un lenguaje para comprimir un programa, entonces se obtendrán mejores logros codificando en función de las palabras claves. En imágenes se pueden tener en cuenta las zonas con igual color, textura, etc.

Esto sugiere que en algunos casos será necesario un estudio previo del objeto a comprimir para así lograr una mejor compresión aplicando el método que más se ajuste a ese caso particular.

Desde el punto de vista de la transmisión de información, se puede considerar al conjunto de datos como una secuencia de símbolos  $S_1, S_2, \dots, S_n$ . Estos símbolos forman un conjunto, posiblemente infinito.

Los distintos métodos se pueden agrupar en tres clases, según el concepto en el que se basen:

- Cantidad de símbolos del conjunto.
- Frecuencia de aparición de los símbolos en el texto.
- Contexto en el que aparecen los símbolos.

### Cantidad de símbolos del conjunto

En algunas aplicaciones los datos pueden reducirse a un conjunto finito expresado en ASCII.

Por ejemplo, en un proyecto de automatización de una empresa que comercializa libros, los títulos de los mismos pueden representarse con un conjunto finito de símbolos.

Si un título estándar tiene  $N$  caracteres, expresado en ASCII requerirá  $7*N$  bits. De las  $2^{7*N}$  combinaciones posibles sólo serán utilizadas, a lo sumo  $2^{25}$  (más de treinta y tres millones de

ejemplares). Si a cada libro se le asigna un número de secuencia, la cantidad de bits necesarios por ejemplar se reduce de  $7 \cdot N$  a veinticinco o menos. Nótese que  $N$  (cantidad de caracteres por título) es del orden de 10 a 100.

## **Frecuencia de aparición de los símbolos en el texto.**

En la mayoría de los textos algunos símbolos o letras aparecen más a menudo que otros. Esta observación sugiere un esquema de codificación en el cual a los símbolos más comunes les sean asignados códigos cortos, y a los símbolos menos frecuentes, códigos largos.

## **Contexto en el que aparecen los símbolos**

El método de Huffman supone que la probabilidad de que aparezca un símbolo es independiente de su predecesor inmediato. Un esquema más complicado podría ser determinar la probabilidad condicional de cada símbolo dado su predecesor. Para el alfabeto, esto implica tener 26 tablas, una para la frecuencia de distribución de cada letra seguida de una A, otra para la B, etc.

Si hay una fuerte relación entre los símbolos y sus sucesores, el uso de este método logrará buenos resultados, aún si la distribución de los símbolos es plana (sin tener en cuenta su predecesor).

La desventaja del método de la probabilidad condicional es la gran cantidad de tablas necesarias. Si hay  $k$  símbolos, las tablas tendrán  $k^2$  entradas. Debido a que para cada símbolo se debe indicar la probabilidad condicional.

Existe otro método que puede ser usado para codificar cadenas de bits que contienen mayor cantidad de ceros que de unos. Esta situación se da en transmisión de matrices ralas y de imágenes.

Cada código representa la cantidad de ceros que hay entre dos unos consecutivos. Para manejar sucesiones largas de ceros, el símbolo que consiste en todos unos significa que la verdadera distancia es  $2^{k-1}$  más el valor del o los símbolos siguientes.

Los algoritmos clásicos de compresión son [Russo 1998]:

- Algoritmo LZW.
- Run-Length.
- Algoritmo Huffman.



## Compresión con pérdida vs. compresión sin pérdida

Las técnicas de compresión pueden causar una pérdida o no de la integridad de los datos. Los esquemas de compresión con pérdida interpretan el archivo de datos, hacen suposiciones algorítmicas y eliminan realmente datos del archivo. A pesar que técnicamente se produce una pérdida de datos, el espectador percibe poca pérdida de calidad. Esta técnica es aconsejable para imágenes en las que perder un cierto nivel de sus características es aceptable para el espectador. Los esquemas con pérdida son necesarios, por ejemplo, para comprimir archivos de video con el fin de almacenarlos.

Los esquemas de compresión sin pérdida conservan todos los datos del archivo de datos. La compresión sin pérdida se recomienda para aplicaciones como hojas electrónicas en las que la pérdida de cualquier dato puede convertir el archivo en inutilizable o impreciso. Además, los ratios de compresión para los esquemas sin pérdida son mucho más pequeños que los de los esquemas con pérdida, habitualmente de 2 ó 3 a 1 como máximo.

También hay técnicas que caen bajo la categoría de compresión de movimiento. Estas técnicas se aplican a secuencias de imágenes relacionadas como las encontradas en dibujos en movimiento y producción de vídeo, secuencias de imágenes médicas y aplicaciones de teleconferencia. Con ellas se elimina la redundancia entre cuadros de imágenes secuenciales.

Las metodologías de compresión de texto, audio, gráficos y vídeo disponibles constan de estándares internacionales o esquemas patentados. La industria de computadoras está desarrollando constantemente algoritmos de compresión que son más eficientes, tienen menos pérdida y son más rápidos para comprimir y descomprimir. También se están realizando esfuerzos para proporcionar metodologías sofisticadas de compresión, descompresión y reproducción que puedan ser fácilmente transportadas a diversas plataformas de computadoras personales.

## Preservación de la información

El primer punto a tener en cuenta es si se desea preservar o no el cien por cien de la información o, lo que es lo mismo, si se admite una degradación parcial de la imagen durante el proceso de compresión o no.

Así, en un primer momento se pueden clasificar los algoritmos de compresión según su fidelidad. Para aquellos que no preservan la totalidad de la información presente en la imagen, es preciso establecer un criterio de fidelidad. En general se usan **criterios de fidelidad objetivos** que

suelen definirse como una función matemática de un pixel antes y después de ser comprimido y vuelto a descomprimir. De esta forma puede definirse el error absoluto por pixel, el error cuadrático y el error cuadrático medio o rms; esto se verá en la sección “Clases de Mediciones”. Análogamente, pueden definirse **criterios de fidelidad subjetivos** que, aunque menos rigurosos, suelen ofrecer resultados más aceptables cuando el propósito fundamental es visualizar la imagen. Dadas las peculiaridades del sistema visual humano, la calidad subjetiva de dos imágenes puede ser muy distinta aunque ambas presenten una relación señal-ruido similar. Esto puede permitir utilizar en algunos casos técnicas de compresión que, aunque objetivamente tengan un porcentaje de pérdida importante, resulten en un impacto menor sobre la calidad subjetiva de la imagen; esto también se discutirá en el apartado “Clases de Mediciones”.

## Compresión de Imágenes

El término compresión de datos se refiere al proceso de reducir la cantidad de datos requeridos para representar una cantidad de información. Se debe realizar una clara diferencia entre *dato* e *información*. Se pueden utilizar cantidades distintas de datos para representar la misma información. Como ejemplo, el relato de una historia puede ser narrada de distintas maneras, la información es la historia; las palabras que las describen pueden ser distintas. Distintos individuos pueden contar la misma historia con distintas cantidad de palabras. Las diferentes versiones contienen datos (palabras) que proveen información no relevante diciendo de esta manera que existe redundancia. Este concepto es fundamental para la compresión de imágenes. Es un concepto abstracto matemáticamente cuantificable.

Si  $n_1$  y  $n_2$  dos conjuntos de datos representando la misma información, la *redundancia relativa*  $R_d$  del conjunto  $n_1$  puede definirse como lo indica la Fórmula 7:

$$R_d = 1 - \frac{1}{C_r}$$

**Fórmula 7: Redundancia de Datos Relativa**

donde  $C_r$  se llama radio de compresión definido como en la Fórmula 8:

$$C_r = \frac{n_1}{n_2}$$

### Fórmula 8: Radio de Compresión

El caso particular  $n_2 = n_1$  implica que  $C_r = 1$  y  $R_d = 0$ , indicando la primera representación no tiene redundancia de información.

Si  $n_2 \ll n_1$  implica que  $C_r \rightarrow \infty$  y  $R_d \rightarrow 1$ , esto indica alta compresión mucha redundancia en los datos.

Si  $n_2 \gg n_1$  implica que  $C_r \rightarrow 0$  y  $R_d \rightarrow \infty$ , indicando que el segundo conjunto contiene muchos mas datos que la representación original.

La compresión de imágenes trata de identificar y explotar tres tipos de redundancias:

- Redundancia de Código
- Redundancia interpixels
- Redundancia psicovisual

## Redundancia de Código

Se puede utilizar el histograma para tener detalle de los códigos de información para poder reducir la cantidad de datos utilizados para representarlos.

Sea una variable aleatoria discreta  $r_k$  en el intervalo  $[0,1]$  representando los niveles de gris de una imagen y que  $r_k$  tiene probabilidad  $P_r(r_k)$ , Fórmula 9.

$$P_r(r_k) = \frac{n_k}{n} \quad k = 0,1,2,\dots,L-1$$

### Fórmula 9: Probabilidad del nivel de Gris en una Imagen

$L$  es el número de niveles de gris y  $n_k$  es el número de veces que el  $k$ -ésimo nivel de gris aparece en la imagen y  $n$  es el número total de pixels en la imagen.

Si el número de bits para representar cada valor de  $r_k$  es  $l(r_k)$ , el promedio de bits requeridos para representar cada pixel es el que muestra la Fórmula 10:

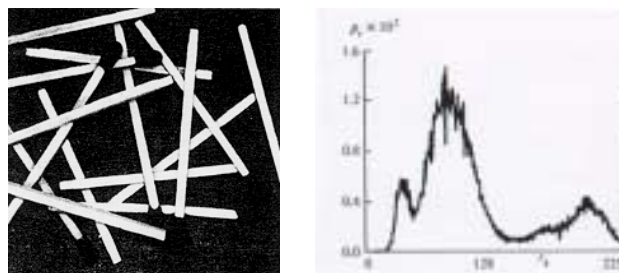
$$L_{avr} = \sum_{k=0}^{L-1} l(r_k) p_r(r_k)$$

**Fórmula 10: Promedio de bits**

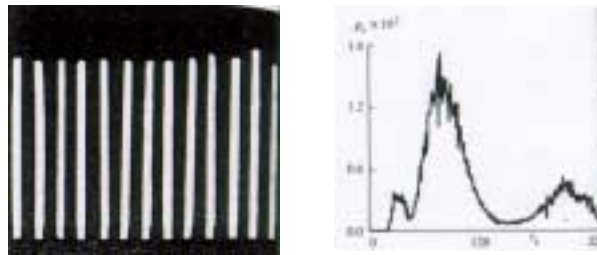
El total de bits requeridos para codificar una imagen de  $M \times N$  es  $M \times N \times L_{avr}$ . Se logrará compresión si se asignan menos bits a los niveles de grises más probables, obteniendo de esta manera compresión. Este proceso es llamado comúnmente como códigos de longitud variable.

## Redundancia interpixels

La Imagen 1 y la Imagen 2 tienen casi el mismo histograma. Estos histogramas indican la presencia de tres rangos predominantes de niveles de grises. Dado que los niveles de grises en estas imágenes no son igualmente probables, se podrían utilizar códigos de longitud variable para reducir la redundancia. Este proceso de codificación no altera la correlación entre los pixels. Los códigos utilizados para representar los niveles de grises no tienen nada que ver con la correlación de los pixels. Esta correlación resulta de las relaciones geométricas o estructurales entre los objetos de la imagen.

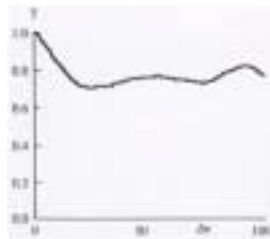


**Imagen 1: Fósforos 1**

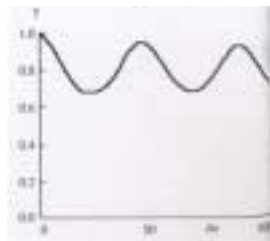


**Imagen 2: Fósforos 2**

La Función 1 y la Función 2 muestran los coeficientes de autocorrelación calculados a través de una línea de cada imagen.



**Función 1: Autocorrelación de Imagen 2**



**Función 2: Autocorrelación de Imagen 3**

Los coeficientes se calculan utilizando la Fórmula 11 de correlación discreta [González 1996] utilizando una versión normalizada como lo indica la Fórmula 12.

$$f_e(x) \circ g_e(x) = \frac{1}{M} \sum_{m=0}^{M-1} f_e^*(m) g_e(x+m)$$

**Fórmula 11: Correlación discreta**

$$\gamma(\Delta n) = \frac{A(\Delta n)}{A(0)}$$

$$\text{donde } A(\Delta n) = \frac{1}{N - \Delta n} \sum_{y=0}^{M-1-\Delta n} f(x, y)f(x, y + \Delta n)$$

### Fórmula 12: Correlación Discreta Normalizada

Esto muestra otra forma de redundancia de datos, relacionada directamente a la interrelación de los pixels en la imagen. El valor de cualquier pixel puede ser predecible desde el valor de sus vecinos, la información de cada pixel pequeña. La información visual de un solo pixel a una imagen es redundante dado que se calcula desde el valor de sus vecinos. Para reducir este tipo de redundancia en una imagen, el arreglo de dos dimensiones utilizado normalmente para la visualización e interpretación debe ser transformado en un formato no-visual más eficiente. Se puede utilizar la diferencia entre pixels adyacentes. Las transformaciones de este tipo (remueven la redundancia interpixel) se denominan *mapeos* (mappings), pueden llegar ser reversibles.

## Redundancia psicovisual

El brillo de una región percibido por el ojo depende de más factores que la luz reflejada por dicha región. Por ejemplo se ha visto que las variaciones de intensidad se perciben en áreas en donde realmente hay intensidad constante. Este tipo de fenómenos se deben a que el ojo no responde con igual sensibilidad a toda la información visual. En el sistema visual cierta información tiene menos importancia que otra. A este tipo de información se le llama redundancia psicovisual. Puede eliminarse sin producir impacto en la percepción de la calidad de la imagen.

La percepción de la información en una imagen normalmente no involucra análisis cuantitativo sobre cada pixel. En general el observador busca distinguir aspectos como bordes, regiones texturales y mentalmente las combina en grupos conocidos. La mente relaciona estos grupos con conocimiento anterior para completar el proceso de interpretación

Esta redundancia es distinta a la anterior dado que la psicovisual esta asociada a la información visual real o cuantificable.

La eliminación de la redundancia psicovisual resulta en la pérdida de información cuantitativa, es una operación no reversible y resulta en compresión con pérdida. La eliminación de esta redundancia se la llama *cuantization*.

## VII - Joint Photographics Expert Group (JPEG)

El objetivo de esta sección es presentar al lector una técnica de compresión con pérdida muy utilizada actualmente y en la cual se ha basado el autor para las experiencias de la tesis.

JPEG (Joint Photographic Experts Group) es uno de los estándares más ampliamente conocidos para la compresión de imágenes con pérdida. El mismo fue el resultado de la colaboración de la International Standards Organizations (ISO) y lo que fue CCITT (hoy ITU\_T), una división de las Naciones Unidas. Fue diseñado para comprimir imágenes, full color y en escala de grises, de escenas naturales del mundo real. Se obtienen buenos resultados al aplicarlo sobre fotografías, trabajos de arte y material similar, aunque no ocurre así con letras, dibujos simples o dibujos de líneas [Wallace 1991]. El JPEG maneja sólo imágenes quietas, pero hay un estándar relacionado llamado MPEG para imágenes en movimiento [Lindley 1995].

El JPEG tiene pérdida, es decir, la imagen descomprimida no es exactamente la misma que aquella con la que se empezó. Hay algoritmos para la compresión de imágenes sin pérdida, pero el JPEG logra una compresión mucho mayor de lo que es posible con los métodos sin pérdida.

El JPEG está diseñado para explotar las limitaciones conocidas del ojo humano, particularmente el hecho de que pequeñas variaciones de color se perciben con menos precisión que pequeñas variaciones en brillo o luminosidad. De esta manera, el JPEG está destinado a comprimir imágenes que serán miradas por seres humanos. Si se planea analizar estas imágenes con una máquina, los pequeños errores introducidos por el JPEG pueden ser un problema, aunque sean invisibles al ojo humano.

El JPEG estándar define tres sistemas de codificación diferentes:

- Un sistema de codificación baseline con pérdida, que utiliza como base la Transformada Discreta del Coseno (DCT).
- Un sistema de codificación extendido para aplicaciones con requerimientos de más precisión, de reconstrucción progresiva, etc.
- Un sistema de codificación independiente sin pérdida para compresión reversible.



Para ser compatible con JPEG un producto debe incluir soporte para el sistema baseline. El estándar propone una sintaxis que cualquier archivo o secuencia de bits debe cumplir para llamarse JPEG, dejando amplias libertades en las etapas de cuantificación y codificación, de manera que las personas encargadas del desarrollo puedan efectuar mejoras y optimizaciones. No se especifica ningún formato de imagen, resolución espacial o particularidades sobre el color.

Una propiedad útil del JPEG es que el grado de pérdida puede variarse ajustando los parámetros de compresión. Esto significa que el autor de la imagen puede negociar el tamaño del archivo contra la calidad de la imagen de salida. Se pueden obtener archivos *extremadamente* pequeños si no importa que la calidad sea pobre; esto es útil para aplicaciones como la indexación de archivos de imágenes. Inversamente, si la calidad de la salida en el seteo de compresión por defecto no es satisfactoria, se puede aumentar la calidad hasta el nivel deseado, y aceptar una compresión menor [Sayood 1996].

## ¿Por qué usar JPEG?

Hacer que los archivos de imágenes sean más chicos es una ganancia cuando se quiere transmitir archivos a través de redes o armar bibliotecas de imágenes. La posibilidad de comprimir un archivo color de 2 Mbyte a 100 Kbytes hace una gran diferencia en espacio de disco y en tiempo de transmisión. El JPEG puede proporcionar fácilmente una compresión de 20:1 de datos color. Si se compara GIF y JPEG, la relación de tamaño es normalmente de 4:1.

No obstante, el JPEG también tiene sus desventajas. Si un software de visualización no soporta JPEG directamente, tendrá que convertir el JPEG a otro formato para *ver* la imagen. Incluso con un visualizador capaz de ver JPEG, lleva más tiempo decodificar y ver una imagen JPEG que ver una imagen de un formato más simple como GIF. De esta manera, usar JPEG es esencialmente un trueque entre tiempo y espacio: se resigna algo de tiempo para poder almacenar o transmitir una imagen de forma más barata. Pero vale la pena notar que cuando hay una transmisión por red o telefónica, los ahorros de tiempo de transferir un archivo más chico pueden ser mayores que el tiempo necesario para descomprimir el archivo.

La segunda ventaja fundamental del JPEG es que almacena información color 24 bits/pixel (16 millones de colores). GIF, el otro formato de imágenes ampliamente usado en la red, sólo puede almacenar 8 bits/pixel (256 colores). GIF se ajusta razonablemente a pantallas baratas de computadora, la mayoría de las computadoras comunes no pueden mostrar más de 256 colores distintos a la vez. Pero el hardware full color se está abaratando constantemente, y las imágenes JPEG se ven *mucho*

mejor que las GIF en este tipo de hardware. Dentro de un tiempo el GIF probablemente parecerá tan obsoleto como parece hoy el formato MacPaint en blanco y negro. Más aun, el JPEG es mucho más útil que el GIF para intercambiar imágenes entre personas que tienen hardware distinto, porque evita que se prejuzgue cuántos colores usar. Por lo tanto el JPEG es considerablemente más apropiado que el GIF para usar como formato estándar por ejemplo en la WWW.

Otro punto que debe mencionarse es el temor generalizado de los usuarios ante el término *compresión con pérdida* [Cygansky 1998]. Se debe considerar que cuando hay que representar escenas del mundo real, *ninguna* imagen digital puede retener toda la información que llega al ojo. En comparación con la escena del mundo real, el JPEG pierde mucha menos información que el GIF. La verdadera desventaja de la compresión con pérdida es que si una imagen se comprime y descomprime repetidas veces, se pierde un poco de calidad cada vez. Esta es una objeción seria para algunas aplicaciones, pero no tiene la menor importancia para muchas otras.

## Descripción del algoritmo JPEG baseline

El sistema recomendado por JPEG es una modificación del esquema original propuesto por Chen y Pratt y se basa en la técnica de codificación utilizando la DCT. En esta sección se describe el algoritmo JPEG baseline, que se ilustra en cada una de sus partes con el bloque de 8x8 de la imagen LENA (Imagen 1) mostrada en la Tabla 1.

**Tabla 1: Bloque de 8x8 de LENA**

121	121	120	119	119	120	120	118
126	124	123	122	121	121	120	110
124	124	125	125	126	125	124	124
127	127	128	129	130	128	127	125
143	142	143	142	140	139	139	139
150	148	152	152	152	152	150	151
156	159	158	155	158	158	157	156
160	162	161	160	162	161	159	160

## ¿Cómo funciona el JPEG?

La compresión JPEG es un proceso en varias etapas. Igual que otras formas de compresión con pérdida se toman los datos de la imagen y se realiza un proceso regularizador que hace que contenga más patrones repetitivos que los que originalmente contenía, debido a esta repetitividad, en un paso posterior, se pueden aplicar un esquema de compresión sin pérdida, como la codificación de Huffman, utilizada por JPEG.

Las imágenes color primero se traducen a un espacio de color, como ser YUV o CIELAB, que almacena la información de brillo, o luminancia, separada de la información de color, o crominancia. Esto aprovecha el hecho de que el ojo humano es mucho más sensible a variaciones pequeñas de luminancia que a variaciones pequeñas de crominancia, en particular en el extremo azul del espectro.

También explota el hecho de que la mayoría de las imágenes contienen áreas bastante grandes en donde los valores de los píxeles adyacentes son muy similares en los canales de crominancia.

El siguiente paso, opcional, es hacer un submuestreo de los canales de crominancia. Esta es una de las dos partes del proceso en la que se pierden datos, y típicamente se hace sólo con seteos de JPEG de alta compresión y baja calidad. Básicamente, submuestreo significa que sistemáticamente se descarta información de color por filas o columnas de píxeles a una razón dada; si se descarta la información de color cada dos filas y cada dos columnas de píxeles, se reducen los datos de crominancia en un 75%. Cuando se descomprime la imagen, los valores de los píxeles faltantes se interpolan a partir de los que están presentes. En la mayoría de las imágenes el efecto no se nota, ya que el canal de luminancia contiene la mayor parte de la información visualmente importante. Puede llegar a observarse una leve pérdida de saturación, pero muy poco más cambia en forma visible.

Luego se registran los datos usando una función matemática llamada Transformada Discreta de Coseno, o DCT. Trabajando con bloques de 8x8 píxeles, la DCT analiza las frecuencias espaciales de la imagen tanto horizontal como verticalmente. La parte con pérdida del proceso es la cuantización, que es la forma en que la ingeniería se refiere a la división que reduce la cantidad de bits necesarios para expresar cada valor de frecuencia. Estos valores luego se comprimen sin pérdida, primero con RLE y luego con la codificación de Huffman. Esta es la parte del JPEG que hace que sean posibles altas razones de compresión.

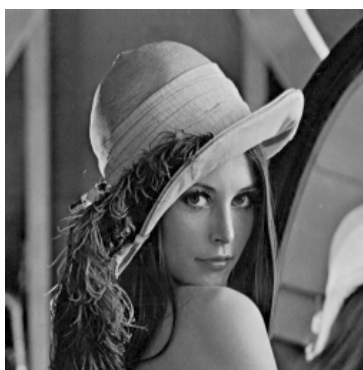
El especialista en JPEG define dos módulos diferentes “*back end*” para la salida final de los datos comprimidos: se permite la codificación Huffman o se permite la codificación aritmética. La

elección no afecta la calidad de la imagen, pero la codificación aritmética normalmente produce un archivo comprimido más pequeño. En imágenes típicas, la codificación aritmética produce un archivo un 5 a un 10 % menor que la codificación Huffman (todos los números de tamaños de archivos aquí citados corresponden a la codificación Huffman).

La variante particular de codificación aritmética especificada por el estándar JPEG está sujeta a patentes. De esta manera, no se puede usar legalmente la codificación aritmética del JPEG a menos que se obtengan las licencias de estas compañías. (La excepción uso experimental de la Ley de Patentes permite que la gente pruebe un método patentado en el contexto de la investigación científica, pero cualquier uso comercial o personal de rutina es una infracción).

La descompresión simplemente invierte este proceso. Los datos se expanden usando la decodificación de Huffman, los valores resultantes se multiplican, se aplica una DCT inversa, y los valores de luminancia y crominancia se traducen de vuelta al RGB.

Se verá con un ejemplo el proceso anteriormente mencionado para una imagen tradicional en la compresión de imágenes, en escala de grises, Imagen 3.



**Imagen 3: lena.bmp**

## Transformación

Como ya se dijo, la transformada utilizada por el esquema JPEG es la transformada DCT. Los valores de los píxeles de la imagen de entrada son reducidos en  $2^{(P-1)}$ , donde P es el número de bits utilizados para representar cada píxel. En nuestro caso se utilizan imágenes de 8 bits cuyos píxeles toman valores entre 0 y 255 de manera que al sustraerles 128, sus valores finales de entrada variarán entre -128 y 127.

El algoritmo comienza con el particionamiento de la imagen en bloques de 8x8 pixeles. Si alguna de las dimensiones de la imagen no es múltiplo de ocho, el codificador replica la última columna y/o fila hasta que el tamaño final sea múltiplo de 8. Esas filas y columnas adicionales se eliminan durante el proceso de decodificación. Luego cada bloque es transformado independientemente usando la DCT, dando como resultado otra matriz de 8x8, pero cuyos coeficientes varían en el rango de -1023 . . 1023.

Si se toma el bloque de 8x8 pixeles de la Tabla 1, se sustraen 128 de sus valores y se aplica la transformada DCT a la matriz resultante, se obtienen los coeficientes DCT de la Tabla 2. Aquí puede notarse que los coeficientes de frecuencia más baja cercanos al vértice superior izquierdo toman valores absolutos más grandes que los coeficientes de frecuencia más alta. Esto es lo que generalmente ocurre, aunque existen excepciones, como en los casos en que se tiene gran cantidad de variaciones en el bloque (por ejemplo bordes).

**Tabla 2: Coeficientes DCT del bloque de pixeles de LENA**

39.88	6.56	-2.24	1.22	-0.3	-1.08	0.79	1.13
-102.43	4.56	2.26	1.12	0.35	-0.63	-1.05	-0.48
37.77	1.31	1.77	0.25	-1.5	-2.21	-0.10	0.23
-5.67	2.24	-1.32	-0.81	1.41	0.22	-0.13	0.17
-3.37	-0.74	-1.75	0.77	1.99	-0.26	-1.30	0.76
5.98	-0.13	-0.45	-0.77	1.99	-0.26	1.46	0.00
3.97	5.52	2.39	-0.55	-0.05	-0.84	-0.52	-0.13
-3.43	0.51	-1.07	0.87	0.96	0.09	0.33	0.01

## Cuantificación

JPEG utiliza una cuantificación *midtread* uniforme para cuantificar los diversos coeficientes de un bloque transformado.

Las longitudes de los intervalos de cuantificación están organizadas en una tabla de cuantificación y pueden verse como una parte fija de esta etapa del proceso. En la Tabla 3 se muestra

un ejemplo de tabla de cuantificación recomendada por JPEG. Cada valor cuantificado se representa con una etiqueta. La etiqueta correspondiente al valor del coeficiente transformado  $\theta_{ij}$  es:

$$l_{ij} = \left[ \frac{\theta_{ij}}{Q_{ij}} + 0.5 \right]$$

donde  $\theta_{ij}$  es el elemento  $ij$  de la tabla de cuantificación y  $[x]$  es la parte entera de  $x$ . Como ejemplo si se considera el coeficiente  $\theta_{00}$  de la Tabla 2 cuyo valor es 39.88. De la Tabla 3 se tiene que  $Q_{00}$  es 16 y por lo tanto

$$l_{00} = \left[ \frac{39.88}{16} + 0.5 \right] = [2.9925] = 2$$

El valor reconstruido es obtenido del valor cuantificado multiplicado por la entrada correspondiente en la tabla de cuantificación. Para el caso anterior, el valor reconstruido del coeficiente es

$$l_{00} \times Q_{00} = 2 \times 16 = 32 \quad l_{00} = \left[ \frac{39.88}{16} + 0.5 \right] = [2.9925] = 2$$

**Tabla 3: Matriz de cuantificación por defecto de JPEG**

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

El error de cuantificación en este caso es  $39.88 - 32 = 7.88$ . Siguiendo los mismos pasos para los demás coeficientes se obtienen las etiquetas de la Tabla 4.

**Tabla 4: Cálculo de etiquetas**

2	1	0	0	0	0	0	0
-9	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

Por la Tabla 3 se puede ver que las longitudes de los intervalos de cuantificación se incrementan a medida que se baja desde el coeficiente DC (posición 0,0) hacia los coeficientes de alto orden. Debido a que el error de cuantificación es una función creciente de las longitudes de los intervalos, los errores más grandes de cuantificación ocurrirán en los coeficientes de frecuencia más alta. La decisión de cuál debe ser la longitud de los intervalos dependerá de cómo el sistema visual humano perciba los errores en esos coeficientes.

Coefficientes en posiciones diferentes tienen importancia perceptual diferente. Los errores de cuantificación en los coeficientes DC y en los AC (coeficientes de posiciones distintas de 0,0) de frecuencia baja se detectan con mayor facilidad que los errores en coeficientes de más alta frecuencia. Por lo tanto se podrían utilizar longitudes de intervalos más grandes para los coeficientes menos importantes.

Se logran distintos grados de pérdida multiplicando la matriz de la Tabla 3 por un escalar. La flexibilidad dada para esta etapa permite construir matrices de cuantificación adaptivas acordes a las características de la imagen que se está comprimiendo. Por ejemplo, en zonas donde se producen variaciones bruscas (como el pelo o los bordes del sombrero en la imagen LENA) podrían aparecer coeficientes de frecuencia alta con un peso importante a la hora de reconstruir la imagen y que podrían no ser tenidos en cuenta si se utilizara una matriz no adaptiva.

## Codificación

En este paso los coeficientes DC y AC resultantes de la cuantificación son codificados utilizando códigos de longitud variable, pero ambos en forma diferente.

Los coeficientes DC (coeficientes de la posición 0,0) se codifican en forma diferencial con respecto al mismo elemento del bloque previo. La razón de esto es que el elemento de esta posición es un múltiplo del valor promedio de los pixels del bloque y, que, en general, estos promedios no difieren substancialmente entre bloques vecinos. Es de esperar que esto ocurra en la mayoría de las situaciones.

Por lo tanto, para codificar las diferencias se asigna un código Huffman al conjunto de categorías en las que pueden caer dichos valores. Una diferencia en particular se codifica grabando primero la palabra de código correspondiente a su categoría y agregando luego bits extras al final de esa palabra de código para especificar su valor en una cantidad indicada por el número de categoría. Las categorías y sus correspondientes rangos de valores se muestran en la Tabla 6.

Si se generaliza para una categoría de diferencia (por ejemplo K) se utilizan K bits para codificar, ya sea K bits menos significativos de la diferencia positiva, o bien K bits menos significativos de la diferencia negativa menos 1.

La Tabla 5 muestra un código default provisto por JPEG donde la primer columna contiene categorías de diferencia, la segunda, la palabra de código para esa categoría y la tercera, la cantidad total de bits necesaria para codificar una diferencia.

**Tabla 5: Código por defecto de JPEG**

Categoría	Palabra de Código	Longitud
0	010	3
1	011	4
2	100	5
3	00	5
4	101	7
5	110	8
6	1110	10
7	11110	12
8	111110	14
9	1111110	16
A	11111110	18
B	111111110	20

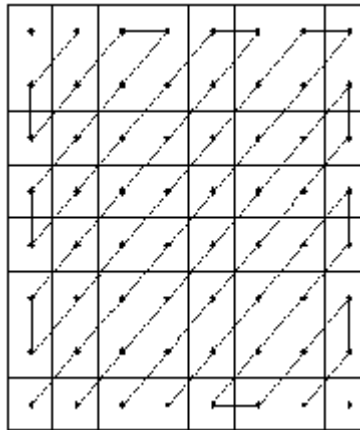


Dado que la categoría 0 contiene solamente un elemento, no se necesitan bits extras para especificar el valor. La categoría 1 contiene dos elementos de manera que sólo se necesita agregar 1 bit al final del código Huffman para especificar alguno de los dos elementos. De igual forma, se necesitan 2 bits para especificar un elemento en la categoría 2, 3 bits para la categoría 3 y  $n$  bits para la categoría  $n$ .

**Tabla 6: Categorías de coeficientes y sus respectivos rangos**

Categoría	Rango de Valores
0	0
1	-1 1
2	-3.....-2 2.....3
3	-7.....-4 4.....7
4	-15.....-8 8.....15
5	-31.....-16 16.....31
6	-63.....-32 32.....63
7	-127.....-64 64.....127
8	-255.....-128 128.....255
9	-511.....-256 256.....511
10	-1023.....-512 512.....1023
11	-2047.....-1024 1014.....2047
12	-4095.....-2048 2048.....4095
13	-8191.....-4096 4096.....8191
14	-16383.....-8192 8192.....16383
15	-32767.....-16384 16384.....32767
16	32768

El código binario para los coeficientes AC se genera en una forma diferente. La matriz transformada y cuantificada se reordena siguiendo un trayecto en zigzag (ver la Figura 9) para obtener una secuencia unidimensional de los coeficientes cuantificados, de manera que la secuencia resultante se encuentre dispuesta de acuerdo al crecimiento de las frecuencias. El mayor beneficio que resulta del nuevo reordenamiento es que se puede tomar ventaja de una mayor longitud en las corridas de ceros.



**Figura 9: Secuencia en zigzag para un bloque de 8x8 de coeficientes**

Concretamente, los coeficientes AC se codifican usando un código de longitud variable como el de la Tabla 7, donde se define  $C$  como la categoría en la que cayó un coeficiente y  $Z$  como el número de coeficientes con valor cero que ocurrieron en la secuencia después del último coeficiente distinto de cero. De acuerdo con la bibliografía, para bloques de 8x8 y píxeles de 8 bits,  $C$  puede tomar valores que varían de 0 a 10. Para los diferentes valores de  $Z$ , el JPEG establece en su tabla de codificación por defecto un valor máximo de 15 ceros consecutivos. De esta manera, cuando en la secuencia en zigzag se encuentra un coeficiente cuyo valor cae en una cierta categoría  $C$  precedido por  $Z$  ceros se graba el código correspondiente a la entrada  $Z/C$  en la tabla. Seguidamente se graba el valor del coeficiente en  $C$  bits de la misma forma en la que se graban las diferencias entre coeficientes DC.

**Tabla 7: Algunas palabras de código de la tabla de codificación por defecto**

Z/C	Palabra de Código	Longitud
0/0	1010(=EOB)	4
0/1	00	3
0/2	01	4
0/3	100	6
0/4	1011	8
0/5	11010	10
.....	.....	.....
F/0 (ZRL)	111111110111	12
F/1	1111111111110101	17
F/2	1111111111110110	18
F/3	1111111111110111	19
F/4	1111111111111000	20
F/5	1111111111111001	21
.....	.....	.....

Existen dos palabras de código especiales llamadas EOB, utilizadas para indicar que el resto de los coeficientes en la secuencia de zigzag son ceros, y ZRL para el caso en que el número de ceros consecutivos en la secuencia exceda 15.

Como se ve, al igual que en la cuantificación, también se proveen tablas de codificación por defecto aunque el usuario puede construir las propias de manera de que se puedan adaptar a las características de la imagen que se está comprimiendo.

## VIII - Clases de Mediciones

### Criterio de Fidelidad

La experiencia muestra que cada esquema de codificación esta sujeto a un conjunto de defectos que son dificultosos de caracterizar. Esto se debe a la forma en que se diseñan los mecanismos de codificación para modificar la precisión de la representación de los datos, esto implica directamente a que el usuario puede ser insensible a errores en algunas partes de la imagen, pero no serlo en otras.

La eliminación de la redundancia resulta en la eliminación de información visual. De acuerdo a este proceso la información de interés puede perderse y es interesante cuantificar la información perdida.

La calidad en sí es un atributo con muchas definiciones e interpretaciones posibles, dependiendo del uso que se le dará a las imágenes. Una *buena* imagen procesada puede ser una que sea perceptualmente agradable o útil en una aplicación específica. Existen dos criterios básicos

- Mediciones fidelidad objetiva.
- Fidelidad subjetiva medida por tests o cuestionarios psicológicos con evaluaciones numéricas.

### Medidas Objetivas de la Calidad

Algo natural para hacer si interesa la fidelidad de una secuencia reconstruida es observar en las diferencias entre los valores originales y los reconstruidos; en otras palabras la distorsión introducida en el proceso de compresión. Si  $f(x, y)$  es una imagen y  $\hat{w}(x, y)$  denota una aproximación de  $f(x, y)$  obtenida por algún método de compresión aplicado a  $f(x, y)$  y descomprimida. Para cualquier valor de  $x$  e  $y$  se puede denotar el error  $e(x, y)$  calculado como indica la Fórmula 13.

$$e(x, y) = \hat{w}(x, y) - f(x, y)$$

**Fórmula 13: error entre f y en (x,y)**

De acuerdo a lo anterior, el error total entre las dos imágenes esta dado por el termino definido en la Fórmula 14 siguiente, llamado *error medio*.

$$em = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \left[ \hat{w}(x, y) - f(x, y) \right]$$

**Fórmula 14: Error Total**

Las imágenes tienen ambas tamaño de  $M \times N$ . El *error cuadrático medio* ( $e_{rms}$ ) entre las imágenes  $f(x, y)$  y  $\hat{w}(x, y)$  está definido por la raíz cuadrada promediada sobre  $M \times N$  como indica la Fórmula 15..

$$e_{rms} = \sqrt{\left[ \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{w}(x, y) - f(x, y)]^2 \right]}$$

**Fórmula 15: Error cuadrático medio**

Un criterio de fidelidad muy utilizado es la relación *signal-to-noise* (*SNR*). *SNR* es la relación entre el ruido y la información útil de la imagen (señal) y se suele expresar en dB. Si se considera que  $\hat{w}(x, y)$  es la suma de  $f(x, y)$  más cierto error  $e(x, y)$ , se puede calcular  $SNR_{ms}$  de acuerdo a la fórmula. También se puede calcular el  $SNR_{rms}$  calculando la raíz cuadrada de la Fórmula 16.

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

**Fórmula 16:  $SNR_{ms}$**

Otra medida objetiva que puede calcularse es la entropía. La entropía es una medida de la variabilidad de los niveles de gris, independiente de su distribución espacial. La probabilidad del nivel  $k$  en un histograma se puede calcular como lo indica a siguiente Fórmula 17 y Fórmula 18.

$$p_r(r_k) = \frac{h(r_k)}{\sum_{k=0}^{L-1} h(r_k)} \quad \text{para } k = 0 \dots L-1$$

**Fórmula 17: Probabilidad del nivel k**

$$\log_2 \frac{1}{p(r_k)}$$

**Fórmula 18: Información del nivel k**

Y la entropía como indica la Fórmula 19.

$$H = -\sum_{k=0}^{L-1} p(r_k) \log_2 p(r_k)$$

**Fórmula 19: Entropía**

## Mediciones Subjetivas de la calidad

Se puede evaluar el *error cuadrático medio* (*mse*) como una medición de la similitud percibida entre un original u una imagen procesada, comparando los valores del *mse* con juicios de similitud. Los resultados obtenidos mediante investigaciones, indican que el *mse* no es válido como criterio de calidad para imágenes procesadas para el sistema de la visión humana, y que se pueden derivar mejores mediciones pesando el error de acuerdo a las propiedades supuestas del sistema visual humano [Marmolin 1986].

La calidad subjetiva de una imagen reconstruida puede verificarse de varias formas. Un conjunto de imágenes aleatorias puede presentarse a los expertos o usuarios típicos quienes las califican de acuerdo a una escala preestablecida.

Luego se puede realizar un análisis estadístico para sobresaltar promedios, variaciones y otras tendencias en los datos.

Son comunes los tests formales subjetivos en el área de speech y compresión de audio. Uno de ellos es el Mean Opinion Score (MOS) y el Diagnostic Acceptability Measure (DAM) [Quackenbush 1988].

Existen considerables variaciones en cuanto a la escala numérica de las respuestas, en relación de cuando proveer o no de frases descriptivas para cada número y en el intento de que, el rango subjetivo, indique la utilidad de la imagen en una aplicación específica. En general se utiliza una clasificación en donde el usuario marca de acuerdo al grado de similitud o no similitud entre dos imágenes, una original y otra la procesada (en este caso comprimida).

Otros experimentos se realizan en tres etapas. La primera de ellas es la *presentación*, en donde se le muestra al usuario un conjunto de las imágenes incluyendo las peores y las mejores imágenes del conjunto completo. En la segunda etapa, la *evaluación*, se les presentan al usuario las imágenes en forma aleatoria y que las califique de acuerdo a la tabla seleccionada. Puede revisar las imágenes tantas veces, o el tiempo que guste. Luego de la evaluación, se le presentan todas las imágenes ordenadas en forma ascendente según la tabla. Esto le da la oportunidad de comparar imágenes sucesivas y hacer una revisión de sus respuestas, esta última etapa es la *revisión*.

La escala no está estandarizada, es común encontrar tabulaciones de 0 a 10. Es claro que mientras más categorías se permitan, se puede realizar una precisión más fina de la calidad. Sin embargo de acuerdo a la experiencia, los usuarios se adecuan a varios tipos de escalas como muestran la Tabla 8, Tabla 9 y Tabla 10

<b>Ranking</b>	<b>Descripción</b>
10-9	Muy Buena
8-7	Buena
6-5-4	Clara
3-2	Mala
1-0	Muy Mala

**Tabla 8: Escala de calidad**

<b>Ranking</b>	<b>Descripción</b>
5	Deterioro muy molesto
4	Deterioro molesto
3	Deterioro poco molesto
2	Deterioro perceptible pero no molesto
1	Imperceptible

**Tabla 9: Otra posible escala**

<b>Valor</b>	<b>Palabra Clave</b>	<b>Descripción</b>
1	Excelente	Imagen de muy alta calidad
2	Correcta	Imagen de alta calidad. La interferencia no es objetable
3	Pasable	Imagen de calidad aceptable, La interferencia no es objetable
4	Marginal	Imagen de pobre calidad. La interferencia es objetable
5	Inferior	Imagen de muy pobre calidad. La interferencia esta claramente presente
6	Inservible	Imagen muy mala. El usuario no quiere verla

**Tabla 10: Escala para imágenes televisivas**



## IX - Experiencia Desarrollada

### Resumen

Se selecciono un conjunto de imágenes a las cuales se les aplico estándar JPEG. Sobre estas imágenes se hicieron cálculos objetivos, como relación señal-ruido, error cuadrático medio, entropía. Se selecciono dos grupos de sujetos evaluadores, uno experto en el tratamiento de imágenes, con trabajos y publicaciones relacionadas al tema y otro inexpertos, sin experiencia previa al trabajo con imágenes.

Con esta información se trato de correlacionar con los promedios obtenidos de la evaluación subjetiva para poder predecir calidad subjetiva a partir de alguna medida objetiva. Es importante tratar de determinar si esta predicción cambia entre los expertos e inexpertos o entre los diferentes tipos de imágenes.

Los resultados aquí obtenidos se deberían verificar definiendo otra experiencia más amplia tanto en personas como imágenes.

### Objetivos

La experiencia planteada tiene como objetivo obtener conclusiones a partir del análisis estadístico sobre 3 aspectos fundamentales:

- 1- La correlación existente entre las medidas objetivas y los resultados de la evaluación subjetiva de expertos e inexpertos.
- 2- Verificar si dentro de las diferentes clases de imágenes hay diferencias en la apreciación subjetiva.
- 3- Estudiar la incidencia de la categoría *experto* y de la categoría *inexperto* en la apreciación subjetiva de la calidad.

El primer objetivo esta orientado a verificar la correlación si es que existe, entre alguna medida objetiva y la evaluación subjetiva de los expertos. También interesa verificar esta relación con la evaluación subjetiva de los inexpertos.

El segundo objetivo esta orientado a verificar la variación asignada de calidad entre las diferentes imágenes dentro de una misma clase. Diferentes imágenes significan imágenes de la misma clase comprimida con diferentes radios de compresión. Es interesante ver la variación de la calificación de calidad subjetiva.

El tercer objetivo esta orientado a estudiar la incidencia de la experiencia en la evaluación y comparar estos resultados con los de los inexpertos.

## **Imágenes utilizadas para la experiencia**

Las clases de estas imágenes corresponden a dominios distintos, los cuales son independientes entre sí. Esto indica que no hay aspectos comunes (recuerdos) mientras se realizar la experiencia.

La primer clase corresponde a caras. Son imágenes de caras sobre diferentes fondos. En general son imágenes con mucha luminosidad y detalles propios de los perfiles humanos. Los diferentes fondos se requieren para ejercitar el algoritmo de codificación.

La segunda clase corresponde a fórmulas del tipo de circuitos electrónicos y diagramas de experimentos físicos, en donde hay un fondo oscuro y los diagramas esta hechos en blanco. Básicamente son imágenes en blanco y negro en donde los algoritmos de compresión con pérdida muestra sus aspectos negativos relacionados a la codificación, cuando se aumenta el radio de compresión.

La última clase son imágenes de paisajes, con mucho detalle y luminosidad. Se percibe claramente las diferentes texturas en dos bandas (arbustos-roca, agua bosque, cielo-arena).

Estas imágenes fueron seleccionadas debido a la dificultad explícita que tienen para mostrar el error de *cuadrículado* que tiene los algoritmos de compresión con pérdida. Obviamente ninguna de las seleccionadas poseen características de *engaño* como los presentados en la sección “*Modelo Biológico de la Visión Humana*”. El número de imágenes utilizadas esta relacionado a la psicología del número  $7\pm 2$  [Brooks 1995] el cual indica un rango numérico adecuado para el manejo de la complejidad.

En el CD que acompaña al trabajo se encuentra la aplicación para instalar, los datos recolectados, además el conjunto de imágenes. A en la página siguiente se ven las originales.



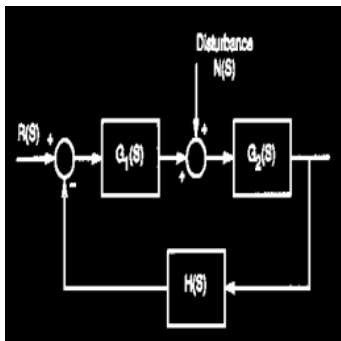
**Imagen 4: Cara1**



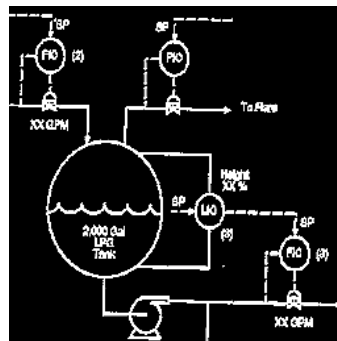
**Imagen 5: Cara2**



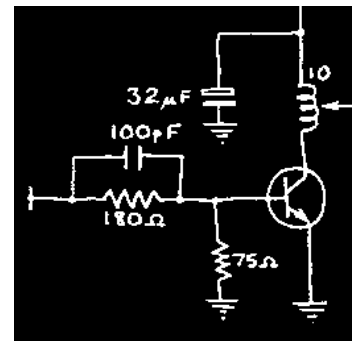
**Imagen 6: Cara3**



**Imagen 7: Fòrmula1**



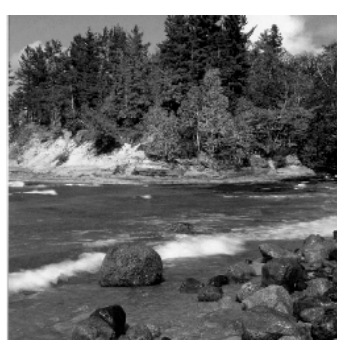
**Imagen 8: Fòrmula2**



**Imagen 9: Fòrmula3**



**Imagen 10: Paisaje1**



**Imagen 12: Paisaje2**



**Imagen 11 Paisaje3**

## Evaluación Objetiva

Se desarrollaron diferentes aplicaciones para calcular valores de apreciaciones objetivas.

Los cálculos seleccionados se refieren al cálculo de la entropía, el error cuadrático medio, la relación señal ruido.

Estas aplicaciones por lo general toman como parámetro dos imágenes, la original y la comprimida a evaluar. Estas aplicaciones se corren manualmente y se cargan en la tabla los resultados objetivos. En la sección X se describe esta aplicación.

## Evaluación Subjetiva

La experiencia desarrollada corresponde a una evaluación subjetiva en donde existe un conjunto de imágenes que se presentan a usuarios expertos e inexpertos quienes la califican en una escala. Esta escala definida de acuerdo al deterioro observado por el sujeto se define como:

1. Deterioro muy molesto.
2. Deterioro Molesto.
3. Poco molesto.
4. Perceptible pero no molesto.
5. Imperceptible.

Las variaciones del rango numérico utilizado para calificar la respuesta esta definida en [Quackenbush 1988]. Este puntaje permite la asignación de un rango que va desde 1 a 5, indicando el grado del deterioro apreciado por el usuario, varia desde 5 donde el deterioro es imperceptible, hasta 1 en el cual indica un deterioro de la calidad de la imagen muy molesta a la percepción visual.

Luego con los resultados de esta evaluación se realiza un estudio estadístico para obtener conclusiones sobre los 3 objetivos indicados precedentemente.

Para soportar esta evaluación subjetiva se desarrollo un Sistema de Evaluación Subjetiva de Imágenes, en donde el usuario tiene la posibilidad de navegar entre las imágenes a evaluar y categorizarlas de acuerdo a la tabla. En la sección X se describe esta aplicación.

Los datos han sido procesados con el paquete estadístico System for Statistics (SYSTAT).

## Desarrollo de la Evaluación

En el desarrollo de esta experiencia se consideraron tres clases de imágenes correspondiente a caras, fórmulas y paisajes. A su vez en cada clase hay tres imágenes base las cuales fueron comprimidas con el algoritmo JPEG con Particionamiento fijo de 8x8 [Wallace 1991] [Russo 1998] utilizando un factor de threshold correspondiente al nombre del archivo con la relación, si el nombre termina con D1 la relación es 8:1, D2 la relación es 8:2, D3 la relación es 8:3 y para D4 la relación es 8:4.

Los radios de compresión no son exactamente iguales debido a que se mantuvo una relación constante entre los radios de las diferentes imágenes. Se intento que la degradación entre las imágenes fuese lo mas constante posible, evitando variaciones abruptas que el evaluador pudiese detectar rápidamente.

La Tabla 11 que se muestra a continuación muestra los datos de compresión por cada clase, dando el tamaño de la original, columna *tamaño*, como así también el de las imágenes comprimidas y los radios de compresión alcanzados, columna *radio*. Las imágenes son archivos BMP y los nombres físicos de las imágenes comprimidas se construye concatenando el nombre de la imagen con el string que se encuentra en la columna *comp*.

Clases											
Caras				Fórmulas				Paisajes			
Imagen	Comp.	Tamaño	Radio	Imagen	Comp.	Tamaño	Radio	Imagen	Comp.	Tamaño	Radio
<b>CARA1</b>		66614		<b>FORM1</b>		66614		<b>PAIS1</b>		66614	
	_D1	4482	14.86		_D1	5213	12.78		_D1	8636	7.71
	_D2	3068	21.71		_D3	3353	19.87		_D3	4056	16.42
	_D4	2169	30.71		_D4	2975	22.39		_D4	3240	20.56
<b>CARA2</b>		66614		<b>FORM2</b>		66614		<b>PAIS2</b>		66614	
	_D1	5771	11.54		_D1	9528	7		_D1	7201	9.25
	_D2	3839	17.35		_D3	5402	12.33		_D3	3269	20.38
	_D4	2553	26.09		_D4	4642	14.35		_D4	2619	25.43
<b>CARA3</b>		66614		<b>FORM3</b>		66614		<b>PAIS3</b>		66614	
	_D1	4813	13.84		_D1	6863	9.71		_D1	10884	6.12
	_D2	3225	20.65		_D3	4000	16.65		_D3	5110	13.03
	_D4	2246	29.66		_D4	3505	19.00		_D4	4064	16.39

**Tabla 11: Tamaños y Radios de las Imágenes**

Las imágenes originales (CARA1, CARA2, CARA3, FORM1, FORM2, FORM3, PAIS1, PAIS2 y PAIS3) tiene le mismo tamaño debido a que son imágenes con el mismo alto y ancho.

Queda claro que las imágenes reducen su tamaño a medida que crece el radio de compresión. El tamaño de las comprimidas es entre el 14% y el 4% del tamaño de la imagen original. Para las CARAS esto va desde el 8% al 3%. Para las FORMULAS esta reducción va desde el 14% al 4% y para los PAISAJES va desde el 12% al 4%.

## **X - Sistema de Evaluación Subjetiva de Imágenes**

En esta sección se describe el software desarrollado para poder cumplir con los objetivos del trabajo. Esto incluye también la descripción de la herramienta estadística utilizada. En el CD que acompaña a la tesis se encuentra la aplicación para su instalación.

### **Software Desarrollado**

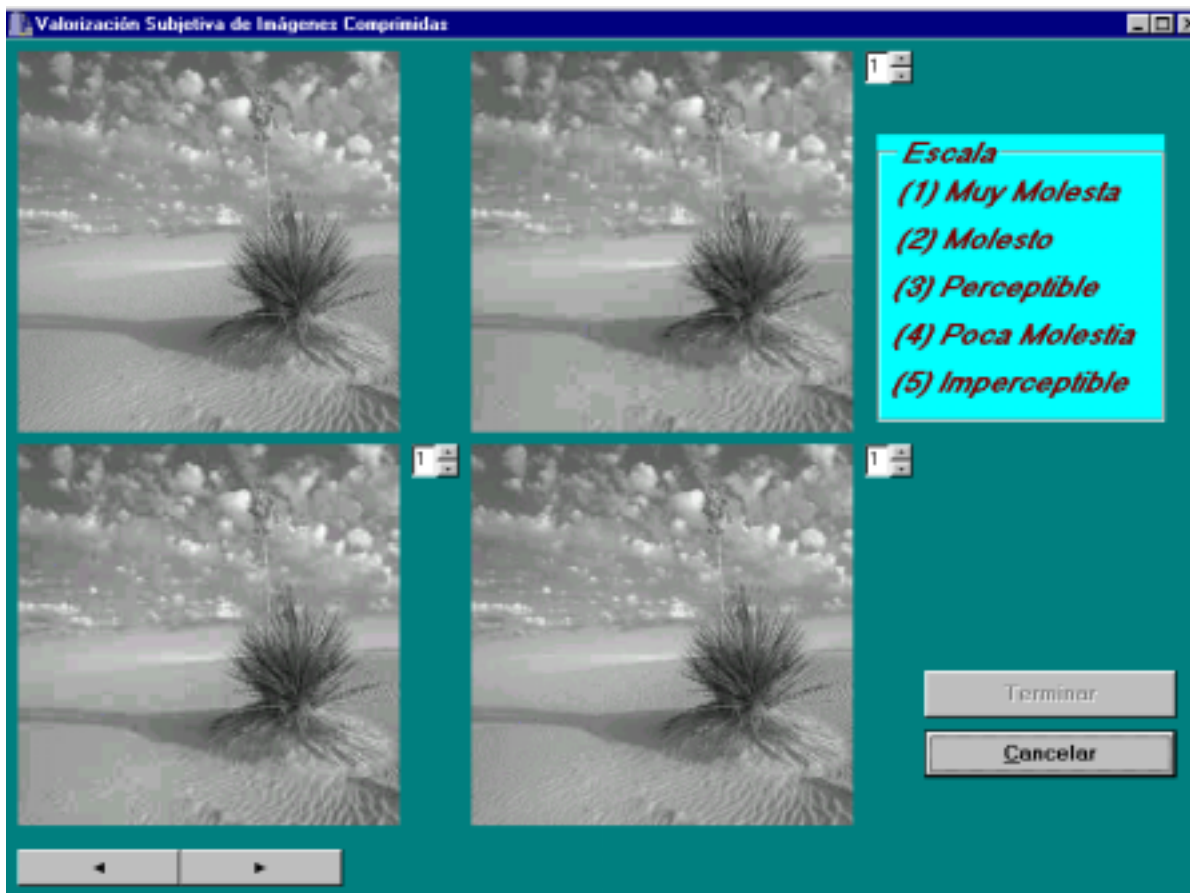
Se desarrollo en C++ Builder versión 4.0 para la construcción de un Sistema de Evaluación de Imágenes. Este sistema consta de dos módulos integrados. El primero de ellos permitió la recolección de los datos de la evaluación subjetiva, este modulo se llama *valorización* y el otro para la presentación de diferentes reportes orientados al análisis estadístico, llamado *evaluación*. A continuación se detalla cada uno de ellos en particular. Las otras aplicaciones estaban relacionadas al cálculo de diferentes medidas objetivas, estas aplicaciones en su conjunto se llaman *aplicaciones de soporte*.

### ***Valorización***

Esta aplicación le permite a los evaluadores a calificar el conjunto de imágenes seleccionadas para el experimento. Cuando arranca el sistema se le presenta al usuario la Pantalla 1.

Se le presentan al usuarios las imágenes originales con las correspondientes imágenes comprimidas. La imagen original se presenta arriba a la izquierda y es la que no tiene asociado una lista de selección de la clasificación.

Las otras tres imágenes son comprimidas con diferentes radios ubicadas en la pantalla en forma aleatoria para tener independencia posicional entre imágenes y el radio de compresión.



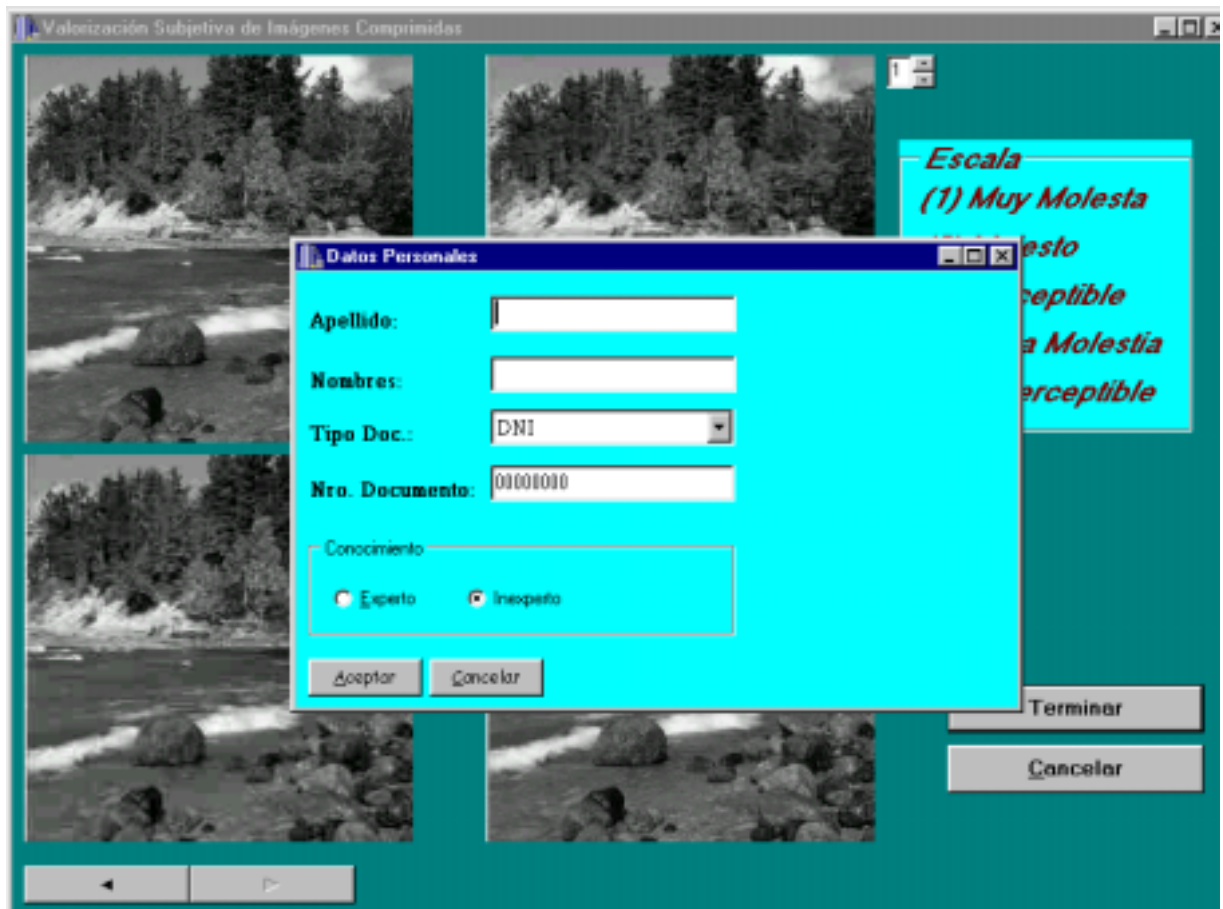
### Pantalla 1: Valorización

También en la esquina superior derecha se muestra la tabla de clasificación utilizada en este experimento. El usuario puede navegar y reclasificar todas las veces que quiera hasta, cuando hace por lo menos una recorrida de todas las imágenes, se habilita el botón *Terminar* con el cual el usuario confirma su evaluación ingresando los datos que se muestra a continuación para registrarlos.

Una vez que el usuario asigna un valor de calidad a la imagen, no tiene efecto el algoritmo de posicionamiento aleatorio.

Cuando el sujeto evalúa todas las imágenes presentadas aparece la Pantalla 2, en la cual se ingresan atributos relacionados a evaluador para guardarse en una tabla de la base de datos para su posterior utilización.



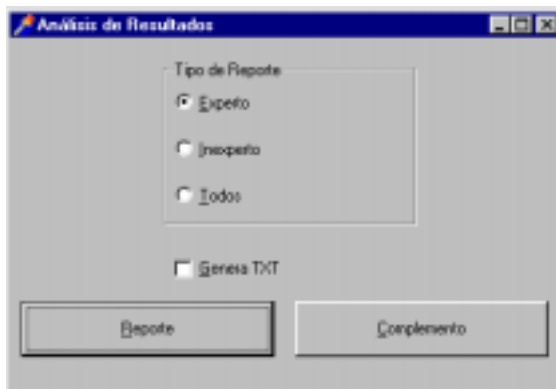


**Pantalla 2: Fin de la Evaluación**

## *Evaluación*

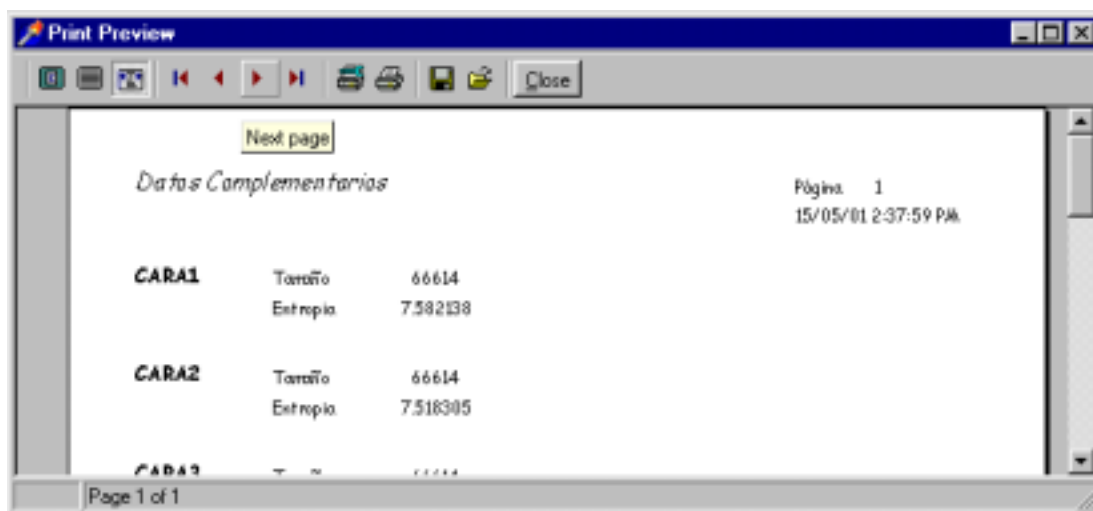
Esta aplicación permite hacer distintos reportes de los datos obtenidos a través de la valorización, ver Pantalla 3. Estos se presentan en forma de reportes, pudiéndose generar archivos de texto para su posterior procesamiento, en nuestro caso por la aplicación SYSTAT.

Permite seleccionar el perfil del usuario (experto, inexperto o ambos) y si el usuario lo desea, generar un archivo de texto para su posterior procesamiento.



**Pantalla 3: Pantalla de Evaluación**

En los párrafos siguientes se describirá la funcionalidad de este formulario. La funcionalidad del botón *complemento* no da información de las imágenes utilizadas en la experiencia como se muestra en la Pantalla 4. Esta información está relacionada a valores de cálculo objetivos como el tamaño y entropía de las imágenes base.



**Pantalla 4: Reporte de Datos Complementarios**

La funcionalidad del botón *Reporte* permite generar de acuerdo a los datos filtrados un reporte como el siguiente en donde para cada imagen calificada se dan los datos de las calificaciones subjetivas además de la información de los cálculos objetivos como relación señal ruido, error cuadrático medio y la entropía, también indica por imagen el total de sujetos que calificaron y el promedio de la calificación. Esto se ve en la Pantalla 5.

**Resultado Evaluación Subjetiva**  
 Insuperior  
 Página 1  
 02/07/01 11:14:01

**CARAI\_D1**

Índice 14,84    Entropía 7,582    MSE 5,506

SNR 35,851

Acuña	Daniel Ruben	4
Barbieri	Andres	4
Bertone	Rodolfo	4
Blesa	María Fernanda	4
Boracchini	Marcos	5
Costas	Marcela	3
Di Gusti	Laura Cristina	3
Galimberti	Carla	3
Garza	Gledys	5
Matton	Vaná	5
Pizarro	Ms. Alejandra	4
Rai	Champrondón	4
Romero	Fernando	4
Rusconi	Sebastián Mariano	4
Samsano	Narciso	5
Schwaers	Robert Jochen	4
Tambo	Diego Fernando	4
Thomas	Pablo Javier	4

**Pantalla 5: Reporte de Calificación**

La opción de generación del archivo TXT permite tener interfaces de este software con otros de análisis de tendencia, por ejemplo el SYSTAT.

### ***Aplicaciones de soporte***

Se desarrollaron un conjunto de programas de soporte que básicamente tenían como parámetro una o dos imágenes y retornaban resultados que se referían a medidas objetivas.

Todos estos programas son en modo consola (no hay interface con los usuarios) y se hicieron para calcular:

- Resta.cpp: Mide la distancia pixel a pixel entre dos imágenes.
- Entropia.cpp: Retorna el valor de la entropía de una imagen.

- Medida.cpp: Realiza el cálculo del Raíz del Error Cuadrático, la relación Signal-to-Noise Ratio (SNR).

## **Aplicación Estadística SYSTAT**

SYSTAT presenta una estructura básica compuesta por 17 módulo que realizan tres grandes grupos de operaciones [ATO 1994].

- Operaciones dedicadas a la gestión de archivos
- Operaciones dedicadas al análisis estadístico.
- Operaciones dedicadas al análisis y visualización gráfica.

La información se debe ingresar en forma de dataset en forma matricial. En este caso toma como entrada los archivos generados desde los reportes.

Antes de introducir datos se debe asignar nombres a las variables (columnas), la única restricción en esta aplicación es que las alfanuméricas deben tener obligatoriamente un \$ en el nombre.

Las técnicas gráficas son un instrumento estadístico sirven para estudiar la estructura de los datos. Persiguen básicamente dos propósitos

- Estudiar la estructura de los datos
- Comunicar resultados a otros usuarios

El módulo SYGRAPH de SYSTAT ofrece una cantidad muy completa de métodos gráficos, desde las representaciones muy sencillas de univariantes a los gráficos multivariantes.

Esta aplicación se utilizo para el análisis de varianza (ANOVA), analizando la estructura de los datos, también se utilizo SYGRAPH para la presentación de resultados como los que se muestra en la sección *XI – Presentación de Resultados*.

## XI - Presentación de los Resultados

En esta sección se presentan los resultados y conclusiones obtenidas tanto de las evaluaciones objetivas como las subjetivas. Las presentaciones se realizan en forma separada para después reunir las conclusiones.

### Resultados Objetivos

Estos resultados presentan la información de la entropía (columna Ent) de cada una de las imágenes base de las clases y las diferentes compresiones, para cada tipo de imagen comprimida se calcula el error cuadrático medio (columna ECM), la relación señal ruido (columna SNR). En las tablas a continuación se presentan los resultados separados por clase

#### *Clase Caras*

	ENT	ECM	SNR		ENT	ECM	SNR		ENT	ECM	SNR
CARA1	7.58			CARA2	7.51			CARA3	7.45		
_D1	7.53	5.50	35.85	_D1	7.52	8.86	31.76	_D1	7.37	7.47	35.26
_D2	7.26	11.52	32.64	_D2	7.51	17.77	28.74	_D2	7.08	14.80	32.29
_D4	6.91	26.38	29.04	_D4	7.33	38.44	25.39	_D4	6.62	31.63	28.99

**Tabla 12: Valores Objetivos Caras**

Los valores de la **Tabla 12** indican que las tres medidas objetivas reflejan el aumento del radio de compresión. Estos valores corresponden las caras.

La columna de la entropía (*ent*) de la **Tabla 12** va en disminución. Cada vez que se comprime, se eliminan niveles de grises. La diferencia absoluta indica una variación de 0.14 Esto indica que no hay variaciones bruscas entre las imágenes comprimidas, esta característica fue buscada para que la evaluación subjetiva no fuese obvia.

La columna del error cuadrático medio (*ECM*) de la **Tabla 12** como es de esperar va en aumento, siendo este brusco debido a que calcula el error (o distancia) desde la imagen original a la comprimida que se analiza. Mientras más compresión se tiene, mas error introducimos.

La columna del error cuadrático medio (*SNR*) de la **Tabla 12** decrece debido a que el denominador de la fórmula 10, crece a medida que aumenta el radio de compresión.

### *Clase Formulas*

	ENT	ECM	SNR		ENT	ECM	SNR		ENT	ECM	SNR
FORM1	1.02			FORM2	0.57			FORM3	0.34		
_D1	1.79	31.63	28.99	_D1	1.99	55.89	20.15	_D1	1.44	37.18	20.54
_D3	2.25	68.25	17.55	_D3	2.69	283.3	13.10	_D3	1.99	185.1	13.56
_D4	1.93	89.69	16.36	_D4	2.27	359.5	12.07	_D4	1.62	226.1	12.70

**Tabla 13: Valores Objetivos Fórmulas**

Los valores de la Tabla 13 indican que las tres medidas objetivas reflejan el aumento del radio de compresión. Estos valores corresponden las formulas.

En este caso la entropía se comporta de modo extraño pero justificable. Se debe recordar que las imágenes de las fórmulas están en blanco y negro. En la primera compresión, la entropía crece con respecto a la imagen original. Esto se debe a que cuando se comprimen en los bordes, nítidos, aparecen nuevos niveles de grises de acuerdo al programa de compresión. Luego de esto, la entropía se comporta en forma decreciente.

Al igual que en caras, la columna del error cuadrático medio (*ECM*) de la Tabla 13 como es de esperar va en aumento, siendo este brusco debido a que calcula el error (o distancia) desde la imagen original a la comprimida que se analiza. Mientras más compresión se tiene, más error se introduce. Es interesante notar el valor absoluto de las diferencias.

Al igual que en caras, la columna del error cuadrático medio (*SNR*) de la Tabla 13 decrece debido a que el denominador de la fórmula 10, crece a medida que aumenta el radio de compresión.

## Clase Paisajes

	ENT	ECM	SNR		ENT	ECM	SNR		ENT	ECM	SNR
<b>PAIS1</b>	6.96			<b>PAIS 2</b>	6.93			<b>PAIS 3</b>	7.49		
<b>_D1</b>	6.96	50.74	20.38	<b>_D1</b>	6.91	39.68	20.39	<b>_D1</b>	7.50	98.27	20.47
<b>_D3</b>	6.95	103.9	17.27	<b>_D3</b>	6.77	78.65	24.79	<b>_D3</b>	7.45	185.8	17.70
<b>_D4</b>	6.90	125.7	16.44	<b>_D4</b>	6.61	95.11	23.97	<b>_D4</b>	7.39	216.3	17.04

**Tabla 14: Valores Objetivos Paisajes**

Los valores de la Tabla 14 indican que las tres medidas objetivas reflejan el aumento del radio de compresión. Estos valores corresponden las caras.

La columna de la entropía (*ent*) de la Tabla 14 va en general en lenta disminución. Cada vez que se comprime, se eliminan niveles de grises. La lenta disminución indica poca eliminación de niveles. Esto se debe al detalle de iluminación que tienen las diferentes imágenes de paisajes. El histograma esta totalmente ecualizado. La diferencia absoluta indica una variación de 0.04. Esto indica que no hay variaciones bruscas entre las imágenes comprimidas, esta característica fue buscada para que la evaluación subjetiva no fuese obvia.

Al igual que en caras, la columna del error cuadrático medio (*ECM*) de la Tabla 14 como es de esperar va en aumento, siendo este brusco debido a que calcula el error (o distancia) desde la imagen original a la comprimida que se analiza. Mientras más compresión se tiene, mas error se introduce.

Al igual que en caras, la columna del error cuadrático medio (*SNR*) de la Tabla 14 decrece debido a que el denominador de la formula 10, crece a medida que aumenta el radio de compresión.

## Resultados Subjetivos

Se realizo la evaluación estadística de los promedios de la variable de salida de interés llamada *calidad de imagen*, evaluado sobre el total de sujetos (expertos e inexpertos). La utilización del promedio fue seleccionada a los fines de normalizar la variable de respuesta y así poder utilizar, respetando los presupuestos involucrados, análisis de varianza mediante un modelo lineal [Montgomery 1991].

El ANOVA es una prueba de comparación entre grupos pero no es a través de la media y su desviación estándar, sino a través de comparar la variabilidad existente en cada variable exploratoria (variable de entrada) con respecto a la variabilidad del error. De este modo si la variabilidad de una variable de exploración es consistentemente mayor que la del error. Se puede concluir que esa variable tiene efectos significativos en la o las variables de respuesta.

El método de análisis de varianza para comparación de promedios parte del supuesto inicial de que no existe diferencia entre los promedios y que los resultados de la muestra son producto exclusivamente del azar. A este supuesto inicial se le conoce como la hipótesis nula y se le designa con  $H_0$ . Un análisis de varianza es un test de hipótesis, por lo tanto se plantean, como en cualquier test de hipótesis, dos hipótesis denominadas hipótesis nula ( $h_0$ ) y hipótesis alterna ( $h_a$ ).

La hipótesis nula significa igualdad al nivel medio del *tratamiento* (variable de entrada considerada en cada muestreo), mientras que la alterna plantea la posibilidad de diferencias significativas. Es decir

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_n$$

$$H_a : \exists u_i \neq \mu_j$$

**Fórmula 20: Hipotesis de ANOVA**

Pero como el modelo de nuestro experimento se puede ver como

$$\hat{y}_{ijk} = \mu + \text{Efectivo } V_1 + \dots + \text{error}$$

**Fórmula 21: Modelo Estadístico**

Las hipótesis se pueden reescribir como

$$H_0 : EF_1 = EF_2 = \dots = EF_n = 0$$

$$H_a : \exists EF_i \neq 0$$

**Fórmula 22: Hipótesis Modificadas**

Básicamente el análisis de Varianza, se utiliza para corroborar si la significación de diferencias entre medias de dos o más grupos, son o no debidas al azar. La cifra estadística obtenida con el ANOVA es la razón  $F_0$  que se compara (lo compara el programa SYSTAT) con el valor F de la tabla para  $\alpha=0.05$  y grados de libertad iguales a *tratamiento* para numerador y del error para el denominador.

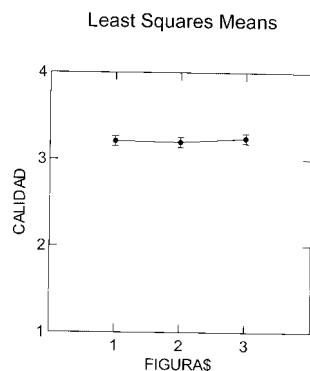


Si el valor obtenido es mayor que el de la tabla, el programa muestra un valor de  $p < 0.05$  en este caso, indicando con esto la significancia del efecto de la variable de entrada considerada en la variable de salida. Ver Tabla 15

Variable de Entrada	Suma de Cuad. (SS)	Grados de Libertad (GL)	Mean Square (MS)	$F_0$	P
Clase	$SS_{clase}$	2	$SS_{clase} / 2$	$MS_{clase}$	$< 0.05$
Compresión	$SS_{compresion}$	2	$SS_{clase} / 2$	$MS_{compresion}$	$< 0.05$
Error	$SS_{error}$	GL	$MS_{error}$		

**Tabla 15: ANOVA**

Las variables independientes consideradas en un modelo de prueba fueron todas de naturaleza categorial: experto/inexperto, la clase de la imagen (caras, formulas, paisajes), la figura (3 tipos de caras distintas, tres tipos de formulas distintas, tres tipos de paisajes distintos) y la compresión. Una primer conclusión que se desprende de los resultados del primer ANOVA realizado, es que dentro de una clase, no es significativo cambiar el tipo de la imagen. En la Figura 10 nos muestra Least Square Differences (LSD) en donde se ve que los intervalos se solapan indicando que no existen diferencias significativas en la variable de salida calidad. Debido a esta razón se descartó esta variable y se consideraron las restantes y sus interacciones



**Figura 10: Eliminación de la Variable Figura**

La tabla ANOVA muestra los cocientes entre cada variable de entrada y el error, repitiendo en cada caso la búsqueda en la tabla.

Luego de eliminada la variable figura, se realizaron varios modelos para analizar y descartar las interrelaciones de las variables de entrada consideradas. Se utilizo para estos modelos todos los datos, sin distinguir entre expertos e inexpertos. A posteriori para cada modelo se analizaron los datos y la determinación de p que nos indica la interacción de las diferentes variables.

Cuando se filtraron las interrelaciones no importantes, se realizo un modelo, el cual se analizo de acuerdo a los dos grupos de usuarios involucrados en el experimento. Primeramente se analizan los datos y la determinación de p que involucran a los usuarios inexpertos y luego al grupo de los expertos.

En las figuras que se muestran a continuación, cuando se ve en un eje, OBSERVADOR\$, significa el tipo de observador que se esta analizando, corresponde al título del análisis y COMPRESION\$ son los distintos 3 radios de compresión que se utilizaron en el experimento. Además las clases están relacionadas con: clase CARAS es el número 1, clase FORMULAS es el número 2 y clase PAISAJES es el número 3.

### ***Modelo 1***

En este modelo se consideraron las variables CLASE, COMPRESION, OBSERVADOR, COMPRESION\*CLASE, OBSERVADOR\*CLASE y OBSERVADOR\*COMPRESION.

Variable de Entrada	SS	GL	MS	F <sub>0</sub>	P
CLASE	5.108	2	2.554	56.067	0
COMPRESION	30.887	2	15.444	339.040	0
OBSERVADOR	0.848	1	0.848	18.617	0
COMPRESIO*CLASE	4.000	4	1.000	21.953	0
OBSERVADOR*CLASE	0.807	2	0.403	8.855	0.001
OBSERVADOR*COMPRESION	0.126	2	0.063	1.380	0.263
Error	1.822	40	0.046		

**Tabla 16: ANOVA Modelo 1**

El valor de  $p > 0.05$  nos indica que la interacción OBSERVADOR\*COMPRESION no importa.

## ***Modelo 2***

En este modelo se elimino la relación OBSERVADOR\*COMPRESION y se consideraron las variables CLASE, COMPRESION, OBSERVADOR, COMPRESION\*CLASE, OBSERVADOR\*CLASE y OBSERVADOR\*COMPRESION\*CLASE.

Variable de Entrada	SS	GL	MS	F <sub>0</sub>	P
CLASE	5.108	2	2.554	52.810	0
COMPRESION	30.887	2	15.444	319.346	0
OBSERVADOR	0.848	1	0.848	17.535	0
COMPRESIO*CLASE	4.000	4	1.000	20.678	0
OBSERVADOR*CLASE	0.807	2	0.403	8.340	0.001
OBSERVADOR*COMPRESION*CLASE	0.110	4	0.028	0.569	0.687
Error	1.838	38	0.048		

**Tabla 17: ANOVA Modelo 2**

El valor de  $p > 0.05$  nos indica que la interacción OBSERVADOR\*COMPRESION\*CLASE no importa.

## ***Modelo 3***

En este modelo se elimino la relación OBSERVADOR\*COMPRESION\*CLASE y se consideraron las variables CLASE, COMPRESION, OBSERVADOR, COMPRESION\*CLASE, OBSERVADOR\*CLASE

Variable de Entrada	SS	GL	MS	F <sub>0</sub>	P
CLASE	5.108	2	2.554	55.071	0
COMPRESION	30.887	2	15.444	333.019	0
OBSERVADOR	0.848	1	0.848	18.286	0
COMPRESIO*CLASE	4.000	4	1.000	21.563	0
OBSERVADOR*CLASE	0.807	2	0.403	8.698	0.001
Error	1.948	42	0.046		

**Tabla 18: ANOVA Modelo 3**

### ***Modelo 4***

Se consideraron las variables CLASE, COMPRESION, OBSERVADOR, COMPRESION\*CLASE. Este análisis se dividió para expertos e inexpertos.

Para los expertos se muestra los resultados en la Tabla 19 y en Tabla 20 para los inexpertos.

Variable de Entrada	SS	GL	MS	F <sub>0</sub>	P
CLASE	4.107	2	2.054	47.473	0
COMPRESION	13.766	2	6.883	159.113	0
COMPRESION*CLASE	2.041	4	0.510	11.792	0
Error	0.779	18	0.043		

**Tabla 19: ANOVA Modelo 4 para EXPERTOS**

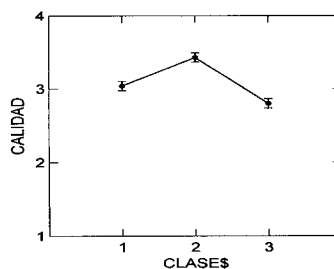
<b>Variable de Entrada</b>	<b>SS</b>	<b>GL</b>	<b>MS</b>	<b>F<sub>0</sub></b>	<b>P</b>
<b>CLASE</b>	1.807	2	0.904	17.427	0
<b>COMPRESION</b>	17.247	2	8.623	166.307	0
<b>COMPRESION*CLASE</b>	2.069	4	0.517	9.978	0
<b>Error</b>	0.933	18	0.052		

**Tabla 20:ANOVA Modelo 4 para INEXPERTOS**

## Análisis de los Inexpertos

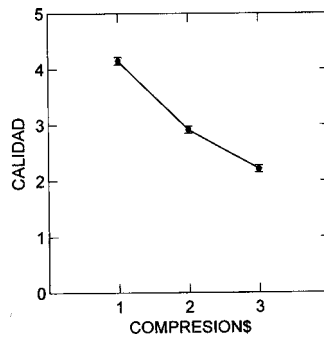
En LSD 1 se ve el promedio de la calidad lograda según la clase. Se ven diferencias significativas pues no hay solapamiento en cuanto al valor de calidad logrado al cambiar la clase. Los valores de calidad menores se corresponden con las caras y los paisajes. De acuerdo a los resultados objetivos que se muestra en las tablas 12, 13 y 14, las caras y paisajes tienen alta entropía. Esto significa que poseen mucha información de detalle que se pierde cuando se comprimen las imágenes.

Se debe destacar que las formulas en promedio muestran la mayor calidad en la imagen de salida observada, aunque son imágenes difíciles de tratar por los algoritmos de compresión con pérdida. Esta dificultad se debe al problema de codificación cuando toma bordes en donde hay niveles de grises muy distintos, el algoritmo promedia y aparecen los azulejos propios de la compresión con pérdida. Gran parte de la imagen es de fondo oscuro y las imágenes originales no son nítidas. Esto disimula la cuadrícula de los bordes, y a simple vista parecen todas imágenes iguales o muy parecidas. Esto justifica que a pesar de tener poco detalle y ser difíciles de tratar. La clasificación subjetiva es buena.



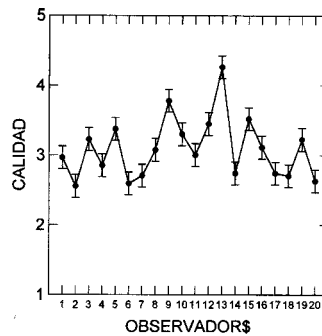
**LSD 1: Calidad-Clase para Inexpertos**

En LSD 2, la calidad de la imagen de salida para el grupo de inexpertos va decreciendo cuando se aumenta el radio de compresión. Esta relación la marca el sentido común puesto que mientras aumenta el radio de compresión se detecta más fácilmente la pérdida de información producida por el algoritmo de compresión. En las medidas objetivas, tanto la entropía y la relación señal ruido disminuyen debido a la pérdida de información, cada vez hay más distancia a la imagen original.



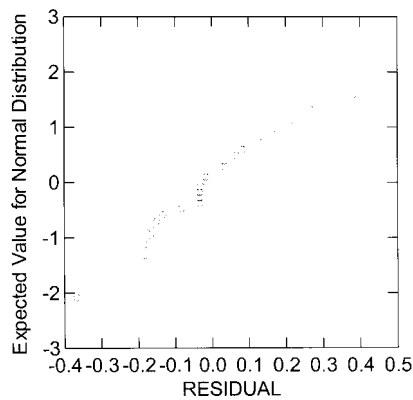
**LSD 2: Calidad-Compresión para Inexpertos**

En LSD 3 se muestra la calidad visual en el grupo de los inexpertos promediando por sujeto para todas las clases. Se observan diferencias naturales en la calificación de la calidad de la imagen entre los distintos observadores inexpertos.

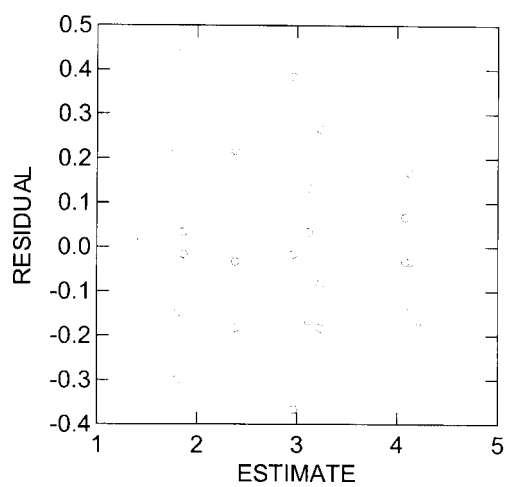


**LSD 3:Calidad-Observador para Inexpertos**

La Función 3 muestra que los residuos aproximan correctamente con la distribución Normal. Mientras que la Función 4 muestra que los residuos están esparcidos en forma de nube, no existiendo datos anormales.



**Función 3: Residuos - Distribución Normal Inexpertos**



**Función 4: Residuos – Estimados Inexpertos**

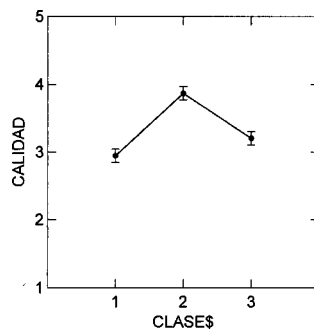


## Análisis de los Expertos

En LSD 4 se ve el promedio de la calidad lograda según la clase. Se ven diferencias significativas debido a que no hay solapamiento en cuanto al valor de calidad lograda al cambiar la clase. Los valores de calidad menores se corresponden con las caras y los paisajes. De acuerdo a los resultados objetivos que se muestra en las tablas 12, 13 y 14, las caras y paisajes tienen alta entropía. Esto significa que poseen mucha información de detalle que se pierde cuando se comprimen las imágenes.

Al igual que en los inexpertos, se debe destacar que las formulas en promedio están muestran la mayor calidad en la imagen de salida del observador, aunque son imágenes difíciles de tratar por los algoritmos de compresión con pérdida. Esta dificultad se debe al problema de codificación cuando toma bordes en donde hay niveles de grises muy distintos, el algoritmo promedia y aparecen los azulejos propios de la compresión con pérdida. Gran parte de la imagen es de fondo oscuro y las imágenes originales no son nítidas. Esto disimula la cuadrícula de los bordes, y a simple vista parecen todas imágenes iguales o muy parecidas. Esto justifica que a pesar de tener poco detalle y ser difíciles de tratar. La clasificación subjetiva es buena.

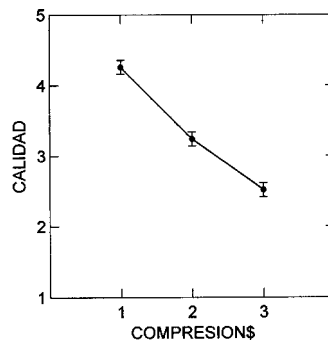
La diferencia con los inexpertos se da en los valores de la clasificación. El área de bajo la curva es menor dado que los expertos clasifican en forma más ajustada. Esto significa que *ven* mejor que los inexpertos.



### LSD 4: Calidad-Clase para Expertos

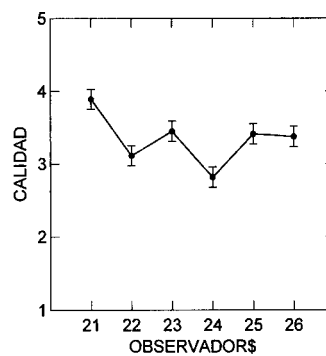
En LSD 5, la calidad de la imagen de salida para el grupo de los expertos va decreciendo cuando se aumenta el radio de compresión. Esta relación la marca el sentido común puesto que mientras aumenta el radio de compresión se detecta más fácilmente la pérdida de información producida por el algoritmo de compresión. En las medidas objetivas, tanto la entropía y la relación

señal ruido disminuyen debido a la pérdida de información, cada vez hay más distancia a la imagen original.



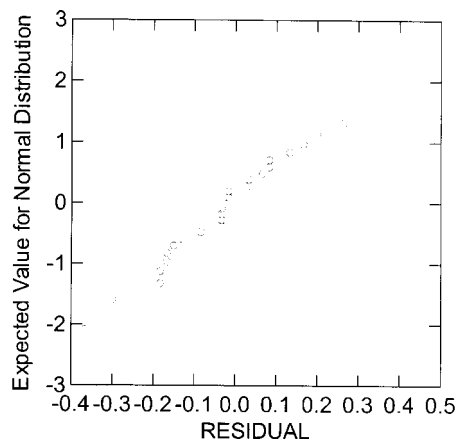
### LSD 5: Calidad-Compresión para Expertos

En LSD 6 se observa que las diferencias en el grupo de expertos son menores los solapamientos, esto es lógico debido a que comparten mas experiencia en la observación de imágenes.

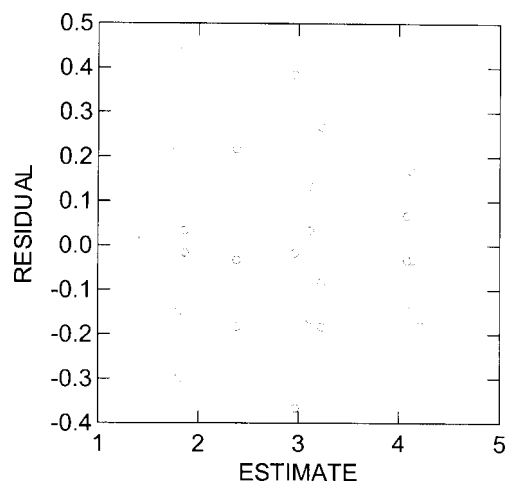


### LSD 6: Calidad-Observador para Expertos

La Función 5 muestra que los residuos aproximan correctamente con la distribución Normal. Mientras que la Función 6 muestra que los residuos están esparcidos en forma de nube, no existiendo datos anormales.



**Función 5: Residuos - Distribución Normal Inexpertos**



**Función 6: Residuos – Estimados Expertos**

### **III - Introducción**

La compresión de una imagen digital puede facilitar su procesamiento, almacenamiento y transmisión. A medida que los Departamentos de las grandes Organizaciones se vuelven cada vez más digitales y distribuidos, la cantidad de datos multimediales (en particular imágenes) que tienen, obligan a considerar su compresión para su almacenamiento y transmisión.

El objetivo general de la compresión es representar una imagen con la menor cantidad posible de bits con la menor pérdida de calidad posible, acelerando así la transmisión y minimizando los requerimientos de almacenamiento. De manera alternativa, el objetivo es lograr la mejor fidelidad posible para una capacidad de almacenamiento disponible limitada.

En el campo de las comunicaciones se ha destinado mucho trabajo a la codificación de imágenes digitales, dirigido a reducir los requerimientos de la tasa de bits para la transmisión de las imágenes. La experiencia ha mostrado que cada esquema de codificación está sujeto a su propio y único conjunto de causas de pérdida que a menudo son difíciles de caracterizar [González 1996]. Esto se debe a la forma en la que se diseñan los esquemas de codificación para variar selectivamente la precisión de la representación; el observador puede ser insensible a los errores en algunas partes de la imagen pero no en otras. Se puede lograr una compresión considerable sólo mediante algoritmos con pérdida, los que no permiten la recuperación exacta de la imagen original. Esta pérdida de información hace que la compresión y otros algoritmos de procesamiento de imágenes con pérdida tengan problemas de confiabilidad debido a la potencial pérdida de calidad [Cosman 1994].

Hay una necesidad de contar con mediciones precisas de las pérdidas subjetivas que puedan usarse para predecir la calidad de una imagen. El objetivo de este trabajo es determinar estas mediciones de la distorsión y probar que se pueden usar diferentes medidas objetivas para predecir la calidad subjetiva de una imagen.

La necesidad de este tipo de mediciones está particularmente reconocida en el área de la codificación de imágenes digitales. La existencia de mediciones de la distorsión subjetivamente relevantes que reflejen las evaluaciones de la calidad de una imagen que hace un espectador, harían considerablemente más fácil la tarea de diseñar y optimizar los esquemas de codificación.

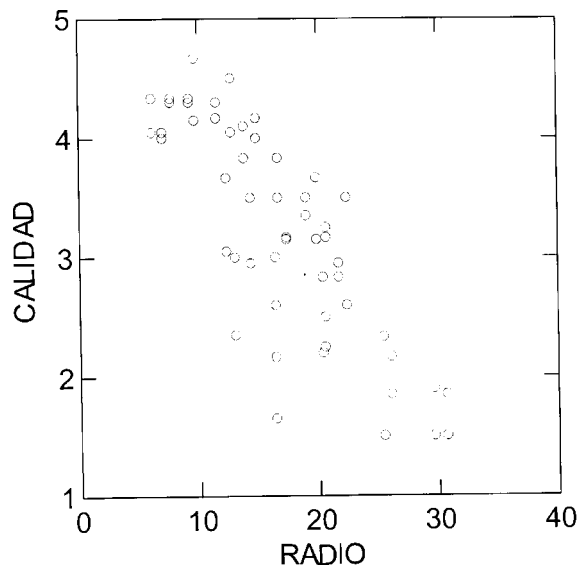
Estas distorsiones se comparan y contrastan con un conjunto de imágenes representativas de varios dominios de aplicación, y se examina cuán buenas son las mediciones de distorsiones que se

pueden obtener fácilmente para predecir las evaluaciones subjetivas, que son más caras en tiempo. Los ejemplos son de imágenes tradicionales que se utilizan en el área de procesamiento y compresión de imágenes comprimidas utilizando JPEG estándar [Wallace 1991] [Russo 1998].

## XII - Conclusiones

Se han analizado las correlaciones entre la variable CALIDAD subjetiva y los valores objetivos incluyendo el radio de compresión, buscando responder a los objetivos planteados.

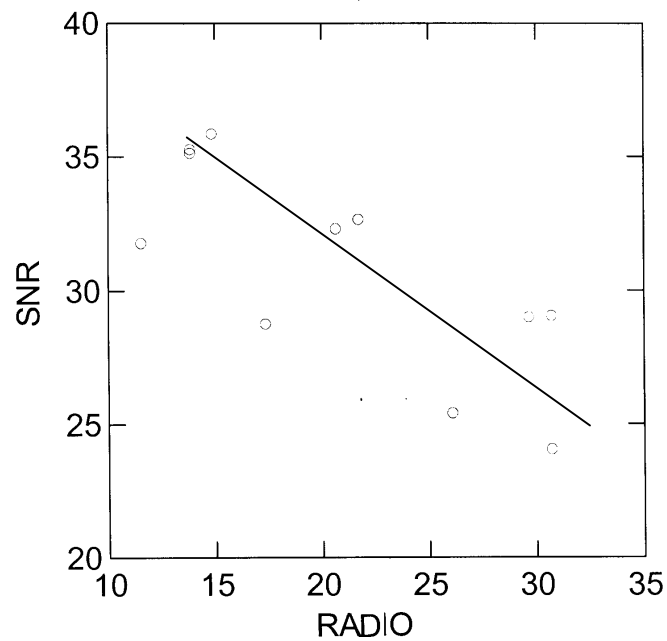
En la Figura 11 muestra que para este experimento, hay una correlación entre la variable de salida, CALIDAD y el radio de compresión, dado que la nube de puntos puede ser encerrada por una elipse cuyo radio tiende a 1. Esto indica que existe una recta con pendiente negativa pues a mayor radio de compresión, la calidad observada disminuye. En esta figura se ven los datos de expertos como inexpertos, también se cumple si se divide la información por tipo de observador.



**Figura 11: Correlación Calidad-Radio para Expertos-Inexpertos**

La Figura 12 muestra la correlación que hay entre la relación señal ruido y el radio de compresión. Si se miran los datos no se ve correlación directa pues los puntos están dispersos.

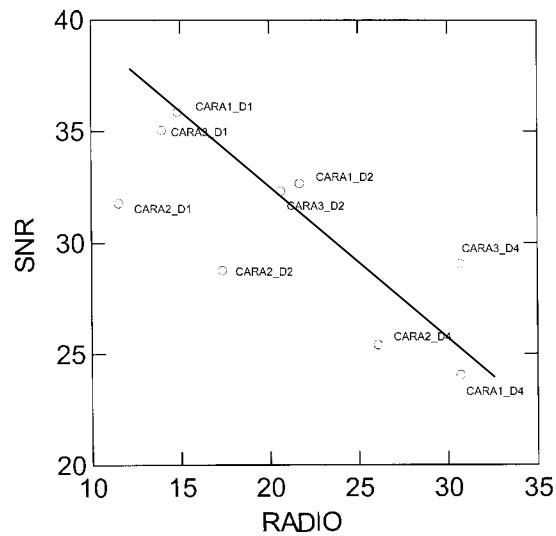
De todos modos la recta que interpola los puntos sigue una ley similar a la de la calidad subjetiva



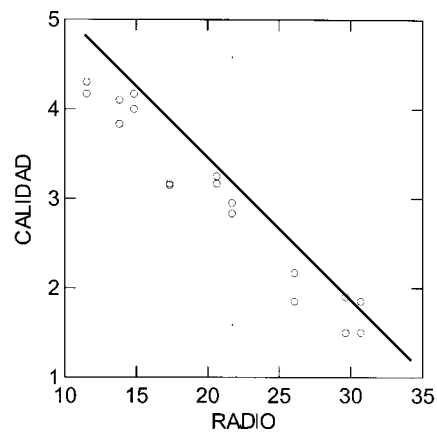
**Figura 12: Correlación SNR - Radio**

Ahora si se detalla la información de los puntos, clasificando por cada imagen comprimida, se ve como correlacionan directamente. Se muestra esto en las siguientes páginas para las CARAS y los PAISAJES.

En la Figura 13 se ve como las compresiones de CARA1, llamadas CARA1\_D1, CARA1\_D2 y CARA1\_D4 forman una recta. Lo mismo ocurre con CARA2 y CARA3. La Figura 15 muestra que para las CARAS las medidas subjetivas y objetivas siguen una variación similar en función del radio de compresión. Se marcan con líneas de puntos la recta interpolante de las medidas objetivas



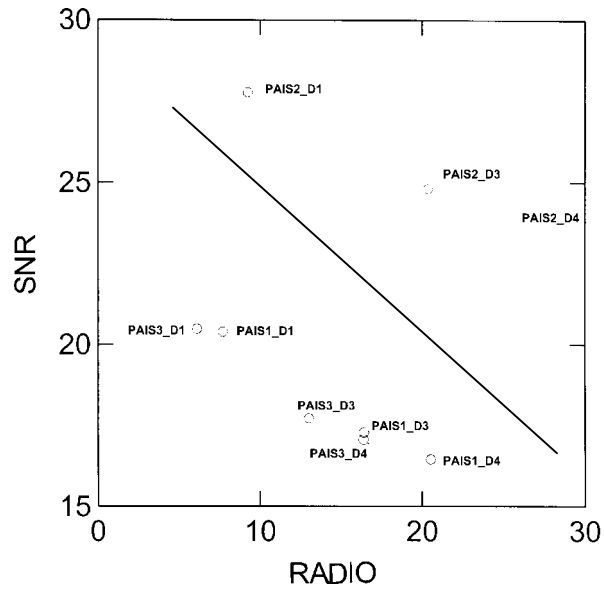
**Figura 13: SNR - RADIO para CARAS**



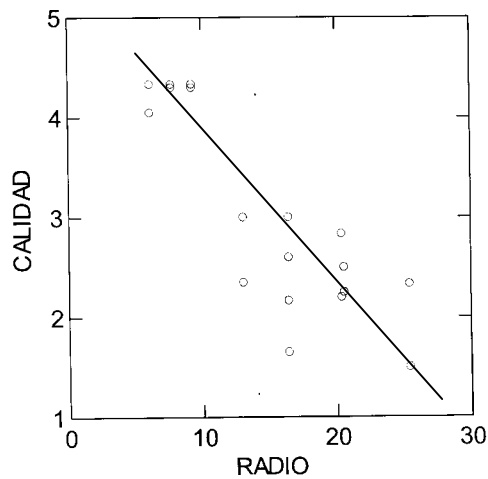
**Figura 14: Calidad - Radio para Caras**



En la Figura 15 se ve como las compresiones de PAISAJE1, llamadas PAIS1\_D1, PAIS1\_D3 y PAIS1\_D4 forman una recta. Lo mismo ocurre con PAISAJE2 y PAISAJE 3. La Figura 14 muestra que para los PAISAJES las medidas subjetivas y objetivas siguen una variación similar en función del radio de compresión.



**Figura 15: SNR - Radio para los paisajes**



**Figura 16: .Calidad - Radio para Paisajes**

Con análisis de los datos experimentales se puede concluir que existe una correlación entre la calidad subjetiva y la relación señal ruido. Los resultados se confirman extendiendo el experimento con mayor número de clases de imágenes y mayor número de radios de compresión.

Con relación a los objetivos planteados y el análisis realizado se puede concluir que:

## **Objetivo 1**

*La correlación existente entre las medidas objetivas y los resultados de la evaluación subjetiva de expertos e inexpertos*

Este estudio de correlaciones indica que para este conjunto de datos experimentales, existe relación entre las apreciaciones de calidad subjetiva y de la relación señal ruidos respecto del radio de compresión (y por lo tanto e la pérdida) de las imágenes.

Un resultado que no se determina en este trabajo es el máximo radio de compresión en donde no sigue valiendo esta correlación. Los problemas están dados en que el aumento de los radios de compresión modifica la interfaz hombre-máquina para la evaluación de la calidad subjetiva. De todos modos no interesaba trabajar con radios de compresión que degradaran la imagen hasta hacerla ininteligible.

## **Objetivo 2**

*Verificar si dentro de las diferentes clases de imágenes hay diferencias en la apreciación subjetiva.*

El análisis de los datos y la correlación con el radio indican que el sujeto evaluador ve las diferencias de compresión y en la calidad percibida.

La pendiente de la recta de degradación de la calidad subjetiva depende levemente de la clase de imagen, tal como se ve en la Figura 13.

## **Objetivo 3**

*Estudiar la incidencia de la categoría experto y de la categoría inexperto en la apreciación subjetiva de la calidad.*

No aparecen diferencias significativas en la calidad apreciada, esto se ve claramente en los LSD 1 y LSD 4 en donde la calificación asignada tiene poca variabilidad. Esto se resume en la Figura 17.

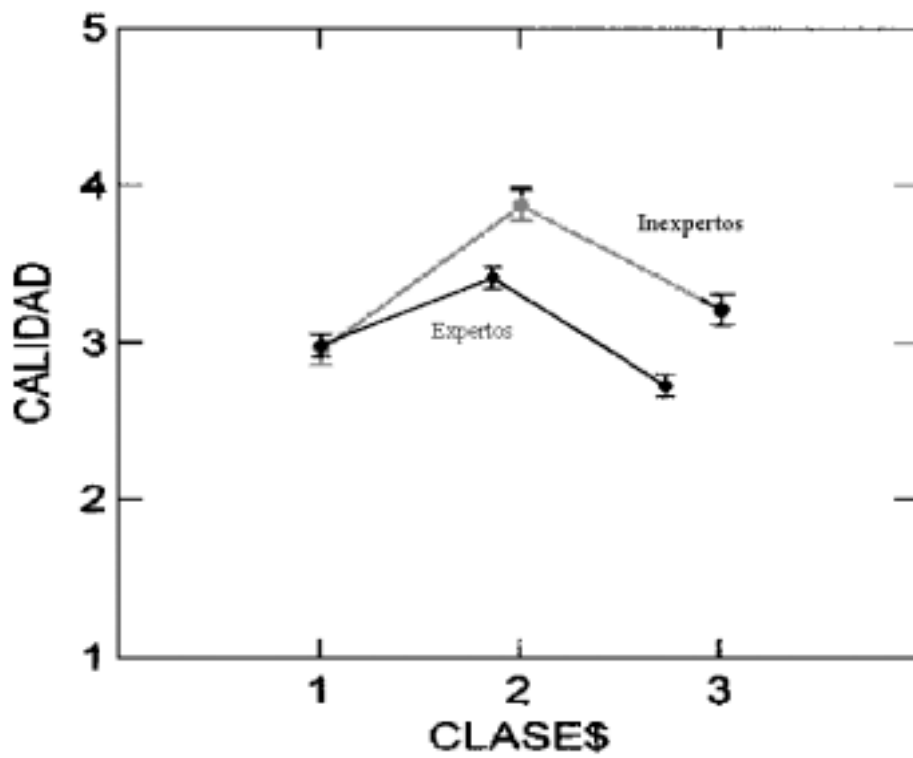


Figura 17: Calidad Subjetiva Expertos-Inexpertos

## XIV - Bibliografía

- [Agresti 1990] *Categorical Data Analysis*. A. Agresti. Wiley & Sons Inc., 1990.
- [Algazi 1992] *Comparasion of Image Coding Techniques with a Picture Quality Scale*. V. R. Algazi, Y. Kato, M. Miyahara, Kotani. SPIE Vol 1771, Applications of Digital Image Processing XV, 1992.
- [Ato 1994] *Fundamentos Estadísticos con SYSTAT*. Manuel Ato, J. Lopez, Addison-Wesley Iberoamericana, 1994.
- [Brooks 1995] *The Mythical Man Month. Essays on Software Engineering*. P Brooks, Addison-Wesley Pub Co. 1995.
- [Budrikis 1972] *Visual Fidelity Criterion and Modeling*. Z. L. Budrikis. Processing of the IEEE Vol 60 Nro 7, July 1972.
- [Campbell 1980] *The physics of visual perception*. F. W. Campbell. Phil. Trans. Roy. Soc., London, series B vol. 290, 1980.
- [Cosman 1994] *Evaluatig Quality of Compressed Medical Images: SNR, Subjetive Rating and Diagnostic Accuracy*. P. Cosman, R. Gray, R. Olshen. Proceedings of the IEEE VOL 82, Nro. 6, Junio 1994.
- [Cygansky 1998] *Information Codign*. David Cygansky. [www.ece.wpi.edu](http://www.ece.wpi.edu)
- [Davisson 1981] *Efficient Universal Noiseless Source Codes*. L. D. Davisson, R. J. McElice, M. B. Pursley, M. S. Wallse. IEEE Transaction Information Theory, VOL IT-27, 1981
- [Fish 1991] *A Subjetive Visual Quality Comparasion of NTSC, VHS and Compressed DSI-compatible Video*. R. Fish, T. Judd. Proc. SID (Society for Information Display), vol. 32 nro.2, 1991.
- [Glassner 1995, VOL 1] *Principles of Digital Image Synthesis*. A. S. Glassner. Morga Kaufman Publishers, INC., VOL 1 1995.
- [Glassner 1995, VOL 2] *Principles of Digital Image Synthesis*. A. S. Glassner. Morga Kaufman Publishers, INC., VOL 2 1995.

- 
- [González 1996] *Tratamiento Digital de Imágenes*. R. González & R. Woods. Addison-Wesley/Diaz de Santos, 1996.
- [Gray 1990] *Entropy and Information Theory*. R. M. Gray. Springer-Verlang, 1990.
- [Hall 1982] *Evaluating Image Quality: A new formulation computer*. E. Hall & S. Dwyer. Bio. Med. Vol 2, 1982.
- [Held 1991] *Data Compression. Techniques and Applications. Hardware a Software Considerations*. Gilbert Held. John Wiley & Sons, 1991.
- [Huang 1971] *Image Processing*. T. S. Huang. Proc. IEEE vol. 59 nro. 11, 1971.
- [Jain 1989] *Fundamentals of Digital Image Processing*. A. Jain. Prentice Hall, 1989.
- [Jense 1996] *Introductory Digital Image Processing. A Remote Sensing Perspective*. J. Jense. Prentice Hall, 1996.
- [Luckiesh 1965] *Visual Illusions*. Luckiesh M. Dover Mineola, 1965.
- [Lukas 1982] *Picture Quality Prediction Based on a Visual Model*. F. Lukas, Z. Budrikis. IEEE Transactions on Communications, VOL COM-20, Nro. 7, Julio 1982.
- [Lynch 1985] *Data Compression: Techniques and Applications*. T. J. Lynch. Lifetime Learning Wadsworth, 1985.
- [Marmolin 1986] *Subjective mse Measures*. IEEE Transaction System. Vol SMC-16. Junio 1986.
- [Montgomery 1991] *Diseño y Análisis de Experimentos*. D. C. Montgomery. Grupo Editorial Iberoamérica, 1991.
- [Nelson 1991] *The Data Compression Book*. Mark Nelson. Prentice Hall, 1991.
- [Netravali 1988] *Digital Pictures: Representation and Compression*. A. N. Netravali, B. G. Haskell. Plenum, 1988.
- [Quackenbush 1988] *Objective Measures of Speech Quality*. S. Quackenbush, T. Barnwell & M. Clements. Prentice Hall Signal Processing Series, 1988.

- [Russo 1998] *Técnicas de Compresión de Imágenes Aplicables a un Ambiente de Oficinas Distribuido*. C. Russo, 1997. Departamento de Informática, Facultad de Ciencias Exactas, Universidad Nacional de La Plata.
- [Santer 1989] *The Statical Analysis of Discrete Data*. T. J. Santer & D. E. Duffy. Springer Verlag, 1989.
- [Sayood 1996] *Introduction to Data Compression*. Khalid Sayood. Morgan Kaufman Publishers, Inc., 1996.
- [Shannon 1948] *A Mathematical Theory of Communications*. C. E. Shannon. Bell System Technologies Journal, VOL 27, 1948.
- [Stockham 1972] *Image Precessing in the Context of a Visual Model*. T. Stockham. Proceeding of the IEEE VOL 60, Nro. 7, Julio 1972.
- [Wallace 1991] *The JPEG Still Picture Compression Standard*. G. Wallace. Multimedia Engineering. Digital Equipment Corporation. Maynard, MASS, 1991.
- [Wang 1992] *An Objetive Measure for predicting Subjetive Quality of Speech Coders*. S. Wang, A. Sekey & A. Gersho. IEEE Journal Sel. Areas Commun. Vol. 10, Junio 1992.
- [Wasserman 1978] *Color Vision: An Historical Perspective*. G. S. Wasserman. Jhon Wiley & Sons, 1978.
- [Lindley 1995] "JPEG-Like Image Compression", Craig A. Lindley, Dr. Dobb's. 1995.