



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
CAMPUS ARAGÓN

ORIGEN DESCRIPCIÓN Y
TRATAMIENTO DE LA IMAGEN PARA
EDICIÓN DIGITAL EN TELEVISIÓN

290242

T E S I S
QUE PARA OBTENER EL TÍTULO DE:
INGENIERO EN COMPUTACIÓN
P R E S E N T A:
JOSÉ OMAR RENDÓN CATAÑO

ASESOR: ING. BLANCA ESTELA CRUZ LUEVANO

MÉXICO, D. F.

FEBRERO 2001



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



**UNIVERSIDAD NACIONAL AUTÓNOMA
DE MÉXICO**

JOSÉ OMAR RENDÓN CATAÑO.

**ORIGEN DESCRIPCIÓN Y TRATAMIENTO DE LA
IMAGEN PARA EDICIÓN DIGITAL EN TELEVISIÓN.**

A mí madre y a mí padre, gracias a los dos por haberme transmitido lo aprendido para volverme lentamente sabio, para saber que al fin el mundo es esto y que en su mejor momento tendré que elegir otros juegos y que los jugaré en serio, porque siempre puedo acudirlos para contarles la verdad, defenderé su alegría como un destino defenderla de los ingenuos y de los canallas, y nuevamente gracias por llevarlas a las cosas cotidianas que solamente ustedes saben que se volverán gigantescos soles la suma de corazón y corazón.

A mi hermano Uriel, porque es increíble no estar solo, porque siempre he podido recurrir a tus brazos, convertirlos en cómplices de mis logros y fracasos, siempre podré recurrir a tí para que cuando haya lagrimas pueda escudarte para mí.

A mis abuelos, el principio de todo, ellos se quedaron por tener fé, algo ganaron y algo se fué pero todas esas huellas se quedaron en la piel.

Pongo estos versos en mi botella al mar,
con el secreto deseo de que algún día llegue
a una playa desierta y un niño la encuentre
y la destape y en lugar de versos extraiga
piedritas, socorros, esperanzas y alertas y caracoles.

Para la creadora de todas mis adicciones mi aire.

Mariana.

A Blanca y a Fernando por haber creído en mí y en éste proyecto,
por ayudarme a dar un paso más.

A mis amigos, por haber estado siempre allí incondicionales.

Y por último a Vincenzo.

ÍNDICE.

INTRODUCCIÓN.....	I
I. CONCEPTOS GENERALES.....	1
I.1 Percepción Visual.....	2
I.2 Concepto de imagen.....	2
I.3 Concepto de imagen electrónica.....	3
I.4 La imagen de televisión.....	3
I.5 Cualidades de la Imagen.....	5
I.5.1 La cadena de televisión como un sistema lineal.....	6
II. VISIÓN HUMANA.....	8
II.1 Instrumentos y partes de óptica.....	9
II.2 Importancia de la luz.....	10
II.3 Visión Humana.....	11
II.3.1 Los caminos visuales al cerebro.....	18
II.4 La Luz.....	21
II.4.1 Colorimetría.....	22
II.4.2 Fundamentos del Color.....	22
II.4.3 Luz visible.....	26
III. FUNDAMENTOS DE VIDEO Y TELEVISIÓN.....	30
III.1 Imagen estática e imagen dinámica. Movimiento.....	31
III.2 Cámara de Televisión.....	31
III.2.1 Dispositivos Acoplados por Carga (CCD).....	33
III.2.2 Características y deficiencias de los primeros CCDS.....	35
III.2.3 Exploración y barrido de una imagen.....	36
III.2.4 Entrelazado.....	39
III.3 Propiedades especiales de la cámara.....	40
III.3.1 El Tubo de Imagen (Tubo de Rayos Catódicos).....	41
III.3.2 Sincronización de la señal.....	41
III.3.3 Tiempo de borrado Horizontal.....	44
III.3.4 Tiempo de borrado vertical.....	45
III.4 Resolución de una imagen.....	46
III.4.1 Determinación del ancho de banda.....	47
III.5 Características de la información de Color.....	47
III.6 Señales de Video.....	49
III.6.1 Señal de Luminancia.....	49
III.6.2 Señal de Crominancia.....	51
III.6.3 Señales I y Q en el sistema NTSC.....	52
III.6.4 Señal de Barras de Color (E.I.A.).....	53
III.7 Sistema de monitoreo de señales.....	55
III.7.1 Monitor Vectorscopio.....	55
III.7.2 Mediciones de Video. (Wave Form Monitored WFM).....	57
III.7.3 Burst o sincronía de color.....	58

III.8	Transmisión NTSC.....	58
III.9	Defectos de la Señal de Video.....	59
III.10	Estándares de Televisión.....	61
III.10.1	Estándar NTSC.....	63
III.11	Señal de Audio.....	63
III.11.1	Características del Sonido.....	64
III.11.2	Medición y monitoreo de Audio.....	64
III.11.3	Problemas y defectos de la señal de Audio.....	65
III.12	Sistema de Cuenta de Tiempo Drop Frame y Full Frame.....	67
III.12.1	Señal de código de Tiempo.....	68
III.12.2	Ventajas del Código de Tiempo (TC).....	69
III.12.3	Código de Tiempo Longitudinal (LTC).....	70
IV.	IMÁGENES DIGITALES.....	72
IV.1	La imagen analógica.....	73
IV.2	La imagen digital.....	73
IV.2.1	Digitalizador o Digitizer.....	73
IV.2.2	Señal Digital de video.....	76
IV.2.3	Conversión digital - analoga.....	79
IV.2.4	Procesamiento Digital de Imágenes.....	80
IV.3	Compresión de datos.....	81
IV.3.1	En qué se basa la compresión de datos.....	82
IV.3.2	Métodos de compresión de datos.....	83
IV.4	Video Compresión.....	84
IV.4.1	Compresión espacial.....	84
IV.4.2	Codificación Huffman.....	86
IV.4.3	¿Por qué la Compresión de Video?.....	87
IV.4.4	Métodos de compresión de imágenes.....	89
IV.5	Redundancia.....	90
IV.5.1	Redundancia de codificación.....	90
IV.5.2	Redundancia entre píxeles (información).....	91
IV.5.3	Redundancia psicovisual.....	93
IV.6	Archivos de Video.....	95
IV.6.1	Estándares de Producción de Video.....	95
IV.6.2	Características de los Archivos de Video.....	96
IV.7	Sistema de representación de imágenes.....	97
IV.7.1	Muestreo y cuantificación.....	98
IV.7.2	Relaciones entre píxeles.....	98
IV.7.2.1	Vecindad y contorno.....	99
IV.7.2.2	Conectividad.....	99
IV.8	Técnicas para la representación de la imagen digital. (Formatos gráficos).....	100
IV.8.1	¿Qué son los formatos gráficos?.....	100
IV.8.2	Estructura General de un Formato gráfico.....	100
IV.8.2.1	Cabecera.....	101
IV.8.2.2	Datos.....	101
IV.8.2.3	Paleta de Colores.....	101
IV.8.2.4	Paleta EGA/VGA 16 Colores.....	101

IV.8.2.5 Paleta VGA 256 Colores.....	101
IV.8.2.6 24 bits de Color (True Color).....	102
IV.8.2.7 Identificador.....	102
IV.8.2.8 Descriptor de pantalla.....	102
IV.8.2.9 Mapa Global de Color.....	102
IV.8.2.10 Descriptor de imagen.....	103
IV.8.2.11 Mapa de color local.....	103
IV.8.2.12 Datos raster.....	103
IV.8.2.13 Terminador.....	103
IV.8.2.14 Bloques de extensiones.....	104
IV.9 Técnicas de Formatos Gráficos.....	104
IV.9.1 BMP (Bitmapped File Header Format).....	104
IV.9.1.1 Tipos de bitmaps.....	105
IV.9.2 Imagen vectorial.....	106
IV.9.3 Tecnología fractal.....	107
IV.10 Modelos de Color.....	108
IV.10.1 El modelo de color RGB.....	108
IV.10.2 El modelo de color CMY.....	111
IV.11 Captura digital.....	112
IV.12 Formatos gráficos más importantes.....	112
IV.12.1 Formato de archivos de imágenes.....	112
IV.12.2 Formato de archivos de video.....	115
IV.13 Formato gráfico JPEG.....	116
IV.13.1 Introducción.....	116
IV.13.2 Codificación.....	117
IV.13.3 Pasos para una compresión JPEG.....	117
IV.13.4 Pasos para la descompresión JPEG.....	118
IV.13.5 Más acerca de JPEG.....	118
IV.14 Formato gráfico MPEG.....	119
IV.14.1 Desafío técnico.....	120
IV.14.2 Niveles y perfiles.....	121
IV.14.3 Las capas y la escalabilidad.....	122
IV.14.4 Flujo de bits de video.....	122
IV.14.5 Bloques.....	123
IV.14.6 Macrobloques.....	124
IV.14.7 Tajadas.....	125
IV.14.8 Compresión temporal e imágenes.....	125
IV.15 Flujos de programa y flujos de transporte.....	127
IV.16 Los beneficios.....	128
IV.17 Almacenamiento.....	128
IV.18 Protección de Señales Digitales en Almacenamiento y Transito.....	129
V. SISTEMA DE EDICIÓN NO LINEAL.....	131
V.1 ¿Qué es un sistema de edición lineal?.....	132
V.2 ¿Qué es un sistema de edición no lineal?.....	133

V.3 Modelo de Edición.....	133
V.3.1 Manejo de Media.....	134
V.3.2 Composición y efectos integrados.....	141
V.4 Calidad de la imagen.....	143
V.5 Características del Audio.....	144
V.5.1 Compatibilidad con otros sistemas.....	144
V.6 Opciones de Almacenamiento.....	145
V.7 Sistema de edición tipo DNG.....	148
V.7.1 Los componentes del sistema.....	148
V.7.2 Grabar y reproducir en tiempo real.....	150
V.7.3 Consolidar.....	151
V.7.4 La producción en las estaciones de trabajo.....	151
V.7.5 Servidor de producción.....	152
V.7.6 Grabación a RAID.....	153
V.7.6.1 Reproducción de material desde el RAID3.....	153
V.7.7 Control de acceso.....	153
V.7.8 Base de datos del Servidor.....	153
V.7.9 Almacenamiento Central en línea.....	154
V.7.10 El Canal de fibra.....	155
V.8 Red ATM.....	155
V.9 Arquitectura de las estaciones de trabajo.....	156
V.10 Flujo de los datos.....	156
V.11 Video Compresión.....	158
V.12 Protección de almacenamiento RAID3.....	158
V.13 Interfaces del sistema.....	160
CONCLUSIONES.....	162
GLOSARIO DE TERMINOS.....	165
BIBLIOGRAFIA.....	183

INTRODUCCIÓN.

INTRODUCCIÓN.

Detrás de todo el brillo y glamour de la producción de programas de Televisión, está una tecnología especializada que debe ser cuidadosamente estudiada para proveer la más alta calidad de imagen y sonido; tal como sería en una obra de arte al ser manipulados por un artista tratando precisamente de comunicar sus ideas. La Tecnología de la Televisión debe ser empleada creativa y operativamente por el personal de Producción.

El hecho de que el personal de producción conozca los términos técnicos le permitirá entender el funcionamiento correcto de la señal de TV y las transformaciones de ésta al tener presente una migración muy importante, "la migración analógica - digital", así como entender los procesos utilizados que serán los mismos en cualquier parte del mundo.

La Edición Digital tiene asociado una enorme cantidad de conceptos relacionados con los componentes de hardware y software, y estos a su vez de desarrollos teóricos. En el presente trabajo se tratarán los pasos fundamentales para llevar a cabo el funcionamiento de un sistema de este tipo.

En el primer capítulo se establecerán los conceptos generales de la definición de una imagen, la cual puede tener diferentes perspectivas dependiendo de los medios de captación existentes, tomando como base que toda imagen es visual, lo que cambia es el tipo de sistema que es utilizado para representarla, por ejemplo sistemas ópticos, electrónicos, matemáticos entre otros. Así mismo, se podrán identificar las características fundamentales de las imágenes y su representación.

Una vez comprendida la definición de una imagen el segundo capítulo recorrerá los conceptos teóricos de los cuales se desprende la obtención de una imagen desde sus más íntimos procesos en la conversión de luz y la manera en como el mejor sistema visual que existe, el Sistema Visual Humano, funciona.

Esto nos lleva a un campo de acción muy extenso, debido a que una imagen es visual, nos obliga a empezar por conceptos físicos en donde los fenómenos de la luz y los porque de la visión son estudiados, esta parte de la física es conocida como: óptica.

Dentro de este ámbito se destacará la importancia de la luz y sus componentes dentro de la percepción de imágenes a través del Sistema Visual Humano, se tratarán los procesos internos de la anatomía humana así como los caminos de representación visual en el cerebro. Dentro de esta representación se destacará la importancia de los elementos de la luz y sus características, elementos sin los cuales ningún tipo de imagen pudiera ser percibida.

Establecidos los conceptos de imagen, los procesos internos del Sistema Visual Humano, y las características de la luz, en el capítulo tercero entraremos a la ingeniería del hombre en su afán de conservar, producir y difundir imágenes, todo ello a través de un sistema que pretende emular a nuestro sistema humano de visión, este sistema es la televisión.

El primer paso es adquirir la imagen. Para ello se necesitan sensores, dispositivos y la capacidad para producir una señal por sus propios medios.

El sensor puede ser una cámara a color que produce una imagen completa cada 1/30 de segundo, pero sería conveniente saber como estos sensores trabajan.

En televisión, las imágenes que contemplamos no son un proceso tan sencillo todas las imágenes están hechas por cuadros para ser exacto 30 por segundo y lo que observamos no es un cuadro completo cada vez, esto se encuentra fundamentado en la persistencia de la visión de los ojos para crear la ilusión de que la imagen completa está en movimiento, cuando en realidad lo que se mueve son los tres haces electrónicos en el tubo de rayos catódicos o CCD dentro de la cámara. La representación y características de la señal de video y audio así como su uso serán mencionados a detalle en este capítulo.

La representación analógica de una imagen originada de la señal de video ahora debe ser transformada en una señal digital, esta migración será descrita dentro de las líneas del cuarto capítulo. Ahora llegó el tiempo de adquirir la imagen digital, como la mayoría de las cámaras que existen dentro de la industria de T.V. aún no cuentan con una salida en formato digital se necesitara usar un conversor analógico - digital. La cual puede ser una tarjeta de digitalización de video.

Con la obtención de la señal digital se debe considerar que el manejo de ésta como tal no es práctico debido a que las imágenes contienen demasiados datos, aunque esta señal está libre de ruidos y es muy sólida, ocupa demasiado ancho de banda para lograr una grabación en cinta que resulte económica, al igual que la grabación en disco o la transmisión por el aire. La solución se encuentra en la compresión de video, para disminuir la información redundante en los cuadros de video. A su vez esto trae una tecnología de formatos gráficos que sirven para identificar el tipo de proceso aplicado a la imagen y el tipo de archivo que lo identifica.

La señal codificada o comprimida es decodificada o descomprimida en el aparato de televisión o en el procesador de una PC, después de obtener la información desde una memoria secundaria o de almacenamiento (como un Disco Duro, Sy Quest, DVD o CD-ROM), y así ser manipulada o transformada con fines amateur o profesionales.

En el quinto capítulo se podrá revisar que una vez que la señal digital es considerada dentro de un formato gráfico, es manipulable, para ello existen sistemas desarrollados para poder modificarla, estos sistemas son llamados Sistemas de Edición No Lineales debido a la facilidad con que diferentes procesos de edición son llevados a cabo sin seguir un procedimiento inevitable para la realización de cada uno de ellos tales como

cortes, efectos, musica o sonidos así como el almacenamiento y movilidad de los mismos.

A todo ello el personal de producción debe entender el porqué de los ajustes que se realizan al equipo y el nuevo funcionamiento de los sistema no lineales, y que se ha adoptado mundialmente como el estándar a seguir en la transformación de la televisión y sus transmisiones directas por Satélite (DTH), en la futura televisión de distribución digital terrestre (DTTV) y de cable, incluyendo la televisión de alta definición (HDTV) y de esta manera estar al tanto de la tecnología de punta generada hasta nuestros días.

CAPITULO I
CONCEPTOS GENERALES

I. CONCEPTOS GENERALES.

La definición de una imagen puede tener diferentes perspectivas dependiendo de los medios de captación existentes.

Antes de entrar al estudio de las imágenes digitales, es conveniente dejar lo más claro posible los conceptos básicos sobre los que se fundamenta la teoría que se obtendrá posteriormente "...Una imagen, al fin y al cabo, no es sino una de las posibles percepciones sensoriales de un aspecto de la realidad...".¹

El proceso de percepción visual humana es un paso fundamental para conocer los parámetros básicos utilizados para describir una imagen, el cual se describe brevemente a continuación.

I.1 Percepción Visual.

Cualquiera de los sistemas de captación de imagen que existen, tienen en común un principio básico sobre el cual se han desarrollado y evolucionado: *La imagen óptica*.

"Una *imagen óptica* es aquella que se genera a través de las lentes. Incluida la visión humana, puesto que el cristalino del ojo no es sino una lente convergente, y la retina el receptor de dicha imagen".²

En realidad no son los objetos los que percibimos a través del sentido de la vista, sino la luz que reflejan los mismos. De hecho, el ojo solo puede captar e interpretar la luz, y no otro tipo de materia o energía.

Por lo tanto la primera condición para que podamos ver un objeto, es que esté iluminado, y a partir de esta condición se percibe la luz que dicho objeto es capaz de reflejar en dirección de nuestros ojos.

I.2 Concepto de imagen.

Se definirá el concepto de imagen de una manera simple y fácil de entender "...consideraremos como imagen a cualquier función bidimensional de la intensidad de luz que, dado un par de coordenadas que definan un único punto (x, y) le haga corresponder el valor de la intensidad lumínica de la imagen en ese punto $f(x, y)$ donde, la intensidad lumínica esta dada entre el rango de cero (ausencia total de luz) e infinito."³

¹Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, pagina 17.

²Vicente Llorens, Fundamentos Tecnológicos de Video y Televisión, Editorial Paidós, España 1995, Objetivo 1.1.

³Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, pagina 22.

Esta definición básica de imagen es aplicada al termino imagen monocromática, "...una imagen monocromática, donde se ignoran las componentes espectrales de luz (esto es el color). En las imágenes monocromáticas, el valor de la luminancia en un punto, $f(x, y)$ recibe el nombre de nivel de gris dentro de una escala de valores que la misma puede asumir, conocida como escala de grises. La escala de grises, en principio restringida entre niveles de cero (ausencia total de luz) e infinito (máxima luminosidad), suele limitarse por razones practicas entre cero y un valor máximo L , que garantiza el blanco de mayor energía posible o representable..."⁴

Teniendo en cuenta estos conceptos se puede observar que la definición de "Imagen" se toma según el tipo de aplicación o el tipo de sistema en que ésta se utiliza dependiendo de propósitos particulares, de esta forma y para los fines de este trabajo se tendrá que tomar en cuenta diferentes tipos de definición dependiendo de la parte del proceso en que nos encontremos.

1.3 Concepto de imagen electrónica

Se denomina así a imágenes capturadas, almacenadas y visualizadas a través de medios electrónicos. Este tipo de imágenes pueden ser "creaciones" o imágenes reales digitalizadas. Las primeras se realizan a través de técnicas de diseño por computadora. Las segundas son imágenes ya existentes que son transformadas a un formato comprensible por la computadora o dispositivo electrónico. El proceso de pasar una imagen real a un formato comprensible para el dispositivo se denomina captura.

Las imágenes electrónicas tienen que representarse de forma que el dispositivo electrónico utilizado para su tratamiento las pueda entender y manejar, por lo que deben encontrarse en formatos digitales o analógicos.

Infografía es la técnica del grafismo⁵ por computadora. Los campos de aplicación de esta técnica son múltiples, podemos citar a modo de ejemplo: teledetección, cartografía, imágenes médicas, bancos de imágenes periódicas, restauración documental, etc.

1.4 Imagen de televisión.

La televisión es básicamente un sistema para reproducir una serie de imágenes fijas, las cuales en una serie sucesivamente rápida, dan sensación de movimiento.

"La palabra televisión significa literalmente: Visión a distancia."⁶

⁴ Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, pagina 23.

⁵ Grafismo: graffa, empleo de signos determinados para expresar una idea. Diccionario Larousse.

⁶ Manual de Servicio Sony, pagina 5.

La señal de televisión esta formada por imágenes y sonidos. La señal eléctrica correspondiente a la información de luz de la imagen de una escena a transmitirse, se le conoce como **señal de video**, y a la señal eléctrica que corresponde al sonido de dicha escena se le llama **señal de audio**.

La señal de T.V. se ha diseñado para convertir todos los elementos de una escena en señales eléctricas y transmitirlos en una sucesión seriada rápida, mas no instantánea, como en cinematografía o fotografía, ni simultanea, esto dentro de un sistema electrónico, grabación magnética, satélite para T.V., circuito cerrado, CATV y transmisión al aire.

Para lograr tal efecto, la TV depende en gran parte de la persistencia de la visión humana para recrear en la pantalla de un cinescopio o tubo de rayos catódicos la escena enfocada por la cámara de TV.

Toda imagen esta creada por un grupo de áreas de luz y sombra. A estas superficies que contienen información de luz o de sombra se les conoce como detalle de una imagen, elemento de imagen, **pixel** ó pel. Todos estos elementos juntos contienen la información visual en la escena. Si estos elementos se transmiten y reproducen en el mismo grado de luz y sombra y posición como el original entonces se produce una imagen.

En fotografía dichos elementos lo constituyen los granos de plata ennegrecidos por la luz, en la cámara de TV son cargas eléctricas diminutas.

El tamaño de los pixeles es lo que determina la cantidad de detalle de la escena que el sistema es capaz de resolver. En cuanto más pequeños sean los pixeles, mejor será la resolución o definición de la imagen reproducida.

Figura de resolución. (figura 1.1)



Una imagen debe tener ciertos requisitos para poder ser captada por una cámara de TV y pueda así ser reproducida con calidad, ya que de lo contrario, la señal obtenida puede ser distorsionada.

1.5 Cualidades de la Imagen.

Los requisitos que deben observarse en una imagen son los siguientes:

- a) **Brillantez:** Es el nivel promedio de luz que refleja una imagen. Cabe mencionar que existen dos tipos de luz: La luz incidental que es aquella que cae en los objetos de la escena y la luz reflejada es la luz que refleja o rebota de los objetos. La luz reflejada es la porción de la luz incidental que permite la percepción de los detalles visuales de los objetos.
- b) **Contraste:** Es la diferencia de intensidades entre las áreas oscuras y las claras de una imagen, a diferencia del brillo que es un valor promedio.

Un buen contraste debe producir una imagen recia, cuyo blanco sea brillante y su negro oscuro.

Para esto hay que tener en cuenta la forma como se ha iluminado la escena, ya que dicha iluminación ambiental debe ser suficiente para hacer que las partes negras (si las hay) se vean oscuras. Por otra parte, si existe una iluminación excesiva, habrá poco contraste por la gran cantidad de luz reflejada, lo que impide tener partes oscuras.

- c) **Detalle:** Es la resolución o definición de la imagen y depende del número de elementos de imagen (píxeles) que puedan ser reproducidos.

A mayor claridad de imagen se tiene más calidad, y esto da una mejor profundidad de campo obteniendo también los detalles de la escena del fondo (Background).

También dependerá, como veremos después del número de líneas de barrido, del ancho de banda del video y de la cámara.

- d) **Nivel de color:** Esta señal se superpone a la información eléctrica de Blanco y Negro es llamada Luminancia y es la cantidad de color o croma y varía desde un tono sin color pálido hasta colores intensos, por eso se le llama también Intensidad de color.
- e) **Matiz:** Llamado también tinte o HUE y corresponde al color indicado.
- f) **Relación de aspecto:** Esto será en la parte de la pantalla donde se reproduzca la imagen y consiste en la relación de las partes anchas a las altas de la pantalla y lo normal es de 4 unidades de ancho por 3 unidades de altura (4:3), es decir, más ancha por un factor de 1.33 que de altura y es casi la misma proporción que en cinematografía y esto es para observar el movimiento, el cual usualmente es en dirección horizontal.

Estas características que se deben observar en las imágenes captadas por una cámara tienen valores que deben identificarse, para ello existen unidades de medida que sirven

como estándares en sistemas métricos. Se mencionan algunas de estas unidades de medida en cuestiones de luz y óptica dentro del contenido del siguiente capítulo.

Pero que hay de las representaciones de una imagen, estas pueden ser vistas de una manera integral en funciones de tres dimensiones y transformadas a sistemas de dos dimensiones, el detalle a continuación.

1.5.1 La Cadena de Televisión como un sistema lineal.

La luz de una escena origina que se forme una imagen $i(x, y, t)$ en el plano focal; a dicha imagen la llamaremos "La función de Video", y note que la función es un vector para imágenes coloreadas. Es la tarea del sistema la de proveer una versión (modificada de la monocromática) $i'(x, y, t)$ en la pantalla para observarla.

La función de Video⁷ se convierte en una señal de video, $v(t)$, en un proceso de barrido de la imagen en la cámara. La señal producida desde cada punto del plano focal es proporcional a la energía de luz integrada la cual cae en el punto en cada tiempo de muestreo.

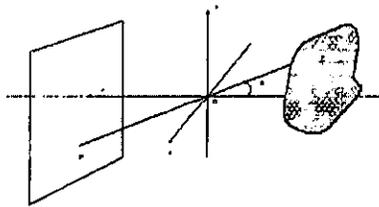


Figura 1.2

En la figura anterior (figura 1.2) podemos observar que cualquier imagen de un objeto de tres dimensiones puede ser representado dentro de un sistema cartesiano de dos dimensiones, en donde el tiempo se representará con el número de repeticiones de esa imagen en un segundo.

La señal de video es posteriormente procesada por el canal (Modulación, Filtración, Digitalización, Transmisión, etc.) produciendo una señal modificada $v'(t)$ para aplicarse a la pantalla.

El proceso en la formación de una imagen en la pantalla es como un trazado en la superficie visible de fósforo y un rastreador patrón como en la cámara donde una cantidad de energía que se emite en cada punto del haz siendo proporcional a la energía de la luz obtenida en el elemento que la corresponde en el plano focal de la cámara. En

⁷ Manual de Servicio Sony, pagina 5.

la práctica la energía emitida, es desparramada en algún intervalo de tiempo, casi siempre más corto que un tiempo de cuadro.

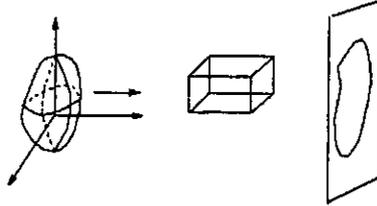


Figura 1.3

Esta descripción revela una gran diferencia entre las funciones de video original y reproducidas. La primera es continua en espacio y tiempo, mientras que la última no lo es. Si así fuese, el sistema sería caracterizado simplemente por su respuesta de frecuencia espacio - temporal, la cual puede ser comparada con la sensibilidad correspondiente del sistema visual humano (HVS).

De la misma manera en la figura 1.3 se puede observar que cualquier imagen de cualquier objeto de tres dimensiones se puede transportar a un eje coordenado y proveerle de movimiento en su relación con el tiempo, es decir, determinando el número de repeticiones por segundo.

El factor de que $i(x, y, t)$ es altamente discontinuo en espacio y en tiempo, es la causa de ineficiencia en la utilización de la capacidad del canal. Con la simple eliminación de la estructura de muestreo por empañamiento (Blurring), el uso de fósforos de larga persistencia o viendo la pantalla desde lejos, atenúa componentes importantes de la señal transmitida de la estructura visual.

Después de ver que la función es discontinua en espacio - tiempo es necesario revisar el porque de esta estructura visual desde el origen del proceso, en el siguiente capítulo se hará revisión de el proceso completo de la misma.

CAPITULO II
VISIÓN HUMANA.

II. VISIÓN HUMANA.

En este segundo capítulo se describirán los conceptos básicos del sistema de visión humano, si bien no se hará a fondo desde el punto de vista fisiológico si resultara un apoyo indispensable para comprender los parámetros característicos que engloban a una imagen. En principio la óptica y la luz.

II.1 Instrumentos y partes de óptica.

La Óptica es la parte de la Física que estudia los fenómenos de la luz y de la visión a través de diversos instrumentos como espejos, lentes etc.⁸

La Fotometría es la parte de la óptica que mide la luminosidad o intensidad de una fuente luminosa y las iluminaciones de las superficies.⁹ Utiliza dispositivos fundamentados principalmente en las propiedades de las lentes, que se emplean para lograr la observación de los objetos y poder procesarlos en un determinado medio.

Iluminación: Es la medida de la cantidad de energía luminosa que recibe o refleja una superficie. Esta varía de acuerdo con la intensidad de la fuente luminosa y de la distancia de los cuerpos a iluminar. Su unidad es el LUX.

Lux (Lx): Equivale a la iluminación de una superficie que recibe un Lumen por m². La unidad de intensidad de luz que ilumina un objeto en unidades métricas es el LUX, y en unidades inglesas es el FOOT-CANDLE (f - c).

$$1 \text{ f-c} = 10.76 \text{ Lux} \quad 1 \text{ Lux} = 0.0929 \text{ f-c}$$

Un lux será directamente proporcional a la intensidad de la fuente luminosa (Cd) e inversamente proporcional al cuadrado de la distancia.

$$1 \text{ Lux} = \frac{1 \text{ Cd}}{\text{m}^2}$$

De lo que se deduce que la iluminación disminuye en razón inversa al cuadrado de la distancia, porque la cantidad de luz emitida por la fuente tiene que repartirse cada vez en mayor superficie.

Luminosidad: Es la intensidad de una fuente luminosa. Su unidad es la candela (Candle).

Candela (Candle, Cd): es 1/60 de la intensidad luminosa emitida por un centímetro cuadrado de un cuerpo absolutamente negro (el que absorbe y emite perfectamente la radiación luminosa) a la temperatura de fusión del platino.

⁸ Diccionario Larousse.

⁹ Definiciones según el Manual de Servicio Sony, paginas 7y 8.

Candela: Equivale casi a la intensidad luminosa de una vela de 2 cm de diámetro y con una flama de 5 cm de altura.

1 Cd = 0.909 Watt

1 Watt = 1.1 Cd.

Por ejemplo: 40 W = 44 Cd.

La intensidad de luz que refleja un objeto determina el brillo del objeto y se mide en LAMBERTS.

Lamberts: 1 Lambert = 0.318 Candles cm²

Flujo luminoso: La energía luminosa sale o fluye continuamente de una fuente luminosa, llamado flujo luminoso y es el total de energía radiante visible que emite la fuente en un segundo. Su unidad es el LUMEN.

Lumen (Lm): Es el flujo luminoso que en un segundo fluye de la fuente luminosa de una candela a través de un ángulo sólido.

El equivalente al LUMEN (flujo luminoso) es de 1 f-e = 1 Lm (incidental) /Foot².

Después de mencionar algunas de las unidades de medida con respecto a la luz y propiamente de la iluminación, a continuación se da a conocer el porque de esta consideración. En este caso la luz es un punto primordial.

11.2 Importancia de la luz.

Cuando la luz cae en un objeto, este la refleja en toda su superficie en diferentes direcciones. Un ojo recoge algunas de esas luces reflejadas y la percibe como una escena visual, y el ojo verá la combinación final de brillos de todas las reflexiones de las pequeñas superficies del objeto.

El ojo no puede enfocar agudamente las imágenes pero tiene sensación del promedio total de las luces incidente y reflejada, si se desea enfocar agudamente el objeto, la luz debe controlarse cuidadosamente para asegurarse que la luz reflejada en un sólo punto en su superficie alcance la parte sensible del ojo.

El ojo usa las propiedades de desviación y concentración de luz en las lentes naturales.

El ojo humano, que se puede comparar con una cámara fotográfica, es un globo formado en la parte exterior por una membrana blanca llamada esclerótica que lo protege, su parte interior está cubierta por la coroides que es una membrana muy opaca

para evitar la reflexión de la luz en el interior del ojo. La lente del ojo es el crystalino, el cual, es elástico y varía su curvatura para enfocar con rapidez los objetos.

El iris, por su color determina el color de los ojos, tiene una abertura central llamada pupila que se contrae o agranda para regular la entrada de la luz en el ojo.

Para poder entender de una manera más amplia estos conceptos se tiene a continuación la descripción de los procesos que ocurren en el cuerpo humano de manera menos general.

II.3 Visión Humana.

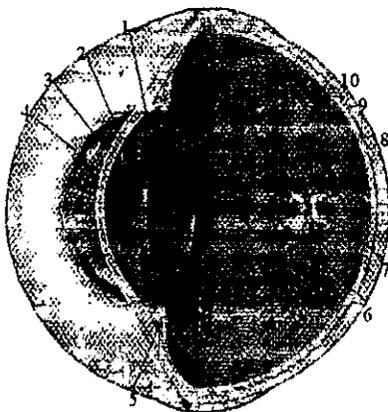
La visión humana consiste de las siguientes partes: El ojo, el nervio óptico que va al cerebro donde se decodifican las señales enviadas.

Teniendo al ojo humano como un instrumento eficiente, trabaja de la siguiente manera:

El ojo, en sus elementos básicos esenciales, es un sistema de lentes cristalinos, que reúne la luz de todas las frecuencias del espectro visible y la enfoca sobre una pantalla sensible a la luz (retina) situada en una concavidad alrededor de la parte posterior del ojo.

En la siguiente figura (figura 2.1) se puede apreciar parte de la anatomía del ojo humano, en donde todos estos procesos se realizan.

La retina se forma por millones de fibras nerviosas, cada una sensible a la luz en una frecuencia distinta y que se conectan en grupos al nervio óptico, el cual va conectado al cerebro, y así se crea la sensación de la visión.



1. Canal de Schlemm.
2. Cornea.
3. Iris.
4. Cristalino.
5. Músculo ciliar.
6. Macula.
7. Nervio óptico.
8. Retina.
9. Coroides.
10. Esclera.

Figura 2.1

Mediante la visión, el ojo es capaz de enviar al cerebro una imagen completa del objeto, su contenido de tono, tinte e intensidad de color y gracias al efecto estereoscópico sus tres dimensiones; pero el ojo tiene también otro atributo que es la persistencia de la visión.

La persistencia de la visión consiste en que el ojo humano posee la capacidad de retener una impresión de la forma, color y brillo de una imagen durante una fracción de segundo, después de que ha dejado de recibirse la luz de la imagen.

Esto se aprovecha en los procedimientos cinematográficos y de TV; ya que si la serie de imágenes se presentan al ojo durante este intervalo de persistencia de la visión, el ojo las integrará y el espectador tendrá la sensación de ver todas las imágenes al mismo tiempo como una imagen completa pero aparte en adición se tiene la ilusión del movimiento, siempre y cuando se tenga, como se dijo, imágenes completas en un rango de repetición mayor a 16 FRAMES/seg. (Cuadros por segundo). El rango de 24 FRAMES /seg. usado en cinematografía es suficiente para tener ilusión de movimiento.

El Parpadeo (FLICKER) en las imágenes en movimiento.

El rango de 24 Hz., no es tan rápido para permitir que el brillo de una imagen se mezcle suavemente al del siguiente cuando entre los cuadros hay una parte negra; por lo que el resultado es un parpadeo de luz definido (Flicker) en la pantalla tipo luz estroboscópica y es peor con imágenes con alta iluminación.

El sistema visual humano, como resultado de la evolución, es muy adaptable y rápidamente es capaz de captar gran cantidad de información útil ante el observador. Esta escena en 3-D, iluminada de manera diferente, moviéndose, produce imágenes ligeramente diferentes en 2-D en las retinas de ambos ojos.

Esto último está en constante movimiento voluntario, sobre la escena, ambos con movimiento de cabeza y por rotación en sus órbitas, y también realizan movimientos involuntarios pequeños los cuales tienen un importante y esencial papel en la visión.

La retina consiste de una matriz de receptores de dos clases:

Las células tipo conos, exclusivamente cubre los 2 grados centrales (la fovea) cuya densidad disminuye lejos del eje.

Las células tipo bastones cuya densidad es de máximo 15 grados del centro.

Los conos son responsables de la alta agudeza visual en el eje y para la visión a color en niveles normales (fotópico).

Los bastones, los cuales son mucho más sensibles, proveen sensibilidad de nivel de baja luz (escotópico) fuera del eje pero tienen mucha más baja resolución espacial. (Figura 2.2)

Cada célula se caracteriza por una cierta sensibilidad la cual depende de su estado de adaptación y de la excitación de sus vecindades. La sensibilidad se caracteriza en ambos por una función estática (entrada - salida) y una respuesta de frecuencia.

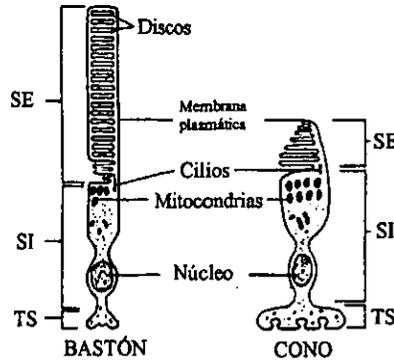


Figura 2.2

La resolución espacial de los puntos del objeto es rígidamente igual al espaciado celular, pero a causa de la cooperación de los receptores retinales, la resolución de líneas largas paralelas es mucho más fina que el espaciado celular. La naturaleza discreta del mosaico retinal no es obvio en visión normal es decir no es interpretado desde el primer momento, mucho del procesamiento visual es efectuado en la retina en sí, pero el procesamiento adicional ocurre en los niveles más altos del sistema nervioso visual.

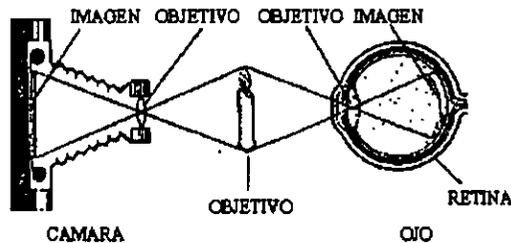


Figura 2.3

OBJETIVO IMAGEN
*Formación de una imagen en el ojo
 y en la cámara fotográfica*

El parpadeo (Flicker) y la rendición del movimiento están asociados a la respuesta de frecuencia temporal. El Flicker es evidente que en 25 o 30 Hz. está casi siempre presente; en 50 o 60 Hz., y siempre está presente en imágenes muy brillantes.

Para imágenes en movimiento, ya sea en cinematografía o en TV, el problema del Flicker se resuelve de la siguiente manera:

Como se dijo, la sucesión entre cuadro y cuadro debe ser en el tiempo de persistencia, ya que si es mayor a éste, entonces el ojo no integrará el brillo de una imagen y el brillo de la otra. Experimentalmente se ha encontrado que el rango de repetición o variación de brillos debe ser mayor a 40 Hertz (veces por segundo).

Por lo que su período será de:

$$T = 1/F = 1 / 40 = 0.025 \text{ Seg. (25 miliseg) donde T es el tiempo.}$$

Mayor a 40 Hz para eliminar el Flicker

Así que el tiempo de transición entre imágenes no debe ser mayor a 25 ms,

En cinematografía, se tiene una frecuencia de 24 FRM/Seg. su período será de 41.6 ms, la persistencia visual no compensa esta variación de brillos, por lo que tendremos Flicker. La solución se dio duplicando la exposición del cuadro esto es, se proyecta dos veces cada cuadro, lo que al final tendremos 48 proyecciones por segundo en la pantalla. Para esto se utiliza un disco obturador (Shutter) para oscurecer la imagen dos veces y permitir el cambio de cuadros y que este movimiento no se vea.

La retención o persistencia de la visión se efectúa específicamente en la retina del ojo. Esto ocasiona que se tengan 48 proyecciones de escenas en cada segundo, y a su vez, la pantalla se oscurezca otras 48 veces por seg. por lo que se siguen teniendo los 24 FRM/Seg. originales. La conclusión es que se incremento el rango de oscurecimiento, disminuyo el tiempo de éste, por lo que está dentro del lapso de persistencia visual, (T de 48 proyecciones = 20.8 ms) por lo que el Flicker se elimina.

El ojo en general es el sistema óptico más complejo y perfecto conocido y al cual todo sistema artificial tiende a emular. Una buena razón por la que los investigadores en visión artificial solamente emulan la visión humana es porque lo que se conoce del sistema de visión humano más allá del propio ojo es principalmente diájunto, especulativo y escaso. Pero hay más, aunque el sistema de visión humano es adecuado para muchas tareas, es obvio, como veremos que adecuado no equivale a infalible. Se demuestra que el sistema no es infalible por la existencia de ilusiones visuales.

Aunque es obvio que nadie puede negar el papel fundamental que juega la percepción en la adquisición de información por los humanos, podría argumentarse que ver es un acto mecánico y que no genera nada. Todo lo que el acto de ver hace es inferir el estado del mundo en la medida de lo que se permite por los datos sentidos. Es decir, produce alimento para el pensamiento, para conceptualizar, clasificar y asignar pertenencia a una

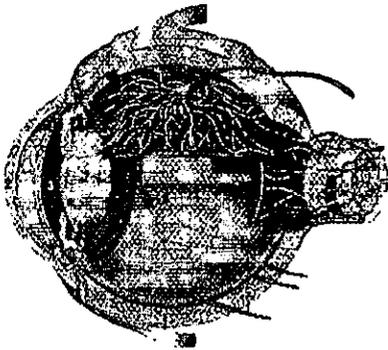
clase de equivalencia basada en formas o funciones, por tanto, asigna propiedades que no son percibidas y tan solo postuladas.¹⁰

Con independencia de sí se quiere emular o no la visión humana bien en la forma o en el funcionamiento- sería bueno preguntarse qué se conoce del sistema de visión humano.

"...El ojo humano viene a ser un globo cuasi - esférico, recubierto exteriormente por la *cornea* y la *esclerótica*, e interiormente por el *coroides* y la *retina*. El coroides termina en el cuerpo ciliar y en la *iris*, que funciona a modo de diafragma limitando la cantidad de luz que entra en el globo ocular a través del orificio central que deja libre llamado *pupila*."¹¹

Por detrás del iris queda el cuerpo lenticular, compuesto de agua, grasas y proteínas que, funcionando a modo de lente convergente (modificable gracias a la tensión de las fibras ciliares) enfoca la luz procedente del exterior sobre de la retina, en la parte posterior del globo (en realidad, la córnea presenta un primer efecto de convergencia a la radiación, actuando como una superficie óptica activa).

El ojo humano es aproximadamente un globo de 2 cm de diámetro que puede rotar en su órbita bajo el control de seis músculos extrínsecos. La luz entra a través de la córnea transparente, pasa a través del humor acuoso que rellena la cámara anterior, llega al cristalino y entonces a través del gelatinoso humor vítreo forma una imagen invertida en la retina.



1. Esclera.
2. Cornea.
3. Iris.
4. Músculo ciliar.
5. Cámara anterior.
6. Aparato Lagrimal.
7. Cristalino.
8. Humor vítreo.
9. Retina.
10. Nervio óptico.
11. Coroides.

Figura 2.4

Delante del cristalino tenemos una membrana opaca llamada iris que da al ojo su color. La luz sólo puede entrar al ojo a través de una apertura circular del iris que se llama

¹⁰ Sobre éste tema en particular se encuentra un tratado en la siguiente dirección:
http://wwwetsi2.urg.es/depar/ccia/public_html/robotica/indice_apuntes.html

¹¹ Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, paginas 18 y 19.

pupila, cuyo tamaño se controla por la expansión y contracción del iris. El cristalino se mantiene en su sitio por los ligamentos suspensorios a través de los cuales el músculo ciliar ajusta la curvatura del cristalino.

El ajuste de la curvatura del cristalino de un ojo por su músculo ciliar se llama acomodación. La acomodación adapta el ojo para obtener la mejor visión a una distancia concreta, (en los ojos sanos, los objetos a esa distancia se forman en la retina, ni delante ni detrás). Cuando los objetos se forman delante o detrás de la retina aparecen borrosos al observador. La miopía describe la incapacidad del ojo para enfocar en la retina objetos distantes del ojo, en este caso los objetos se forman delante de la retina.

La hipermetropía describe la incapacidad de un ojo para enfocar en la retina objetos que están próximos al ojo, estos objetos se forman detrás de la retina.

La retina es una membrana nerviosa compleja con un mosaico de fotorreceptores que cuando es estimulada por la luz produce señales eléctricas nerviosas. Estas señales se transmiten al cerebro a través del nervio óptico, evocando la experiencia de la visión. El lugar de la retina donde todas las fibras nerviosas individuales que constituyen el nervio óptico se juntan, recibe el nombre de disco óptico. Esta región no tiene fotorreceptores y por eso se llama a menudo punto ciego. No lejos del punto ciego existe una zona con una alta concentración de fotorreceptores llamada la fóvea, con un diámetro menor que un milímetro, la fóvea tiene una importancia enorme ya que proporciona la agudeza visual en escenas fuertemente iluminadas. Los objetos se forman en la fóvea cuando dirigimos nuestra mirada hacia ellos.(Figura 2.5)

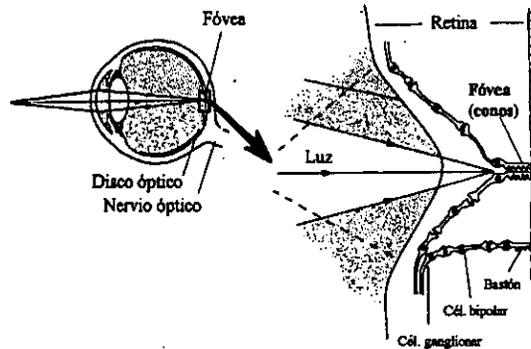


Figura 2.5

La retina, que es menos de un milímetro de gruesa, ha sido el objeto de numerosos estudios. El primer punto importante a notar es que las fibras del nervio óptico se colocan cerca del humor vítreo y, sin embargo, los bastones y conos, los dos tipos de receptores que reciben su nombre de su forma, están localizados cerca de la coroides. Esta disposición es obviamente contra intuitiva ya que obliga a que la luz pase a través de casi toda la profundidad del tejido de la retina que esta lleno de capilares sanguíneos,

células y fibras nerviosas antes de que pueda ser sensada. Sin embargo, en la fovea, al contrario que en el resto de la retina, mucho del tejido retinal entre los fotorreceptores y el humor vítreo está desplazado hacia un lado creando una zona con acceso directo a la luz.

La luz llega a la retina a través del humor vítreo que está en la parte inferior de la imagen y luego pasa a través del tejido retinal para ser detectado por los fotorreceptores que son los bastones y los conos. Los bastones, que son capaces de detectar luz mucho más débil, facilitan la visión en la oscuridad. Por otra parte, los conos, que al contrario de los bastones tienen diferentes variedades, cada variedad con una sensibilidad preferente hacia la luz verde, roja o azul nos proporcionan la visión en colores. La siguiente figura (figura 2.6) muestra la distribución de conos y bastones en la retina humana. En la capa sináptica más exterior, los conos y bastones terminales forman sinapsis con las células bipolares. Una sinapsis es un lugar en el cual una célula del sistema nervioso transmite señal a otra. Las células bipolares llevan a las células ganglionares con las que forman sinapsis en la capa sináptica interna. Es la fibra nerviosa de las células ganglionares la que llega a formar el nervio óptico que lleva la información al cerebro.¹²

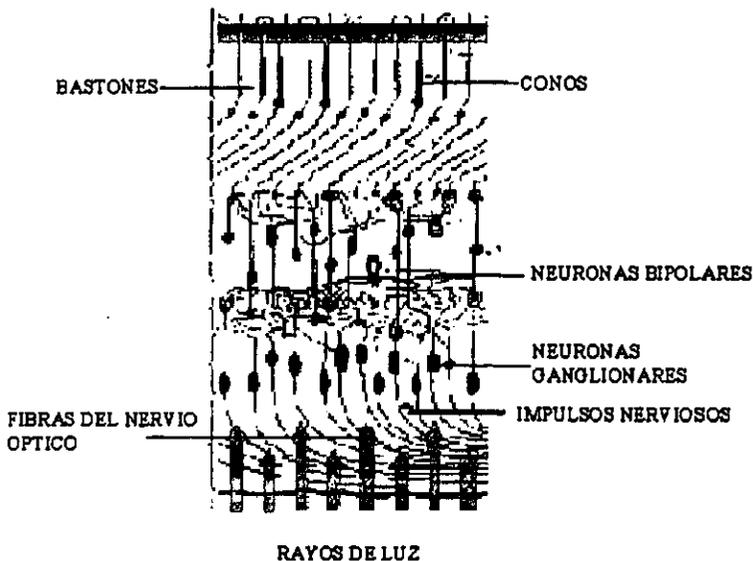


Figura 2.6

El número de fibras ópticas es alrededor de un millón, "... los conos, en cantidades que rondan los seis o siete millones por globo ocular, son sensibles a la longitud de onda de

¹² En la dirección siguiente se encuentra información extensa sobre este tema:
http://www-etsi2.urg.es/depar/ccia/public_html/robotica/indice_apuntes.html

la luz que les llega, esto es, al color. Los bastones (en número mucho mayor, que ronda los 100 millones por ojo) por su parte, son sensibles a los niveles de luminosidad, y su sensibilidad es mucho mayor que la de los conos. Esta es la razón por la que, en condiciones de luz muy débiles, aun se distinguen formas, y no colores...".¹³ Es obvio que la señal generada en los conos y bastones es codificada y compactada antes de llegar al cerebro. Esta compactación se lleva a cabo con la ayuda de células amacrinas y horizontales que proporcionan conexiones sinápticas laterales en las capas interna y externa, respectivamente.

II.3.1 Los caminos visuales al cerebro

Aunque en los dos siglos pasados se ha visto una sustancial mejora en nuestra comprensión de la estructura y funcionamiento del ojo humano, las partes centrales del sistema visual tienen muy poco que no sea un misterio de momento. Es la conversión de las imágenes retinianas en conocimiento del mundo lo que forma la auténtica barrera para nuestro entendimiento de la visión humana. Aunque tenemos idea de los caminos visuales del ojo al cerebro, sabemos poco de lo que ocurre en el cerebro.

La figura a continuación (figura 2.7) ilustra los mayores caminos visuales del ojo al cerebro. De cada uno de nuestros ojos emerge un nervio óptico que transporta señales eléctricas nerviosas del ojo al cerebro. Las fibras que constituyen cada nervio óptico pueden dividirse en dos grupos: las más cercanas a las fosas nasales y las más exteriores.

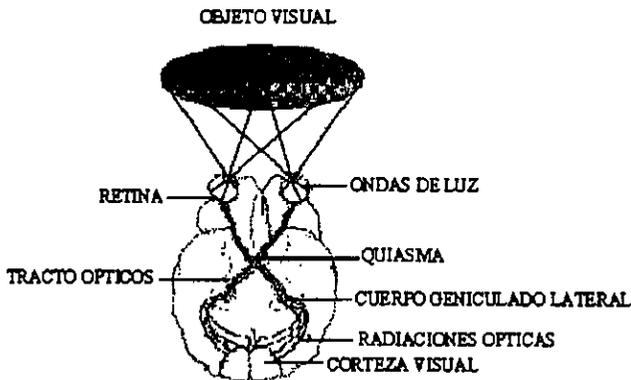


Figura 2.7

¹³ Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, páginas 18 y 19.

Las más exteriores van al mismo lugar del cerebro que en el que se originaron mientras que las más próximas a la nariz cruzan en el quiasma óptico para ir al lado contrario del cerebro al que estaban.

Así pues como se ve en la figura 2.8 el campo visual izquierdo va al lado derecho del cerebro y el derecho al izquierdo. Lo que es importante es que las dos imágenes retinianas de cualquier punto de la escena que es visible a ambos ojos se localizan en la misma región del cerebro. Es la disparidad entre las dos imágenes retinianas la que hace posible la percepción de la profundidad estereoscópica dentro del campo de visión que comparten los dos ojos (el término estereoscópico significa ver los objetos en tres dimensiones debido a la diferencia entre las imágenes formadas en los dos ojos; es importante que tengamos en cuenta que es posible percibir un objeto en tres dimensiones sin el uso de ambos ojos).

Para entender el significado de la visión estereoscópica basta con intentar colocarle el tapón, con una mano, a un bolígrafo que sujetamos con la otra mano a una cierta distancia delante nuestra. Veremos que es mucho más difícil cuando lo intentamos con un ojo cerrado.

Volviendo a los caminos visuales, a partir del quiasma óptico las fibras ópticas siguen en dos grupos llamados los tractos ópticos. Cada tracto comprende fibras que proceden de las mismas partes de las dos retinas.

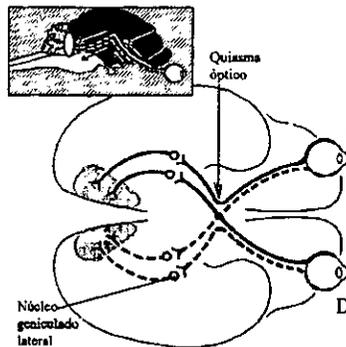


Figura 2.8

La mayoría de las fibras proceden hasta los núcleos geniculados externos que es una región donde se producen las últimas sinapsis. De los dos núcleos geniculados externos las señales son transmitidas vía las radiaciones ópticas hacia la corteza estriada que recibe este nombre por su apariencia a rayas en el cerebro humano. La importancia de la

fóvea se manifiesta en la corteza estriada por el tamaño del área que le dedica en ella. No se sabe mucho de lo que pasa en las regiones de la mente dedicadas a la visión.

Como ya hemos visto la fiabilidad del sistema de visión humano ha sido demostrada ampliamente por la existencia de ilusiones visuales, ambigüedades e inconsistencias.

Uno de los ejemplos de ilusión es la espiral de Fraser que ya hemos visto.

Existen otros ejemplos, como los que se muestran en las figuras de ilusiones ópticas siguientes.(Figura 2.9)

Seis ilusiones ópticas clásicas. En la ilusión de Zöllner las líneas diagonales son paralelas pero no lo parecen, en la de Poggendorff, la línea recta no lo parece, en la de Helmholtz los cuadrados parecen rectangulares, en la de Müller-Lyer las líneas horizontales parecen de diferente longitud y en los de Hering y Wundt las líneas rectas parecen curvas.

Otro problema es que una figura puede tener más de una posible interpretación, en otras palabras, la figura puede ser ambigua. Múltiples interpretaciones pueden coexistir o una puede dominar a la(s) otra(s).

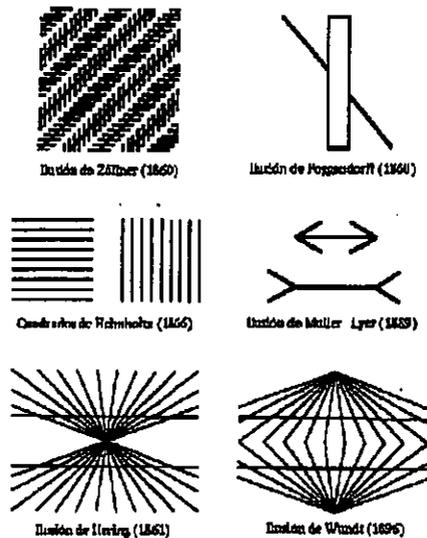


Figura 2.9

Finalmente, es posible que, aunque percibida como una figura que ni es ambigua ni corresponde a una ilusión óptica, la figura sea completamente irrealizable en el sentido de que no podemos construir el objeto 3-D percibido en el espacio real 3-D.

Las diferentes ilusiones, ambigüedades e inconsistencias que hemos analizado son mucho más que curiosidades. Nos llevan a preguntarnos si inferimos de nuestras imágenes en la retina más de lo que de verdad está soportado por la geometría y física de la formación de imágenes. Helmholtz en su libro *Handbook of Physiological Optics* publicado a mediados del siglo pasado expresó que cada imagen es una imagen de algo sólo para aquel que conoce como leerla, y que está capacitado con la ayuda de la imagen a formar una idea de la cosa. La implicación de esta afirmación es que no es como si el sistema visual humano estuviera haciendo inferencias precisas y exactas basadas en la física de la formación de imágenes en el ojo, sino que el sistema visual invoca reglas que se obtienen y están sesgadas por la experiencia previa del individuo y tal vez por la especie. Como resultado los humanos podrían ver lo que no hay y no ver lo que hay. Esto nos lleva como humanidad a preguntarnos, ¿le permitiríamos a una máquina hacer lo mismo?¹⁴

Aunque esta última pregunta no tenga aún una respuesta existen otras que si la tienen, por ejemplo ¿Qué es la Luz?, Se ha hablado de ella y de su importancia pero no se ha dado una definición formal en éste trabajo, ya que todo el proceso parte desde ahí.

II.4 La Luz.

La luz es la suma de todas las radiaciones monocromáticas del espectro visible. "...La luz es una radiación electromagnética que viene caracterizada por su frecuencia, en hertzios (Hz), o por su longitud de onda λ , en nanómetros (nm) o en micrómetros(mm). En concreto, la luz visible se distribuye en torno al rango de longitudes de onda:

$$350 \text{ nm} \leq \lambda \leq 780 \text{ nm}$$

donde cada λ representa un color distinto...".¹⁵

La sensación que conocemos como *luz blanca* no es mas que un efecto puramente fisiológico. Es decir, "... el estímulo se recibe por el sentido de la vista y se procesa en el cerebro creando en éste la sensación de luz blanca...".¹⁶

La luz blanca es una mezcla de luz de color, "... lo que se demuestra descomponiendo los rayos solares mediante un prisma. Los colores de la descomposición son el rojo, naranja, amarillo, verde, azul y violeta...".¹⁷

¹⁴ Acerca de este tema en particular se puede encontrar la información completa en la siguiente dirección:
http://www-etsi2.urj.es/depar/ccia/public_html/robotica/indice_apuntes.html

¹⁵ Alberto Domingo Ajenjo, *Tratamiento Digital de Imágenes*, Anaya Multimedia 1994, pagina 19.

¹⁶ Vicente Llorens, *Fundamentos Tecnológicos de Video y Televisión*, Editorial Paidós, España 1995, Objetivo 5.1.

¹⁷ Otto Cimann, *Fundamentos de Televisión*, Ed. Marcombo Boixareu, España 1989, pagina 17.

“... El ojo percibe cualquier objeto gracias a la luz que refleja del total que sobre él incide. En general, la fracción reflejada depende de nosotros lo que determina el color del objeto tal y como lo percibimos...”¹⁸

Esto quiere decir, que cada color tiene una longitud de onda diferente y que el ojo es capaz de distinguir cada una de ellas por lo que es aplicable el siguiente concepto: “...La diferenciación en frecuencias persiste más allá del rango de radiación visible, donde carece de sentido hablar de color, asumiremos de aquí en adelante la equivalencia frecuencia - color para nuestros propósitos...”¹⁹

Para poder medir estas longitudes de onda y explicar los fenómenos que se desprenden del color se analizará un poco más a detalle a continuación.

II.4.1 Colorimetría.

En el intento del hombre por explicar la naturaleza de los colores y sus relaciones mutuas, se estudiaron las propiedades naturales, medicinales, físicas, estéticas, etc. del mezclado de los mismos, observando los colores que pueden obtenerse con la mezcla de pigmentos.

Esto lo podemos constatar desde los tiempos de las cavernas en sus pinturas a los restos arqueológicos de Egipto, India, América Central y del Norte, China y Grecia, entre otros, con la maravillosa sensación de color que tenían estos hombres.

“...Prácticamente todos los colores pueden ser producidos como las combinaciones de tres colores llamados Primarios Aditivos y son: El Rojo, El Verde y El Azul...”²⁰

Pero de nada servirían los colores si en realidad no sabemos que son o de donde provienen, y la respuesta es la Luz.

II.4.2 Fundamentos del Color

En 1666 Isaac Newton descubrió que cuando un rayo de luz solar pasa a través de un prisma de cristal, el rayo de luz que sale no es blanco sino que está formado por un espectro continuo de colores que van desde el color violeta al rojo. Podría decirse que el espectro de color puede dividirse en seis amplias regiones: violeta, azul, verde, amarillo, naranja y rojo. Sin embargo, cuando se ven todos los colores juntos ningún color del espectro termina bruscamente, más bien se mezcla suavemente con el siguiente.

¹⁸ Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, página 19.

¹⁹ Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, página 20.

²⁰ Manual de Servicio Sony, página 35.

La idea del color es inicialmente muy simple, los colores que los seres humanos percibimos en un objeto vienen marcados por la naturaleza de la luz reflejada del objeto. Un cuerpo que refleja luz relativamente equilibrada en todas las longitudes de onda del visible aparece blanco. Un objeto que refleja en un rango limitado del espectro visible muestra algún tipo de sombra de color.

La caracterización de la luz es fundamental para la ciencia del color. Si la luz no tiene color su único atributo es la intensidad. Luz sin color es lo que vemos en los televisores en blanco y negro.

Para una luz con color, el espectro electromagnético de luz visible va de 400 a 700 nm (aproximadamente). Se usan tres cantidades para describir una luz cromática: radiancia, luminancia y brillo. La luminancia se mide en lúmenes y es una medida de la cantidad de energía que el observador percibe. Por ejemplo, la luz emitida por una fuente que opera en el infrarrojo puede tener mucha energía, radiancia, pero no es visible por el observador y, por tanto, su luminancia es nula. Por último, el brillo es un descriptor subjetivo que se puede medir difícilmente, lleva asociado el concepto de luz sin color y es el responsable básico de la sensación de color.

Debido a la estructura del ojo humano, todos los colores se ven como una combinación de los tres llamados colores básicos: Rojo (R), verde (G) y azul (B). En esta obra usaremos la notación inglesa (R, G y B) para referirnos a cada uno de los tres planos de color. Con la idea de producir un estándar la Comisión Internacional de Iluminación (CIE) designó en 1931 las siguientes longitudes de onda para los tres colores primarios: azul = 435.8 nm, verde = 546.1 nm y rojo=700 nm, sin embargo, es obvio que un solo color no puede ser llamado rojo, verde o azul. Por tanto, tener tres longitudes de onda específicas para los tres colores no significa que estas tres componentes trabajando solas puedan generar todos los colores. Esto es importante, ya que se cree, erróneamente, que estos tres colores mezclados pueden producir todos los colores visibles. Esto no es cierto salvo que también permitamos a las longitudes de onda que varíen.

Si sumamos los colores primarios produciremos los colores secundarios, magenta (rojo más azul), cian (verde más azul) y amarillo (rojo más verde). Si mezclamos los tres colores primarios o un secundario con su color primario opuesto en las intensidades correctas se tendrá luz blanca. Ver la figura 2.10 que muestra los colores primarios y sus combinaciones para obtener los secundarios.

Colores primarios.

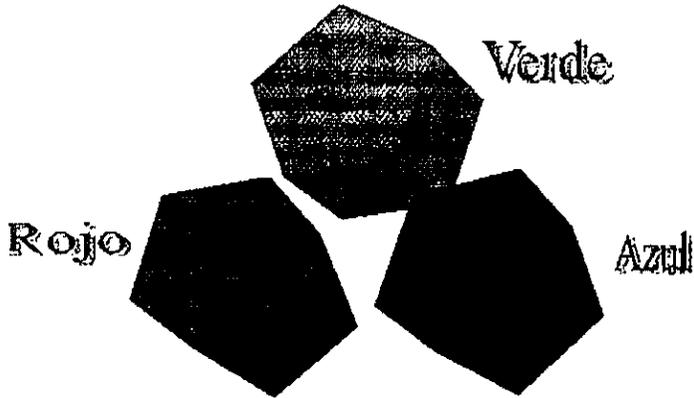


Figura 2.10

Diferenciar entre los colores primarios de luz y los colores primarios de pigmentación o colorantes es importante. En este caso, un color primario se define como uno que absorbe un color primario de luz y transmite los otros dos. En este caso, los colores básicos de pigmentación son magenta, cian y amarillo y los secundarios son rojo, verde y azul. Una combinación apropiada de las tres pigmentaciones primarias produce el negro.

Colores secundarios.

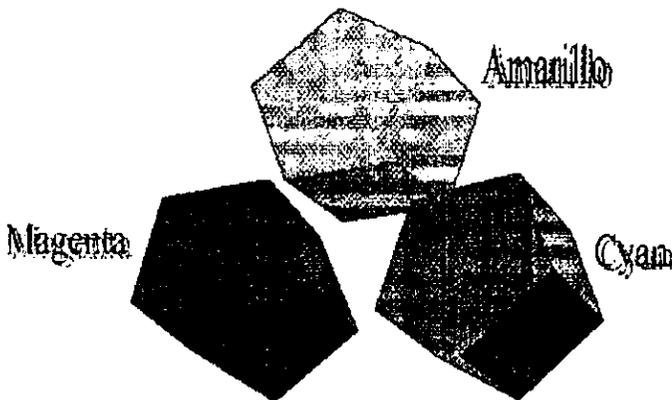


Figura 2.11

La recepción de televisión en color es, como sabemos, un ejemplo de la naturaleza aditiva de los colores.

Las características que generalmente se usan para distinguir un color de otro son el brillo, el color (matiz) y la saturación. El brillo lleva asociado el concepto de intensidad. El matiz es un atributo asociado con la longitud de onda dominante en la mezcla de las ondas luminosas. La saturación se refiere a la cantidad de luz blanca mezclada con el matiz.

Los colores puros del espectro están completamente saturados. Colores como el rosa están menos saturados, siendo la saturación inversamente proporcional a la cantidad de luz blanca que se añade. Un ejercicio interesante es editar en el entorno de ventanas las propiedades de los mismos y ver como estos tres conceptos se relacionan.²¹

El color y la luz son dos cosas distintas, el color en los objetos no es luz, "... La luz visible se compone de oscilaciones electromagnéticas con longitudes de onda de 780 hasta 390 nm (nano metros).

380 ...450 nm - violeta
 450 ...482 nm - azul
 497 ...540 nm - verde
 575 ...580 nm - amarillo
 585 ...595 nm - naranja
 600 ...780 nm - rojo
 nm = 10 exp -9 metros..."²²

En la televisión en color se utilizan tres señales cromáticas: azul, rojo y verde. Mezclando estos tres colores primarios se forman todos los demás y a la inversa se puede descomponer la luz de color en estos tres colores.

Otro concepto que esta ligado a la percepción de imágenes en el sistema visual humano es la **luminancia**, "... La luminancia, o energía emitida por una unidad de área en una unidad de ángulo sólido, es la magnitud fotométrica que corresponde con la sensación de claridad o *brillo*, cuando esa energía se percibe en el ojo..."²³

A su vez de la luminancia nace el concepto de **contraste**, "... El concepto de contraste nace; por su parte de la diferencia entre los valores de luminancia de dos objetos contiguos. Aunque dos objetos presenten igual luminancia, los brillos aparentes (luminancia subjetiva) que de ellos se aprehenden pueden ser considerablemente distintos si se advierten diferencias importantes con las luminancias del entorno..."²⁴

²¹ Fundamentos de Luz y Color: [Http://www-etsi2.ugr.es/depar/ccia...html/robotica/indice_apuntes.html](http://www-etsi2.ugr.es/depar/ccia...html/robotica/indice_apuntes.html)

²² Otto Cimann, Fundamentos de Televisión, Ed. Marcombo Boixareu, España 1989, pagina 17.

²³ Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, pagina 20.

²⁴ Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, pagina 21.

Estos conceptos, la luminancia, el contraste, el brillo etc. se enlazarán a otros en el capítulo siguiente, en donde se expondrán de manera más amplia ya que forman parte importante de la señal de video.

11.4.3 Luz Visible.

La luz es la radiación de energía electromagnética de frecuencia media que el ojo humano interpreta como tal. La sensación de color está directamente relacionada con la frecuencia de la luz observada. Así, al cambiar las frecuencias de las ondas electromagnéticas que constituyen la luz, aparece un cambio de color visto por el ojo.

La luz del sol, que contiene todas las frecuencias del espectro visible, se considera por esta razón que es de luz blanca, ya que la suma de todos los colores nos da el blanco. Esto puede observarse a través de un prisma.

Espectro Electromagnético.

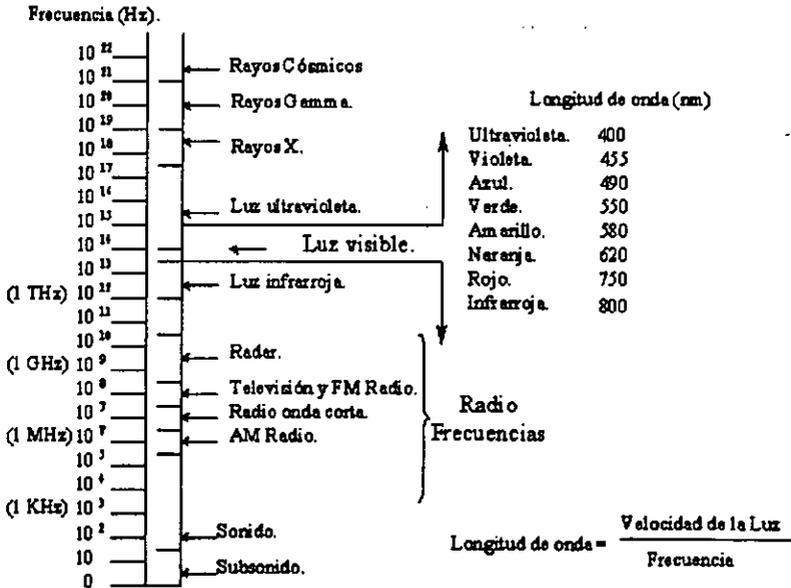


Figura 2.12

Cada color sale del prisma formando ángulos distintos, y esto depende de la frecuencia individual del color. A mayor frecuencia, mayor será el ángulo de refracción esto se conoce también como "Descomposición de la luz".

Los colores en la naturaleza no son puros, ya que no están formados por una sola frecuencia. Una lámpara roja emite luz roja, aparentemente, pero la mayoría de las frecuencias emitidas se encuentran en la banda Roja del espectro. Un color puro estaría formado por una frecuencia, pero tales condiciones son raras y así permite suponer que realmente los colores están formados o contaminados por otras frecuencias.

Una luz monocromática está formada por ondas de la misma longitud de onda.

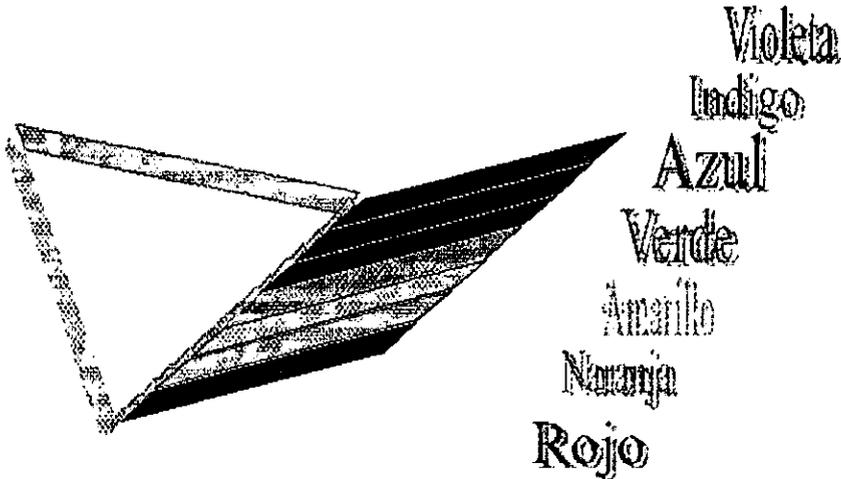


Figura 2.13 Descomposición de la luz.

Una luz policromática es la luz blanca que produce el sol ó lámparas incandescentes.

En la figura 2.13, se tiene la Descomposición de la luz al pasar a través de un prisma y los colores que la componen.

El color de un cuerpo iluminado depende de su constitución molecular, que determina su capacidad para reflejar distintas longitudes de onda luminosas y de la clase de luz empleada para iluminarlo. Entonces un objeto refleja la porción del espectro cuyas frecuencias son características de su color, mientras que absorbe todas las demás. Una hoja verde refleja el verde si se ilumina con luz blanca y absorbe las demás frecuencias, pero no reflejará su color característico si las frecuencias correspondientes a ese color (verde) no se hallan en la luz. Por ejemplo, si se ilumina con luz roja la hoja verde, se verá negra, al no tener componentes verdes en la luz roja y esto es un proceso sustractivo. Un objeto no presentará su color característico si las frecuencias correspondientes a ese color no se hallan presentes en la luz emitida; dicho objeto presentará una tonalidad negra, ya que no habrá componentes correspondientes al verde en la luz roja emitida.

Otro proceso sustractivo lo presentan los cuerpos transparentes, que dejan pasar un color y absorbe los demás.

En un proceso aditivo, la sensación de luz blanca se forma con la suma de las radiaciones de los colores aditivos primarios rojo verde y azul.

Si se sobrepone un filtro amarillo y otro turquesa o cian, en el área de empalme se ve el color verde, lo que se interpreta como la absorción de todos los colores de la luz blanca excepto el verde. Al sobreponer el magenta y el turquesa o cian, se ve el color Azul; al empalmar el Magenta y el Amarillo se verá solamente el Rojo. Donde se empalman los tres colores se verá una superficie **NEGRA**, lo que nos indica que todas las componentes de luz han sido absorbidas. Debido a este proceso "... los colores **AMARILLO, MAGENTA Y CIAN (TURQUESA)** se les conoce como colores **PRIMARIOS SUSTRATIVOS...**"²⁵ (figura 2.14).

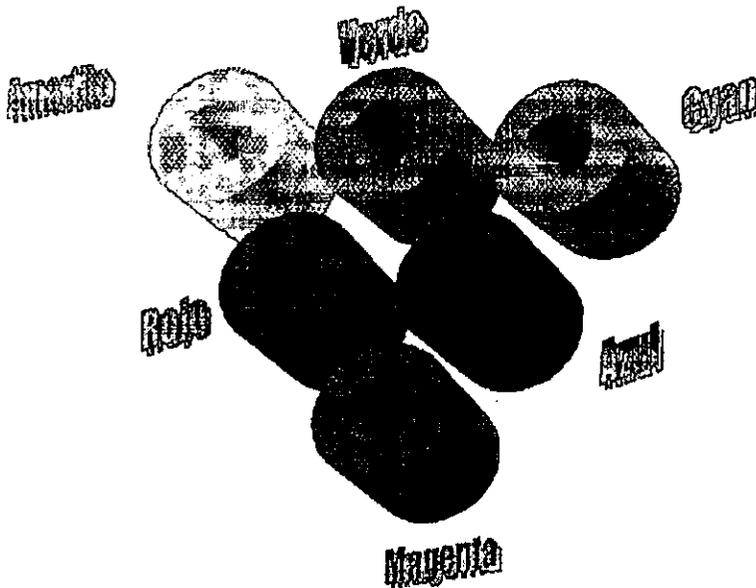


Figura 2.14

Las reglas para un sistema aditivo, donde las longitudes de onda de los tres colores primarios generados, son opuestos a los colores primarios sustractivos. Al proyectar en una pantalla blanca los tres colores que son el VERDE, el ROJO y el AZUL, se observa lo siguiente:

²⁵ Manual de Servicio Sony, pagina 38.

En donde se empalman el Rojo y el Verde se ve el color Amarillo; al empalmarse el Rojo y el Azul, se obtiene el Magenta, y donde se empalman el Azul y el Verde se ve el Cian, y donde coinciden los tres colores se observa un área BLANCA que como se sabe es la reunión de todos los colores del espectro, por lo que reciben el nombre de colores PRIMARIOS ADITIVOS.

Variando la cantidad de luz de un color aditivo, respecto al otro, se obtienen diferentes tonos; al combinar en diferentes proporciones los tres, se obtienen todos los colores que el ojo humano puede apreciar, incluyendo el Blanco, para el cual las proporciones son:

30% ROJO + 59% VERDE + 11% AZUL = 100 % BLANCO

Como se pudo observar, la luz blanca esta compuesta por frecuencias que se denominan: colores, estos a su vez tienen características que el ojo humano es capaz de percibir y con esto diferenciar tonos, que tan intensos o puros pueden ser.

Para adentrar en la revisión de los orígenes en la generación de una imagen, en este caso como lo menciona el nombre de este trabajo de una imagen de televisión en el capítulo siguiente se tratará la descripción completa de la generación de una imagen en la pantalla de un televisor.

CAPITULO III
FUNDAMENTOS DE VIDEO Y TELEVISIÓN.

III. FUNDAMENTOS DE VIDEO Y TELEVISIÓN.

III.1 Imagen estática e imagen dinámica. Movimiento.

Existe otra consideración dada con el manejo de imágenes: *el movimiento*. "...Una imagen representa, un mapa de la función intensidad de luz en una zona concreta de un espacio bidimensional, mapa que viene dado por la función $f(x, y)$. Esta imagen se denomina estática, por cuanto representa dicho mapa en un único instante de tiempo.

Para lograr un efecto de movimiento, el mapa debe de actualizarse en momentos sucesivos para mostrar la variación de la función en el tiempo, esto es, la función pasa de bidimensional a tridimensional, considerando la variable t (tiempo), en la función siguiente: $f(x, y, t)$.²⁶

Conocido es el hecho de que el ojo humano retiene la última impresión visual recibida durante un tiempo cercano a 1/24 segundos, por lo cual, para evitar efectos de parpadeo, la frecuencia de actualización de la función de intensidad lumínica debe ser, cuando menos mayor de 24 Hz.

Esta retención lumínica es uno de los principios que utiliza una cámara de Video o Cine para que se pueda apreciar un movimiento continuo al generar una imagen pero, no aunado a esto existen muchos conceptos para poder observar una imagen en la televisión, la parte inicial de un sistema de imágenes de este tipo es precisamente "la cámara".

A continuación se tratará el funcionamiento a fondo de la cámara de TV y las señales que produce y así obtener una imagen analógica.

III.2 Cámara de Televisión.

Es el componente más importante del equipo de TV debido a que origina la señal de video correspondiente a la imagen.

Sabemos que un tipo de energía, por ejemplo el calor, se utiliza para crear otra forma de energía, la luz, y a su vez, esta energía lumínica se convierte, ya codificada en energía eléctrica.

"...En la naturaleza existen ciertos elementos que reaccionan al incidir sobre ellos rayos de luz, por lo que son fotosensibles y tenemos al cesio. La energía de la luz disloca a los electrones del material y los deja en libertad, pudiéndose generar una corriente eléctrica.

²⁶ Alberto Domingo Ajenjo. Tratamiento Digital de Imágenes, Anaya Multimedia 1994, paginas 23, 24..

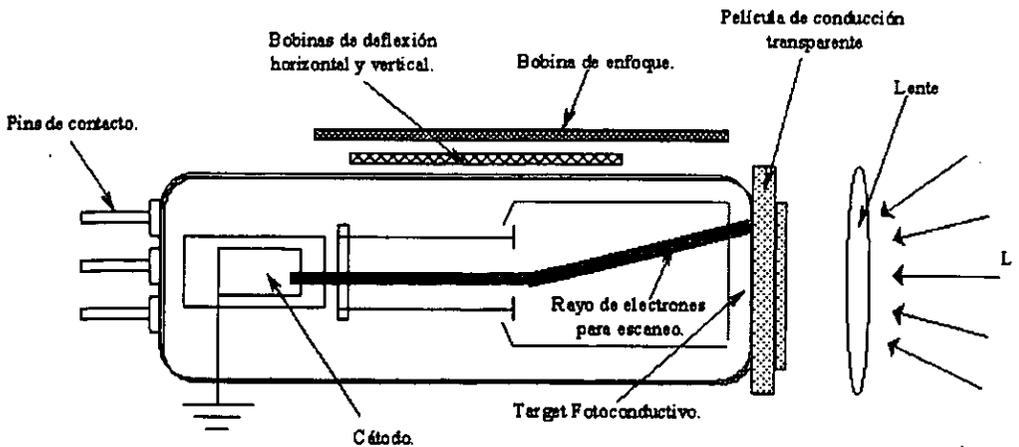
Otros elementos como el selenio, al exponerse a la luz varían su resistencia a la corriente eléctrica y en consecuencia varía su conductividad (Fotoconductividad)...²⁷

Antes de 1986, las cámaras de TV utilizaban tubos electrónicos llamados tubos de cámara que reunían las condiciones de ser fotoeléctricos y fotoconductivos. Transforman las señales ópticas concentradas en una superficie fotoconductiva llamada target gracias a las lentes, y los pixeles de ésta, se convierten en muchas cargas diminutas eléctricas individuales. En este patrón de imagen de carga, el Target, es explorado en líneas horizontales de izquierda a derecha y de arriba hacia abajo por un rayo llamado haz de electrones, los que "leen" la información contenida en el target y la convierte en una serie consecutiva de señales eléctricas, proporcional en amplitud al brillo de los elementos de imagen; lo que en otras palabras significa que la conversión de la imagen completa en una señal de video se realiza por un barrido de ésta en el target, pixel por pixel en forma secuencial, lo que tarda un tiempo pequeño en barrer toda la escena que contendrá un "cuadro" de imagen.

La función del haz de electrones es el de descargar cada punto (pixel) en el target de la imagen, para producir una corriente eléctrica a la salida del tubo. Esta corriente es muy pequeña (décimas de Amper), por lo que se utiliza un preamplificador a la salida de éste, de alta ganancia y bajo ruido. El barrido del haz se corta durante el retorno de cada línea horizontal y al llegar al punto inferior de la imagen a esto se le conoce como espacio de borrado (Blanking) por lo que no será visible.

Después del preamplificado, sigue un procesador de señal y un sumador de sincronía.

Figura 3.1 Operación de un tubo vidicón.



²⁷ Manual de Servicio Sony, pagina 13.

El procesamiento de la señal corrige sombreados indeseados y provee una relación de contraste deseada. Este sombreado se debe a las características no lineales o uniformes de la placa fotosensible del tubo, por lo que se le da una corrección para un contraste deseado llamada relación de Gamma (refuerza el blanco en la imagen reproducida, y esto es similar a la expansión y compresión del volumen en audio). También existen reductores de ruido y amplificadores de detalle.

En la parte final del procesamiento de la señal se preveió mantener el voltaje (clamping) de las partes borradas (del Blanking) para tener niveles de referencia; y por último se tiene la señal de video compuesta resultante con sus variaciones, pulsos de borrado y sincronía.

Los tubos de cámara eran de varios tipos en la construcción del target y son:

1. El vidicón, estaba hecho de trisulfuro de antimonio.
2. El plumbicón (Philips), similar al anterior y estaba fabricado con óxido de plomo.
3. El saticón (Hitachi), su placa es de selenio, arsénico y telurio.
4. El silicon-vidicon, Utiliza una unión semiconductor de silicio en el target.
5. Chalnicon (Toshiba), tiene un arreglo multicapa de óxido de estaño, selenuro de cadmio.
6. El newvicón (Matsushita), el target tiene capas de zinc - selenio amorfo cubiertas de 1 trisulfuro de antimonio.

Todas estas variantes de tipos de tubos, se realizó para tener más sensibilidad a la luz, ya que los tubos requieren mucha intensidad de luz y se construyen similar al vidicón. (Ver figura 3.1).

III.2.1 Dispositivos Acoplados por Carga (CCD).

Son dispositivos de estado sólido para generar o convertir las imágenes en señal de video, por lo que se desarrollaron en los principios de la década de los años ochenta y ya a la venta desde 1986; y sustituyeron a los tubos de cámara, por lo que su ventaja con respecto a los tubos es de disminuir considerablemente el consumo de energía, peso, tamaño, errores geométricos y de resolución.

Existen tres tipos de CCDS, según su funcionamiento y son:

a) CCD de transferencia de cuadro (FT).

El CCD usa miles de elementos sensibles a la luz en un arreglo tipo matriz para detectar la intensidad de luz en varios puntos de la imagen enfocada. El brillo de una imagen dada, determina la cantidad de cargas eléctricas generadas por un elemento de imagen, las que permanecen en cada elemento hasta que una señal de control comande a las cargas de video para transferirlas a unos elementos de "transporte" adyacentes. Los que

emplean la misma señal de control para comandar cualquier carga existente para evacuarla y hacer un hueco para la siguiente carga.

La técnica de transferencia de cargas se le conoce como "la brigada de cubetas" porque es similar a la de bomberos al apagar un fuego. Las cubetas de agua se transfieren de un miembro de la brigada a otro. Las transferencias del CCD llevan las cargas de un elemento de la matriz al siguiente para crear la señal de video.

Esta transferencia de cargas dentro del CCD crea un "barrido" de la imagen enfocada. Durante la transferencia de carga, el CCD debe de borrarse o ser insensible a la luz porque los elementos de la matriz usados para transferir las cargas a veces se usan como elementos de imagen.

Las cargas que representan la información visual se transfieren rápidamente desde la matriz de imagen a una segunda matriz de almacenamiento temporal; donde las cargas se mueven "lentamente" fuera de la matriz a otra de transferencia de salida adyacente antes de salir del dispositivo.

La transferencia de la matriz de almacenaje a la matriz de salida sucede en un rango de frecuencia de barrido vertical, o sea, un grupo de cargas correspondiente a un barrido horizontal de video, se transfiere a la matriz de salida a la vez. Todas las cargas en la matriz de almacenaje se transfieren a la matriz de salida en fracciones de 1/60 de segundo para tener el efecto de barrido vertical.

Esto puede ayudar a visualizar el proceso de transferencia de cargas considerando que las cargas de un barrido horizontal de video se envían, las cargas del siguiente barrido horizontal se transfieren como se necesiten.

El ciclo de transferencia de cuadro es como sigue:

1. La matriz sensible a la luz se expone a la imagen en un tiempo.
2. Durante el borrado vertical, la matriz visual se protege de la luz por un medio mecánico o electrónico llamado Shutter (obturador) y la información se lleva rápidamente a una matriz de almacenaje temporal.
3. Después del borrado vertical, la matriz visual es expuesta otra vez a la información de luz, mientras que la información almacenada en la matriz de almacenaje temporal es leída "lentamente", un barrido a la vez para crear 262.5 trayectorias horizontales efectivas, barridas 60 veces por segundo.

b) CCD por Transferencia de Entrelínea (IT)

Este CCD también usa el concepto de la brigada de las cubetas, pero los elementos no sensibles a la luz de transferencia se colocan en la matriz de imagen adyacentes a los elementos de imagen. Las cargas que representan y los brillos de imagen se mueven a los elementos adyacentes de transferencia y bajan en un registro de salida.

Los elementos de transferencia adyacentes mejoran la respuesta para cambiar rápidamente el brillo de la escena pero reduce la habilidad del CCD para detectar detalles de imagen en el eje horizontal de la misma.

Esta técnica de barrido de transferencia de carga permite al operador variar el tiempo de exposición del área sensible a la luz sin alterar el rango de barridos de TV; esta variación de tiempo de exposición puede variar la velocidad del Shutter, es decir, de obturación.

La habilidad de los CCDs le permite manejar de mejor manera grandes relaciones de contrastes a altos niveles de luz, sin problema, cosa que en los tubos no pueden aplicarse. No le provoca daños permanentes, pero genera una línea blanca que se extiende verticalmente, debido al sobreflujo de cargas.

Estructura simplificada de un CCD.

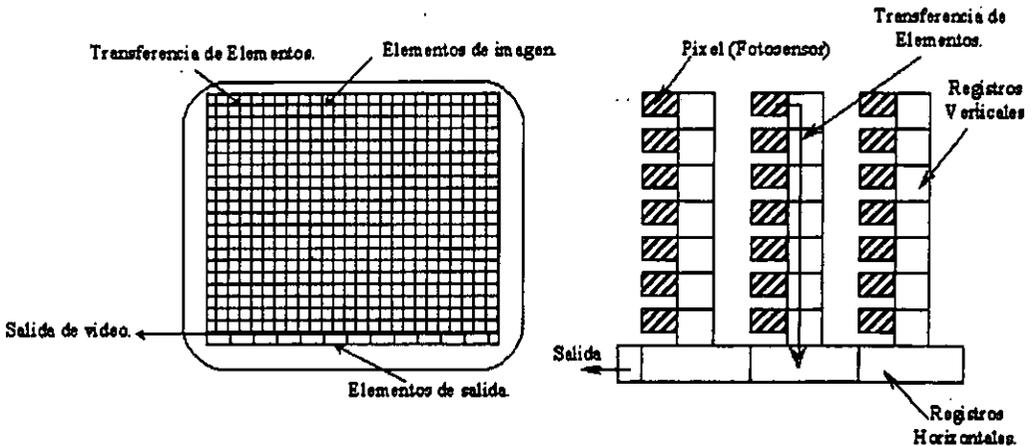


Figura 3.2

En la figura 3.2 se observa la transferencia de elementos hacia la salida del CCD.

c) CCD por Transferencia de Entrelínea y Cuadro (FIT)

Este resulta de la combinación de los anteriores CCDS FT e IT.

III.2.2 Características y Deficiencias de los primeros CCDS.

Las principales son: Blooming, Smear y Lag.

1. **Blooming:** Es la línea blanca que aparece en las imágenes con destellos muy luminosos como se dijo, no puede almacenar el fotosensor una cantidad demasiado grande de cargas, por lo que va a parar al registro vertical contaminándolo; se solucionó agregándole un drenaje de rebasamiento al lado del fotosensor, lo que le disminuye la fotosensibilidad por el espacio del drenaje.
2. **Smear.** Se presenta cuando incide luz infrarroja en el fotosensor (longitud de onda grande) y crea cargas dentro de éste, las que se difunden en el registro vertical. Se soluciona con filtros infrarrojos, lo que afecta la sensibilidad de luz por ser éstos algo gruesos, por lo que se implementa un drenaje eliminando el smear.
3. **Lag:** Se produce por la mezcla de información de dos campos. Debido a la permanencia de algunos electrones se quedan ahí, por lo que se deben eliminar por completo.

Estos casos se aplicaban para los diferentes tipos de CCD incluyendo CCD por Transferencia de Entrelínea y Cuadro mencionado en el inciso C del párrafo anterior.

III.2.3 Exploración y barrido de una imagen.²⁸

Este proceso comienza cuando el haz de electrones barre la imagen enfocada en el target o en el caso de los CCDS, la imagen se convierte en cargas eléctricas las que son transferidas en forma de "barrido horizontal" de manera similar al tubo.

Después de completar cada barrido horizontal de izquierda a derecha en el target el haz se oscurece o "borra" momentáneamente en el retorno de derecha a izquierda para comenzar la nueva línea horizontal, y de similar manera, el haz se oscurece al ir de la parte inferior a la parte superior de la imagen.

La cámara de TV idealmente debería producir una imagen instantánea, al igual que una fotografía, pero no es así, igual que el ojo humano ante una hoja impresa, si éste permanece fijo, la información es muy pequeña e imprecisa, pero si se pone en movimiento por la página, recogerá toda la información contenida en ésta, y se le conoce como exploración rectilínea. Los sistemas de TV utilizan entonces esta exploración rectilínea.

El sistema Norteamericano de TV que empleamos se llama NTSC 525/60.

NTSC significa Comité Nacional para el Sistema de TV y nos indica que para la exploración de la imagen tanto en el transmisor como en el receptor se tienen los siguientes estándares:

²⁸ Descripción del proceso, Manual de Servicio Sony, páginas 21,22 y 23.

Hay dos procesos de exploración seguidos simultáneamente:

Uno que se mueve en sentido Horizontal y, el otro se mueve en sentido Vertical. Estas exploraciones son lineales, esto es a velocidad constante.

El estándar del sistema M del Comité Consultativo Internacional de Radio (CCIR) que se ha adoptado, para un sistema Blanco y Negro es el siguiente:

- 525 Barridas o trayectorias Horizontales llamadas Líneas H por Cuadro de imagen (imagen completa).
- 262.5 Barridas o Líneas H por Campo.
- 30 Cuadros de imagen completa por segundo.
- 60 Campos (medio cuadro o imagen) por segundo.

Estos valores variarán con respecto a la información de color.

Escaneo necesario para el entrelazado de líneas en los dos campos.

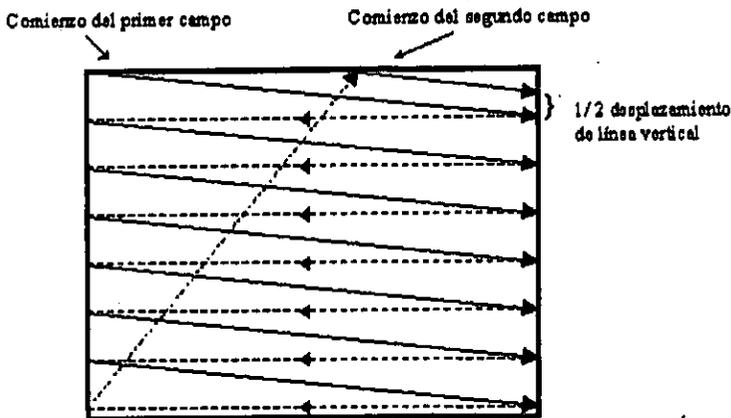


Figura 3.3

En la figura anterior (figura 3.3), podemos observar que las líneas continuas representan la señal eléctrica correspondiente a una parte de la imagen (Líneas H), y las líneas punteadas indican el retorno del haz electrónico, el cual es más rápido que el trazo.

El barrido H (correspondiente a las líneas H) es más rápido que el Vertical. Esto es mejor para que el haz pueda efectuar un número adecuado de barridos H, pero no debe ser muy baja o aparecerá un Flicker indeseable.

Un Cuadro (FRAME) de imagen es la cantidad mínima de movimiento de una imagen completa de una escena.

Este cuadro de imagen estará formado de 525 líneas H y cada cuadro se explorará dos veces para evitar el Flicker, ya que la frecuencia de Cuadros por segundo es de 30 Hz, lo que provocará Flicker. Así, cada exploración Vertical (112 cuadros) será de 262.5 Hz y se le llamará Campo.

De lo anterior se tiene que la frecuencia de exploración vertical o de campo es de 60 Campos/seg. y coincide con la frecuencia de la corriente eléctrica de la Compañía de luz, lo que disminuye la interferencia de zumbido (HUM) en la imagen.

La frecuencia Horizontal es de:

$$525 \text{ H / Frames} \times 30 \text{ Frames / SEG} = 15,750 \text{ H / SEG}$$

$$262.5 \text{ H / Field} \times 60 \text{ Field / SEG} = 15,750 \text{ H/SEG}$$

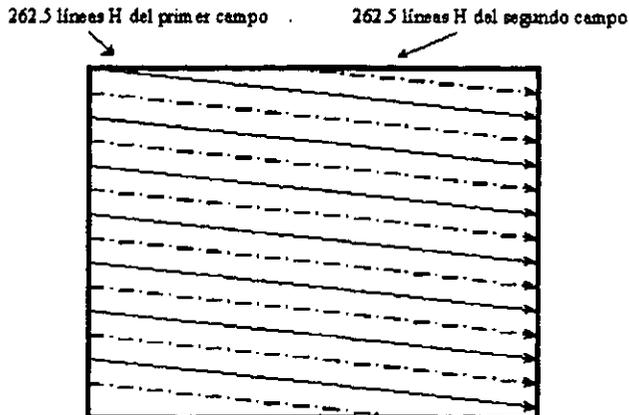
Field = Campo.

Frame = Cuadro.

Como el número total de líneas H por campo (262.5 H) no es número entero, se puede pensar que al sobreponer un campo con el anterior, puede no existir una correspondencia en la situación de las líneas H.

Esto no es así, ya que el número de líneas obliga a que cada campo comience con media línea de diferencia con respecto al anterior, lo que provoca que haya un entrelazado entre las líneas correspondientes a dos campos consecutivos.

Cada campo, leerá la mitad de las líneas de un cuadro, pero un campo leerá las líneas numeradas como nones, por lo que se llamará Campo Non y el otro campo leerá las líneas pares, llamandose Campo Par. Las líneas trazadas por el campo 2 (par) cubren los espacios entre las líneas del campo 1 (non) realizándose así el entrelazado. A continuación la Figura 3.4.



Veamos un ejemplo sencillo de esto:

- (1) Las líneas de barrido H se entrelazan a las pares.
- (2) Todas las líneas impares son barridas, omitiendo las líneas pares. Entonces.
 - (1) el sistema de TV con el fin de proveer dos campos comienza antes.
 - (2) Cuando las líneas pares son barridas para completar el campo,
 - (1) produce vistas de la imagen para cada cuadro de imagen.
 - (2) Terminando se puede ver un cuadro completo sin perder cualquier información de imagen

Como se vio en la figura, al principio del campo 1, éste principia en la parte superior izquierda y termina en el centro de la parte inferior y el campo 2 empieza en la parte central superior y termina en la parte derecha inferior, tanto en la cámara como en el receptor.

Los tiempos de duración de Un Campo es de $T = 1 / 60 \text{ Hz} = 0.016 \text{ Seg. (16 mseg)}$.
 Los tiempos de duración de Un Cuadro es de $T_f = 1 / 30 \text{ Hz} = 0.033 \text{ Seg. (33 mseg)}$.
 El período de 1 H es: $1 / 15,750 \text{ Hz} = 63.49 \text{ useg}$.

III.2.4 Entrelazado.

Ya que 30 cuadros por segundo (Hertz), progresivamente barridos producirían un parpadeo (Flicker) inaceptable en un área grande, la estrategia del entrelazado fue introducido en un principio, doblando el rango de Flicker a 60 Hz y así se preserva el número completo de líneas en el cuadro.

La sola condición bajo el cual el Flicker de 30 Hz de área grande puede resultar con el entrelazado es cuando el brillo promedio de los campos par y non es significativamente diferente, lo cual es un raro evento. Se sabe que la persistencia en la carga del fósforo es lo suficientemente corta para no causar embarramiento (blurring) entre cuadro, el movimiento vertical puede a veces producir una pantalla con la mitad del número de líneas de barrido, lo que puede ocurrir angustiosamente de forma continua.

De igual manera, el movimiento horizontal puede producir bordes verticales aserrados (aliasing), lo cual no lo realiza de manera ordinaria a causa de la integración de la cámara. Además, las distancias de visión en las cuales las líneas pueden ser claramente resueltas, el rango de Flicker de 30 Hz y es fácilmente visto como un brillo o resplandor. Una comparación lado a lado de imágenes entrelazadas y no entrelazadas (La segunda toma el doble del ancho de banda) hace esta diferencia obvia.

Aún en distancias de vista grandes, en el cual la estructura de línea no puede resolverse, en áreas que tienen detalles significativos verticalmente a 30 Hz el Flicker es claramente visible en entrelazados en ciertas circunstancias. El Flicker ocurre cuando las líneas non y par son lo suficientemente diferentes en cualquier frecuencia de resolución espacial y que puede eliminarse reduciendo la resolución vertical de la cámara y/o la pantalla, o

integrando a un cuadro completo con algún dispositivo temporal promedio. Con cualquier método, la resolución espacio - temporal se reduce. El entrelazado produce problemas los cuales vienen a ser más obvios como la resolución vertical se incremento y como la imagen es vista más cercanamente mientras la resolución vertical se incremento muy ligeramente.

Este proceso debe de estar sincronizado al hacer los barridos en los campos non y par, esta es una de las propiedades que se aprecian desde la cámara, enseguida serán descritas algunas propiedades que las cámaras ya sean de tubos o de CCD tienen para integrar la luz en los puntos donde toca el haz de electrones cuando barre el target.

III.3 Propiedades especiales de la cámara.

En las cámaras de tubos, el target, el cual integra efectivamente la luz incidente en cada punto cada vez que barre el haz en forma sucesiva, el cual casi se descarga completamente en cada, campo. En el área de integración, en la dirección vertical, comprime por lo menos dos de las 525 líneas de barrido nominales. Así un patrón Vertical de 262.5 ciclos por altura de imagen (p. h.) es traducido o reproducido con respuesta cero y una frecuencia de la mitad la cual es atenuada sustancialmente, a un grado el cual depende de su fase. Aunque el teorema de muestreo nos dice que se podría usar un ancho de banda completo de 262.5 c/ p.h., la respuesta vertical de la cámara se incrementa.²⁹

En algunas cámaras CCD las que tienen una fila de detectores para cada línea de barrido. El apareo de dos líneas de datos para cada línea de salida es hecho deliberadamente. En los T.R.C. (Tubos de Rayos Catódicos) de la cámara, el proceso es considerablemente más complejo debido a la física de la descarga de la target y de la forma del haz de electrones.

Así, las áreas oscuras son completamente descargadas por el borde principal del haz (leading edge), mientras que las áreas brillantes no son completamente descargadas hasta que pase el borde posterior (trailing edge) del haz. La distorsión geométrica resultante y la distorsión de escala de tono de área pequeña no son serias. Es más importante el hecho de que la resolución vertical de los tonos de cámara es mucho menor que la resolución horizontal (líneas/mm en la target); con el entrelazado hemos sacrificado una porción significativa de la definición vertical teórica y con esto el beneficio supuesto.

²⁹ Manual de Servicio Sony, pagina 24.

III.3.1 El Tubo de Imagen (Tubo de Rayos Catódicos).

El Tubo de Rayos Catódicos esencialmente es lineal, tal que la salida de luz integrada en cada punto puede hallarse envolviendo un rastro ideal (diámetro de grano cero) con la sección de cruce de haz la cual es generalmente Gaussiana esto se puede apreciar en una pantalla sin señal, ed decir son todos los pequeños puntitos que hacen ver a la pantalla como si fuera ruido. Con tal forma, la eliminación de la estructura de la línea de desenfoco (o empañamiento en el ojo) también se empaña la imagen. En los tubos de colores, un factor adicional es la estructura de los granos pequeños de fósforo (ruido espacial) el cual no es muy sensible para nosotros.

En la figura siguiente (figura 3.5) se puede ver de manera general el funcionamiento del proceso de escaneo por un Tubo de Rayos Catódicos el cual se lleva a cabo por el haz de electrones de arriba hacia abajo y de izquierda a derecha y que a su vez es reiniciado por las bobinas de deflexión .

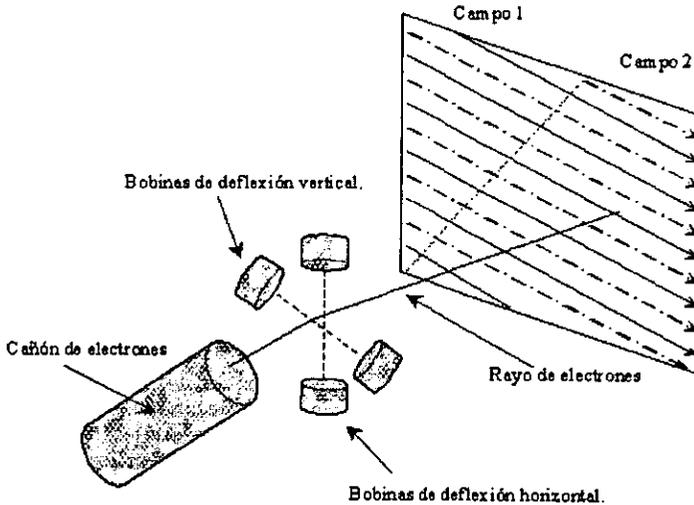


Figura 3.5

III.3.2 Sincronización de la señal.

Existen conceptos a continuación que no deben de confundirse, se tiene un intervalo denominado como H el cual significa el tiempo que tarda o que necesita el haz de luz en barrer una línea completa incluyendo el trazo de la línea y el retraso para colocarse en posición para comenzar una nueva.

En cada campo existen 262.5 líneas Horizontales, las que al pasar de una a otra, deben de estar a cierta secuencia, así como de un campo a otro, [entrelazado 2:1].

Pero en el tubo de imagen del receptor, el haz que barre la pantalla debe de reensamblar los pixeles de la imagen en cada línea H con la misma posición que en el tubo de cámara o CCD; y de igual manera de forma vertical.

Por lo que a la señal de imagen generada en la cámara, se le insertan señales de borrado y sincronía para formar la señal de video, final y completa.

La señal de negro de imagen debe hallarse arriba del nivel de borrado. Ya que este nivel de borrado debe encontrarse a un nivel de negro mayor que el de la imagen, siendo su misión cortar el haz electrónico del cinescopio.

Así que los pulsos de sincronización Horizontal se transmiten en cada "Línea H" para guardar el barrido H sincronizado, el cual sirve para el retorno del haz. La señal de video que aparece antes del borrado (Blanking) es el que está en la parte derecha de la pantalla del monitor de imagen.

Posterior a esta porción, en el borrado, está una zona del Blanking H llamada *Pórtico Frontal* (Front Porch) y con él se inicia el borrado para asegurar que el haz está cortado antes de comenzar el retorno. Luego sigue la sincronía H, tiempo de retorno, y después sigue el *Pórtico Posterior* (Back Porch) ya en el comienzo de la línea y es importante para que comience antes de restaurar el haz en la parte visible de la pantalla.

Interfaz 2:1.

En la figura siguiente (figura 3.6), se aprecia la localización física de las áreas ubicadas en la pantalla, las líneas punteadas indican las zonas que no son visibles normalmente y las líneas continuas las que sí lo son.

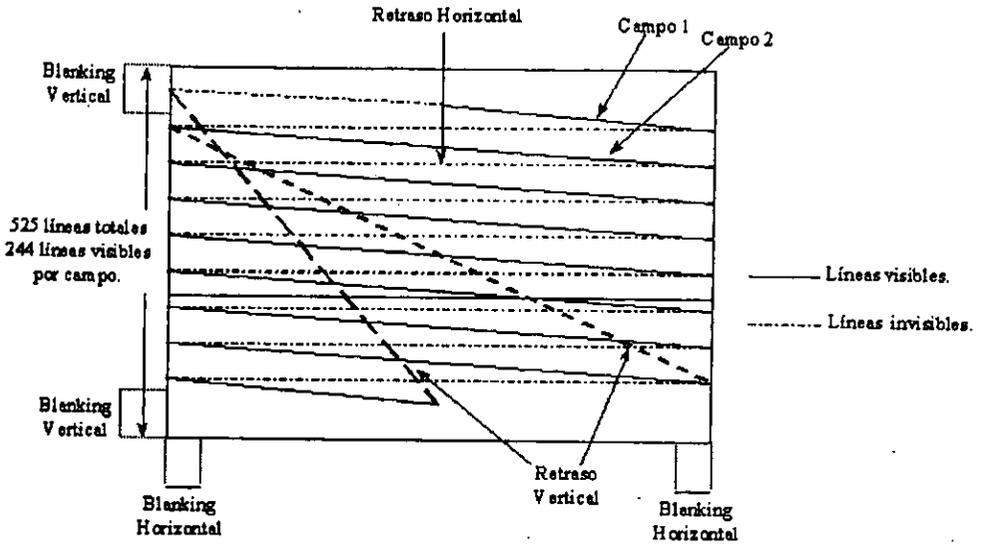


Figura 3.6

Voltaje Gráfico de un escaneo Horizontal.

140 IRE = 1.000 Volt

100 IRE = 0.714 Volt

40 IRE = 0.286 Volt

1 IRE = 0.007 Volt

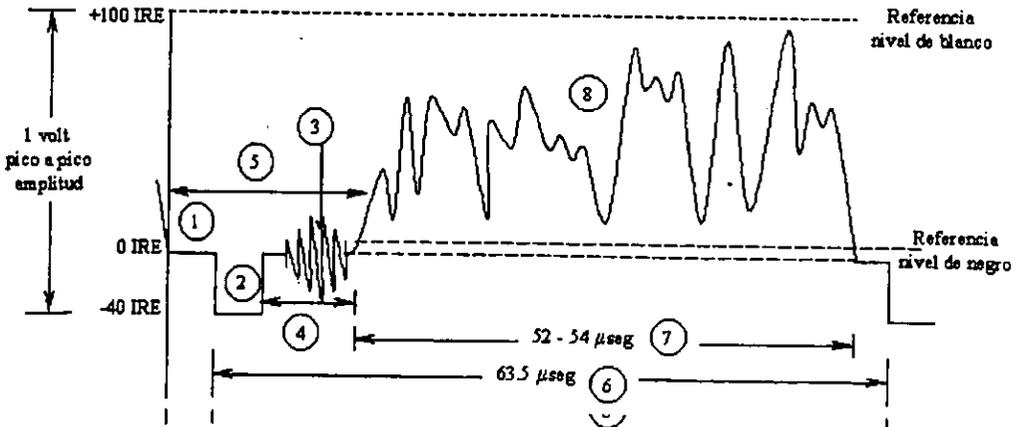
Una unidad IRE equivale a 0.714 Volt.

(IRE= Instituto de Ingenieros de la Radio, hoy llamada IEEE).

Esta figura (figura 3.7), muestra una señal completa de una línea horizontal "Línea H", las partes que la componen y su duración correspondiente..

Figura 3.7

La amplitud del video (8) es una señal analógica que continuamente esta variando.



Front Porch = 1.5 μ seg.

1. Pulso Horizontal de Sincronía = 4.7 μ seg.
2. Burst de color de 8 a 10 ciclos a 3.5795 Mhz.
3. Back Porch = 4.7 μ seg.
4. Blanking de imagen = 11.0 μ seg.
5. Una línea de video = 63.5 μ seg.
6. Tiempo de una línea activa de video = 52 - 54 μ seg.
7. Señal analógica de video.

III.3.3 Tiempo de borrado Horizontal.³⁰

En las figuras anteriores se observan los detalles del período del Blanking H. El intervalo marcado H es el tiempo necesario para barrer una línea completa, incluyendo trazo y retraso. Como se sabe, el tiempo de una línea H es de 63.49 μ seg.

El tiempo del borrado H es de 0.17 H de ancho, por lo que será de 10.9 μ seg.

Por lo que el tiempo de video será de $63.49 - 10.9 = 52.59$ μ seg.

El Front Porch dura 1.5 +/- 0.1 μ seg y esto es 0.02 H de ancho.

El Back Porch dura, 4.7 +/- 0.1 μ seg y esto es 0.07 H de ancho.

La Sincronía H dura igual que el Back Porch.

El pulso de sincronización vertical se transmite para cada Campo, para sincronizar el movimiento de barrido vertical. En la siguiente figura puede observarse la porción del borrado vertical del video compuesto, la que se caracteriza por pulsos negativos más anchos.

El Blanking vertical consiste de 6 pulsos anchos de duración de 0.5 H cada uno de ellos llamados pulsos Preigualadores (Pre - equalizers) que sirven para el retorno del haz de la parte inferior a la parte superior; después siguen otros 6 pulsos de duración de 0.5 H angostos o seriados (dentados) que son la sincronía Vertical, es decir, el retorno del haz; y después otros 6 pulsos de duración de 0.5 H similares a los igualadores previos, pero estos se llaman Postigualadores (Post - equalizers) y es la preparación del nuevo campo [6 pulsos de 0.5 dará 3 H]. Por lo que con estos pulsos de 0.5 H tenemos un total de 9 H de pulsos de 1/2 H, (Preigualadores, Sincronía Vertical y Postigualadores).

³⁰ Descripción de los procesos de Tiempos de borrado horizontal y vertical, Manual de Servicio Sony, paginas 24 - 31.

Los pulsos igualadores sirven para que la sincronización Horizontal durante el borrado Vertical sea estable y no se pierda en ningún momento.

III.3.4 Tiempo de Borrado Vertical.

Es el tiempo que tarda en cambiar la amplitud de la señal a niveles de video negro y es conocido como borrado V que a su vez cuenta con los siguiente pulsos para su sincronía conocida también como sincronía V.

Los pulsos del Blanking Vertical cambian la amplitud de la señal de video al negro. El ancho de este borrado es de $0.8 V = 0.8 \times 0.0166 = 1.3333$ mseg. (1.33 milisegundos) y esto será 21 H por cada campo, lo que serán 42 H borradas en un cuadro ($0.08 \times 525 = 42$).

El que la frecuencia de los pulsos de media línea H sea del doble de la Frecuencia Horizontal, se debe a que la relación de entrelazado es de 2: 1.

A continuación se tienen los diferentes aspectos de dos intervalos de borrado verticales consecutivos (Non y Par).

Campo Non:

- El principio del borrado V y de la sincronía V coincide con la Frecuencia H.
- Termina en media línea H.
- Tiene 6 pulsos Postigualadores de la misma duración.
- El borrado se enumera con los dos primeros Preigualadores.
- Tiene 12 H no visibles en la imagen llamadas inactivas.
- El borrado V dura 21 H.

Campo Par:

- El inicio del campo coincide con el final de media línea H, por lo que no coincide con la sincronía H.
- Termina en Línea H completa.
- Tiene 5 pulsos Postigualadores de duración de $1/2$ H y el sexto es de un H de duración.

- El borrado se enumera a partir del segundo Postigualador (y tercero).
- Tiene 12 H no visibles en la imagen llamadas inactivas.
- El borrado V dura 21 H.

El funcionamiento de los pulsos de sincronía son interpretados de manera descrita en la figura 3.8.

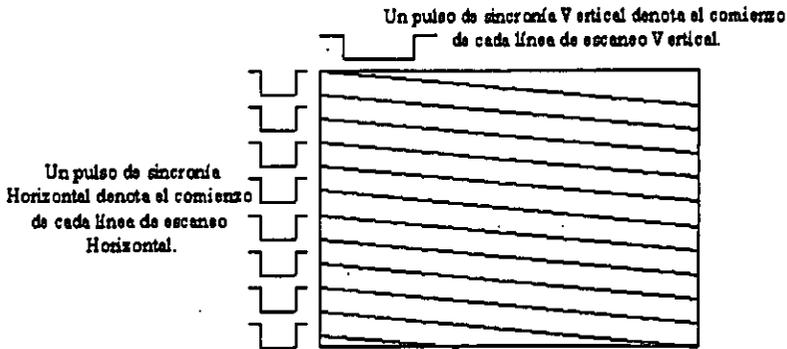


Figura 3.8

III.4 Resolución de una Imagen.

“...El sistema de TV está basado en que el observador humano en promedio es capaz de resolver o solucionar cuando está colocado a una distancia de 8 veces la altura de la imagen de la pantalla...”³¹

En esta distancia nosotros podemos tener la información suficiente (resolver patrones) de líneas alternadas blancas y negras y que consisten en 330 a 350 líneas paralelas.

Considerando la resolución vertical (Líneas horizontales apiladas una sobre la otra), se tiene que considerar el hecho de que el rastreo está hecho de líneas horizontales y no se utilizarán todas en forma completa mostrando los patrones de líneas horizontales. Algunas, caerán entre las líneas o fuera de la parte inferior o superior de la imagen.

Las pruebas indican que una relación de 0.7 de las líneas en total en el rastreo contribuye a la resolución vertical. Entonces el rastreo debe contener $337/0.7$ H, o sea, 483 H, esto

³¹ Manual de Servicio Sony, pagina 33.

es para la parte visible de la imagen y se necesita sólo el 8% para el borrado vertical tal que el número de líneas totales del rastreo será $483/1 - 0.08 = 525$ H.

III.4.1 Determinación del ancho de banda.

El rango de líneas es de 15750 Hz, si tiene que resolver 330 Líneas verticales en la imagen, por lo que para expresar la resolución Vertical y Horizontal en los mismos términos se tiene que tomar en cuenta el hecho de que la pantalla es más ancha que su altura.

La relación de aspecto es 4:3 por lo que 237 H cubren solamente 3/4 del ancho de la imagen y debemos multiplicar por 4/3 para obtener el número de líneas total a través de la imagen.

Entonces $337 \times 4/3 = 449$. La duración de la línea Horizontal es de 63.5 μ seg, por lo que la parte activa o visible es de 52.6 useg (sin borrado). Para dibujar 449 Líneas alternadas blanco y negro durante 52.6 useg. se debe de tomar la mitad de 449, y esto será 224.5 y se divide entre 52.6 useg lo que da 4.2 MHz. Su periodo será 0.234 useg. por lo que el ancho de banda para una resolución completa de video es de 4.2 MHz.

Este ancho de banda se requiere para tener una señal completa de brillo ó luminancia que es conocido como FACTOR 80.

Una regla para la resolución en H con respecto al ancho de banda es el factor 80 MHz.

Ancho de banda [MHz] = Líneas resolución H / 80.

Res. H [Líneas] = Ancho B x 80.

Por. Ej. $320 \text{ H} / 80 = 4 \text{ MHz}$.

Se ha observado, que el proceso de llevar una imagen a una pantalla de TV no es un proceso sencillo ya que a ella se encuentran ligados muchos conceptos técnicos. Se ha mostrado parte de los componentes de la señal de Video y el comportamiento de los elementos físicos tales como el haz de electrones y sus procesos de sincronización así como las partes que componen el sistema, ahora toca el turno del color y la manera en como se compone la señal con este elemento.

III.5 Características de la información de Color.

El color se caracteriza por su Brillantez, su Saturación y por su Matiz.

BRILLANTEZ: Es la cantidad de luz contenida en un color dado.

Por ejemplo un color Café, que en realidad es un amarillo con poca brillantez; lo que es lo mismo si a un blanco se le disminuye la cantidad de luz se apreciará gris, al hacerlo.

SATURACION: Es la cantidad y pureza de color, considerada tomando como base la luz blanca.

Por ejemplo, un tono pálido de un color rojo (el rosa) es un color contaminado con cierta cantidad de blanco. Una saturación al 100% indica un matiz o tinte espectral puro, lo que prácticamente no es posible; por otro lado, una saturación de 0 % representa al Blanco.

MATIZ: Es el tono del color. Es decir, es el color de que se trate: el rojo, el amarillo, etc.

Todos estas características entre otras se encuentran reguladas por instituciones las cuales son las responsables de establecer estándares para todas ellas, por ejemplo, para nuestro país y muchos más de América, estos estándares son regidos por el NTSC.

El NTSC (Comité Nacional para el Sistema de TV) propuso un requisito de compatibilidad que cumpliera la Televisión a color.³²

- La señal de color debe ser compatible con la información Blanco y Negro. La información de brillo de una imagen a color, será la información Blanco y negro; y por otro lado la información de color no debe causar componentes visibles indeseados en un receptor Blanco y Negro.
- No debe transmitirse información de color cuando la cámara esté explorando superficies blancas, negras o grises.
- El sistema debe de ser capaz de reproducir la gama completa de colores que un observador normal pueda recibir.

La imagen a color en TV aprovecha ampliamente los comportamientos del ojo ante los colores, los cuales son:

- a) El ojo aprecia mayor brillantez en el color Verde.
- b) El rojo es el color que se distingue más fácilmente.
- c) El Azul es el color menos luminoso.
- d) El ojo aprecia perfectamente todas las tonalidades en objetos grandes vistos en la pantalla; no es muy crítica la diferencia de tonos en áreas medianas y no aprecia el color en objetos menores.

³² K. Blair Benson, HDTV Advanced Television for the 1990s, Mc. Graw Hill 1991, paginas 46 y 47.

Las señales que están estandarizadas por organismos como NTSC , SECAM, PAL, etc. son de gran importancia tal es el caso de la señal de luminancia, la cual corresponde a la información de brillo en las imágenes, descrita a continuación.

III.6 Señales de video.

III.6.1 Señal de Luminancia.

Esta señal es la correspondiente a la información de brillo de las imágenes.

Esta información puede utilizarse en un receptor - monitor blanco y negro. Esto es para tener la compatibilidad con ese sistema, y una señal de ancho de banda de 4. 2 MHz. que representa todas las componentes de brillo, detalles y contraste, en una escena a color y se derivará en la cámara de TV basado en la sensibilidad al color de las componentes de los colores primarios Rojo (R), Verde (G) y Azul (B) la que será en las siguientes proporciones:

$$Y = 0.30 R + 0.59 G + 0.11 B$$

Donde Y = luminancia.

R= rojo.

G= verde.

B= azul.

Como se ha indicado, en base a que la sensibilidad a la cantidad de brillo en los tres colores anteriores no es la misma. En la cámara después del prisma que separa los colores básicos, tendremos los transductores (tubos de cámara o elementos CCD) para derivar la información de brillo y posteriormente de color también y dichos elementos son monocromáticos, esto es, solamente reciben un tipo de luz y la ausencia de éste (negro).

Para que se tenga una idea de que es la luminancia, es la información de brillo de un color, ya que unos colores aparecen más brillantes que otros y éstos a su vez tienen sombras y / o partes blancas y grises y la luminancia indicará como se verá el color en una reproducción blanco y negro.

El blanco obtenido por la cámara de TV es por la combinación de los tres colores primarios aditivos, separados por un prisma, en las proporciones apropiadas.

El blanco en TV de referencia se especifica como un color a una temperatura de 6500 grados °K y esto corresponde, a un Blanco Azuloso, como la luz del día.

La influencia de esta "temperatura" no es en realidad que se encuentre afectado por esa escala propiamente como si fuera clima sino que la cámara tiene que identificar un ajuste para balancear las proporciones de los tres colores básicos. Este balance es de 6500 grados

Kelvin, se debe de recordar que la cámara no puede percibir los colores como lo hace el ojo humano.

El símbolo °K indica los grados Kelvin de la temperatura absoluta, y en esta escala °K corresponde a -273 grados C. Esta escala también indica la temperatura de color y está basado en el concepto del cuerpo negro, el cual absorbe todos los rayos de luz incidente y no refleja nada (concepto teórico solamente). Si el cuerpo se calienta, comienza a emitir luz visible, primero luz roja de tono oscuro y de poca intensidad, luego roja, después naranja hacia el blanco "caliente", como los metales. Si un metal se calentara lo suficiente sin derretirse, tendría una radiación blanca azulosa. De lo anterior, se dice que una alta temperatura K, se acercará al espectro solar. Un rango bajo de °K indica que una fuente de luz tendrá su radiación en la zona Roja visible. Las luces incandescentes caseras tienen bajo valor K. Para operaciones de color, el rango de K de las lamparas viene a ser un factor muy importante.

Cuando una cámara se ajusta, se requiere que haya una escena o imagen que contenga blanco para referencia para ajustar la salida de los colores R, G y B.

En la figura siguiente se observan los niveles de luminancia y crominancia en las barras de color que se usan para el ajuste de la señal de video, las cuales se tratarán con detenimiento más adelante.

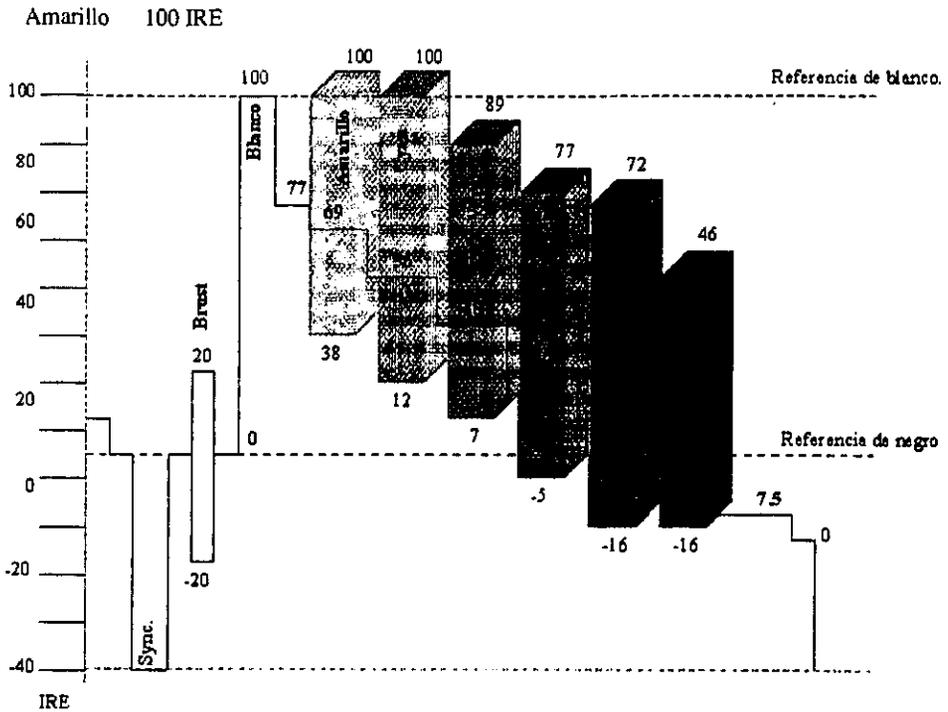
Una unidad IRE equivale a 0.714V (IRE= Instituto de Ingenieros de la Radio, hoy llamada IEEE).

Luminancia: Se observa como una especie de escalera en donde están acomodadas en forma descendente.

Negro	7.5 IRE
Azul	15 IRE
Rojo	28 IRE
Magenta	36 IRE
Verde	48 IRE
Cyan	56 IRE
Amarillo	69 IRE
Gris	77 IRE

Crominancia: Se muestran los niveles ideales dentro de los cuales estos colores pueden ser identificados como tales.

Azul	46 IRE
Rojo	72 IRE
Magenta	77 IRE
Verde	89 IRE
Cyan	100 IRE



Niveles de luminancia y crominancia en las barras de color que se usan para el ajuste de la señal de video.

Figura 3.9

III.6.2 Señal de Crominancia

Este término se utiliza para combinar el tinte (HUE) y la saturación, y esta información se transmite sin información de blanco y negro, consistiendo en una señal modulada en cuadratura con portadora suprimida por dos señales R-Y y B-Y, las que se conocen como señales de diferencia de color.

Se le sustrae a estas señales la luminancia (Y) para así eliminar la información de brillo quedando solamente la información de color y posteriormente estas dos señales se modulan estando una de otra 90 grados fuera de fase (ángulo), por lo que se dice que están en cuadratura. Estas señales se modulan con una portadora de valor de 3.579545 MHz posteriormente.

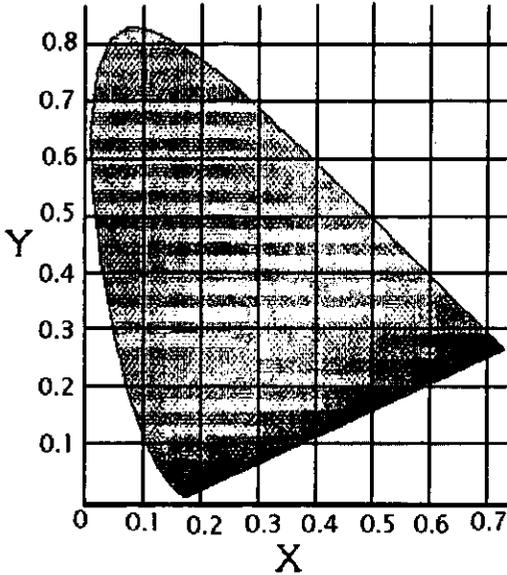


Figura 3.10

En esta figura (figura 3.10) se tiene un diagrama de cromicidad para un observador, es decir, muestra el espacio de color que el sistema visual humano es capaz de captar, este diagrama es conocido como "Espacio de color CIE" (International Commission of Illumination).

III.6.3 Señales I y Q en el sistema NTSC.

De los estudios realizados por NTSC en 1950 sobre agudeza visual a ciertas longitudes de onda, se descubrió que el ojo humano puede distinguir diferencias de color de fino detalle que coinciden correspondientes a los colores naranja - cian, mientras que en el magenta - verde no se percibían claramente.

Sobre estas bases se decidió considerar dos ejes, el eje I y el eje Q como ejes de modulación de señal de crominancia los cuales están girados 33 grados respecto a RY y BY, el eje naranja - cian coincide con I, con mayor resolución de color y el magenta-verde con Q con menor resolución en color pero menor ancho de banda

III.6.4 Señal de barras de color. (E.I.A.)³³

Los generadores de barras de color producen unas señales verticales que se usan para prueba y procedimientos de ajuste para unos niveles correctos de color y de luminancia.

Los generadores anteriores que producían barras de color, tenían problemas ya que estaban saturadas a un 100%, lo cual significaba que la luminancia llegaba a un nivel de 100 IRE para el pico blanco y 133 IRE para la crominancia correspondiente al amarillo. Esto origina problemas en el equipo de transmisión, además, en la realidad las cámaras actuales nunca obtendrán colores de la imagen saturados al 100%. Entonces la señal de barras de color se redujo a lo que ahora se le llama barras de color al 75% de saturación.

Cabe aclarar que en estas barras podemos tener el nivel de blanco ya sea al 75 % o al 100%, pero los colores tienen una saturación de solamente 75%.

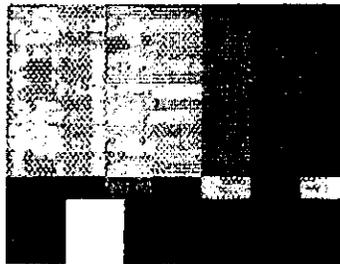


Figura 3.11 Barras de Color

Los valores de luminancia para barras de color se obtienen de la siguiente manera:

Blanco: Se sustraen 7.5 IRE de negro de las 100 IRE lo que resulta en 92.5 de variación de señal, luego se toma el 75% de este valor, dando un valor de 69.375 y finalmente se le agregan las 7.5 unidades dando un valor final aproximado de 77 unidades IRE.

³³ Manual de Servicio Sony, pagina 53.

GRAY (80 UNITS)	YELLOW	CYAN	GREEN	MAGENTA	RED	BLUE
BLUE	BLACK	MAGENTA	BLACK	CYAN	BLACK	GRAY
I	WHITE (100 UNITS)	+Q	BLACK	3.5	7.5	11.5

Figura 3.12

Rojo: Tomamos el 30% de 69.375 unidades de blanco lo que nos da 20.813 unidades a las que les sumamos 7.5 unidades de base y nos da 28.313 IRE. Este mismo procedimiento se utiliza para calcular los valores "Y" de las barras de color.

Valores de crominancia para barras de color.

Para obtener estos valores pico de crominancia, lo que se tiene que hacer es obtener el 75% para colores saturados al 100% por ejemplo:

Amarillo: su valor es de +/- 44 IRE multiplicados por 0.75 nos da 33 IRE y así puede obtenerse el valor de cromina para las demás barras de color.

Las señales arriba mencionadas como se describió tienen rangos establecidos como estandar, para poder verificar que estos rangos no sean sobrepasados ya sea en menor o mayor cantidad deben ser monitoreados, para ello existe un aparato de medición llamado vectorscopio.

III.7 Sistema de monitoreo de señales.

III.7.1 Monitor Vectorscopio.³⁴

La siguiente figura (figura 3.13) muestra los vectores para valores de croma de barras de color al 75% y así mismo nos indica sus amplitudes y fases. Aquí la amplitud significa el barrido pico a pico. Las amplitudes se miden radialmente desde el centro con círculos indicando desde 20 hasta 100 IRE.

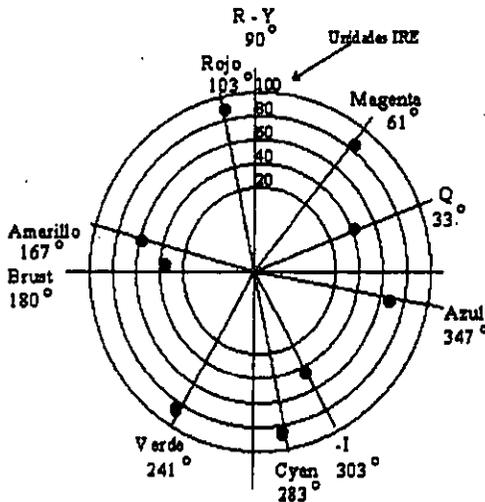


Figura 3.13

Las fases se miden en grados a partir del eje B-Y, el Brust está a 180 grados por ejemplo, el amarillo tiene una amplitud total de 62 IRE con un ángulo de fase de 167 grados.

Las tolerancias que se ven como cuadros pequeños son de ± 2.5 IRE y ± 2.50 en fase.

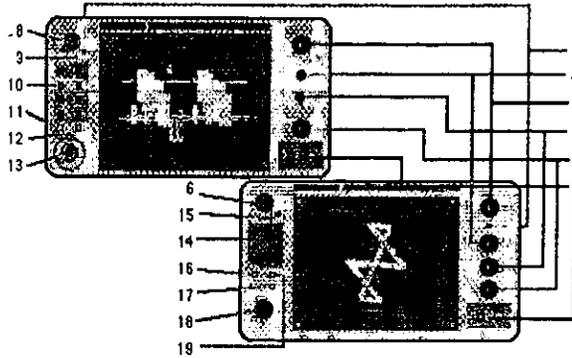
Los bordes de sector que rodean a cada cuadro son tolerancias de ± 20 IRE y ± 100 de fase. También tenemos unas marcas para las señales de BURST, I y Q.

La figura a continuación (figura 3.14) muestra físicamente un vectorscopio y sus partes.

³⁴ Manual de Servicio Sony, página 54.

Monitoreo de la Forma de Onda y Vectorscopio.

Figura 3.14



1. Power.
2. Posición Horizontal.
3. Posición Vertical.
4. Foco.
5. Intensidad de iluminación.
6. Ganancia.
7. Led de calibración.
8. Control de entrada.
9. Botones de respuesta:
 - Flat.
 - IRE.
 - Croma.
 - Ganancia.
10. Sweep o barrido.
11. Sincronía Sync.
12. Restaurador de DC.
13. Selector de Línea.
14. Input.
15. Sub portadora Phase.
16. Saturación.
17. Escala.
18. Control de Fase.
19. Circulo de prueba.

III.7.2 Mediciones de Video. (Wave Form Monitored WFM)

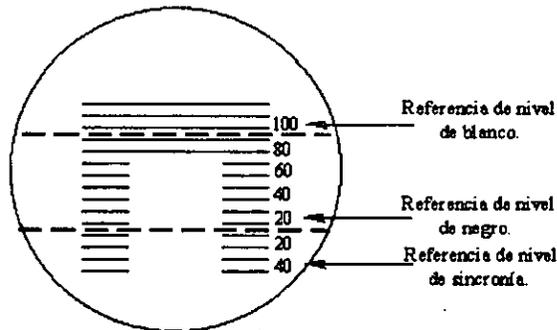
Con el fin de realizar mediciones de amplitud y tiempo de una señal de video, se requiere un mínimo de equipo, ya que se requiere solamente de un monitor forma de onda (Wave Form Monitored).

Para utilizar este dispositivo, es necesario verificar su calibración de amplitud, ya que es generada internamente una señal de onda cuadrada de amplitud de 1 Volt + pico a pico, o sea, 140 unidades IRE, al igual que la señal de video. Una unidad IRE equivale a 0.714V (IRE= Instituto de Ingenieros de la Radio, hoy llamada IEEE).

Si la calibración no es correcta se debe de corregir; al estar bien, se procede a medir la señal de video confiadamente, obteniendo 100 unidades IRE en pico blanco y - 40 IRE en sincronía (1 Volt + pico-pico). Las mediciones de amplitud, se efectúan a la señal de luminancia, ya que es la información de cantidad de luz. Las mediciones de tiempo, en el borrado horizontal, medirán dentro de los límites de +4 IRE y -40 IRE según el Comité Federal de Comunicaciones (FCC).

Así, el Blanking Horizontal se medirá entre los límites de + 4 IRE; la Sincronía H se medirá entre los límites de - 4 IRE (Leading Edge y Trailing y Edge); el Front Porch entre los límites de + 4 IRE del video y - 4 IRE del Leading Edge.

Figura 3.15



Estándar de la escala IEEE, vista en un vectorscopio.

III.7.3 Burst o sincronía de color.³⁵

Las señales recortadas componentes I y Q, se utilizarán para modular la señal de color con portadora suprimida, las que se sumaran vectorialmente para tener finalmente un vector resultante llamado crominancia, donde su magnitud, corresponde a su saturación y su ángulo de fase a su Matiz o Hue.

Debido al argumento anterior, se necesita sincronizar la fase de color del receptor, para así demodular los colores en la misma posición que el original y no se tengan matices distintos. Por lo que es necesario establecer una sincronía de color que se llama Burst, y consiste de 9 ciclos de la frecuencia de 3.58 MHz, con una fase de 180 grados y tendrá una amplitud de pico a pico de 40 IRE colocado en el Back Porch y no interfiere con la sincronización de deflexión.

En otras palabras, la crominancia se modula a 3.58 MHz y el Burst tiene 9 ciclos de esta frecuencia.

El Burst está en el período de Blanking H y no es visible. No existirá señal de Burst en las 9 líneas del periodo vertical de igualadores de sincronía V.

III.8 La Transmisión NTSC.³⁶

Proceso de codificación. NTSC (ENCODER).

Cada cámara incluye un proceso completo, para que las señales de entrada al equipo mezclador (switcher) estén en normas y no tenga problema alguno al procesarlas o a una VTR, etc.

Las salidas de los amplificadores de R, G, B, donde se introduce la corrección de (gama), entran al codificador que consiste de:

1.- Matrización de $Y = 0.3OR + 0.59 G + 0.11B$ (información B/N).

2. - Formación de I de color por matrización y pasa por un filtro pasabajo de 1.5 MHz para entrar al modulador balanceado: $I = 0.6OR - 0.28G - 0.32B$ para la agudeza visual que requiere mayor resolución de color.

³⁵ Manual de Servicio Sony, pagina 67.

³⁶ Manual de Servicio Sony, pagina 65.

3. - Formación de Q de color por matrización y pasa por un filtro pasabajo de 0.5 MHz para entrar al modulador balanceado, el que no requiere mayor ancho de banda: $Q = -0.21R - 0.52G + 0.31B$
4. - Las señales de banda recortada (I, Q) pasan a los moduladores balanceados con portadora suprimida cuyas salidas se suman para dar un vector de crominancia único.
5. -Un generador de sincronía de color NTSC el que genera las siguientes señales:
 - a) La señal FSC (Frecuencia Subportadora Continua), subportadora continua, 3.58 MHz de referencia, la cual se defasará, por un lado a 123 grados (I) y 33 grados (Q), y por otro lado 180 grados para la señal de Burst.
 - b) Señal de Burst Flag, Burst Gate, es una compuerta que permite la introducción de 9 ciclos de FSC a 180 grados en el Back Porch para tener el Burst.
 - c) La señal de Sincronía Compuesta, que consiste de sincronismo horizontal y de sincronismo vertical, para proporcionar una señal compuesta VBS (Video Burst - Sincronía).

Distribución de frecuencias en el espectro electromagnético.

El ancho de banda de frecuencias para la transmisión de video - luminancia debe de ser de 4.2 MHz.

El ancho de banda de la componente de color I de 1.5 MHz es de 3.58 MHz +/- 1.5 MHz lo que ocasionara que la banda lateral superior se salga del limite (4.2 MHz) por lo que se utiliza 0.5 MHz solamente en la banda superior y 1.5 MHz en la banda lateral inferior completa.

III.9 Defectos de la Señal de Video.

La transmisión de señales de TV debe ser lo más nítida posible, ya que entre la cámara y el receptor existe un gran numero de aparatos y circuitos que contribuyen al deterioro de la señal, veamos los defectos para poder atenuarlos.

- a) Provocado por circuitos o Líneas de transmisión.

Moire.- Un tipo de batido coherente normalmente producido por distorsión armónica de la señal FM (un modulador - demodulador mal ajustado) o por tomar con una cámara información a muy alta frecuencia (ópticamente) ya que se confunden las líneas de exploración con una imagen muy detallada.

Flare.- Distribución de luz extraña, por ejemplo, en película seria una luz ajena a la imagen.

Ghost (fantasma): Es una segunda imagen indeseada desplazada de la señal original.

Gobo.- Un decaimiento repentino de luz en la imagen.

Womp.- Un elevamiento repentino de luz en la imagen.

Snow (nieve).- Puntos blancos en imagen ocasionado por señal débil, amplificadores ruidosos, etc. Esto nos indica una baja relación señal a ruido.

Low Level (Niveles Bajos): Al tener niveles bajos, ocasionan disminución de brillantez promedio de imagen y la hace mas susceptible a interferencias y ruidos.

R.F. (Radio Frecuencia): Es la inducción o interferencia de señal de muy alta frecuencia en el video por mal blindaje en cables o equipos, o malas conexiones entre los cables y la tierra. Se observa como malla.

Blurring (Empañamiento): La imagen se ve con efecto niebla (Foggy) por defectos de lentes en una cámara.

Interferencias: Son producidas por efectos electromagnéticos y mecánicos, por mal aislamiento o blindaje de la señal y son los siguientes:

Hum (zumbido): Producido por 60Hz de la línea o armónicos, debida a la señal de corriente alterna; o por un filtrado deficiente de fuentes de alimentación. Son barras horizontales oscuras y anchas que varían de posición vertical.

Micrófonismo: Inducción de audio o bajas frecuencias en la imagen provocadas por mal aislamiento o vibraciones mecánicas del equipo.

Halo: Al efectuarse exploración de objetos muy brillantes, provoca un contorno negro alrededor de dichos puntos brillantes; provocados por mal ajuste de un modulador, demodulador o circuitos cercanos.

Smear: Ruido en las primeras cámaras CCD, provocada por puntos muy brillantes y se observa como una línea blanca vertical.

b) Provocado por cintas magnéticas y VTR.

Drop out.- Consiste en una caída de nivel de RF en una cinta que causa disparidad o pérdida de señal de video. Puede provocarse por contaminantes o cinta defectuosa.

Clogging.- Consiste en obstrucción de la cabeza de video (tapada), ya que se acumula óxido o suciedad en el Gap de la cabeza y provoca degradación o talla en el video reproducido o grabado.

Sobremodulación: Provocado por excesiva desviación de Frecuencia de portadora debido a niveles muy altos o muy bajos de señal original; o también por niveles de corriente de grabación altos. Se observa como halos y saturación en blancos.

Baja Modulación: Debido a bajos niveles de grabación en la cinta magnética y provoca ruido en la señal (señal).

Mistracking (Fuera de Tracking): No lee la cabeza al centro de la pista en la cinta, problemas de compatibilidad entre una VTR y otra, por lo que se verá ruido en imagen característico (en la parte superior o inferior de la imagen).

Skew (Efecto de Persiana): Sucede en formatos de cinta magnética de grabación helicoidal, al conmutarse dos cabezas de video, y se vera que no coinciden los tiempos, debido a diferencias de tensión de cinta entre una VTR y otra.

Rayadas de Cinta.- Se verán como disturbios o líneas variantes en la imagen que varía de posición con respecto al tiempo. Se debe a cintas magnéticas rayadas o dobladas.

III.10 Estándares de Televisión.

Después de haber mencionado las características de la señal de video podemos resumir que los distintos estándares de televisión existentes se basan en los mismos principios:

- a) Barrido de líneas
- b) Repetición de campo.
- c) Transmisión de color como componentes separados de crominancia y luminancia.

A su vez pueden tener variaciones en los siguientes elementos:

Número de líneas de imagen.

Las imágenes de 405, 525, 625,819 líneas están en uso actualmente. La resolución es muy baja para 405 líneas y demasiado alto para 819 líneas. La tendencia es reemplazarlas por el sistema de 625 líneas.

Frecuencia de Campo.

El factor para limitar el parpadeo (Flicker) y el rizo de la línea de corriente alterna, fue utilizar las frecuencias de imagen de 50 y 60 HZ, para 525y 625 líneas, respectivamente.

Transmisión de Color.

Se han desarrollado tres sistemas de T.V., independientemente del barrido de líneas o frecuencia de campo.

NTSC (Comité nacional para sistema de T.V.) 1948.

PAL (Línea de fase alternada) 1961.

SECAM (Sequentiel a mémoire) 1957.

La luminancia para los receptores monocromáticos y los tres colores primarios (rojo, verde y azul) se transmiten en forma de señales de diferencia de color relativo a la señal de luminancia.

Estándares Básicos de T.V. para intercambio de material televisivo.

Se han adaptado dos estándares básicos para el intercambio internacional de programas de T.V.

	Estándar FCC (Comisión Federal de Comunicaciones)	Estándar CCIR (Comité Consultivo Internacional de Radiocomunicaciones)
Líneas por imagen	525	625
Campos por segundo	60	50
Sistema de color	NTSC	PAL / SECAM
Ancho de banda de vídeo	4.2 MHz.	5, 5.5 y 6 MHz.
Subportadora de color	3.58 MHz	4.43 MHz.

El principal problema de conversión de estándares es la conversión de frecuencia de campo de 50HZ a 60HZ y viceversa. La información debe ser almacenada y luego barrida a la nueva frecuencia.

Países con los distintos Estándares.

PAL: Argentina, Australia, Austria, Brasil, Bélgica, China, Dinamarca, Alemania, Hong Kong, España, Italia, Sudáfrica, Suiza.

SECAM: Albania, Bulgaria, Egipto, Francia, Grecia, Hungría, Rusia, Irán, Líbano, Marruecos, Polonia, Siria, Zaire,

NTSC: Canadá, Chile, República Dominicana, Guatemala, Venezuela, Japón, EUA, Corea, México, Panamá, Filipinas, Puerto Rico, Taiwán, Haití.

III.10.1 Estándar NTSC.

"...El estándar de video consiste en 525 líneas horizontales de video y esta información es regenerada 30 veces por segundo. Se tienen 30 cuadros por un segundo, cada campo requiere la mitad de ese tiempo, o lo que es lo mismo 1/60 de segundo para desplegarse. Las 525 líneas están divididas equitativamente entre los dos campos, cada campo comprime 262.5 líneas. El decimal 0.5 es a razón de que los dos campos contienen la mitad de una línea. Al desplegar 262.5 líneas de video en 1/60 de segundo significa que cada línea requiere aproximadamente 63.5 microsegundos. Este número es importante para establecer los parámetros de un digitalizador porque es también la longitud de tiempo entre las posiciones verticales de píxeles en un display NTSC..."³⁷

"... No todas las 262.5 líneas de video en un campo son usadas para desplegar una imagen. Aproximadamente 18.5 líneas son utilizadas para los espacios verticales en blanco y la sincronización para desplegar la imagen. Esto nos deja 244 líneas visibles de información por campo. Estas 244 líneas por cada campo, 244 líneas por cuadro, hacen posible que la imagen sea vista en los monitores tanto de televisión como los de computadoras. De las líneas usadas por los espacios verticales en blanco (*vertical blanking*), de 17 a 20 por cada campo pueden ser utilizadas también para otros propósitos. Las aplicaciones de estas líneas escondidas pueden realizar pruebas, hacer diagnósticos del equipo y controlar la señal de transferencia..."³⁸

Pero la señal de video no solamente esta regulada dentro de estos estándares, para que esta señal sea completamente útil debe de estar acompañada de otra señal muy importante, la señal de Audio que es un complemento indispensable en la televisión. A continuación algunos de los puntos más importantes.

III.11 Señal de Audio.

La energía puede propagarse a través del espacio y de la materia por medio de vibraciones. El sonido, la luz, las ondas de radio, etc., esto es, se comportan como movimientos ondulatorios.

El movimiento ondulatorio se considera como un movimiento periódicamente repetitivo con respecto a la unidad de tiempo.

La señal de audio consiste en las señales eléctricas que corresponden ya transformadas al sonido que se esté considerando. (esto a través de un transductor)

El Sonido, es el fenómeno producido por la vibración de un cuerpo y se propaga a través de un medio elástico como sería, el aire, el agua, los metales, etc., pero no se propaga en el vacío.

³⁷ Craig A.Lindley, Practical Image Processing, Ed. Jhon Wiley & Sons. 1991 USA. Pagina 48.

³⁸ Craig A.Lindley, Practical Image Processing, Ed. Jhon Wiley & Sons. 1991 USA. Pagina 49.

El sonido se dice que son perturbaciones audibles y entre ellos están los instrumentos musicales, la voz humana, ruido acústico, etc. El oído humano solamente puede escuchar sonidos comprendidos entre 50 ciclos por segundo (Hertz) y 18 000 Hertz. Los sonidos por debajo de los límites se les conoce como Infrasonidos; y arriba del límite se les llama Ultrasonidos, los cuales ciertos animales puede captarlos.

III.11.1 Características del Sonido.

Las características del sonido son tres: La Intensidad, el Tono y el Timbre.

a) La Intensidad: Es la magnitud o nivel del sonido. Se dice que el sonido es "débil" ó "fuerte", dependiendo de la amplitud y de la distancia de la fuente sonora.

b) El Tono: Es la frecuencia del sonido que produce la fuente sonora. Un sonido, que es "grave o bajo" si tiene baja frecuencia, (menor a 1 kHz), y "alto o agudo" si son frecuencias mayores a 1 kHz.

c) El Timbre: Es la cualidad por la cual se distinguen las fuentes sonoras, esto depende de la composición de unas ondas llamadas ondas componentes Armónicas que se adicionan a la frecuencia Fundamental.

Frecuencia Fundamental: Consiste en un sonido o tono puro y su frecuencia es pura sin componente alguno.

Armónica: Es un múltiplo de la frecuencia fundamental.

III.11.2 Medición y Monitoreo de Audio.

Debido al gran intervalo de intensidades para las cuales es sensible nuestro oído, un instrumento sonoro que produce sonido a determinada intensidad necesita aumentarla diez veces para que podamos percibirla como una doble intensidad, o sea, el oído humano no responde en proporción directa a los cambios de nivel de sonido como éste lo hace. Para poder graficar esta respuesta, se utiliza una graficación logarítmica, ya que es de esta manera la respuesta del oído humano, para así poder medir las intensidades del sonido.

La forma en que el oído aprecia las diferencias de intensidad de niveles, es con respecto a un nivel dado y su unidad relativa de medición es el DECIBEL.

A continuación tenemos los valores relativos en intensidades de sonidos en Decibeles:

Umbral de la sensación sonora en el oído = 0 dB.

Conversación en voz baja = 20 dB

Radio con volumen regular = 40 dB

Conversación ordinaria = 65 dB

Calle con mucho tráfico = 70 dB

Perforadora neumática = 95 dB

Umbral de sensación desagradable en el oído = 120 dB.

Estas potencias mecánicas, al convertirse en una señal eléctrica por medio de un transductor, se convierte en potencia eléctrica, por lo tanto en voltaje y corriente.

Las características de un amplificador transductor o cualquier dispositivo que maneje la señal de audio, deberá de manejar todos los parámetros de dicha señal en forma transparente, esto es sin distorsión alguna; y a esto se le conoce como un dispositivo de alta fidelidad (Hi - Fi).

La amplitud de audio contra las características de frecuencia de dichos dispositivos o sistemas de audio es una expresión de variación de amplitud como función de frecuencia de audio en forma senoidal al aplicarse a la entrada del sistema y a la salida éste es medido. Dicha variación se mide en Decibeles. Las variaciones del nivel del audio contra la gama de audiofrecuencias deberán de estar dentro de las especificaciones del sistema o equipo que se trate.

III.11.3 Problemas y defectos de la Señal de Audio.

A) Defectos en la señal, ocasionado por el sistema.

Distorsión de frecuencia:

Este problema ocurre cuando cambia la magnitud relativa de las diferentes componentes de frecuencia de una onda de frecuencia, en otras palabras, puede variar su tonalidad.

Distorsión Armónica:

Consiste cuando a una señal senoidal pura se le adicionan señales de componentes o frecuencias armónicas causado por problemas de linealidad del sistema, lo que ocasiona que varíe el timbre de la fuente sonora, se mide en %.

Distorsión de Fase:

Consiste en el retardo variable en que aparezcan las diferentes frecuencias en el sistema de audio

Modulación Cruzada:

Se produce cuando las variaciones en la amplitud de la señal se aplican o inducen por si mismas sobre otra señal de audio de frecuencia distinta que esté cerca, afectando a ésta última y esto depende de la energía de cada señal.

B) Distorsiones provocadas por sistemas Mecánicos.

Diferencia de velocidad en Grabación respecto a reproducción y viceversa.

Variación de la velocidad de grabación o reproducción en forma periódica llamado también WOW (variaciones menores a 1 kHz) y Flutter (fluctuaciones mayores a 1 kHz).

Ruido: Es aquella señal residual existente sin aplicarse señal alguna a la entrada del sistema. Los tipos de ruido más conocidos son los siguientes:

Zumbido o HUM: Es una señal residual de 60 o 120 Hz.

Ruido de fritura o SCRATCH: producido por pastillas o componentes electrónicos defectuosos.

Sonido silbante o HISS: Se aprecia como un "ziceo" constante en la salida del sistema, es un ruido de altas frecuencias.

Microfonismo: Se produce por la vibración mecánica de los componentes electrónicos y mecanismos del sistema.

Todo sistema televisivo está conformado de audio, y por su puesto de video además de estar apoyado en sus sistemas de almacenamiento de imagen y sonido, estos sistemas son los múltiples formatos de cintas que se pueden encontrar tanto de manera comercial como de manera profesional tales como cintas VHS, 8mm, Betacam o formatos digitales etc. Estos formatos deben de tener un orden de almacenamiento, para ello existen registros, estos registros no son otra cosa que pulsos magnéticos dentro de una cinta con un recubrimiento ferroso y para poder medir y cuantificar estos registros se tienen los llamados "Código de tiempo" que serán expuestos a continuación.

III.12 Sistema de Cuenta de Tiempo Drop Frame y Full Frame.³⁹

Originalmente, en la señal monocromática, se tenían 30 Frames/seg. exactamente, por lo que el contador de tiempo de cintas de video lo calculaba de esta manera, sin problema alguno. Por lo que en una hora, se tendrán los siguientes datos:

$$\text{No. Frames/Hr} = 30 \text{ Frm/Seg.} \times 36,00 \text{ Seg. / H} = 108\,000 \text{ Frm/Hr.}$$

Con la aparición de la señal de color, la frecuencia de cuadro varió a 29.97 Frm/Seg. Pero el contador de tiempo tiene que leer cuadros completos en la cinta de video, por lo que habrá un exceso de: $30.00 - 29.97 = 0.03 \text{ Frm}$ esto es, 30 cuadros completos en la señal de color NTSC durará más de un segundo, o sea:

$$30.00 - 29.97 = 0.03 \text{ Frm / Seg. de exceso en tiempo.}$$

Si se cuentan, en una Hora, tenemos:

$$0.30 \text{ Frm exceso / seg.} \times 3600 \text{ seg. / hora} = 108 \text{ Frm exceso / Hr, lo cual corresponde a 3.6 seg.}$$

Dicho de otra manera, 108,000 cuadros de señal de color corresponde a Una Hora y 3.6 segundos en tiempo real.

Debido a este inconveniente, es necesario compensar esta lectura de tiempo de la cinta, y se realiza de la siguiente manera: Como existe un exceso de número de cuadros que no existen, se debe de eliminar de la cuenta de cuadros en el contador de tiempo de cinta los 108 Cuadros, para que así una hora de contador corresponda a una hora real. Si eliminamos 2 cuadros de dicha cuenta en cada minuto tendremos que:

$$2 \text{ frames / Min.} \times 60 \text{ Min./Hora} = 120 \text{ frames / Hora}$$

Por lo que se tendrán 120 Cuadros que eliminar y solamente se necesitan quitar 108 cuadros; por lo que no se eliminan en las decenas de minutos 00, 10, 20, 30, 40 y 50; dando con esto seis decenas, lo que corresponde a 12 cuadros.

$$2 \text{ Frames / Min} \times 6 \text{ Decenas de Min / Hora} = 12 \text{ Frames sin eliminar / Hora.}$$

Así, de los 120 Cuadros / Hora a eliminar, se exceptúan los 12 Frm / Hr correspondientes a las decenas, nos quedarán.

³⁹ Manual de Servicio Sony, pagina 110.

120 Frames a eliminar / Hr - 12 Frm decenas sin eliminar = 108 FRM eliminados / HORA

108 Cuadros que son necesarios excluir, para que el tiempo de cinta sea tiempo Real y efectivo.

La cuenta de tiempo puede ser de manera mecánica, electrónica o una pista especial grabada en la cinta llamada Código de Tiempo (TC).

Este tipo de tiempo de cinta que se elimina dos cuadros en cada minuto, excepto en las decenas de minuto, se conoce como DROP FRAME (DF). Al tiempo de cinta que no se elimina ningún cuadro, y que por lo tanto no corresponde al tiempo real (ya que tendrá un exceso de 108 cuadros / hora) se le conoce como FULL FRAME o NON DROP FRAME (FF o NDF).

III.12.1 Señal de Código de Tiempo.

Entre 1956 y 1964, la edición en video Tape, se efectuaba cortando físicamente la cinta y se unían, los tramos de cinta con las escenas deseadas, para seguir así la secuencia y formar un programa, por lo que era lento, tedioso e inexacto.

Después de 1964, ya se introdujeron los sistemas de unión electrónica, los que permitían editar de una VTR a otra, en un empate y coincidencia más precisa, pero no exacta, y la búsqueda del punto de edición se, marcaba con tonos electrónicos de 400 Hz. Era necesario un equivalente electrónico para que identificara cada cuadro de video en forma única para que en una búsqueda electrónica del punto de edición y sincronización entre VTRs sea rápida y precisa.

La compañía EECO diseñó un sistema de edición con el método de un código de tiempo (TC) en 1967. De esta manera, la búsqueda de un cuadro de video específico a editar fue simplemente materia de ingresar a un sistema computarizado el número correspondiente del tiempo y así permitir a dicho sistema que traslade la cinta hasta dicho tiempo.

En 1969, la Sociedad de Ingenieros de Imagen en Movimiento y Televisión (SMPTE) desarrolló el código estándar que permitiría intercambiar entre todas las VTR dicho código; junto con la Unión de Radiotransmisores europeos (EBU), por lo que dicho código de tiempo se llama CODIGO DE TIEMPO SMPTE/EBU.

Así pues, la exactitud, confiabilidad y repetibilidad son importantes para la utilización del código de tiempo SMPTE/EBU. Antes, los sistemas de edición utilizaban los métodos de cuenta de incremento / decremento de localización automática de puntos de edición. Por lo que, los pulsos eléctricos de un tacómetro de contador de cinta o los pulsos de Control Track (CTL) o las marcas grabadas como tonos en la cinta se utilizaban para la localización automática del video, pero estos pulsos podían variar (se

pueden distorsionar al cambio de dirección o velocidad o por Drop-Outs) en la modalidad de Shuttle o Play; por lo que el TC SMPTE/EBU al grabarse en la misma cinta como una pista de audio (longitudinal) o como parte del video (intervalo vertical) corresponderá de manera biunívoca, y éste identificará (le da una dirección) a cada cuadro por Hora, Min., Segundo y cuadros.

III.12.2 Ventajas del Código de Tiempo (TC)

1.- La referencia de que el tiempo es preciso. La duración de una escena puede medirse y determinarse con exactitud de cuadros.

2.- El TC permite la intercambiabilidad y compatibilidad entre sistemas de edición.

Los puntos de entrada (IN) y de salida (OUT) seleccionados en dicho editor puede enlistarse por sus designaciones de TC, ya sea en una lista escrita o en un medio electrónico y meterlo a un segundo sistema distinto, y distante, en el cual las ediciones serán idénticas que en el primero. Esta compatibilidad es la clave del ahorro de tiempo y dinero en un proceso llamado AUTO-ASSEMBLY.

Con este proceso, los programas se denominan como pistas en master en VTRs de 1 pulgada, Betacam o digital (D-1 D-2) y que simultáneamente se graban con el mismo TC en formatos menores como 3/4" (U-matic), VHS, V-8, etc. Para preparar los carretes de trabajo, los que a su vez se elabora una lista de decisiones de edición sin maltratar las cintas master, por lo que se edita con los carretes de trabajo al igual que se va a efectuar posteriormente con dichas cintas master. A esta edición se le conoce ON-LINE.

Como el tiempo indicado en el TC de la lista de decisiones de edición OFF-LINE corresponde al tiempo de las cintas master, las ediciones off-line se duplicarán en las cintas master y las ediciones On-Line serán rápidas lo que evitará dañar las cintas (casi no son tocadas) y evitar borrados accidentales de las señales y mucho más rápida dicha edición si es del tipo Auto-Assembly.

3.- El TC permite la sincronización precisa de las VTR reproductoras (Play) con la VTR grabadora (Recorder). El TC actúa como los Sprockets electrónicos de la cinta. Esto se logra con circuitos sincronizadores - comparadores de cada TC en movimiento. Calculan el tiempo real de edición, verificará la correcta secuencia de cuadro de color (C.F.) y adelantará o atrasará las cintas para que empaten en forma correcta. A esto se debe que a veces no entre en punto exacto la edición si no se observa la secuencia de C. F. Para poder realizar la sincronización, se arrancan dichas cintas 5 segundos previos al punto (Cue) de entrada (In) de edición y a este tiempo se le llama PRE-ROLL.

III.12.3 Código de Tiempo Longitudinal (LTC).⁴⁰

Es una señal eléctrica formada por pulsos digitales llamados Bits. Se divide en palabras de 80 bits en cada cuadro de imagen. Por lo que en un segundo tenemos que:

NTSC: $80 \text{ Bits / Frm} \times 30 \text{ Frm / Seg.} = 2,400 \text{ Bits / Seg.}$

Los 30 Bits/Frm del LTC se dividen como sigue:

Bit 0-3	Corresponde a la información de Unidades de Frames.
8 -9	Decenas de Frames.
16-19	Unidades de segundo.
24-16	Decenas de segundo
32-35	Unidades de minutos.
40-42	Decenas de minutos.
48-51	Unidades de Horas.
56-57	Decenas de Horas.
64 -79	Corresponde a la información de Sincronía, la cual indica el fin de cuadro el principio del siguiente, así como la dirección de cinta.

Esta información es igual en todos los cuadros de video.

BIT 10 Indica si la cuenta de tiempo es tiempo tipo DROP-FRAME o FULL-FRAME (tiempo real o no) .

BIT 11 Corresponde a la identificación del cuadro " A " para una secuencia C.F. correcta (campo 1 del cuadro "A").

BIT 4-7, 12-15, 20-23, 28-31, 36-39, 44-47, 52-55 y 60-63, Estos 32 bits se refieren a los Bits de usuario (User Bits), los que se utilizan para datos de 4 caracteres alfanuméricos para indicar numero de carrete, fecha de grabación, nombre del editor, etc.

Bit 58 No está asignado.

Bit 59 es un bit de paridad.

El LTC se utiliza a una velocidad mayor a 1/5 seg. de velocidad normal en adelante.

Una vez conociendo como se origina y se compone una señal de video con todas sus características, el paso siguiente encontrar la manera de cambiar esta señal a unos y ceros, es decir, hacer una conversión de una señal analógica a una señal digital.

⁴⁰ Manual de Servicio Sony, pagina 119.

Este tema así como todas las variantes de interpretación de esta señal digitalizada tales como formatos y características de luminancia y crominancia entre otros serán tocados en el siguiente capítulo.

CAPITULO IV
IMÁGENES DIGITALES.

IV. IMÁGENES DIGITALES

IV.1 La imagen analógica

La técnica a grandes rasgos consiste en representar la imagen analógica a través de una onda senoidal. Las variaciones en la amplitud o en la frecuencia de esta onda determinan la información a partir de la cual se reconstruye la imagen.

Para la transmisión de las imágenes se emplean dos sistemas:

1. Modulación de amplitud (AM). La amplitud de la señal depende de la densidad de la fotografía, y el nivel o frecuencia de la señal no varía.
2. Modulación de frecuencia (FM). La amplitud de la señal es fija y la frecuencia de la señal varía entre dos límites predeterminados que representan el blanco y el negro.

IV.2 La imagen digital.

Son imágenes representadas de forma comprensible para la computadora, a través de un formato digital (mediante unos y ceros).

IV.2.1 Digitalizador o Digitizer.

Toda imagen es la representación de un aspecto de la realidad, que debe de ser capturada mediante sensores que transformen la información en ella contenida (intensidad luminosa, color, temperatura o cualquier otro parámetro), en señales (eléctricas o no), que serán posteriormente manipuladas. Esta es una etapa de captura, si es en un sistema digital, recibe el nombre de *digitalización*.

"Un digitalizador convierte una imagen en código de una representación numérica inteligible para una computadora digital..."⁴¹

"La computadora exige como característica previa para aceptar cualquier tipo de información que ésta venga dada en forma digital, o sea, como *arrays* finitos de palabras binarias de longitud definida..."⁴²

⁴¹ Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, página 25.

⁴² Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, página 27.

En la actualidad los digitalizadores más usados son los *scanners*, las cámaras de video y los CCD de estado sólido y en si de lo que se trata es convertir una imagen analógica y continua en otra digital y discreta.

Anteriormente se mencionó como un rayo de electrones es usado para crear una imagen en un CTR (tubo de rayos catódicos), así mismo se mencionaron las que son las señales monocromáticas y de color, y se puso especial atención en el tipo, número y duración de varios pulsos de sincronía encontrados en la señal de video. Estos pulsos juegan un papel muy importante para el desarrollo del digitalizador.

"A grandes rasgos, la función de un digitalizador es remover la porción de sincronización de los componentes de la señal de video antes de convertir la amplitud de la señal análoga en una representación digital..."⁴³

El digitalizador o la computadora debe de tener la suficiente inteligencia para entender la señal de video para que la formación de las imágenes pueda ser correctamente interpretada.

Los digitalizadores de video tienen muchas formas y tamaños con variaciones en sus alcances y funcionalidad. Los más importantes atributos son:

- .Capacidad de digitalización Cuadro contra Campo.
- .Capacidad de digitalización de video Still - Frame (cuadro sin movimiento) contra tiempo real.
- .Resolución.
- .Capacidades monocromáticas contra color.
- .Capacidades de procesamiento de imagen On - Board (a bordo).
- .Capacidades de salida de video (video output).
- .Prueba de ancho de bit (sample bit width).

Los digitalizadores también son llamados *grabadores de cuadros*, y permiten que una señal de video se digitalice en tiempo real al digitalizarse los dos campos que forman un cuadro de video. La digitalización en tiempo real permite que los objetos captados por una cámara puedan tener movimiento sin que afecte la misma digitalización. Esto es posible como larga es la digitalización de un cuadro de video en menos de 1/30 de segundo y el objeto no se mueva durante el proceso de digitalización. La digitalización desde una fuente de televisión o un VCR (video cassette recorder) es posible con un digitalizador en tiempo real porque el digitalizador puede mantener actualizado la velocidad de la imagen en video. No se podría usar un digitalizador de Still - frame ya que es un digitalizador estacionario y cualquier movimiento en la imagen resultará como una imagen distorsionada.

⁴³ Craig A. Lindley, Practical Image Processing Inc, Ed. Jhon Wiley & Sons. 1991 pagina 54.

La resolución define el número de píxeles (*picture elements*) que el digitalizador es capaz de producir. Un digitalizador en *tiempo real* utiliza una matriz cuadrada de píxeles que es de 256 por 256 o de 512 por 512. La imagen será entonces de 256 píxeles horizontales por 256 líneas verticales. Los números binarios son escogidos por las dimensiones de la matriz por el requerimiento de memoria en el digitalizador.

Los digitalizadores pueden muy bien ser adaptadores ya que tienen compuestos números de bits usados para representar cada muestra de imagen o píxel. Cada muestra representa la intensidad de la imagen en la posición donde el píxel fue digitalizado. Los demás bits en una muestra, tienen una precisión en la representación digital de la imagen.

Los digitalizadores de color generalmente tienen tres sets de A - D (conversión análoga - digital) y estos sets tienen tres memorias de imagen, una para el rojo, otra para el verde y otra para el azul que son los componentes de la imagen. Los tres sets de digitalización trabajan para cada elemento de la imagen.

Una característica de algunos digitalizadores es que cuentan con un procesador integrado, esto hace que se ahorre tiempo sin que tarde en hacer la transferencia a la computadora.

Muchos digitalizadores tienen una salida de video para sacar la señal directamente a un monitor o un televisor.⁴⁴

Los digitalizadores externos se conectan a la computadora a través de los puertos de impresora. Estos puertos de impresión fueron escogidos en lugar de los puertos seriales o las tarjetas de buses de interface por:

- Los datos pueden ser transmitidos a mayor velocidad en un puerto paralelo que con un puerto serial.
- Una sola tarjeta de bus de interface puede ser requerida para la línea original de la PC los micro canales PS/2. Esta tarjeta puede estar integrada en algún slot dentro de la computadora.
- El puerto de impresión tipo Centronix desarrollado por IBM se ha convertido en un standard para todas las computadoras.
- La interfaz de impresión es la más compatible en todas las computadoras.
- El puerto paralelo de conexión requiere menos circuiteria.

Los puertos paralelos tienen en total 12 líneas de salida y 5 líneas de entrada de 25 pins. Cuando el digitalizador es conectado a un puerto paralelo las funciones que tenía el puerto cambian. Uno de los cambios más grandes es que el digitalizador tiene que ser

⁴⁴ Craig A. Lindley, Practical Image Processing Inc, Ed. Jhon Wiley & Sons. 1991 paginas 54 y 55 .

capaz de pasar el pixel digitalizado y toda la información sincronizada a la computadora usando solamente 5 líneas de entrada disponibles.

Considerando que el convertidor analógico - digital tenía seis bits de ancho y al menos tres bits de estatus eran requeridos, fue necesario multiplexar la función de las líneas de entrada y hacer el puerto bidireccional.⁴⁵

IV.2.2 Señal Digital de video.

La señal digital de video esta hecha de dígitos binarios o bits, señales que están presentes o ausentes. El bit generalmente es representado por una señal de onda rectangular, pero puede tener cualquier forma que defina que esta presente o ausente. En términos eléctricos el bit tiene tres formas: "on - off", "plus - minus", y "1 - 0".

En lenguaje de computadoras una secuencia de 8 bits quiere decir byte. Este número de bits es suficiente para representar todos los caracteres (letras minúsculas, mayúsculas, números, signos de puntuación etc.) definido por el American Standard for Code Information Interchange (ASCII). La capacidad de memoria en circuitos integrados esta dada en bytes. Como los "256 kilobytes" de memoria RAM que es un chip que puede almacenar 262,144 bytes o 2,097,152 bits.

El termino general de un conjunto de bits es la *palabra*, usualmente de una longitud determinada. Para las computadoras la longitud de una palabra está dada por los bits que tiene cada posición de memoria, en varias formas 16 bits en una computadora personal hasta 60 en una super computadora. En televisión las longitudes de palabra por lo regular son de 7, 8 ó 9 bits, hay tres formas de la palabra de 8 bits, on-off, plus - minus, con su correspondencia con los ya designados 1's y 0's.

Desde que cada bit en una palabra de 8 bits puede tener combinaciones de 1's ó 0's, hay $2^8 = 256$ diferentes. Por lo tanto una palabra puede representar 256 niveles de señal diferentes. Con cada nivel se representa $1/256 = 0.39$ por ciento del total del rango.

El error en la amplitud es del 0.4 % después de la cuantización. Esta pequeña diferencia en la amplitud es suficiente para tener alta calidad en la representación de una señal de video, y en la palabra de 8 bits sigue las especificaciones según los estándares de CCIR. Las palabras de 7 bits, primeramente usadas en la recepción para el procesado, tiene 128 niveles, el error es equivalente a un rango entre 42 y 48 décimas de Byte debajo de el máximo nivel de la señal. Este rango de 42 a 48 dB se compara con los 48 dB de la señal de ruido que comúnmente es aceptada sin afectar el buen desarrollo en la salida de una

⁴⁵ Craig A. Lindley, Practical Image Processing Inc, Ed. Jhon Wiley & Sons. 1991 pagina 59.

cámara a color. Las palabras de 9 bits son usadas para dar mayor precisión en el proceso con equipo digital de video.⁴⁶

La señal producida por una cámara es una señal analógica de la incidencia de luz sobre una superficie sensitiva, distribuida a lo ancho y a lo alto, y discernida por rayos con movimiento, en tiempo como una secuencia de elementos de imagen en arreglos de campos o cuadros. La salida de la cámara está continuamente variando voltajes que de un tiempo a otro puede tener valores dentro de un rango de operación. El mismo tipo de variación de voltaje continuo es requerido en las tres terminales de entrada del tubo de imagen. Entre la cámara y la pantalla existen procesos para mantener la señal en una forma analógica.

Las señales analógicas tienen muchas formas de distorsión y de interferencia que a veces son imposibles de remover. Si a esto se le agrega el ruido a la señal o si el rango de la amplitud es distorsionado por no tener linealidad, estos cambios se convierten en parte de la señal como si estos fueran generados por la cámara. Como una señal analógica pasa a través de todos los circuitos de la cámara hacia la pantalla, esta cae presa de cambios no deseados.

Si de cualquier forma la señal analógica ha cambiado después de salir de la cámara a una forma digital, la señal digital puede ser regenerada antes de que se convierta en una señal débil y ser afectada por ruido, la señal digital no es objeto de distorsión. En teoría una señal digital puede ser transmitida sin error al dispositivo de despliegue o la pantalla. Además puede ser convertida a una señal analógica que puede ser una copia exacta de la salida de una cámara.

Hay tipos de interferencias y otras formas de daño que afecten al digitalizar, pero por protección esto se puede evitar añadiendo bits extras, no necesariamente en el contenido de la señal, sino solo para proveer de redundancia y corrección de errores. Para llevar a cabo este proceso, la señal analógica debe de ser convertida a una forma digital (conversión A - D) y después transmitida y procesada en una señal analógica de nuevo (conversión D - A).⁴⁷

Para convertir una señal analógica a una forma digital, es necesario una sucesión de palabras digitales que contengan solo dos valores discretos, 0 y 1. La señal analógica de entrada puede ser confinada al límite del espectro, entonces pasa por un filtro pasa - bajo (low - pass) anterior al convertidor. El convertidor propiamente primero efectúa muestreos de la señal analógica de entrada, midiendo la amplitud y regularizándola en

⁴⁶ K. Blair Benson Donald G. Fink. HDTV Advanced Television for the 90s. Ed. Mc. Graw Hill 1991 objetivos 4.2 y 4.1.

⁴⁷ K. Blair Benson Donald G. Fink. HDTV Advanced Television for the 90s. Ed. Mc. Graw Hill 1991 objetivo 4.2.

intervalos discretos de tiempo. Éstas amplitudes individuales son igualadas en el *quantizer* (contador), contra un largo número de niveles discretos de amplitud (256 niveles para convertidores de palabras de 8 bits). Cada uno de esos niveles discretos puede ser representado por una palabra de 8 bits, en su equivalente digital. El proceso de igualación de cada amplitud discreta con una única palabra es llevado al *encoder* (codificador), aquí escanea la lista de palabras, las recoge y las presenta. El codificador pasa las series de palabras codificadas en secuencias que corresponden a la señal analógica que fue probada. Ésta "corriente de bits" es la versión digital de la entrada analógica.

La lista de palabras digitales correspondiente a las muestras de amplitudes se conoce como *code* (código).⁴⁸ La siguiente tabla representa un ejemplo de los niveles de amplitud y sus rangos en palabras de 8 bits, son tres rangos: 0 - 15, 120 - 135 y 240 - 255. Las señales codificadas en esta forma están moduladas por pulsos de código (*pulse-code-modulated PCM*). Leyendo los valores de derecha a izquierda, los valores son 1, 2, 4, 8, 16, 32, 64 y 128. El nivel de la amplitud 120 en la tabla tiene un valor binario de 01111000, con 1's en los lugares 8, 16, 32 y 64, por lo tanto su amplitud vale $8 + 16 + 32 + 64 = 120$.

**Valores Binarios para niveles de Amplitud.
(Palabras de 8 bits).**

Amplitud	Binario	Amplitud	Binario	Amplitud	Binario
0	0	120	1111000	240	11110000
1	1	121	1111001	241	11110001
2	10	122	1111010	242	11110010
3	11	123	1111011	243	11110011
4	100	124	1111100	244	11110100
5	101	125	1111101	245	11110101
6	110	126	1111110	246	11110110
7	111	127	1111111	247	11110111
8	1000	128	1000000	248	11111000
9	1001	129	1000001	249	11111001
10	1010	130	1000010	250	11111010
11	1011	131	1000011	251	11111011
12	1100	132	1000100	252	11111100
13	1101	133	1000101	253	11111101
14	1110	134	1000110	254	11111110
15	1111	135	1000111	255	11111111

⁴⁸ K. Blair Benson Donald G. Fink. HDTV Advanced Television for the 90s. Ed. Mc. Graw Hill 1991 objetivo 4.1.

El código PCM es aún utilizado pero existen otros códigos más elaborados.

La gran velocidad con la que el convertidor debe de igualar las muestras analógicas a las palabras con código digital (cientos de millones de bits por segundo), no puede ser manejada por los circuitos que operan secuencialmente. En su lugar, son manejados canales paralelos, uno por cada uno de los 256 códigos de palabras en el muestreo de 8 bits. Evidentemente este es un gran problema ya que los circuitos integrados deben de funcionar con un preciso tiempo de sincronía. Esto se resolvió agregando un reloj. Todos los convertidores requieren un reloj para tener el completo control de muestreo de los canales paralelos.

Una referencia de voltaje es aplicada a través de los 255 resistores en serie. Muestras pequeñas de voltaje son tomadas de cada resistor. El fondo de las muestras es cero voltaje (el nivel 256), y cada muestra sucesiva presenta un voltaje mayor que el anterior, se tiene que cada uno representa $1/255$ de esos voltajes de referencia. Cada muestra de voltaje llega a la entrada de un comparador donde otra señal de entrada analógica está conectada. Como la señal varía, cada comparador determina si la señal está arriba o abajo del nivel original de voltaje aplicado. Si está arriba el comparador genera un 0 y si está abajo genera un 1. Las salidas del comparador están alimentando al convertidor, que está sincronizado para aceptar los dígitos del comparador en paralelo y convertirlos en una forma serial. Produciendo una palabra de 8 bits por cada muestra, los convertidores están conectados en cascada para poder manejar ambos rangos de referencia el mayor y el menor, estas dos salidas determinan donde aparecen primero y al último los 0's y los 1's...⁴⁹

IV.2.3 Conversión digital - analógica.

La conversión digital - analógica (D-A), de palabras de 8 bits tiene 8 entradas, una por cada bit en la palabra. Su función es traducir cada palabra a su correspondiente nivel de amplitud que fue cuantificado, procediendo según la tabla vista anteriormente. La sucesión de palabras en la corriente de bits produce una sucesión de amplitudes que trazan una onda que representa una señal analógica. Las amplitudes generadas de esta manera no son idénticas al tamaño de las muestras de la señal de entrada, tienen un error de cuantización, así que la onda analógica de salida despliega una serie de pasos por minuto. El contenido espectral de estos pasos está fuera del espectro de la señal analógica de entrada y estos pasos son removidos pasando al convertidor D - A por un filtro pasa - bajo así como pasó en la señal de entrada del convertidor A - D.

⁴⁹ K. Blair Benson Donald G. Fink, HDTV Advanced Television for the 90s. Ed. Mc. Graw Hill 1991 objetivo 4.4.

ESTA TESIS NO SALE
DE LA BIBLIOTECA

El convertidor D-A esta orientado con los mismos requerimientos para la operación a gran velocidad como el convertidor A - D, las palabras de entrada llegan a velocidades de millones de bits por segundo .

Un convertidor D -A opera creando corrientes (que iguala el lugar de cada 1 en la palabra). Sumando las corrientes para crear su correspondiente nivel de amplitud, el cual sirve como salida de la amplitud de esa palabra.

Una forma física de un generador de corrientes consiste en grupos de resistencias en paralelo, de 1, 2, 4, 8, 16, 32, 64 y 128. Cuando un voltaje es aplicado a través de los resistores, la corriente pasa por los rangos entre 1 y 128, cada valor de la corriente es seleccionado por 8 switches los cuales corresponden al 0 ó 1 del bit que reciben, la suma de los switches prendidos (on) representa la salida de la amplitud correspondiente a la palabra de entrada, la corriente es convertida a voltaje al pasar a través de un amplificador operacional estabilizador.⁵⁰

IV.2.4 Procesamiento Digital de Imágenes.

El termino "procesamiento digital de imágenes" generalmente se refiere al proceso de imágenes bidimensionales en una computadora digital. Una imagen digital es un arreglo de números reales o complejos representados por un número finito de bits.

Una imagen dada por una transparencia, diapositiva, fotografía o video es digitalizada y almacenada como una matriz de dígitos binarios en la memoria de la computadora. La imagen digitalizada puede ser procesada y/o desplegada sobre un monitor de televisión de alta resolución. Para desplegarla, la imagen es almacenada en un buffer de memoria de acceso rápido el cual repasa el monitor a 30 cuadros por segundo para producir una imagen visible en tiempo real.

Mini o micro computadoras se usan algunas veces para comunicar y controlar la digitalización completa, el almacenamiento, procesado y las operaciones de desplegado en redes de computadoras.

Pasos a seguir en el sistema:

Objetivo → observar → digitalizar → comprimir → almacenar → procesar → almacenar → desplegar o grabar...⁵¹

⁵⁰ K. Blair Benson Donald G. Fink. HDTV Advanced Television for the 90s. Ed. Mc. Graw Hill 1991 objetivos 4.7.

⁵¹ Anil K. Jain. Fundamentals of Digital Image Processing. Prentice Hall 1989. Paginas 1 y 2.

Así, como se menciona en los pasos anteriores, nos lleva a un punto muy importante, la Compresión de datos, el digitalizar es un proceso que puede consumir toda la memoria de una computadora, lo que hace indispensable este proceso. En seguida los detalles.

IV.3 Compresión de datos

A medida que incrementa el rendimiento en los sistemas de computadoras y de comunicación digital, la implementación de mejores técnicas de compresión de datos son necesarias para la tecnología moderna de multimedia.

Compresión de datos es la ciencia de remover información redundante de un mensaje para minimizar la cantidad de bits que se necesita para guardar información en disco o transferirla a través de un "network" (estación de televisión). Cuando tenemos sistemas con anchos de banda ilimitado la compresión de datos no hace mucha diferencia. Pero cuando visualizamos la realidad del presente, vemos que si hace bastante diferencia la utilización de compresión de datos.

Un ejemplo de un esquema de compresión de datos son los sistemas de comunicación digital diseñado por Samuel Morse. Al asignar representaciones cortas a los caracteres mas frecuentemente usados, los telégrafos podían enviar mensajes más rápidos que si se representaran los caracteres con su tamaño normal, como en el caso de los códigos ASCII. Un texto codificado en ASCII se puede reducir a la mitad del tamaño normal al comprimirlo utilizando métodos de compresión de datos. Un dato importante de los esquemas de compresión de textos es que no hay perdida de datos. No se pierde información en el proceso de compresión y descompresión.

Otro ejemplo es el caso de compresión de imágenes. Cuando se comprimen imágenes la virtud de "non-lossiness" o no perdida de datos no es como la característica del sistema visual humano para no utilizar grandes radios de compresión. Un ejemplo es el siguiente: el ojo humano es menos sensitivo al color azul que al color verde. Si tomamos ventaja de esto cuando comprimimos una imagen quitamos alguna información del color azul y probablemente nuestros ojos no lo detectarían. Algunos de los esquemas recientes para compresión de imágenes lo son JPEG, MPEG y más recientemente Philips y SONY han introducido audio- compresión en DCC(Digital Compact Cassette) y MD(Mini Disk).

Los métodos de compresión de datos se dividen en dos categorías: perdida y no- perdida. La primera categoría es útil donde se permite la perdida de calidad de transmisión y almacenamiento mientras se retenga el contenido general de los datos como imágenes y transmisión de audio. La segunda categoría, método de no-perdida, usa el factor de que algunos símbolos de datos son usados mas frecuentemente que otros. Códigos de largo variable son asignados a los datos entrados con códigos cortos que son asignados a los símbolos o secuencia de símbolos mas frecuentes y códigos largos a símbolos raros.

Algunos de los métodos y técnicas utilizadas para la compresión de datos lo son LZ, Huffman, Markov, ACB, PPM, y DMS.

IV.3.1 En qué se basa la compresión de datos

El término compresión de datos se refiere al proceso de reducción del volumen de datos necesario para poder representar una determinada información (debemos aclarar que los términos datos e información no son sinónimos, es decir, los datos son el medio a través del cual se conduce la información).

La redundancia de los datos es un punto clave en la compresión de datos digitales. Este concepto es una entidad matemática cuantificable. Si se tienen los conjuntos de datos que representan la misma información; entonces podemos conocer la relación de compresión.

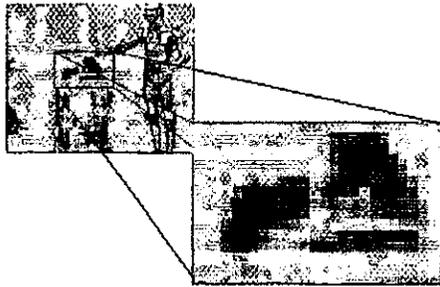


Figura 4.1

En esta imagen (figura 4.1) se toma solo una parte del histograma de color que forma una imagen cada cuadro representa un pixel el cual contiene su propia información de color, cuando estos pixeles se repiten también se repite la información que contiene, a esto se le llama redundancia relativa.

La redundancia relativa de los datos del conjunto de datos representado por las imágenes se puede ver de la siguiente forma, tomando un ejemplo de como 10:1, esto indica que respecto al segundo conjunto de datos, en el primero no existen datos redundantes. Una relación de compresión práctica como 10:1 significa que el primer conjunto de datos contiene 10 unidades de transporte de información (bits) por cada unidad del segundo conjunto de datos (comprimidos). La redundancia correspondiente, 0,9, implica que el 90 por 100 de los datos del primer conjunto son redundantes.

En la compresión de imágenes podemos aprovechar tres tipos de redundancias: redundancia de codificación, redundancia de entre pixeles y redundancia psicovisual, temas que se tocarán más adelante.

IV.3.2 Métodos de compresión de datos.

LZ - La compresión LZ substituye el texto inicial con referencias a un diccionario; esto se asemeja a la metodología usada por dos interlocutores. Con el crecimiento del diccionario, el numero de bits necesario para formular la referencia crece proporcionalmente por un logaritmo binario al tamaño del diccionario. El tamaño de las frases en el diccionario(en el caso mas simple, un árbol binario) al principio aumenta el tamaño de la referencia, pero luego de cierta saturación, la velocidad del crecimiento tiende a una dependencia logarítmica. El tamaño optimo del diccionario varia por diferentes tipos de datos; mientras mas variable los datos, mas pequeño el tamaño optimo del diccionario.

PPM - La idea de modelo de contexto esta basado en el factor de que la distribución de posibilidades en el alfabeto depende en el contexto más cercano, en otras palabras, las letras son mas comunes que aparezcan en un patrón particular, o sea, cerca de otras letras particulares. Para esta tecnología existen una serie de restricciones para "sliding frame" o movimiento del procesamiento de datos o letras, del cual un modelo de contexto es creado. Con el incremento de tamaño de "sliding frame", mas y mas variedad de información es procesada , con el resultado que la distribución de posibilidades aumenta y la efectividad del modelo de contexto decrece rápido. Con el incremento del largo del contexto, los costos incrementan exponencialmente.

ACB - El principio general de la compresión ACB es el siguiente: el método ACB no visualiza el marco total pero si fragmentos de este, los cuales están cerca por contenido al contexto mas cercano. Fragmentos o frases de la imagen recibida difiere mucho en calidad; frases cerca por contenido son distintas en su relación a otras frases, frases mas lejanas son menos distintas. En este caso, el método ACB tiene condiciones ideales para trabajar, en estas este trabaja más rápido considerando unidades mas largas, mas aun, este puede distinguir el valor de las frases, considerando la probabilidad de aparición o uso. En caso de incremento ilimitado de un marco o tamaño de imagen, la compresión ACB siempre asegura un tamaño optimo del diccionario de contexto.

DMC - Modelos probabilisticos con numero finito de estados puede ser descrito por un automatizador finito. Un conjunto de estados $S(i)$ y un conjunto de probabilidades de transición $P(i,j)$ del estado i al estado j son llamados modelos de Markov. Marcov's Dinamic Coding(DMC) es un interés practico, es este el DMC trabaja adaptándose, comenzando desde un modelo inicial simple y añadiendo, de ser necesario, nuevos

estados. Tecnología PPM es un caso particular de DMC. DMC aloja la construcción de modelos de contexto no solo para símbolos simples, también para frases o líneas. En este sentido ACB puede ser clasificado como una variedad de DMC.

IV.4 Video Compresión

La televisión digital es ya algo común en muchos estudios. Sin embargo, la transmisión de la señal digital no es practica debido a que las imágenes contienen demasiados datos. La solución se encuentra en la compresión de video.

Desde los primeros días de la industria cinematográfica, lo que han contemplado los espectadores como imágenes en movimiento realmente han sido secuencias de imágenes fijas. Por supuesto que el truco radica en que las imágenes cambian lo suficientemente rápido para convencer al ojo que lo que esta observando es un movimiento real. En televisión incluso, lo que contemplamos no llega a ser un cuadro completo cada vez. Los diseñadores se apoyan en la persistencia de la visión de los ojos para crear la ilusión de que la imagen completa esta en movimiento, cuando en realidad lo que se mueve son los tres haces electrónicos en el tubo de rayos catódicos. No obstante, los resultados son bastante aceptables.

Con la llegada de la televisión digital, los ingenieros se han sentido tentados a extender sus poderes ilusionistas todavía más. La señal de video digital esta libre de ruidos y es muy sólida, pero por otra parte, ocupa demasiado ancho de banda para lograr una grabación en cinta que resulte económica, al igual que la grabación en disco o la transmisión por el aire.

Afortunadamente, en la mayoría de las imágenes de televisión hay muchos detalles que se repiten en los fondos de dichas imágenes, en el cielo azul y en los cuadros comunes sucesivos, los cuales simplemente pueden ser descartados sin que el ojo se entere que nuevamente lo han engañado. A este proceso se le llama compresión digital del video.

IV.4.1 Compresión espacial.⁵²

La palabra espacial se refiere al espacio ocupado por una sola imagen y el objetivo de la compresión espacial es minimizar la duplicación de datos en cada imagen. La reducción de la velocidad de bits en la compresión espacial se logra primeramente transformando los datos a partir del dominio en el espacio y en el tiempo, y llevándolos al dominio de la

⁵² <http://www.udec.cl/investigación/apuntes/video>

frecuencia, usando para ello el método de la Transformada del Coseno Discreto (DCT). Después aplicando las técnicas de cuantificación y de codificación por longitud variable, para reducir de esta manera la velocidad de bits.

El video es normalmente mostrado sobre un dispositivo basado en el tiempo, como puede ser un monitor de formas de ondas en vez de un analizador de espectro basado en la frecuencia, pero para desarrollar la reducción de datos, debemos transformar primero los datos del video al dominio de la frecuencia. Es aquí donde se utiliza el DCT (que es una fórmula trigonométrica derivada a partir del análisis teórico de Fourier) para transformar los datos de cada bloque de 8 x 8 píxeles en bloques de coeficientes de frecuencia de 8 x 8. En el dominio de la frecuencia, la mayor parte de los elementos de imagen de alta

8	4	2	1	0	0	0	0
4	2	1	0	0	0	0	0
2	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	1	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

Figura 4.2

energía (y, por tanto, los más notables) están representados por las bajas frecuencias en la esquina superior izquierda del bloque como se muestra en la figura anterior (figura 4.2), y los detalles menos importantes están representados como las frecuencias más altas en la esquina inferior derecha. Donde hasta este momento no se ha descartado ningún bit.

Después de la codificación DCT, el dato está sujeto a un proceso de cuantificación donde es analizado para reducir los datos en las altas frecuencias que es donde el ojo es menos sensible. Utilizamos más bits por píxel para cuantificar los importantes coeficientes de baja frecuencia, y menos bits por píxel para los coeficientes de alta frecuencia. Las componentes de DC son cuantificadas normalmente con 10 bits, porque si emplea una cuantificación más gruesa de muy baja frecuencia, los bloques en sí mismos se harían visibles en las imágenes. Hasta se ha logrado el primer paso en la reducción espacial de los bits.

Para crear el flujo de bits de video comprimido, los 64 coeficientes de frecuencia son explorados en forma de zig-zag como en la figura siguiente (figura 4.3), desde la esquina izquierda superior hasta la esquina derecha inferior. A su vez, como puede verse arriba en la figura, las áreas de alta frecuencia están representadas por flujos de ceros. Una

reducción de datos adicional puede lograrse ahora transmitiendo solamente los números de ceros, en vez de los valores usuales de los coeficientes.

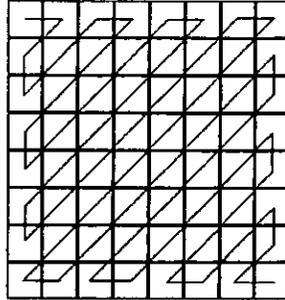


Figura 4.3

La última etapa del proceso de compresión espacial, utiliza la Codificación por Longitud Variable (VLC). La VLC, asigna palabras de códigos más cortos para los eventos que ocurren con más frecuencia y palabras con códigos más largos, para los eventos menos frecuentes; ello es también reversible. Los sistemas JPEG y MPEG emplean estos métodos de compresión espacial para la reducción de la velocidad de los bits, estos sistemas serán analizados más adelante.

Existen diferentes técnicas para llevar a cabo una compresión, como ya se ha mencionado es necesario la utilización de códigos para ello, uno de los más recurridos es la Codificación Huffman.

IV.4.2 Codificación de Huffman

La técnica más popular para la eliminación de la redundancia de codificación es la de Huffman. Cuando codificamos individualmente los símbolos de una fuente de información, usando esta técnica se obtienen el número más pequeño posible de símbolos de código por símbolo de la fuente.

El primer paso del método de Huffman consiste en realizar una serie de reducciones partiendo de los símbolos que consideramos y sus correspondientes probabilidades (los cuales tenemos ordenados). Los símbolos de menor probabilidad son combinados en un único símbolo que los sustituye.

Reducciones de una fuente. El Método de Huffman se refiere a:

En la columna situada más a la izquierda se representa un hipotético conjunto de símbolos fuente y sus probabilidades (ordenados de mayor a menor probabilidad). Para realizar la primera reducción de la fuente, se combinan las dos últimas probabilidades, 0,06 y 0,04, formando un símbolo compuesto (que representa a las anteriores) de probabilidad 0,1. Éste pasará a la primera columna de reducción de manera estas nuevas probabilidades se encuentren también ordenadas. Este proceso lo repetiremos hasta que consigamos una fuente reducida con dos símbolos.

La segunda etapa del procedimiento consiste en codificar cada fuente reducida, empezando por la más pequeña hasta llegar a la fuente original. El código binario de longitud mínima para una fuente de dos símbolos, está compuesto por los símbolos 0 y 1. Estos símbolos se asignan a los dos símbolos de la derecha. Este símbolo de la fuente reducida cuya probabilidad es 0,6 se generó combinando dos símbolos de la fuente reducida situada a su izquierda, el 0 utilizado para codificarlo se asigna ahora a ambos símbolos y se añade arbitrariamente un 1 y un 0 a cada uno de ellos para distinguirlos entre sí. Esta operación se repite hasta que lleguemos a la fuente original.

Asignación de códigos. Método de Huffman

Si nos hemos fijado, vemos como este código tiene la restricción de que los símbolos se deben codificar uno a uno. Una vez hecho esto codificar/decodificar, se limita a una simple consulta en una tabla. Si tenemos una cadena de símbolos codificados, éstos son decodificables de manera única, porque cualquier cadena de símbolos de código se puede decodificar de forma única. Por tanto, para toda cadena de símbolos codificados por Huffman se puede decodificar examinando individualmente los elementos de la cadena de izquierda a derecha. Si tenemos como ejemplo la cadena 010100111100; vemos como la primera palabra de código válida es 01010, que es la primera parte del código. La siguiente parte válida es 011.

Este método es usado por los formatos gráficos PCX y GIF. Reducen zonas idénticas dentro de una imagen.

IV.4.3 ¿Por qué la Compresión de Video?

La digitalización y el almacenamiento de una secuencia de diez segundos de vídeo en movimiento a tiempo real requiere de la transferencia de una cantidad enorme de datos en un periodo corto. Para reproducir un solo cuadro de vídeo digital con componentes a 24 bits se necesita cerca de 1MB de datos de computadora; diez segundos de vídeo llenan un disco duro de 300MB. El vídeo de pantalla completa con movimiento a tiempo real necesita que la computadora proporcione datos a una velocidad de 30MB por segundo;

esto es mucho más que lo que puede manejar un PC. Un disco duro típico transfiere datos a sólo 1MB por segundo, y un reproductor de CD-ROM a velocidades pequeñas que van desde 150K por segundo a 1,2 MB por segundo.

Este problema es superando con las técnicas de compresión de imágenes. Los métodos de compresión de imágenes de vídeo a tiempo real, como el JPEG, MPEG, DVI y C-Cube se encuentran disponibles para comprimir información de vídeo digital en relaciones que van de 50:1 a 200:1. Los esquemas de compresión JPEG y MPEG utilizan el método de transformada discreta del coseno (DCT), un método que cuantifica la habilidad del ojo humano para detectar las distorsiones de color e imagen.

La compresión de una señal es un proceso de digitalización de una señal análoga de televisión en ceros y unos. La compresión de vídeo elimina datos redundantes que hacen que disminuyan las cantidades de información en cada cuadro del vídeo. La señal codificada o comprimida es decodificada o descomprimida en el aparato de televisión o en el procesador de un PC, después de obtener la información desde una memoria secundaria o de almacenamiento (como un CD-ROM). En términos simples la función de la compresión de vídeo es hacer más fácil y liviano el transporte de información, y posteriormente procesarla según sea conveniente.

De esta forma la compresión de video, desde una señal análoga a una digital, reduce los costos de distribución, disminuye su tamaño, facilita su transporte, incrementa la calidad y seguridad de la señal, entre otras ventajas.

Existen muchos esquemas de compresión, los cuales son denominados CODEC (compresores/decompresores). Cada uno de estos tiene sus propias ventajas y debilidades, capacitando a algunos para actuar en ciertas condiciones o especializados en ciertas tareas. Cada CODEC determina que tan bien será la imagen y que tan rápido fluirán las imágenes.

La compresión produce una codificación de las imágenes mediante un método de compresión. Básicamente estos métodos dividen una imagen o cuadro de vídeo en bloques de pixeles y examinan los datos redundantes.

Este proceso se conoce como análisis de cuadros, debido a que los métodos realizan comparaciones entre los cuadros de imágenes, eliminando la información redundante, reduciendo así el tamaño del archivo. Esta técnica es conocida como compresión por pérdida y es la técnica utilizada por MPEG. En general, el truco de la digitalización mediante compresión es balancear esta eliminación de información con la calidad de la imagen resultante. Una técnica efectiva eliminará información redundante hasta un nivel imperceptible para el ojo humano.

Un parámetro para comparar la calidad de los distintos CODECs es ver que tan simétricos o asimétricos son estos, aunque CODECs simétricos son utilizados en distintas aplicaciones que los CODECs asimétricos. Un CODEC simétrico es aquel que utiliza una igual cantidad de tiempo tanto en compresión como en descompresión. Estos son comunes en aplicaciones de captura de vídeo en tiempo real, como las videoconferencias.

En cambio, un CODEC asimétrico empleara mucho más tiempo en comprimir una señal de vídeo que en descomprimirla. Ocupa mas tiempo en la etapa de compresión para tratar de alcanzar un nivel mayor en la calidad de la imagen y en el intento de comprimir lo mas posible. Este tipo de compresión funcionan de mejor forma en aplicaciones para CD-ROM.

Otro factor importante para comparar CODECs es su rendimiento, los que pueden variar desde una compresión a la mitad de la información (2:1) hasta otros que comprimen a sorprendentes niveles como 200:1, lo que obviamente es beneficioso a la hora de poder trasladar información desde una unidad de memoria secundaria hasta el procesador.

Aun así existen otras estrategias utilizadas por los CODECs para optimizar su rendimiento, tales como:

Técnicas de compresión intracuadros, originalmente diseñadas para fotografías estáticas de alta calidad (idea en la que se basa MPEG).

Métodos de transformación vectorial, que se basan en la transformación de anchos de onda.

Tecnologías de fractales, la cual se basa en las similitudes entre una agrupación de objetos y otra.

Tomando en cuenta las agrupaciones y las técnicas de compresión existentes es conveniente tocar el tema de la representación de los datos a continuación se describe las relaciones entre los pixeles y la forma en como son tomados en cuenta.

IV.4.4 Métodos de compresión de imágenes.

JPEG - El compresor de imágenes JPEG utiliza métodos matemáticos basados en una función llamada DCT (Discrete Cosine Transform). El DCT define una relación matemática entre valores de pixeles en bloques 8×8 . Los datos que representa la relación DCT es comprimida usando una función de cuantización. La cantidad de datos resultantes puede ser eficientemente comprimida. Un esquema decodificador sigue los

pasos matemáticos en orden inverso para reconstruir la imagen, resultando en un archivo que es casi igual al archivo original.

IMAGE PAC - incorpora una decimación jerárquica (reduciendo el número de píxeles usados para representar la imagen) y una función de interpolación (restaurando el número de píxeles de la imagen), usada para asegurar que los archivos de imágenes puedan ser reconstruidas eficientemente en resoluciones variantes. Los mismos datos son usados para construir diferentes resoluciones para los archivos de Image Pac. Los datos de la imagen Base son usados para construir la imagen llamada Base4, y los datos de la imagen Base4 son usados para construir una imagen Base16. El proceso de compresión Image Pac usa cuantización para comprimir los datos residuales, y por eso se pierden datos que no son restaurados, esta cantidad de datos perdida es mínima.

Los métodos de codificación tienen dentro de sus procesos que utilizar o eliminar información repetida es decir, información redundante las cuales pueden ser entre píxeles, psicovisual etc.

IV.5 Redundancia.

IV.5.1 Redundancia de codificación

El histograma de niveles de gris de una imagen puede servir para la obtención de códigos que permitan reducir la cantidad de datos necesarios para representar información. Esta información está representada como lo mencionan los párrafos anteriores en palabras de bits, es decir, la longitud media de las palabras código asignadas a los valores de los diversos niveles de gris se halla sumando el número de bits empleados para representar cada nivel de gris y la probabilidad de que aparezca este nivel.

Los diferentes niveles de gris están asignados a palabras código con el mismo tamaño. Si ahora por el contrario asignamos palabras código más cortas a los niveles de gris que aparecen con mayor frecuencia y palabras código más largas a los niveles de gris que aparecen con muy poca frecuencia el número medio de bits para codificar la imagen disminuye.

Al asignar menos bits a los niveles de gris más probables y más bits a los menos probables, se puede conseguir la compresión de los datos.

Este proceso es conocido como codificación de longitud variable.

IV.5.2 Redundancia entre pixeles (información).⁵³

Supongamos que tenemos dos imágenes (Imagen1 e Imagen2), cuyos histogramas son prácticamente idénticos (Imagen A e Imagen B). (La diferencia puede encontrarse en la diferente colocación de las mismas figuras en las dos imágenes).

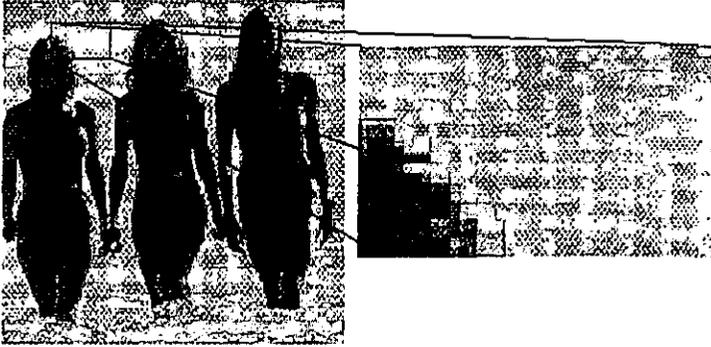
Imagen 1.

Figura 4.5 (a)

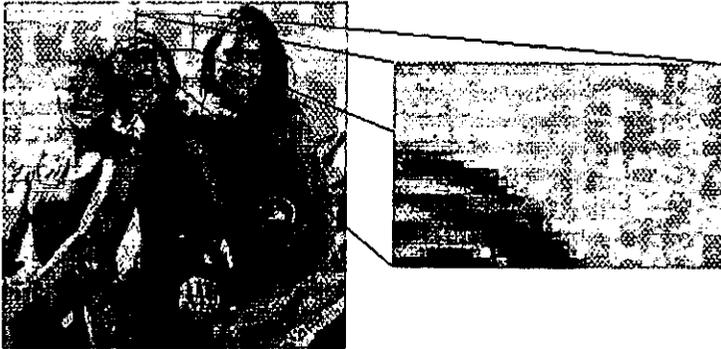
Imagen 2.

Figura 4.5 (b)

⁵³ Información de redundancia:
http://www-etsi2.ugr.es/depar/ccia/public_html/robotica/indice_apuntes

Las probabilidades de los niveles de gris que encontramos en las imágenes son distintas por lo que podríamos usar códigos de longitud variable para reducir la redundancia de codificar en binario natural sus píxeles.

Sin embargo, los códigos empleados para representar los niveles de gris de cada imagen no tienen nada que ver con la correlación entre los píxeles, ya que estas correlaciones provienen de las relaciones estructurales o geométricas entre los objetos de la imagen.

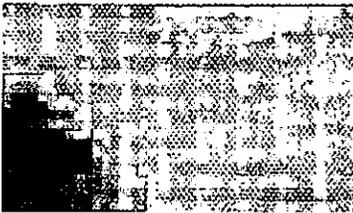


Imagen A.



Imagen B.

Histograma en escala de grises.

Figura 4.6 (a y b)

El valor de la probabilidad en la escala de gris debe ser estrictamente menor que el número de píxeles de una línea. Es posible predecir razonablemente el valor de un determinado píxel a partir de sus valores vecinos. La mayor parte de la contribución visual de un único píxel a una imagen es redundante, ya que, podría haberse inferido de acuerdo con los valores de sus vecinos. Estas dependencias entre píxeles se conocen como redundancia espacial, redundancia geométrica y redundancia interna. Que se engloban como redundancia entre píxeles.

Para reducir las redundancias entre píxeles de una imagen, la distribución bidimensional de píxeles normalmente empleada para la percepción e interpretación humana debe ser transformada a un formato más eficaz.

Se pueden utilizar las diferencias entre píxeles adyacentes para representar una imagen. Las transformaciones de este tipo, que eliminan las redundancias entre píxeles, se denominan genéricamente correspondencias.

En conclusión, la imagen se puede representar de forma más eficaz mediante el valor y la longitud de las secuencias de píxeles con un nivel de gris constante que mediante una distribución bidimensional de píxeles binarios.

Para ello sólo se necesita una secuencia de pares, donde el primer elemento indica el nivel de gris (constante) y el segundo la longitud del recorrido. De esta manera la representación para una imagen sólo son necesarios 88 bits para representar los 1024 bits de datos binarios.

IV.5.3 Redundancia psicovisual⁵⁴

La iluminación de una región, tal como es percibida por el ojo, depende de otros factores además de la luz reflejada por esta región. Este fenómeno se produce porque el ojo humano no responde con la misma sensibilidad a toda la información visual. Cierta información simplemente tiene menor importancia relativa que otra en el proceso visual normal. Se dice que esta información es psicovisualmente redundante, y se puede eliminar sin que se altere significativamente la calidad de la percepción de la imagen.

Que haya redundancia psicovisual no es de extrañar puesto que la percepción del ojo humano sobre la imagen no se realiza píxel a píxel, observando la luminosidad de cada uno. La percepción se centra en bordes o regiones con texturas y colores que la mente combina mentalmente con grupos reconocibles y con el conocimiento previo con el fin de terminar el proceso de interpretación de la imagen.

Al contrario que la redundancia anteriormente vista, está directamente relacionada a la información visual real y no con los datos que representan la información. Podemos sacar en claro de todo esto, que eliminar datos psicovisualmente relativos conlleva una pérdida de información, que no es recuperable (la operación es irreversible). Estamos ante una compresión con pérdida de datos.

El método utilizado para producir estos resultados se conoce como cuantificación con escala de grises mejorada. Este método reconoce la sensibilidad del ojo humano a los bordes y los suaviza, para lo que añade a cada píxel un número pseudoaleatorio que obtiene a partir de los bits de menor peso de los píxeles vecinos.

Primero se suma el valor del nivel de gris actual (de 8 bits) y los cuatro bits menos significativos de la suma anterior (esta suma inicialmente vale 0). Sin embargo, si los cuatro bits más significativos del valor actual son 1111, en su lugar se suma 0000. Se

⁵⁴ Información de redundancia psicovisual:
http://www-etsi2.ugr.es/depar/ccia/public_html/robotica/indice_apuntes

toman los cuatro bits más significativos de la suma resultante como valor del pixel codificado.

La utilización de estos métodos produce disminución de la resolución espacial, falsos contornos, emborronamientos, atenuaciones, etc. De ahí que haya que utilizar técnicas como las de entrelazado 2:1 entre líneas para reducir la frecuencia de barrido de video (sin pérdida de calidad en la imagen).



Imagen I



Imagen II

Figura 4.7

En las imágenes anteriores se puede apreciar fácilmente la disminución en la resolución de la Imagen II con respecto a la Imagen I esto se debe a que a la primera le fue eliminada la redundancia psicovisual por lo que se generó un emborronamiento o blur que no permite apreciarla imagen de una manera nítida.

Este método es usado por el formato gráfico JPEG. Elimina zonas que se sabe que el ojo humano no va ser capaz de reconocer, y además se pueden utilizar como efectos especiales para la televisión. Este formato debió de acoplarse a los estándares internacionales que rigen el tema para ello se incorporaron como archivos de video.

IV.6 Archivos de Video.

IV.6.1 Estándares de Producción de Video.⁵⁵

Los tres estándares de producción de vídeo en el mundo son: NTSC, PAL, SECAM. Estos tres estándares y formatos son intercambiables. Cada sistema se basa en un estándar distinto que determina la forma en que se codifica la información para producir una señal que finalmente genera una imagen de televisión. Actualmente se esta trabajando en la elaboración de un nuevo estándar de vídeo y televisión, el HDTV.

NTSC.

EE.UU, Japón y muchos otros países utilizan un sistema de producción y despliegue de vídeo que se basa en las especificaciones establecidas en 1952 por el National Televisión Standars Committe (NTSC). Estas siglas definen un estándar para registrar la información en una señal electrónica que finalmente genera una imagen de televisión. Como se especificó el estándar NTSC, cada cuadro de vídeo se compone de 525 líneas de barrido sobre la cara interna del tubo de vacío cubierto de fósforo cada treintavo de segundo por un haz de electrones de alta velocidad. La operación ocurre tan rápido que sus ojos perciben la imagen como fija. En realidad, el haz de electrones da dos pasadas para dibujar cada cuadro de vídeo; en la primera se encarga de las líneas impares y , en la segunda de las pares. Cada pasada (que ocurren a una velocidad de 60 ciclos por segundo o 60 Hz) dibuja un campo. El proceso de creación de un cuadro de vídeo se llama Entrelazado (técnica que ayuda a evitar parpadeos en las pantallas de televisión).

PAL.

El sistema de línea de fase alterna (Phase Alternate Line, PAL) se utiliza en el Reino Unido, casi toda Europa, Australia y Sudáfrica. PAL es un método que agrega color a una señal de televisión en blanco y negro que dibuja 625 líneas a una velocidad de 25 cuadros por segundo. Como NTSC, las líneas pares y impares se entrelazan; cada campo toma un cincuentavo de segundo para dibujarse (50 Hz).

⁵⁵ Información de estándares de video:
<http://www.udec.cl/investigacion/depart/public/video>

SECAM.

El sistema secuencial de color y memoria (Sequential Color and Memory, SECAM) se utiliza en Francia, Rusia y algunos otros países. Aunque SECAM es un sistema de 625 líneas a 50 Hz, difiere mucho de los sistemas de color NTSC y PAL en su tecnología de base y su método de producción.

HDTV.

La televisión de alta definición (High Definition Televisión) está programada para ser el próximo paso en la tecnología de televisión. Hasta el momento proporciona 1200 líneas de resolución y una relación de aspecto de 16:9. Esta relación de aspecto permite que se vean películas en Cinemascope y Panavision. Hay tres estándares de HDTV en competencia, dos con formato analógico, uno japonés y el otro europeo, y el desarrollado en EE.UU con formato digital.

IV.6.2 Características de los Archivos de Video.

Antes de analizar los distintos tipos de formatos de vídeo, se tomara en cuenta algunas cifras que nos ayudaran a comprender mejor sus ventajas y características. Estas cifras pueden parecer monstruosas o gigantescas, pero así podremos visualizar las grandes ventajas que nos proporcionan los distintos tipos de formatos de vídeo. Supongamos que se quiere almacenar o digitalizar una imagen o vídeo con las siguientes características:

- Resolución VGA a 640x480 (es decir, toda la pantalla), a todo color.
- 24 bits para el color en cada pixel.
- 30 cuadros por segundo (usando la norma NTSC).

Un solo segundo de esta digitalización pesará 27,6 Mb. Esto significa que un minuto de vídeo digitalizado pesará 1,6 GB (o más de 2 CD). Considerando que se ha digitalizado solo las imágenes y no se ha acompañado de audio.

Para poder disminuir el peso de este archivo podríamos pensar en disminuir la calidad del vídeo. Una de las acciones a tomar podría ser disminuir la cantidad de cuadros por segundo. Así, siguiendo la norma europea de televisión, la PAL, se reducirá a 25 cuadros por segundo. Esto solo logra disminuir a 1,4 GB de información un minuto de vídeo.

Si se considera el costo de US\$ 1 por cada megabyte de información, un minuto de vídeo digitalizado costaría aproximadamente. US\$1400.

Además, suponiendo la disponibilidad de recursos para digitalizar vídeo, existen otras restricciones como la capacidad de los buses de datos de poder mover tal cantidad de bytes por segundo. Un CD-ROM 12x es capaz de mover solo 1,5 MB por segundo, y no los casi 30 MB que pesaría una digitalización con las características antes nombradas.

Obviamente, hay que considerar otras soluciones, o combinaciones de estas, para disminuir el tamaño de una digitalización de vídeo. Entre las más comunes se tiene:

Reducir físicamente el tamaño de la imagen en la pantalla, ya sea a 1/16 (120x160) o a 1/4 (320x240) de ella. Esta solución genera una reducción exponencial de la cantidad de información resultante.

Desplegar menos cuadros por segundo que las normas consideradas, sin embargo, a menos de 24 cuadros por segundo el movimiento se torna discontinuo siendo este efecto notorio al ojo humano.

Reducir la precisión de los colores. Si bien una imagen con color real requiere 24 bits por pixel, desplegar una imagen a 256 colores requiere solo 8 bits por pixel, lo que reduce en 2/3 esta información.

Lamentablemente todas estas soluciones producen un efecto común: todas ellas causan una degradación de la calidad de la imagen original, ya que desechan información, reduciendo el peso o tamaño de los archivos con vídeo digitalizado. Aún así, una imagen a 1/4 de pantalla y a 256 colores, reduce en 1/12 el tamaño del archivo. Estamos hablando, entonces, de 2 MB de información por segundo de vídeo, lo que aún es muy grande de transferir al procesador hasta para un CD-ROM de doceava velocidad.

La solución, entonces, para la digitalización de vídeo, la cual llegará a ser una gran herramienta en diversos ámbitos, como la educación, comunicación o entretenimiento, es la compresión de la información, utilizándose los diversos formatos de vídeo, algunos de los cuales se revisarán a continuación.

IV.7 Sistema de representación de imágenes.

Si bien es importante capturar fielmente la imagen original y procesarla correctamente de acuerdo con los criterios que se hallan elegido, también es de gran importancia presentar los resultados obtenidos de esos procesos de manera comprensible para el usuario.

Para ello el mejor sistema de visualización comprende el *sistema gráfico* basado en la pantalla gráfica. Un sistema gráfico viene definido por dos parámetros básicos: resolución y número de bits por pixel.

La resolución va a indicar cuantos puntos puede tener una imagen al representarse. Obviamente, cuanto mayor sea el número de puntos, mayor será el detalle de la imagen.

El número de bits por pixel, este valor define el número de colores o en caso de un sistema monocromático, el número de niveles de gris que puede tomar un pixel en concreto.

En el caso concreto de las imágenes, se ha comprobado que si bien la resolución es un factor importante, lo es mucho más el número de niveles de gris o el número de colores por punto.

IV.7.1 Muestreo y cuantificación.

El método habitual para digitalizar una imagen es barrer la misma fila por fila, y muestrear cada una de las filas. El número de filas que se barran (m) y el número de pixeles que se muestren por fila (n) determinaran la resolución de la imagen digital ($m \times n$), que viene a ser el grado de discernibilidad de detalles en la imagen.

El siguiente punto a tomar en cuenta es la cuantificación de cada una de las muestras. Si el muestreo representaba una digitalización en espacio, la cuantificación viene a ser una digitalización en amplitud, entendiéndose por amplitud el nivel de intensidad luminosa de cada punto. Al cuantificar cada pixel, por el simple hecho de limitar el número de bits, se reduce el número de valores existentes desde infinito hasta 2^p posibles valores.

La cantidad de la cuantificación viene dada, fundamentalmente, por el número de bits que se asignan a cada una de las muestras. Dado que las computadoras suelen trabajar no con bits aislados sino con agrupaciones de ellos (nibbles, bytes, palabras, etc.), suele ser práctico utilizar un valor menor o igual a ellos, evitando que un pixel desborde de un byte a otro.

Una vez definido el número de bits por pixel, la imagen final queda representada por un mapa tridimensional, de orden $m \times n \times p$, que a cada punto del espacio le hace corresponder un nivel de intensidad luminosa definido como p . Para simplificar p , se puede decir que es el número de bits por pixel.⁵⁶

IV.7.2 Relaciones entre pixeles.

Estas relaciones son necesarias para comprender los casos de la mayoría de las operaciones básicas de las rutinas de procesamiento. Vecindad, contorno, conectividad y distancia son términos de algunas de ellas.

⁵⁶ Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia 1994, Páginas 27,28.

IV.7.2.1 Vecindad y contorno.

Se dice que todo pixel p , de coordenadas (x,y) , tiene cuatro pixeles que establecen con él una relación de vecindad, horizontal o vertical, que son:

Horizontal: $(x - 1, y)$ y $(x + 1, y)$

Vertical: $(x, y - 1)$ y $(x, y + 1)$

Estos cuatro pixeles definen el *contorno de orden cuatro*, denotado por $C4(p)$. (Figura 4.8)

Además de las diagonales:

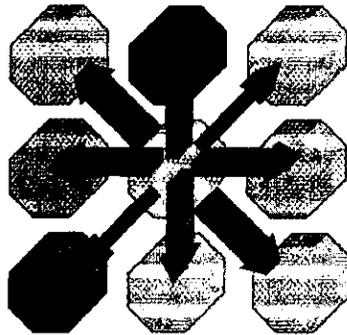


Figura 4.8

Diagonal: $(x - 1, y - 1)$, $(x + 1, y - 1)$
 $(x - 1, y + 1)$ y $(x + 1, y + 1)$

definen el contorno de orden cuatro denotado por $CD(p)$.

Ambos contornos de orden cuatro, juntos, constituyen el *contorno de orden ocho* o $C8(p)$.

IV.7.2.2 Conectividad.

La conectividad es un aspecto de gran importancia a la hora de delimitar o reconocer las fronteras o límites entre los elementos integrados en una misma imagen. Para delimitar si dos pixeles están conectados entre sí, suele recurrirse a

algún tipo de criterio que calibre la similitud entre ambos. Tres de los más frecuentes son:

.Los píxeles p y q están conectados si existe proximidad física entre ambos, esto es, la distancia que los separa es menor que un umbral dado.

.Los píxeles p y q están conectados si existe entre ellos una relación de vecindad, del tipo:

- Conectividad de orden 4.
- Conectividad de orden 8.
- M - conectividad. Esto es, p y q están m - conectados si:
- q pertenece a el contorno de orden 4.
- q pertenece a el contorno de orden 4 de diagonales $CD(p)$ y los números de bits por pixel de orden 4 intersectan a los números de bits por pixel de orden 4 de q .

. Los píxeles p y q están conectados si sus niveles de gris satisfacen un determinado criterio de similitud por ejemplo si son iguales).

IV.8 Técnicas para la representación de la imagen digital. (Formatos gráficos)

IV.8.1 ¿Qué son los formatos gráficos?

Básicamente, los formatos gráficos son archivos en los cuales se guarda información que conforma una imagen. Cada formato es independiente. Las posibilidades que ofrece cada formato con respecto a la gama de colores, a la compatibilidad, a la rapidez de carga, etc., merece ser explicada para determinar cuál de ellos es el más adecuado para la tarea que estamos realizando.

Con respecto a la estructura, la mayoría posee un header que indica al programa que lo solicite las características de la imagen que almacenan; por ejemplo su color, tipo, resolución, etc. Cada formato tiene una organización propia de su estructura.

IV.8.2 Estructura General de un Formato gráfico.

Esta estructura se toma de formatos que no contienen imágenes True Color (16 millones de colores), sino paletadas, es decir, que necesitamos una lista de los colores necesarios para representar la imagen.

IV.8.2.1 Cabecera

En los primeros 128 bytes se encuentra la información que define el tipo y las características del fichero y es la zona que primero se debe consultar antes de acceder a los datos propiamente dichos.

IV.8.2.2 Datos

En esta zona se encuentran los datos codificados de la imagen. La codificación utiliza una simple técnica llamada RLE (Run Length Encoded).

La decodificación/codificación es a nivel de línea y no se puede codificar datos de distintas líneas o rasters. La potencia que ofrece este tipo de compresión es muy limitada, sin embargo, en imágenes sencillas se alcanza un buen índice de compresión.

IV.8.2.3 Paleta de Colores

La información referente a la paleta de colores utilizada, puede estar incluida en la cabecera (si son 16 o menos los colores de la imagen), al final del fichero (imágenes de 256 colores), o simplemente no aparecer, si son imágenes True Color.

- * Paleta EGA/VGA 16 Colores
- * Paleta VGA 256 Colores
- * 24 bits de Color (True Color)

IV.8.2.4 Paleta EGA/VGA 16 Colores

Los datos de los colores leídos en la cabecera deben ser tratados previamente antes de pasárselos al hardware de la PC. Por ejemplo, si se trabaja en el modo de 16 colores EGA, sólo están disponibles 4 niveles de RGB para cada color (dependiendo del formato), y como en el fichero vienen en valores de 0 a 255, se divide entre 64 para obtener el valor adecuado.

IV.8.2.5 Paleta VGA 256 Colores

La paleta de 256 colores está formateada y tratada de la misma manera que la de 16 colores. Para mantener la compatibilidad con versiones anteriores (también dependiendo del formato), se coloca al final del fichero precedida por un número decimal (generalmente el 12). Como el hardware de la VGA trabaja con componentes de color

entre 0 y 63 en vez de entre 0 y 255, se debe dividir los valores leídos entre 4 para adaptarlos convenientemente.

Para acceder a una paleta de 256 colores, primero debe comprobarse el número de versión en la cabecera y después se leen los últimos 768 bytes del fichero.

IV.8.2.6 24 bits de Color (True Color)

Su formato es el siguiente: imágenes de 8 bits y 3 planos, sin información de paleta de colores. Los planos de color son ordenados del siguiente modo: líneas de rojo, verde y azul (RGB).

IV.8.2.7 Identificador

Sirve para identificar a un fichero como un fichero válido de imagen. Consiste en una serie determinada por cada formato de caracteres. Los tres últimos caracteres pueden utilizarse como número de versión.

IV.8.2.8 Descriptor de pantalla

En este apartado se describen los parámetros generales para todas las imágenes contenidas en fichero. Define las dimensiones globales requeridas para la pantalla lógica y para el espacio de las imágenes, la existencia de mapa de colores, los colores del fondo y la información sobre la profundidad de los colores. Estos datos se almacenan en series de 6 bits.

La altura y la anchura de la pantalla lógica pueden ser mayores que la pantalla física. El valor del pixel también define el número máximo de colores dentro de una imagen. Su rango de valores es de 0 a 7 que representa de 1 a 8 bits, con lo que se tiene las posibilidades blanco y negro, y 256 colores. El bit 3 de la palabra 5 está reservado para un uso futuro y debe ser 0.

IV.8.2.9 Mapa Global de Color⁵⁷

El Mapa global de color es opcional pero muy recomendado para imágenes en donde sea necesaria una adaptación de los colores. Un campo determinado como "M" del byte 5 del descriptor de imágenes indica la existencia de este mapa de colores.

Por otro lado, aunque es posible asignar un mapa de colores a cada imagen contenida en el fichero, las restricciones actuales en el hardware suelen obligar a que se utilice habitualmente este mapa global. Si el mapa global de color está presente, su definición se

⁵⁷ <http://www.udec.cl.investigación/apuntes/video>

encontrará después del descriptor de pantalla. El número de entradas en el mapa de color es igual a 2 elevado al número de bits por pixel (8 bit/pixel = 256 entradas o colores) y cada entrada consiste en los tres bytes que representan los valores de rojo, verde o azul (RGB)

Los componentes de color representan una intensidad entre 0 y 255, por lo que habrá que adaptarlos al hardware específico de la PC (en VGA, de 0 a 63, se dividirá entre 4).

IV.8.2.10 Descriptor de imagen

El descriptor de imagen define la posición actual y el tamaño de la siguiente imagen dentro del espacio definido en el descriptor de pantalla.

También define si dicha imagen posee un mapa de color propio y la forma en que se mostrarán los píxeles (entrelazados o no entrelazados). Cada descriptor de imagen es señalado por un carácter separador de imágenes (0, x, 2, C ó ",") , para sincronizar y comprobar su existencia.

IV.8.2.11 Mapa de color local

El mapa de color local es opcional y se incluye para futuros usos. Si el bit "M" del byte 10 del descriptor de imagen está activado, se aplicará el mapa de color que siga al descriptor de imagen sólo en la siguiente imagen. Al final de dicha imagen, el mapa de color volverá a ser el definido en el descriptor de pantalla. Nótese, que el campo pixel del byte 10 del descriptor de imagen se utiliza, sólo si se indica un mapa de color local.

IV.8.2.12 Datos raster

Aquí se definen los datos de la imagen propiamente dichos. El orden es de izquierda a derecha para cada fila. Las filas suelen ir sencillamente de arriba hacia abajo, pero si el bit "T" (entrelazado) del byte 10 del descriptor de imagen está activado, el orden de las filas seguirá un esquema predeterminado, en cuatro pasadas consecutivas:

Esta es la parte que se comprime sobremanera utilizando una variante del método LZW, creando la principal característica de los ficheros, el factor de compresión.

IV.8.2.13 Terminador

Para indicar la terminación de un fichero GIF, se intercala un carácter (0 x 3B o ",") cuando el decompresor se encuentra este carácter después que una imagen haya sido procesada, suele esperarse una acción del usuario, como pulsar una tecla o similar.

IV.8.2.14 Bloques de extensiones

Con estos bloques se provee de un mantenimiento y actualización de la definición, dando la posibilidad de ampliaciones futuras, sin romper la compatibilidad. Los bloques son almacenados de manera similar a los bloques de datos raster, pero no están comprimidos.

Un bloque de extensión puede preceder a cualquier descriptor de imagen o estar antes del terminador.

IV.9 Técnicas de Formatos Gráficos.

Las imágenes pueden ser representadas de diferentes formas. En la actualidad contamos con tres técnicas: bitmaps, vectores, geometría fractal.

IV.9.1 BMP (Bitmapped File Header Format)

Junto con el surgimiento de Windows 3.1 se desarrolla un nuevo formato gráfico bitmap que constituye el standard adoptado por este entorno operativo. Este es el BMP, en el cual están almacenadas las imágenes que constituyen los llamados wallpapers por ejemplo. Este formato guarda las imágenes descomprimidas, lo que significa mayor velocidad de carga y mayor espacio requerido. Con respecto a la resolución, cualquiera es aceptable. Las imágenes pueden ser de 1, 4, 8 y 24 bits. La estructura de los BMPs es sencilla: se trata de un header que contiene varias características de la imagen. Este header está compuesto por información acerca del tamaño, el número de colores, y una paleta de colores (si es necesario) de la imagen. A continuación del header se encuentra la información que constituye la imagen en sí.

Tiene una curiosa forma de almacenarla: comienza desde la última línea inferior. Es por eso que los programas encargados de exhibir los BMPs en pantalla trazan la imagen de abajo hacia arriba.

Contienen imágenes basadas en píxeles (por ejemplo una imagen cuya resolución es de 640 x 480 píxeles, contiene 640 píxeles horizontales y 480 píxeles verticales). Las imágenes generadas por escáner son de tipo bitmap. Cuanto mayor sea la gama de colores, más realismo se consigue con este tipo de formato. Las imágenes bitmap poseen un tamaño natural en el cual se imprimirán perfectamente, pero, a diferencia de las vectoriales, no ofrecen grandes posibilidades con respecto a la variación del tamaño. Al aumentar el tamaño bruscamente, es fácil notar una gran disminución de la calidad.

Otro punto en contra de los formatos bitmap es la cantidad de memoria y espacio que ocupan. Debido a esto, la mayoría de ellos utilizan diversos métodos de compresión de la información que constituye la imagen.

Esta técnica consiste en representar las imágenes a través de múltiples puntos denominados píxeles que al visualizarse todos juntos dan la sensación de mostrar una imagen continua. Un píxel es una información relativa a la posición y valor de luminosidad de un punto de la pantalla de la computadora. Una imagen digital consiste, pues, en un conjunto discreto de posiciones y valores de luminosidad.

Un concepto importante a tener en cuenta es el de resolución. Tenemos dos formas de medirla: en números absolutos (tamaño), viene a ser el número de puntos de largo por el número de puntos de alto que definen una imagen; en número de píxeles por unidad de medida, pulgadas o milímetros. Lo normal es utilizar pulgadas (dpi o ppp = puntos por pulgada). La segunda magnitud está relacionada con las posibilidades del dispositivo de salida (monitor o impresora), hablándose de baja resolución (200 dpi), media resolución (240-600 dpi) y alta resolución (por encima de 600 dpi). A mayor resolución mayor tamaño del archivos de imagen pero también mayor calidad y mayores posibilidades de ampliación de la imagen con calidades aceptables.

IV.9.1.1 Tipos de bitmaps:

1. Imágenes bitonales (blanco y negro): se representan mediante una matriz de bits, los cuales sólo se pueden encontrar en dos estados activados o desactivados. Su ventaja principal es el poco espacio de almacenamiento que requieren, siendo su mayor desventaja la imposibilidad de poder representar imágenes en color y la baja resolución que presentan. Estas imágenes pueden mejorarse a través de la técnica del halftone. Esta técnica consiste en crear una ilusión de variedades de gris mediante la utilización de patrones binarios de blanco y negro.

2. Paleta de colores o niveles de grises: se utiliza un byte para representar cada píxel, con lo que es posible representar hasta 256 distintas tonalidades de grises o colores. En el caso de imágenes en color, cada byte representa un número que a modo de índice es utilizado para buscar el color real de ese píxel en una tabla, en la que están representados todos los colores disponibles para esa imagen. Esta tabla se denomina paleta de colores y cada imagen lleva asociada la suya propia.

3. Color real: son las de mayor calidad y las que más espacio ocupan. Cada píxel contiene información completa sobre su color, que se expresa como una combinación de diferentes intensidades de rojo, verde y azul. El número de colores posible depende del número de bits que se usen para representar cada píxel. Como se representan las imágenes con los tres colores básicos o primarios (verde, rojo y azul), el color de cada píxel está compuesto por la mezcla de estos tres, por lo que el número de bits de almacenamiento para un píxel en color real será generalmente múltiplo de tres. Así un color de 16 bits es realmente un

color de 15 bits, teniendo uno sobrante al que se le da otro uso. Se suelen usar tres tipos de archivos de imagen de color real:

Color real de 16 bits: permiten 32 tonos distintos de cada color primario que combinados proporcionan 32.768 posibles colores.

Color real de 24 bits: se suelen almacenar por planos de color. Cada plano contiene la imagen en un color: una de color rojo, otra de color verde y otra de color azul. La imagen final se obtiene como combinación de esos tres planos. Cada plano se representa mediante píxeles de 8 bits, lo que da 256 colores por plano, con un total de algo más de 16 millones de colores.

Color real de 32 bits: es posible representar cada color primario mediante 10 bits lo que proporciona más de 1000 millones de colores.

Formatos de archivos gráficos: una vez generadas las imágenes se almacenan en archivos con formatos especialmente diseñados para el almacenamiento de la información que permite reconstruir la imagen digital. Los más usuales hoy día son: BMP, PCX, TIFF, GIF, TARGA, EPS, JPEG, PCPAINT, PICT Y PICT2.

Análisis y reconocimiento de imágenes: hay programas especialmente diseñados para el análisis de las imágenes pudiéndose identificar o bien letras (los sistemas OCR o ICR) o bien patrones, los sistemas de reconocimiento de patrones, de análisis de mapas, etc. Son técnicas aún en desarrollo y vinculadas a los avances en el campo de la inteligencia artificial.

Una imagen puede ser representada y descrita desde diferentes perspectivas dependiendo como se le catalogue, tal es el caso de la imagen vectorial.

IV.9.2 Imagen vectorial

Las imágenes se describen mediante objetos. Cada parte de la imagen es un objeto que puede ser tratado de forma independiente, pudiendo estar a su vez constituido por otros objetos que también son independientes. Es decir, cuando se define cada uno de los objetos que componen la imagen y se integra en ella no pierde sus características de objeto, pudiendo ser editado y manipulado en cualquier momento, sin que por ello los demás componentes del dibujo se vean modificados. Los objetos están definidos mediante vectores. Los archivos en los que se almacenan este tipo de imágenes están formados por las coordenadas de los vectores que las describen. Estos archivos en lugar de almacenar píxeles almacenan secuencias de comandos.

Un vector es una línea que se describe con la localización de los puntos de sus extremos. Un rectángulo podría definirse de la siguiente manera: RECT 0,0,200,200. Empleando

coordenadas cartesianas, el programa dibujará ese rectángulo comenzando en el extremo superior izquierdo de la pantalla, alcanzando 200 píxeles a la derecha horizontalmente y 200 píxeles verticalmente hacia abajo.

El problema de estas imágenes es la lentitud a la hora de redibujar una imagen ya definida, acentuada mientras más compleja sea ésta, ya que se tienen que ir haciendo cálculos para cada objeto, y posteriormente dibujarla. Para resolver en parte este problema algunos programas de tratamiento de este tipo de imágenes mezclan esta técnica con la de los bitmaps, tratando a los píxeles como objetos.

Hay programas que permiten la conversión de las imágenes vectoriales en bitmaps y viceversa. El primer caso es posible a través de cualquier programa de captura de pantallas. El segundo caso es más complicado, sin embargo algunos programas son capaces de calcular los bordes de una imagen de mapa de bits, o las formas de las áreas coloreadas en la imagen, e inferir un polígono que la describa este procedimiento es conocido como autotracking.

Son más simples que los bitmap. Consisten en una serie de dibujos lineales basados en una lista de objetos gráficos, por ejemplo: líneas, curvas, triángulos, círculos, rectángulos, etc. Estos objetos, y muchos más, ubicados estratégicamente en la pantalla, forman dibujos lineales complejos. En este tipo de formatos las áreas vacías entre las líneas pueden ser llenadas con colores o con "rellenos", pequeños diseños que se repiten una y otra vez sin dejar espacios vacíos hasta llenar el área en cuestión. El tamaño de las imágenes almacenadas en este tipo de formatos puede ser modificado sin notar pérdida alguna de calidad. Gracias a esta característica son muy útiles a la hora de imprimir imágenes.

IV.9.3 Tecnología fractal

Esta tecnología está basada en la representación y diseño de imágenes a partir de ecuaciones matemáticas complejas. La palabra fractal tiene sus orígenes en el vocablo latino fractus, que significa irregular o roto. Este término fue introducido en la década de los 70 por el matemático Benoit B. Mandelbrot. Un fractal es un objeto geométrico que posee autosemejanzas, es decir, las partes más pequeñas del dibujo son altamente similares a las partes más grandes. Un ejemplo típico es el de una costa. En este ejemplo cualquier nivel de detalle que tomemos observaremos siempre similitudes con el nivel de detalle anterior.

La geometría fractal se aplica a múltiples campos de la ciencia y de la técnica (matemáticas, química, física, biología, medicina, astronomía, economía, etc.), últimamente se viene aplicando al mundo de la infografía donde ya están saliendo al mercado aplicaciones comerciales que utilizan esta técnica para el tratamiento y

compresión de imágenes. Un campo donde está teniendo gran éxito es el de la representación de formas de la naturaleza, especialmente en la creación de paisajes artificiales de planetas imaginarios en películas de ciencia ficción. Por ejemplo, para dibujar una montaña se suele seguir el siguiente proceso iterativo: se toma una figura geométrica (un triángulo), se calcula el punto medio de cada arista y los puntos resultantes se desplazan un factor aleatorio para unirse después entre sí, obteniendo nuevos triángulos irregulares. Este proceso repetido un número de veces necesario da como resultado una imagen fotorrealística de una montaña.

Es dentro del campo de la compresión de imágenes digitales donde se están alcanzando mayores logros, llegándose a alcanzar ratios de compresión de hasta 10.000:1.

IV.10 Modelos de Color

El propósito de un modelo de color es facilitar la especificación de colores en algún formato estándar dentro de un formato gráfico. En esencia, un modelo de color es una especificación de un modelo de coordenadas 3-D y un subespacio dentro de ese sistema donde cada color se representa por un punto único.

La mayoría de los modelos de color que se usan hoy se orientan hacia hardware como monitores o impresoras o aplicaciones de manipulación de color. El modelo orientado a hardware más común es el RGB (rojo - verde - azul), el modelo CMY (cian-magenta-amarillo) para impresoras en color y el YIQ que es el estándar para la difusión de TV, en este caso la Y corresponde a luminancia y la I y Q son dos componentes cromáticas. Para la manipulación del color se usan normalmente los modelos HSI (matiz, saturación e intensidad) y HSV (matiz, saturación y valor).

Los modelos de color más usados en el procesamiento de imágenes son RGB (Red, Green, Blue), YIQ y HSI. En la sección siguiente se detallará el modelo RGB.

IV.10.1 El modelo de color RGB⁵⁸

En el modelo RGB (Red, Green, Blue) cada color aparece en sus componentes primarias espectrales rojo, verde y azul. El modelo está basado en un sistema de coordenadas cartesiano. Por conveniencia todos los colores han sido normalizados de forma que todos los valores de rojo, verde y azul están dentro de un rango.

Las imágenes en el modelo de color RGB están formadas por tres planos de imágenes independientes, cada una de los colores primarios. Cuando son introducidas en un

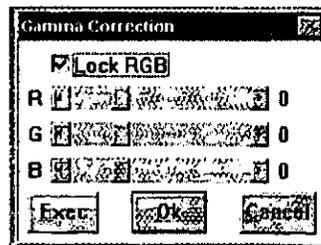
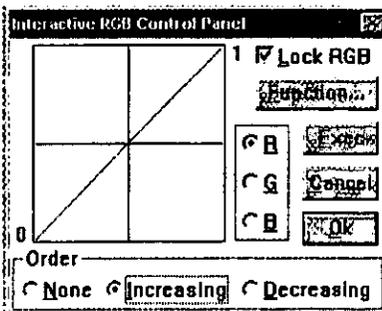
⁵⁸ <http://www.udec.cl/investigación/apuntes/video>

monitor RGB, las tres imágenes se combinan en la pantalla de fósforo para producir una imagen de color compuesta. Por tanto, el uso del modelo RGB para el procesamiento de imágenes tiene sentido cuando las imágenes vienen expresadas en términos de los tres planos de colores. Alternativamente, la mayoría de las cámaras en color que se usan para adquirir imágenes digitales utilizan el formato RGB, lo que hace, si cabe, más interesante este formato.

Uno de los mejores ejemplos de la utilidad del modelo RGB es el procesamiento de imágenes aéreas y de satélites multispectrales. Estas imágenes se obtienen para diferentes rangos espectrales. Por ejemplo, las imágenes LANDSAT se obtienen como mínimo en cuatro ventanas espectrales distintas de la misma escena. Dos ventanas están en el espectro visible y corresponden aproximadamente a verde y rojo; las otras dos están en el infrarrojo. Así pues cada plano de la imagen tiene sentido físico y la combinación de color utilizando el modelo RGB para procesamiento y visualización tiene sentido cuando se ve en una pantalla en color. La figura muestra una imagen en color junto con sus tres bandas de color rojo, verde y azul. Obsérvese que el mayor nivel de color en cada banda corresponde con el color predominante en cada objeto.



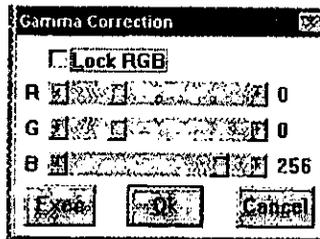
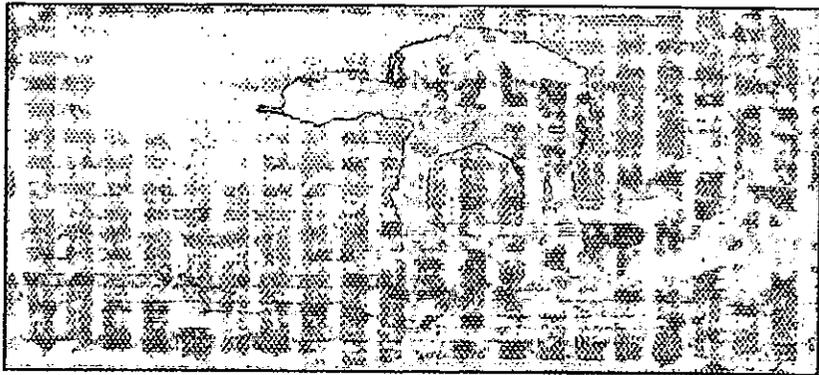
Figura 4.9



En la figura anterior (figura 4.9) se tiene la imagen original en color. Y la representación de sus componentes de color en formato RGB.

Sin embargo, no siempre es éste el mejor modelo de color para el procesamiento de imágenes. Consideremos, por ejemplo que queremos realzar una imagen en color del ejemplo anterior, parte de la cual está oculta por sombra. Este problema puede ser abordado como la igualación del histograma.

Figura 4.10



En esta imagen (figura 4.10) se tiene la misma figura pero con modificaciones en el plano de azul, los cambios son el resultado de la combinación de los tres componentes que han quedado desequilibrados.

Al tener tres bandas y puesto que el histograma trabaja sólo con las intensidades, cuando igualemos el histograma en cada banda, en cada una de ellas tendremos mejora, sin embargo, el resultado global, (cuando se combinen las tres bandas), será impredecible.

Algunos otros modelos de color, por ejemplo el que discutiremos brevemente a continuación son más útiles para este tipo de problemas.

RGB es la elección más común en lo que se refiere a gráficos por computadora ya que los CRT usan fósforo rojo, verde y azul para crear el color deseado. Sin embargo, RGB no es la forma más eficiente de describir las tres componentes de color. El igual ancho de banda da lugar a la misma profundidad del pixel y resolución del display para cada componente. Sin embargo, la sensibilidad de la componente de color del ojo humano es menor que la de la luminancia.

Por estas razones, muchos métodos de codificación de imágenes y sistemas de transmisión utilizan la luminancia y dos diferencias de colores. Estos son los formatos YUV y YIQ.

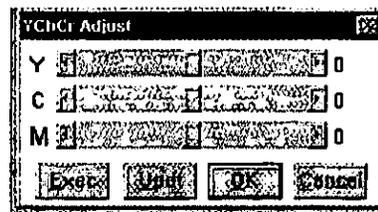
IV.10.2 El modelo de color CMY

Como se ha mecho mención anteriormente el cían, magenta y amarillo son los colores secundarios. Para convertir de colores primarios a secundarios hemos de realizar la siguiente transformación.

La figura muestra (figura 4.11) las representaciones en las correspondientes bandas de la imagen. Para obtener estas imágenes se ha normalizado cada canal a uno independientemente, dividiéndolo por 255, a continuación se le ha restado a uno dicha cantidad y por último se ha multiplicado el resultado por 255.⁵⁹



Figura 4.11



⁵⁹ http://www-etsi2.urg.es/depar/ccia/public_html/robotica/indice_apuntes.html

IV.11 Captura digital

Son varios los métodos que se pueden emplear para capturar y digitalizar imágenes. El empleo de uno u otro depende del objetivo de la captura. Para digitalizar una imagen o un dibujo representado sobre un papel deberemos utilizar un escáner, para digitalizar un objeto tridimensional una cámara digital o una cámara de vídeo estático. Si queremos capturar una imagen que estamos viendo en un momento dado en la pantalla de la computadora deberemos utilizar un programa capturador de pantallas.

Escáner: Son unos dispositivos que, gracias a un programa desarrollado para tal efecto, hacen una réplica digital de la imagen que está impresa en el papel. Hay tres tipos de escáner: de sobremesa, de mano y de película.

Estos últimos se utilizan para digitalizar directamente rollos de película fotográfica negativa o diapositiva. Las casas Kodak, Nikon y Polaroid comercializan este último tipo de escáner.

Cámaras digitales: Son también conocidas como cámaras de vídeo estático.

Sirven para digitalizar fotografías en tiempo real. Su apariencia es como la de una cámara fotográfica normal, pero contienen un pequeño disquete magnético en lugar del tradicional carrete. Junto con la cámara se proporciona una interfaz y una tarjeta digitalizadora que convierte las imágenes analógicas en imágenes digitales, para que la computadora las pueda utilizar.

Digitalización de vídeo: Para capturar imágenes provenientes de una vídeo - cámara o de un magnetoscopio conectado a un televisor. Se necesita una tarjeta digitalizadora de vídeo. Se emplea también para la digitalización de originales en tres dimensiones.

Capturadores de pantalla: Son programas que se utilizan para capturar pantallas de otros programas. Uno de los más conocidos es el Paint Shop. si no se dispone de ningún capturador de pantallas se puede utilizar en Windows la tecla "Impr Pant", quedando la imagen que esté en ese momento en el monitor almacenada en el Portapapeles del Windows, desde allí se puede guardar en un formato de archivo digital BMP.

IV.12 Los formatos gráficos más importantes.

IV .12.1 Formatos de Archivos de Imágenes.

Estos formatos se utilizan en imágenes sin movimiento.

BMP (Bitmapped File Header Format).

CDR (Corel Draw)

Es de tipo vectorial, pero pueden insertarse elementos bitmap en las imágenes. Es uno de los formatos con más posibilidades con respecto al color, a la calidad de los diseños y al manejo de fuentes.

CGM (Computer Graphic Metafile)

Es de tipo vectorial y soporta elementos bitmap para incluir en las imágenes.

DXF Drawing Exchange Format.

Este es el formato vectorial por default de AutoCAD. Soporta hasta 256 colores. Su estructura no contiene información comprimida como en la mayoría de los casos, sino números y órdenes a realizar escritos en ASCII.

EPS Encapsulated PostScript Files

Soporta gráficos de tipo vectorial y posee limitaciones en cuanto al uso de elementos bitmap.

GIF Graphic Interchange Format

Los GIFs utilizan una paleta de entre 2 y 256 colores. Poseen una rutina de compresión muy eficaz que, aunque demora un poco la carga, reduce los archivos a un tamaño mucho menor que otros formatos.

Gracias a esa rutina de compresión que empequeñece los archivos, el GIF es el formato óptimo para ser bajado de BBS o Internet. La resolución máxima alcanzada es la de 1024 x 768 píxeles en 256 colores, pero no hay razón por la cual no pueda crearse una imagen de mayor tamaño.

HPGL Hewlett Packard Graphics Language

Es el formato utilizado por las impresoras Hewlett Packard. Es un formato de gráficos vectoriales y su extensión puede variar (por ejemplo: HGL, HPP, PGL, PLT, etc.).

IMG Aplicación GEM

Es capaz de almacenar imágenes de entre 2 y 256 colores. No hay restricciones con respecto al tamaño de las imágenes y utiliza una rutina de compresión medianamente eficaz.

MSP Microsoft Paint Files

Este formato bitmap solamente utiliza blanco y negro, pero no posee limitaciones acerca del tamaño de las imágenes. Posee un método de compresión normal que está compuesta por una tabla interna que permite que un programa descomprima únicamente las líneas de la imagen que requiere, dejando el resto de la imagen comprimida.

PCX PC Paintbrush

Es un formato bitmap y soporta imágenes de hasta 24 bits en color (unos 16 millones de colores). No hay restricciones con respecto al tamaño de las imágenes.

Su método de compresión apunta a la rapidez de acceso en vez de a la reducción de tamaño de los archivos. Casi todas las aplicaciones que circulan en el mercado soportan este formato, y esa es su mayor virtud: la compatibilidad. La gran mayoría de los programas de desktop publishing y de tratamiento de imágenes en sí soportan este formato, incluso en su última versión de 24 bits.

PIC PC Paint Pictor

La resolución máxima alcanzada por este formato es de 320 x 200 pixeles en 256 colores. Si se desea utilizar una resolución de 640 x 480 pixeles, los colores deberán ser 16. El método de compresión es eficaz siempre y cuando se trabaje con imágenes relativamente simples, y no con pantallas escaneadas.

TGA True Vision Targa

Es el formato utilizado por las tarjetas Targa. El standard está lo suficientemente bien definido como para evitar problemas de incompatibilidad. Las imágenes son bitmap y pueden ser de cualquier tamaño y contener tantos colores como se pueda imaginar (de 2 a 32 bits en colores).

TIFF Tagged Image File Format

TIF contiene una serie de bloques que conforman la imagen. Estos bloques pueden contener cierta información sobre la imagen en sí, su tamaño, su manejo del color, información a las aplicaciones que utilicen ese archivo, texto, y hasta thumb nails. Un thumbnail o miniatura es una pequeña representación de una imagen mucho más extensa.

WPG WordPerfect Graphic Files

Excepto el formato EPS, WordPerfect no importa otro formato que no sea el WPG. Este formato soporta tanto gráficos bitmap como vectoriales de 2 a 256 colores.

IV.12.2 Formatos de Archivos de Video.

Estos formatos se utilizan en imagenes con ó sin movimiento.

JPEG

El formato JPEG ofrece los imprescindibles 16 millones de colores (true color), unido a una compresión realmente asombrosa (valores superiores a 20:1 son habituales).

Sólo tiene una limitación: para obtener esos valores de compresión modifica sutilmente la imagen, descartándose su uso en aplicaciones en las que se desea mantener una calidad bit a bit. El diseño de este formato está pensado para almacenar imágenes del "mundo real", también llamadas imágenes de tono

PIC PC Paint Pictor

La resolución máxima alcanzada por este formato es de 320 x 200 pixeles en 256 colores. Si se desea utilizar una resolución de 640 x 480 pixeles, los colores deberán ser 16. El método de compresión es eficaz siempre y cuando se trabaje con imágenes relativamente simples, y no con pantallas escaneadas.

TGA True Vision Targa

Es el formato utilizado por las tarjetas Targa. El standard está lo suficientemente bien definido como para evitar problemas de incompatibilidad. Las imágenes son bitmap y pueden ser de cualquier tamaño y contener tantos colores como se pueda imaginar (de 2 a 32 bits en colores).

TIFF Tagged Image File Format

TIF contiene una serie de bloques que conforman la imagen. Estos bloques pueden contener cierta información sobre la imagen en sí, su tamaño, su manejo del color, información a las aplicaciones que utilicen ese archivo, texto, y hasta thumb nails. Un thumbnail o miniatura es una pequeña representación de una imagen mucho más extensa.

WPG WordPerfect Graphic Files

Excepto el formato EPS, WordPerfect no importa otro formato que no sea el WPG. Este formato soporta tanto gráficos bitmap como vectoriales de 2 a 256 colores.

IV.12.2 Formatos de Archivos de Video.

Estos formatos se utilizan en imágenes con ó sin movimiento.

JPEG

El formato JPEG ofrece los imprescindibles 16 millones de colores (true color), unido a una compresión realmente asombrosa (valores superiores a 20:1 son habituales).

Sólo tiene una limitación: para obtener esos valores de compresión modifica sutilmente la imagen, descartándose su uso en aplicaciones en las que se desea mantener una calidad bit a bit. El diseño de este formato está pensado para almacenar imágenes del "mundo real", también llamadas imágenes de tono

continuo, como digitalizaciones o renderizaciones (creación de efectos) de alta calidad cuando se utiliza para archivos animados.

AVI

Para las computadoras personales (PC).

MPG

Abreviatura de MPEG que es independiente de la plataforma y necesita su propio visor.

.MOV ó QT

Para películas QuickTime, inicialmente única para Macintosh, que se encuentra ahora disponible para Windows y UNIX.

Dentro de un ambiente profesional se pueden encontrar comunmente formatos de tipo JPEG y MPEG los cuales se tratarán detalladamente a continuación.

IV.13 Formato gráfico JPEG

IV.13.1 Introducción

Es uno de los tipos de compresión utilizados profesionalmente para el procesamiento de imágenes con aplicaciones televisivas, multimedia o videoconferencia entre otros; lo primero que se tiene que destacar de este formato son dos cosas: que ofrece los actuales e imprescindibles 16 millones de colores y que alcanza ratios de compresión superiores a 20:1. La limitación de este formato es que para obtener esos enormes ratios la imagen se debe modificar, descartando las partes en los resultados del método de compresión que el ojo humano no va a poder detectar. Cuando tratamos de comprimir imágenes sencillas (vectoriales, con pocos colores, etc.) se puede ver como la compresión es mínima y las modificaciones hechas en la imagen son visibles tras la compresión.

La abreviatura JPEG viene de la iniciales "Joint Photographic Experts Group". Se trata de un grupo de expertos que definieron las bases de este formato. Debemos tomar la precaución, cuando trabajamos con formatos que modifican la imagen, de guardar la imagen original. Si por ejemplo, convertimos una GIF a JPEG y luego hacemos la conversión inversa, es posible que el fichero GIF ocupe el doble que el original. Esto ocurre debido a que el formato JPEG no funciona correctamente con imágenes que contengan zonas de un sólo color. Existe un formato gráfico que toma lo mejor de estos dos formatos, el HSI JPEG, que desactiva la compresión JPEG cuando encuentra zonas extensas del mismo color. El problema es que es incompatible tanto con el GIF como con el JPEG.

Se intenta almacenar imágenes de tipo vectorial o dibujos sencillos no realísticos, se observará como la compresión disminuye enormemente, y las modificaciones hechas sobre la imagen original por el método de compresión se observan a simple vista.

IV.13.2 Codificación

El formato JPEG sólo puede almacenar imágenes de 24 bits (true Color), utilizando tres canales para su almacenamiento o de escala de grises, usando un solo canal.

La compresión JPEG consiste en una serie de complejas operaciones matemáticas como: conversión del formato de color, transformación separada del coseno (DCT), cuantizaciones y codificación entrópica que no se trataran en este documento ya que sería demasiado extenso.

IV.13.3 Pasos para una compresión JPEG.

- Antes de empezar, se trabajará con datos puros (raw) del tipo RGB, o de escala de grises, por lo que si no están en este formato, habrá que prepararlos previamente.
- Conversión de los datos originales RGB en datos YCbCr (luminancia y crominancia vistos en temas anteriores), también conocidos como datos YUV. Este proceso se denomina "conversión del espacio de color" (color space conversion), y es realizado porque el ojo humano es mucho más sensible a la información sobre luminancia que a las crominancias, y si se separan es posible comprimir más estas últimas sin mermar aparentemente la calidad. Este paso se ignorará si los datos de entrada se encuentran en escala de grises.
- Reducción de las muestras de entrada en componentes de color. Este procedimiento es denominado submuestreo (subsampling). Un muestreo utilizado es: 2h:1v, 1h:1v, e indica que el primer componente (luminancia) tiene el doble de muestras horizontales que los otros dos componentes (crominancias), y el mismo número de muestras verticalmente.

Los pasos siguientes se realizarán una vez para cada "barrido" del fichero de salida JPEG, o lo que es igual, una vez si el fichero es entrelazado o varias si es no-entrelazado:

- Creación de los UMC (unidad mínima de código (gif)), o partición de la imagen en bloques de píxeles.

- Transformación DCT (transformada discreta del coseno) para cada UMC (unidad mínima de código). Este es el verdadero corazón del formato JPEG y es la que se encarga de la compresión propiamente dicha.
- Escalada de cuantificado, y reordenado "zigzag" de los elementos de cada bloque de UMC's.
- Codificación aritmética o de Huffman de los bloques transformados.
- Creación de las cabeceras necesarias para un fichero JPEG.

IV.13.4 Pasos para la descompresión JPEG.

Los pasos necesarios para la descompresión de un fichero JPEG son éstos:

- ◆ Lectura del fichero JPEG, extrayendo los datos de la cabecera.
- ◆ Decodificación de la secuencia de datos, bien sea aritmética o del tipo Huffman.
- ◆ Escalado, cuantificado y reordenado "zigzag" de los elementos de cada bloque UMC.
- ◆ Desensamblado de los bloques UMC (también reconstrucción de los Bloques originales si el fichero fuera entrelazado).
- ◆ Posible aplicación del algoritmo de suavizado (smoothing) a cada bloque.
- ◆ Transformación inversa DCT para cada bloque.
- ◆ Reconstrucción de la imagen original, mezclando los componentes de color.
- ◆ Reversión del tipo de color original (de Y Cb Cr a RGB). Si se trata de una imagen en escala de grises, este paso se omite. Aplicación de un posible ajuste gamma.
- ◆ Con los datos raw se creará el fichero de salida deseado.

IV.13.5 Más acerca de JPEG.

La industria de la computación ha estado utilizando durante muchos años las técnicas de compresión para poder disminuir la cantidad de datos en los discos duros y en la

transmisión a través de los modems. Fue la industria de la computación la que adopto la norma del Grupo de Expertos Fotográficos Asociados o Joint Photographic Experts Group (JPEG) para comprimir las imágenes digitales fijas de alta resolución. No paso mucho tiempo sin que a alguien se le ocurriera mostrar también el video en la computadora, y así apareció el JPEG para movimiento. El Grupo de Expertos de la Imagen en Movimiento o Moving Pictures Experts Group (MPEG) fue formado en 1988 para determinar las normas internacionales de la compresión digital de las imágenes en movimiento, particularmente, para satisfacer el creciente interés en los CD-ROM.

Las técnicas del JPEG y del MPEG para movimiento ahora se usan extensamente en la producción de imágenes computarizadas, y también resultan muy efectivas por su costo para la grabación en disco, en CD-ROM, etc., sin embargo, ha surgido el MPEG-2 (un estándar ratificado por la ISO/IEC); la industria lo ha adoptado con una velocidad sorprendente, apoyándose casi completamente en el fuerte deseo que tienen de proporcionarle a los espectadores una gran capacidad de selección de programas suministrados directamente al hogar (Delivered Direct to Home, DTH), a través del satélite o por cable, utilizando para ello, decodificadores colocados encima del receptor. La codificación MPEG-2 esta basada en un grupo de circuitos integrados, cuyo primer fabricante fue la compañía C-Cube Microsystems. Las técnicas de compresión todavía están evolucionando, pero independientemente que este sistema ofrezca o no el método definitivo de compresión para el video digital, el MPEG-2 ha sido tan ampliamente aceptado que los ingenieros de televisión ya no pueden ignorar este sistema de reducción de datos.

IV.14 Formato gráfico MPEG.

Existen algunos ejemplos de compañías que han adoptado la norma de compresión de video digital MPEG-2. La firma Nethold's Multichoice inauguro en Bélgica sus 20 canales digitales de TV, también: Holanda, Luxemburgo, Escandinavia, el Medio Este y Africa mediante PAnAmSat, el 4 de octubre de 1995. El sistema utilizara un millón de decodificadores MPEG-2, los cuales serán suministrados por Philips, Panasonic y PACE.

La red de televisión Echostar's DISH, esta pensando ofrecer 150 canales de TV. El operador Canadian DTH esta planeando lanzar un servicio de 100 canales MPEG-2. El operador de televisión por pago (Pay TV) australiano Galaxy, esta instalando un sistema DTH de satélite digital de 10 canales empleando dos transponders del satélite Oteq 3B. El sistema utiliza codificadores MPEG-2 con acceso condicional Irdeto (codificado). La compañía Shinawatra Satellite de Tailandia, ha comenzado pruebas de su sistema de radiodifusión MPEG-2 DTH a través del satélite Taicom en banda Ku de Tailandia.

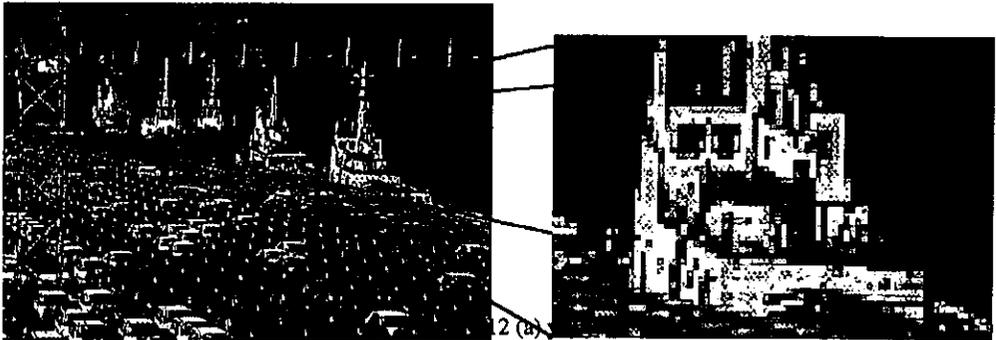
En el momento que se escribe este artículo, al menos diez fabricantes están produciendo decodificadores MPEG-2, y se espera que otros muchos sigan su ejemplo. Tanto el

sistema europeo DVB, como es sistema de Televisión Avanzada (ATV) de los Estados Unidos usaran la compresión MPEG-2 para la distribución de televisión digital terrestre.

IV.14.1 El desafío técnico.

No se puede comprimir el video hasta un punto, si no desecha algo, reduciendo en consecuencia la calidad de la imagen. Afortunadamente, el sistema visual humano es incapaz de absorber todo el material que este presente en una imagen en movimiento compleja, por tanto, con técnicas de compresión que seleccionen hábilmente el descarte de información, el ojo será incapaz de notarlo, pudiéndose lograr resultados impresionantes.

Como se puede observar en las imágenes a continuación (figura 4.12 (a),(b),(c),y (d)) y haciendo un desglose paulatino de un cuadro de televisión cada imagen está formada por un numero finito de pequeños pixeles (elementos de imagen). En la representación convencional 4:2:2 de las señales de televisión NTSC y PAL, hay 720 pixeles a lo largo de la parte activa de cada línea horizontal. En el sistema NTSC hay 483 líneas activas por cuadro (son 576 líneas activas en el sistema PAL) y 30 cuadros por segundo (25 para el PAL).



Cada pixel esta formado por 8 bits para la luminancia y 4 bits para cada una de las dos señales de diferencia de color (R-Y y B-Y, también conocidas como Cr y Cb), para un total de 16 bits.



Figura 4.12 (c) y (d)

El único objetivo del MPEG-2 es reducir estas velocidades de bits hasta un nivel algo más manipulable, y su éxito radica en la reducción de datos, fundamentalmente en dos áreas de las imágenes en movimiento. La primera área es la información contenida en cada cuadro (espacial: la relacionada con el espacio, es decir, las zonas redundantes de un cielo azul, etc.). La segunda área es la del detalle que no cambia de un cuadro a otro cuadro (temporal: la relacionada con el tiempo).

IV.14.2 Niveles y perfiles.

Los documentos publicados para la normalización de MPEG son ISO/IEC 13818 que cubren la compresión del audio, del video y la estructura de multiplexado necesaria para combinar los datos de audio, video y temporización, capaces de obtener la exitosa reproducción de un video con audio sincronizado. El MPEG-2 no es solo un estándar mundial, sino que es capaz de operar con todo, desde datos comprimidos para computadoras con velocidades de menos de 4 Mbits/seg. pasando por la televisión convencional a velocidades entre 10 y 15 Mbits/seg., hasta la televisión de alta definición operando a 80 Mbits/seg. Esto se conoce como los diferentes niveles, y la arquitectura del MPEG-2 soporta todos los niveles.

El MPEG-2 también proporciona flexibilidad en el tipo de compresión utilizada para cada nivel. Los tipos de compresión son conocidos como perfiles, y pueden variar desde un extremo máximo usando la señal total 4:2:2, hasta la eliminación de cuadros completos en el extremo mínimo.

Los codificadores pueden variar considerablemente, fundamentalmente al depender de una aplicación, por consiguiente, los detalles del esquema de codificación deben ser transmitidos junto con los datos para permitir que el decodificador reconstruya la señal. De esta manera, los codificadores pueden diseñarse para manipular los diversos niveles, utilizando diferentes perfiles, al tiempo que mantiene el costo de los decodificadores a un valor mínimo dentro de una aplicación deseada. La mayoría de la radiodifusión de 525 y 625 líneas emplea el perfil principal en el ámbito principal.

IV.14.3 Las capas y la escalabilidad.

Una de las características más ingeniosas del MPEG-2, es su habilidad para transmitir señales de video con un amplio rango de calidad. Un decodificador MPEG-2 relativamente barato puede reconstruir una imagen útil, empleando solamente una parte del flujo de video codificado. El resto de los datos se reserva para la mejora de la calidad. Los datos de video codificados consisten de una serie de flujos de bits de video llamados capas (layers).

La primera capa se conoce como, la capa base, y ella siempre se puede decodificar independientemente. A las otras capas se les llaman capas de mejora (enhancement layers).

Estas capas pueden usarse para extensiones espaciales, temporales y escalares. Si solamente hay una capa, a los datos de video codificado se le dice que tiene un flujo de bits de video no escalable. Si hay dos o más capas, a los datos se les dice que tienen una jerarquía escalable. La escalabilidad presenta un beneficio adicional, porque ayuda al video a ser resistente a los errores en el paso de transmisión. Los pasos de transmisión con la mejor calidad ante el error, pueden reservarse para la información crítica de la capa base, mientras que los datos de la capa de mejora pueden enviarse a través de un canal con inferior calidad ante el error.

IV.14.4 El flujo de bits de video.

El flujo de bits de video esta formado por bloques de pixeles, macrobloques (MB), imágenes, grupos de imágenes (Group Of Pictures GOP) y secuencias de video, resumiendo:

- Bloques
- Macrobloques
- Tajada
- Imagen
- Grupo de imágenes (GOP)
- Secuencia de video

IV.14.5 Bloques.

El elemento más pequeño es un bloque, formado por 8 líneas con 8 píxeles en cada línea. Los bloques se agrupan en macrobloques (MB), según alguno de los perfiles predefinidos del MPEG-2. El formato de macrobloque 4:2:0, tiene 4 bloques para la luminancia, un bloque para Cr y un bloque para Cb. La figura de abajo indica un macrobloque compuesto por 4 bloques.

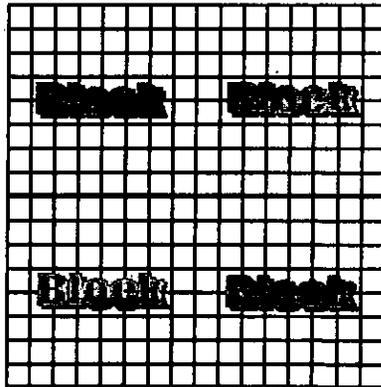


Figura 4.13

IV.14.6 Macrobloques.

El formato del macrobloque 4:2:2, tiene 4 bloques para la luminancia, 2 bloques para Cr y

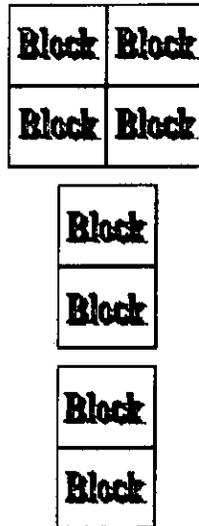


Figura 4.14

2 bloques para Cb (figura 4.14). Finalmente, el formato 4:4:4 presenta 4 bloques para la luminancia y 4 bloques para Cr y Cb respectivamente (figura 4.15 (a)).

Como se puede deducir, un MB 4:2:2 contendrá 8 bloques y, por tanto, $8 \times 8 \times 8$ (512) píxeles (figura 4.15 (a)).

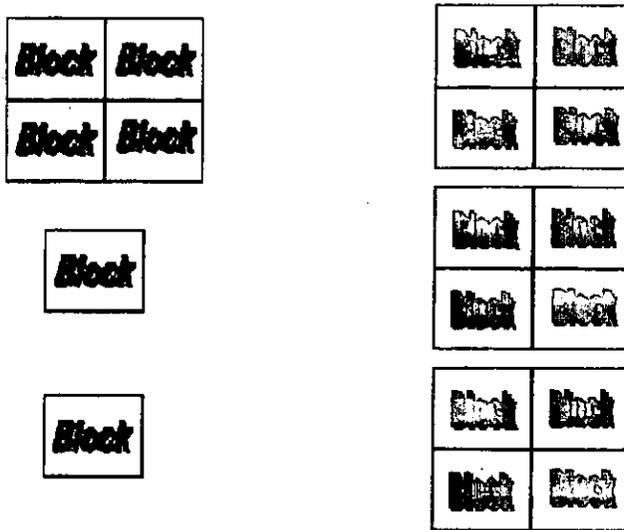


Figura 4.15 (a) y (b)

IV.14.7 Tajadas.

Las tajadas, son flujos de macrobloques arreglados horizontalmente a lo largo del raster. Las tajadas pueden variar en longitud desde un mínimo de un macrobloque hasta un máximo de una línea. Las imágenes y los grupos de imágenes serán analizados durante nuestra discusión acerca de la compresión temporal.

IV.14.8 Compresión temporal e imágenes.

La compresión temporal esta diseñada para minimizar la duplicación de los datos contenidos en imágenes sucesivas.

Esto se logra transmitiendo los datos de vectores de movimiento y alguna información de diferencia en vez de transmitir sucesivamente la imagen completa. Para facilitar la predicción del movimiento, el MPEG-2 separa el video en tres tipos de imágenes:

- Imágenes I (intracodificadas).
- Imágenes P (Codificadas por predicción)

- Imágenes B (Interpoladas bidireccionalmente)

Las imágenes I son la referencia clave para los otros dos tipos de imágenes. Ellas se derivan a partir de compresión de la información en un único cuadro o campo seleccionado (compresión espacial). Las imágenes fijas son mejor preservadas utilizando cuadros completos, pero como la velocidad de campo es 2 veces la velocidad de cuadro, el movimiento es mejor atendido empleando las imágenes que se basan en los campos. Algunos codificadores MPEG-2 son capaces de analizar el video que llega a la entrada para determinar los cambios entre campos sucesivos. Si no existen cambios entre los campos pares e impares, el codificador asume que los dos campos forman parte del mismo cuadro y los codifica como tal.

Los cambios entre campos son anotados y convertidos en vectores de movimiento, los que son codificados en forma de datos para una posterior intercepción por parte del decodificador. De esta manera se logra una reducción sustancial de bits. Los cambios son transmitidos en la forma de imágenes P y B. Las imágenes P son precedidas directamente a partir de la imagen previa I.

Las imágenes B se derivan por el uso de la información contenida tanto en la imagen P o en la imagen I, y estas fuentes de referencia pueden estar lo mismo delante que detrás de la imagen B que esta siendo creada, de ahí, el termino interpolación bidireccional. Tanto las imágenes del tipo P y B son comprimidas espacialmente antes de la transmisión. La técnica de compensación del movimiento utilizando el método anterior se conoce como compresión temporal.

Los tres tipos de imágenes se transmiten de forma secuencial como un grupo de imágenes, siendo siempre la imagen I la primera imagen. Típicamente, hay 12 imágenes en un GOP, pero como ya se mencionó anteriormente, algunos codificadores pueden detectar los cambios entre campos sucesivos, y si estos cambios son sustanciales, los codificadores asumen que se ha producido un cambio en la escena, y ello fuerza a la creación de una nueva imagen I, produciéndose de esta manera que la secuencia vuelva a comenzar. Los GOP se envían en una secuencia de video que contiene los datos que conducen las matrices de cuantificación, las velocidades y las dimensiones enviadas hacia la tajada, proporcionan códigos de inicio únicos que facilitan la detección por parte del codificador.

La única limitante al generar estas imágenes virtuales es que hasta ahora los ingenieros no han encontrado una manera fácil de poder editar con las imágenes B o P. Consecuentemente, las estaciones de televisión aun aceptan trabajar utilizando las técnicas del JPEG de movimiento para el procesamiento dentro de sus instalaciones, hasta tanto se encuentre una solución a este problema. No obstante, las relaciones de compresión de hasta 10:1 pueden todavía lograrse empleando el JPEG. Las relaciones de compresión en el orden de 25:1 pueden lograrse con el MPEG-2.

IV.15 Flujos de programa y flujos de transporte.

Con esto ya se ha comprimido el video, ¿qué sigue a continuación? Antes de poder almacenar o transmitir los datos, se tiene que unir el audio, el video y la información del sistema.

Normalmente existen dos multiplexores de audio/video para ello. Uno conduce el audio y el video, en flujos de empaquetamiento elemental y produce el flujo de programa y el otro utiliza los mismos datos para generar el flujo de transporte. Los flujos de programa están normalmente reservados para pasos de transmisión robustos, donde es difícil que se produzcan errores. El paquete de datos del flujo de programa, puede ser de longitudes diferentes y puede contener un número relativamente grande de bytes. Por otra parte, un flujo de transporte puede contener uno o muchos programas con una o muchas bases de tiempo independientes. Por consiguiente, múltiples canales de televisión pueden ser multiplexados conjuntamente de esta manera. Los paquetes de flujos de transporte tienen siempre 188 bytes de longitud. El flujo de transporte está diseñado para ambientes donde el error pudiera ocurrir.

La velocidad de transmisión de bits para una transmisión de TV o para copiado en cinta debe estandarizarse de la siguiente manera, la parte activa es considerada como el movimiento que debe de representar.

VELOCIDADES DE BITS PARA LA PARTE ACTIVA DEL VIDEO

NTSC 720 x 486 x 29.97 x 16 168 Mbits/seg.

PAL 720 x 576 x 25 x 16 x 166 Mbits/seg.

Las técnicas de compresión JPEG y MPEG continuarán siendo usadas para los requisitos necesarios de compresión en computadoras de bajo costo y en la televisión profesional donde se necesita la edición:

El MPEG-2 se ha convertido en un estándar internacional para la compresión del video con señales que simplemente van a ser almacenadas, distribuidas y vistas. Han sido desarrollados CD-ROM empleando los métodos de compresión del MPEG-2. El MPEG-2 ha sido adoptado mundialmente como el estándar de compresión para la Televisión

Directa por Satélite (DTH) y para la futura televisión de distribución digital terrestre (DTTV) y de cable, incluyendo la televisión de alta definición (HDTV).

IV.16 Los beneficios.

Para aquellos que todavía se preguntan la necesidad de convertir sus facilidades a video digital, pudiera servirles de ayuda el análisis de lo que puede lograrse con el video comprimido, permitiéndole con ello, ver mucho mas claro lo que significa dar este paso. Si las señales de video digital pudieran procesarse de manera tal que ellas se grabaran económicamente en discos duros de computadoras sin ninguna perdida de calidad evidente, entonces las posibilidades de editar, pintar y animar seria ilimitada. Por otra parte, si el video digital pudiera ocupar el mismo ancho de banda que ocupan las señales analógicas convencionales, los espectadores pudieran recibir en sus hogares imágenes con calidad de estudio garantizada. Todo esto y mas es posible empleando las técnicas de compresión, pero para ello lo primero que debe tener es una señal digital.

IV.17 Almacenamiento.

Se puede utilizar cualquier dispositivo para el almacenamiento de la imagen digital; disco duro de la computadora, disquete, disco óptico, cintas magnéticas, etc. Debido al gran tamaño que ocupan los archivos de imagen se recomiendan las tecnologías ópticas para el almacenamiento masivo de éstas.

Dentro de las tecnologías ópticas tenemos: los discos ROM, WORM, Regrabables (magneto-ópticos) y por su puesto los discos duros. Para la lectura de estos discos se utilizan unidades especiales, los más llamativos son los Juke Box o arrays que permiten almacenar varios discos ópticos en un mismo dispositivo, pudiendo acceder al contenido de cualquiera de ellos de una forma automática.

	4bits	5bits	6bits	7bits	8bits
32 pixeles	512	640	768	896	1024
64 pixeles	2048	2560	3072	3584	4096
128 pixeles	8192	10240	12288	14336	16384
256 pixeles	32768	40960	49152	57344	65536
512 pixeles	131072	163840	196608	229376	262144

Cantidad de memoria precisa para almacenar una imagen. Relación bits por pixel - tamaño de la imagen. ⁶⁰

IV.18 Protección de Señales Digitales en Almacenamiento y Tránsito.

Por la regeneración de la corriente de bits en intervalos regulares, muchos de los disturbios a los que las señales análogas están expuestas pueden ser evitados. Pero los errores pueden ser introducidos en cualquier momento de la transmisión, almacenamiento o recuperación de señales digitales, por errores de tiempo o sincronía, por la introducción de bits desde fuentes externas o por la omisión como resultado de una falla en el conteo y reconocimiento.

Para minimizar el efecto de los errores de las funciones de transmisión y procesamiento, existen muchos métodos para corregir o cubrirlos.

En general, la protección y corrección de señales digitales, envuelven la introducción de por lo menos 1 bit y generalmente más a palabras de 4 a 8 bits.

Esto incrementa proporcionalmente el ancho de banda ocupado por una muestra, pero la señal de protección esta considerada dentro del espectro, particularmente cuando la transmisión a larga distancia no esta involucrada.

La palabra con estos bits extra es conocida como block, identificada por el total de bits y los bits de acarreo de la señal de información. De esta manera un block de (10,8) expresa 10 bits, 8bits de información y 2 para protección de error.

Un simple código emplea blocks de (9,8) el bit extra es conocido como "bit de paridad". El codificador cuenta el numero de 1's en la palabra de 8 bits, si esta cuenta con un numero impar, añade un 1 para hacer una palabra de 9 bits, si es otro numero, un cero es sumado.

Así, todas las palabras de 9 bits contienen números 1's de más. El decodificador cuenta de la misma manera los 1's en cada palabra de 9 bits y si el numero es impar, esa palabra puede ser incorrecta en al menos uno de las posiciones de sus bits.

⁶⁰ Alberto Domingo Ajenjo, Tratamiento Digital de Imágenes, Anaya Multimedia, 1994, página 30.

El decodificador debería entonces de esperar en particular una palabra de 9 bits para retransmitirla hasta que no aparezca un error de paridad, pero el retraso involucrado. Generalmente no puede ser tolerado en aplicaciones de video. Preferentemente el decodificador elimina la palabra con el error y forma otra en su lugar por interpolación desde un procedimiento and/or de las siguientes palabras. Una palabra representa un elemento de la imagen, por ejemplo, puede ser repetida para llenar el hueco o las palabras siguientes o anteriores pueden ser promediadas para hacerlo. Para que este proceso se lleve a cabo, las palabras siguientes y las palabras anteriores deben de estar disponibles en la memoria de almacenamiento y el error de corrección involucra un retraso de al menos una línea de escaneado. Tal error oculto es factible solo cuando hay pequeños cambios entre las palabras. En imágenes estacionarias o sin movimiento, los cuadros sucesivos tienen elementos idénticos y un error en la palabra puede ser reemplazado por la palabra correspondiente del cuadro anterior suministrado de la misma palabra que fue guardada.

Una corrección de error más elaborada se requiere cuando muchos bits sucesivos en una palabra están perdidos o cambiados por ruido u otro tipo de interferencia. La corrección de los bits necesitada es más numerosa, y su posición en el bloque se determina por un esquema complejo, conocido como código FEC *forward error correction* (código adelantado de corrección de errores). Este tipo de esquemas puede corregir más de cuatro errores en un block de 12 bits.

De acuerdo con este tema hemos observado las transformaciones de una señal de video analógica a una señal de video digital y esta a su vez ha sufrido transformaciones como compresión y ha sido incluida en formatos gráficos ya sea de datos o de video. Una vez alcanzado este objetivo ahora toca el turno a la manipulación de este material para dar forma a un programa, comercial, o segmento de noticias para ser transmitido por TV. En el siguiente capítulo se podrá analizar el esquema del sistema en general y los elementos que lo componen.

CAPITULO V
SISTEMA DE EDICIÓN NO LINEAL.

V. SISTEMA DE EDICIÓN NO LINEAL.

Una vez observado el origen y la descripción de cada uno de los pasos a seguir para obtener video digitalizado, el siguiente paso es manipularlo. La manipulación de éste tipo de video se lleva a cabo comúnmente en lo que se conoce en términos televisivos como sistemas de edición no lineales, esto debido a la flexibilidad de operación en los procesos de transformación por los que puede pasar un segmento de imagen hasta su transmisión en televisión.

En la migración de las tecnologías análogas por tecnologías digitales, las no lineales han revolucionado todos los aspectos de la producción de televisión, el tratamiento de imágenes y los materiales o Media (el Media es uno de los nombres técnicos que se utilizan en TV para cortos de video digitalizado) que se utilizan para ello. Si se trabaja en cine o video, se especializa en audio o en proyectos de multimedia, produciendo programas de televisión o contenidos de CD - ROM, los comentarios al principio eran si el usar tecnología basada en discos (Disk-based) sería útil, pero ahora más bien sería como implementarlos lo mejor posible.

Existen presiones que los medios de comunicación, en este caso los televisivos, toman como prioridades, por ejemplo si los costos se incrementan, la presión de economizar y producir mas con menos, etc., para ello estas herramientas tienen que ser flexibles, estar completamente probadas y ser rápidas.

Los sistemas de edición no lineal ofrecen a los editores la flexibilidad y funcionalidad que simplemente no tenían en las salas de edición tradicionales. Los editores son liberados de las restricciones de los sistemas basados en cintas, Tape - Based.

La velocidad y la flexibilidad del sistema basado en discos permite producir programación con mas creatividad, con mayor eficiencia que determinan estar en la cima de este tipo de negocios.

La más importante consideración de todo. ¿Que tan productivo puede ser?

V.1 ¿Qué es un sistema de edición lineal? ⁶¹

(Sistemas Basados en Cinta o Tape_Based)

Los sistemas de edición lineal son también conocidos como salas de edición tradicionales en las cuales se tienen ciertas restricciones para llevar a cabo una edición cualquiera que esta sea, por ejemplo y sobre todo la necesidad de recorrer continuamente las cintas de

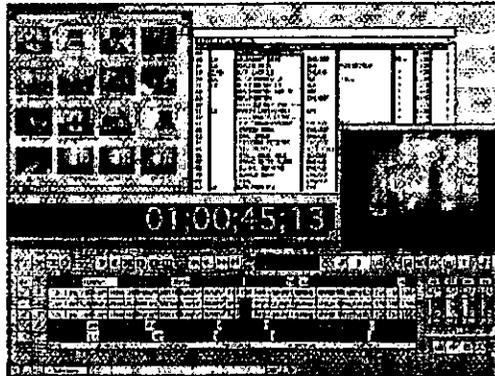
⁶¹ <http://www.avid.com/avidnet/>

video para encontrar las tomas deseadas, realizar vistas previas de la edición o los cortes realizados y aún después la grabación de los mismos, la imposibilidad de crear fácilmente efectos o transiciones entre tomas. Todo esto presenta una problemática muy importante, el tiempo, todas estas restricciones entre muchas otras son llevadas a cabo siguiendo una serie de pasos de edición simple lo cual hace mecánico su funcionamiento y por lo consiguiente es muy tardado para lo cual se buscaron alternativas que fuesen flexibles y funcionales para liberar a los editores de estos problemas.

V.2 ¿Qué es un sistema de edición no lineal?⁶²

Son sistemas de computadoras que permiten una mejor manera de editar ya que están basados en discos con los que se puede tener acceso al azar a cualquier cuadro de video o mover instantáneamente las tomas que son requeridas, eliminando la necesidad de adelantar y regresar cintas muchas veces. Se pueden hacer múltiples versiones de historias ya que es más fácil y rápido utilizando menos puntos de edición (keystrokes) además se puede añadir mas impacto en ellas con efectos. Se pueden agregar títulos, efectos de transición, efectos de movimiento y efectos digitales DVE, aún si la historia ya terminó. La libertad de creatividad y la velocidad son los factores más importantes que tiene este tipo de sistemas lo que le ha dado gran aceptación. Abajo (figura 5.1) un ejemplo general de algunas pantallas utilizadas en éste tipo de sistemas.

Figura 5.1



V.3 Modelo de Edición.

Muchos sistemas de edición no lineal permiten desarrollar tareas similares es decir, son parecidos. El modelo de edición debe de permitir terminar mas proyectos, con el mejor valor de producción en el menor tiempo. El costo de un sistema de este tipo es trivial comparado con el costo en tiempo. Si el producto final es un CD ROM, cinta de video o

⁶² <http://www.avid.com/avidnet/>

transmisión al aire, el modelo de edición debe de permitir llevar a cabo las siguientes tareas de manera rápida y fácil.

- Tareas de empalmar en edición común, sobre escribir, levantar y extraer, trimear, re acomodamiento de tomas, agregar efectos de transición, mezcla de audio y escribir nuevas marcas de tiempo código para localización.
- Efectos comunes de edición, como composición de tres capas de video o crear efectos de Picture-in-picture.
- Manejo común de funciones de Media como multinivel de filtros, clases y búsqueda de clips encontrados en bins (un bin es una ventana donde se encuentran los segmentos de video o media u efectos creados, este concepto se aplicará más adelante).

Este tipo de características les permiten ser una herramienta muy poderosa para adaptarse a diferentes métodos de trabajo. La mejor manera de evaluar un método de ediciones no considerando el lugar o las ventanas específicas o el diseño de los iconos, sino determinar cuanto tarda en completar una tarea. Esto permite un proceso eficiente, el cual permitirá producir más trabajo en menos tiempo.

V.3.1 Manejo de Media.

Estableciendo un eficiente y comprensible modelo de edición, se encuentra el manejo del Media o material de video en bruto, editar se puede resumir como organizar. Se debe de ser capaz de mantener los tracks de video, organizados de una manera lógica y localizar lo que se necesite de una manera rápida y fácil. Sin un buen manejo de Media se pierde el tiempo en demasía.

El manejo del Media acompaña a las funciones de digitalización, es decir, en que espacio se almacena y de que forma están almacenados los clips, la forma de visualizar y organizar, como se hacen las búsquedas y las clases de material, como se mueve y se respalda el material.

La digitalización es un proceso que consume tiempo y se buscan las características que

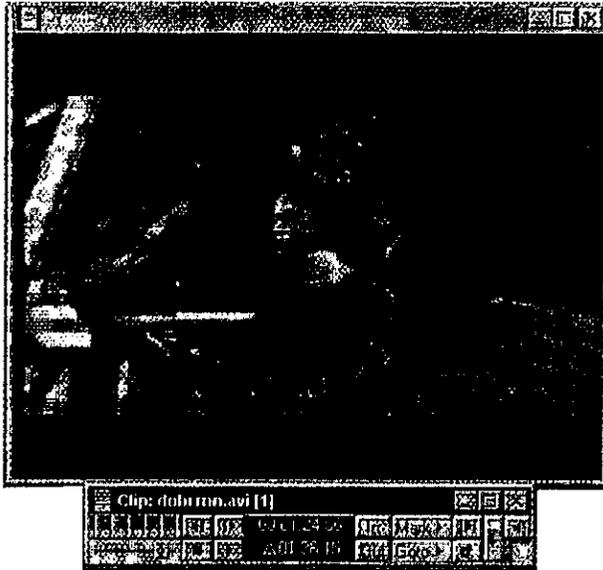


Figura 5.2

ayuden a hacer el proceso más rápido, eficiente y redituable. Las características de digitalización eficientes pueden incluir la creación de clips basados en rompimientos de tiempos código y la capacidad de poner comentarios y nombres a los clips que se graban. Estos clips de duración variable pueden ser capturados en herramientas de captura similares a la imagen anterior (figura 5.2).

El "Batch digitize"⁶³ o digitalización en grupo es otra característica de estos sistemas, permite especificar un grupo de clips que se van a digitalizar, el cual provee dos ventajas de trabajo, la primera, se puede apuntar la selección de campos como mejor nos parezca, en lugar de repetir todo el tiempo que tarda esta tarea en escanear las cintas y elegir las mejores tomas, se puede tener en el sistema la digitalización automática directamente de los puntos que fueron anotados anteriormente.

A continuación se puede observar una pantalla (figura 5.3) que es utilizada por el proceso de digitalización en ella se ingresa el código de tiempo o "time code" en donde se encuentra

⁶³ <http://www.avid.com/avidnet/>

el material a digitalizar proveniente de una cinta de video, en ella se puede indicar tambien el tiempo de salida o la duraci3n que se desea para el clip.

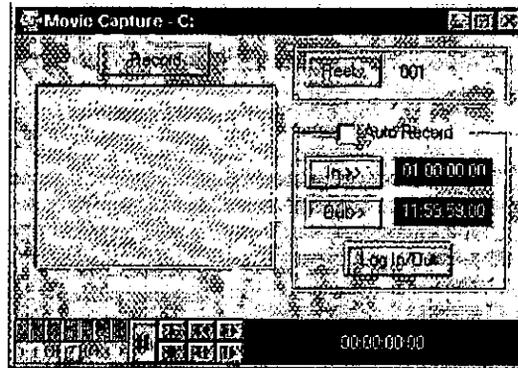


Figura 5.3

Si estas anotaciones de tiempos no existen el sistema avisa que hay un error, e indica que hay que verificar las entradas y las salidas correctas y entonces finalizar con la digitalizaci3n autom1ticamente.

Segundo, la digitalizaci3n en grupo permite autom1ticamente redigitalizar el material en la calidad de resoluci3n de video que se requiera.

Por ejemplo, se puede digitalizar el material fuente a una menor resoluci3n para conservar espacio de almacenamiento, y cuando el proyecto est3 finalizado se puede redigitalizar solo las secciones que se utilizaron en el programa final. El sistema debe de apuntar las cintas fuente y usar los tiempos c3digo para digitalizar solamente los clips necesarios. Los sistemas digitalizan los tiempos c3digo en orden y una cinta a la vez. De esa forma se puede dejar que el sistema trabaje solo a excepci3n de que se tengan que hacer cambios.

Se describen a continuaci3n:

En la primera imagen (figura 5.4) se observa la imagen de una pantalla utilizada para la revisi3n del media del cual se sacarn clips digitalizados.

En la segunda (figura 5.5) nos presenta una imagen con lneas de captura para nombrar el clip, indicandonos sus caracteristicas y comentarios.



Figura 5.4

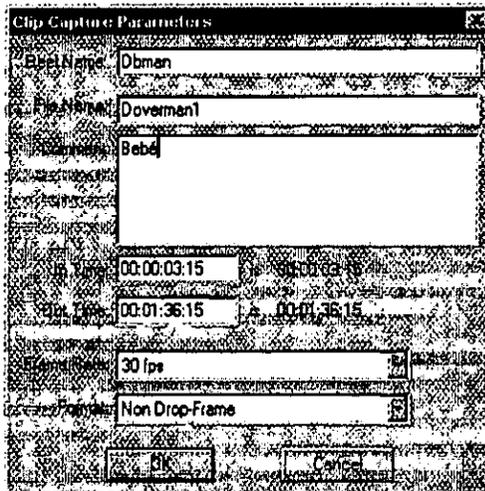


Figura 5.5

En la primera imagen arriba se encuentra la herramienta de digitalización (figura 5.4), en ella se escanean las tomas que serán utilizadas, durante este proceso se toma nota del tiempo en que comienza cada una y el tiempo en el que terminan, las cuales serán anotadas en el bin de tiempo que se muestra abajo (figura 5.5), en donde se le dará un nombre a cada uno de los clips a digitalizar y así reconocer cada una de las tomas o imágenes que serán utilizadas.

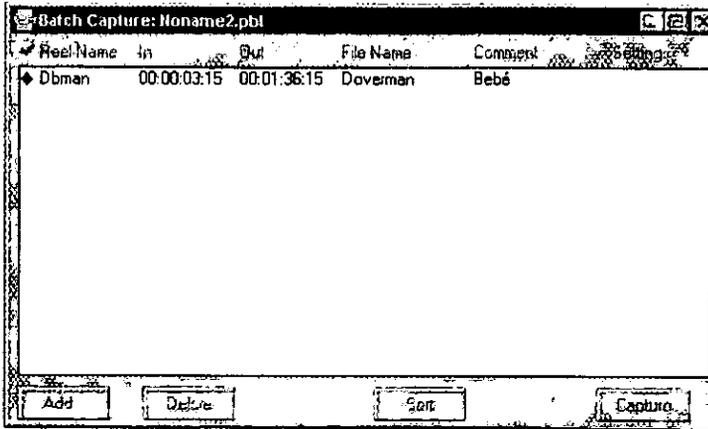


Figura 5.6

En la imagen anterior (figura 5.6) aparecen los clips que han sido digitalizados en orden en un bin, desde donde pueden seleccionarse para ser utilizados como sea conveniente.

Los sistemas deben de tener la habilidad de "consolidar" el material, la consolidación automáticamente copia las secciones de Media utilizadas en la secuencia final y si se quiere, borrar los archivos originales cuando ya se termino. Esto es un método conveniente para hacer copias de respaldo de archivos importantes en otros discos.

La mayoría de los sistemas de edición no lineal utilizan algún tipo de BIN⁶⁴ o ventana en donde se visualizan los clips que han sido almacenados, estas Ventanas facilitan el trabajo.

La siguiente imagen (figura 5.7) muestra un Bin o ventana de construcción este es una de las ventanas más importantes ya que en ella se desarrolla todo el proceso de edición, también se conoce como "timeline" o "línea de tiempo", los clips seleccionados se van colocando en el orden como serán utilizados, en este bin se pueden realizar todos los

⁶⁴ <http://www.avid.com/avidnet/>

movimientos necesarios dentro de una edición de video tales como cortes, trims, ajustes de audio, musicalización, efectos etc.

Esta ventana puede contener hasta 24 canales de video y 16 canales de audio independientes (el número de canales puede variar dependiendo del software utilizado y la capacidad del sistema), canales de efectos que se aplican a la imagen tales como cambios de color en la imagen, posterizaciones, movimientos de imagen, rotaciones, zoom, colocación de gráficos, supers ó títulos etc.

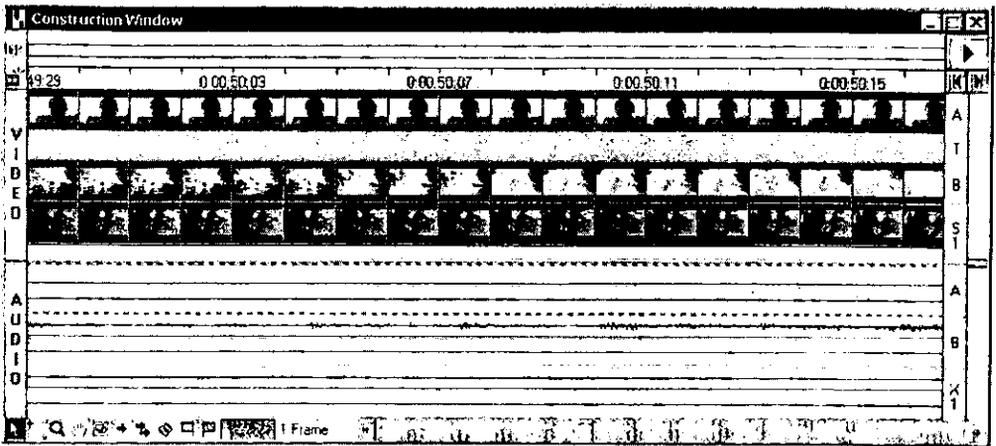


Figura 5.7

Las funciones de base de datos deben de ser capaces de manejar todo el Media para una recuperación rápida, por ejemplo, se deben de catalogar historias por nombre, con una palabra clave o con otra clasificación común, y los comentarios que se agreguen a los clips y hacer búsquedas por palabra creando niveles de clase y desplegar los clips que las contengan según sea el criterio.

Como en la imagen siguiente (figura 5.8) los bins deben de permitir ver los clips como los comentarios y además los cuadros de video que lo conforman algunos sistemas permiten ver el primer cuadro de video que se encuentra en ese material para saber parte de su contenido. Se deben de distinguir con iconos los elementos de un bin para diferenciar los gráficos, los clips fuente, las secuencias editadas, los efectos, etc.

Se debe de ser capaz de especificar cual cuadro representa cada clip y reorganizar los cuadros si es necesario, ya que entre mas organizado esté el bin se editara más rápido. Un buen sistema debe de permitir organizar un bin por storyboard (el termino storyboard es utilizado para describir un bin en donde en lugar de solo contener los nombres y la duración de cada clip se incluyen los primeros cuadros de video, haciendo ver al bin como una serie de fotografías que narran una historia) y pasarlos al timeline (el "timeline" o línea de tiempo es un bin donde se construye la historia o segmento de edición es uno de los bins más importantes en donde se editan directamente las secuencias, este bin esta ejemplificado en la imagen anterior) para crear una secuencia. En la figura siguiente (figura 5.8) se puede observar algunas formas de organización en este bin las imágenes muestran el primer cuadro de video de los clips a editar incluyendo el tamaño de la imagen, el tiempo de duración y el sonido.

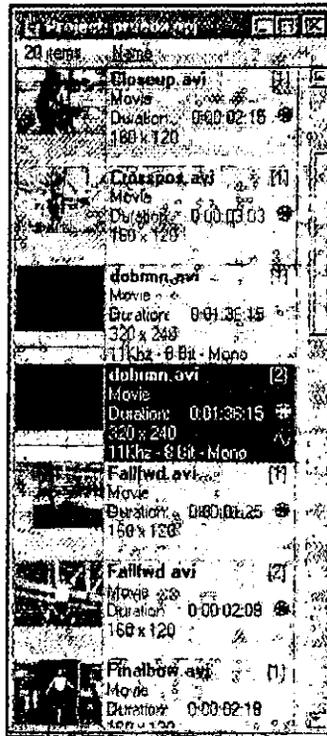


Figura 5.8

V.3.2 Composición y efectos integrados.⁶⁵

Los efectos son componentes críticos en cualquier sistema lineal de video, la habilidad de integrar la composición y las capacidades de los efectos incluyendo los de 3D, verdaderamente distinguen a los sistemas profesionales.

Un sistema de edición no lineal permite crear efectos de transición, efectos multicapa, efectos en 3D y títulos dentro de la misma aplicación y cambiarlos en cualquier momento sin afectar la capas anteriores.

Al contar con estas funciones frecuentes integradas en una misma aplicación se puede trabajar de manera mucho más rápida solo aprendiendo a usar la interface de los fabricantes.

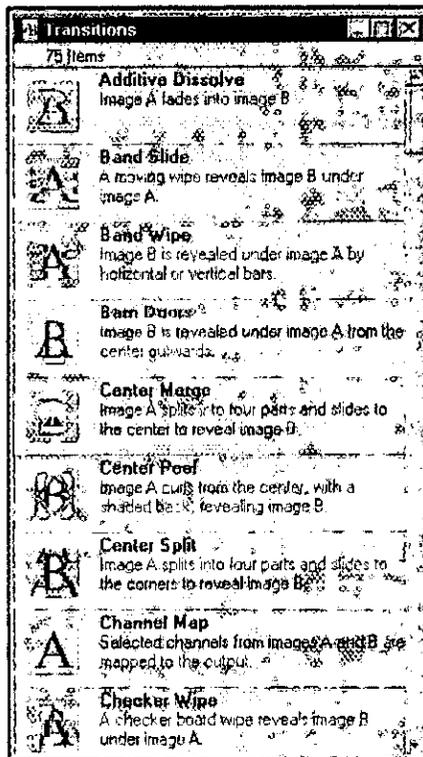


Figura 5.9

⁶⁵ <http://www.avid.com/avidnet/>, las imágenes forman parte de la aplicación Adobe Premiere para PC, Windows 95.

En la figura 5.9 se muestra una parte de un Bin de Efectos en el cual se puede encontrar más de 200 tipos diferentes de efectos dependiendo de la aplicación.

Los mejores sistemas permitirán crear efectos complejos con todo control, una vez creado un efecto, se debe de salvar o guardar para poder ser utilizado después sin necesidad de repetir todo el proceso para crearlo, al menos se debe de estar seguro de que el sistema contiene efectos de transición, efectos de movimiento, DVEs en capas y de llaves.

Por ejemplo en la siguiente imagen (figura 5.10) se muestra la creación de un efecto que comunmente se conoce como disolvencia, el tiempo que lleva hacerlo no tarda más de unos cuantos segundos lo que contrasta en sistemas lineales en donde se debe de realizar un proceso de AB Roll que lleva aproximadamente 7 u 8 minutos dependiendo de las pruebas previas antes de grabar.

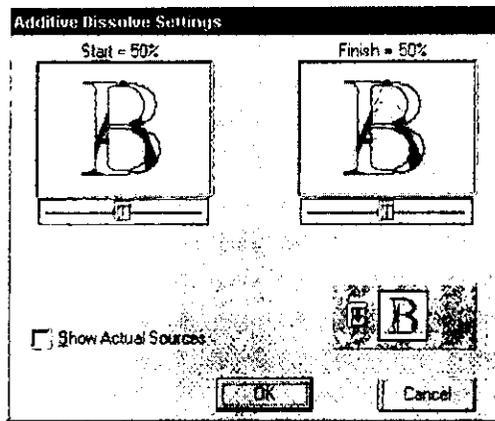


Figura 5.10

Un sistema basado en discos, no solo tiene ventajas para edición horizontal sino también para edición vertical o por capas, es decir se puede tener muchas capas de video y crear efectos de niveles multicapa. Cuando se combinan muchas capas de video y sonido, las herramientas las integran para lograr un intercambio individual de elementos y reensamblar la composición. Las características de tiempo real son esenciales en la experimentación, a partir de que se termina un proyecto se pueden tener cientos de efectos, que se necesita considerar el tiempo que se tardara en renderear o crear el efecto. Los inconvenientes del render es el tiempo que este tarda en crear el o los efectos y el espacio en memoria. Al detener un render se detiene también el flujo de trabajo.

V.4 Calidad de la imagen.⁶⁶

Una vez que se determine que tipo de sistema de edición no lineal se debe de utilizar, el siguiente factor es el de la calidad de la imagen. Se deben de considerar factores importantes. El primero, es por su puesto como aparece la imagen en la pantalla, pero el cómo una imagen aparece en la pantalla es un solo aspecto de la calidad de la imagen. Una señal oculta de calidad provee de criterios objetivos en los cuales se puede calificar a un sistema de este tipo.

Los sistemas deben de capturar todas las 21 líneas de blanking. Las imágenes que tengan niveles de negros debajo de 7.5 IRE o blancos arriba de 100 IRE no deben de ser registrados en clips ya que puede resultar imágenes planas.

Solamente un tamaño de imagen de 720 x 486 captura todas las líneas activas de una imagen y el blanking. El tamaño de 720 x 486 es también parte del desarrollo de las transferencias en tiempo real para formatos digitales. Y Los formatos de 640 x 486 están limitadas para captura analógica y salida de video.

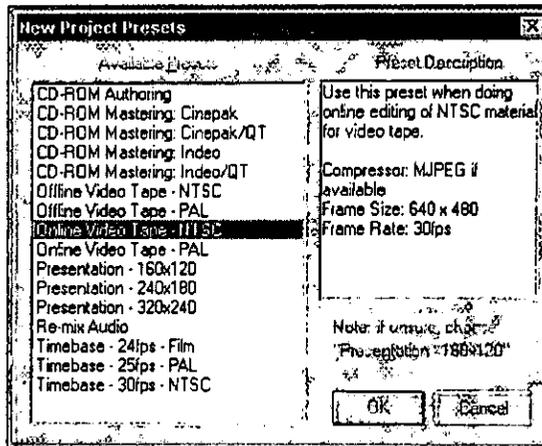


Figura 5.11

La imagen de arriba figura (5.11) muestra las diferentes opciones de calidad que estos sistemas pueden utilizar y los tipos de sistemas de televisión que se utilizan en el mundo.

Los sistemas de edición no lineal han incrementado sus aplicaciones no solo para productos cuya función sea transmitirse por T.V. Desde que los formatos Multimedia están realizados

⁶⁶ <http://www.avid.com/avidnet/>, Manual Sony pag 215.

en un solo campo de video, los productores de CD ROM, los contenidos en Web u otros productos multimedia han determinado si estos sistemas son adecuados para trabajar en un solo campo, al hacer esto se ahorra espacio de almacenamiento lo cual ahorra dinero a los productores.

V.5 Características del Audio.

Las características del audio pueden ser cruciales para el éxito del producto final. En estos sistemas se pueden hacer ajustes directamente a los niveles y la calidad de audio, depuración de cambios en picos, despliegue de la onda y ecualización en tiempo real, mezcla de audio y se pueden manejar por lo menos 4 canales de audio. No se necesita crear varios cortes de cinta para crear efectos ni mezcla de los mismos, se pueden integrar disolvensias de audio y otro tipo de efectos.

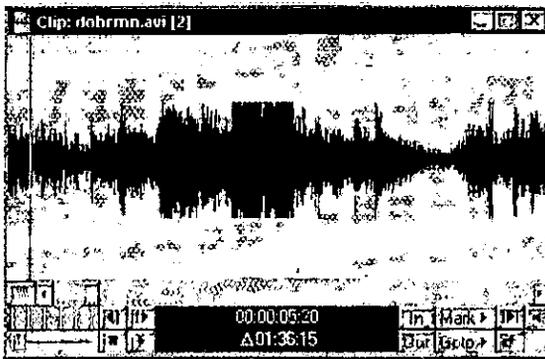


Figura 5.12

Arriba en la imagen (figura 5.12) se encuentra la representación de un segmento de audio digitalizado, es así como se apreciaría en la pantalla de la computadora.

A su vez muchos de los proyectos necesitan de audios externos para musicalización o efectos especiales de sonido, estos pueden ser importados de diferentes formatos como CD externo, algunos de los formatos de audio que están reconocidos como compatibles son: AIFF, WAV y Sound Designer II. Se soporta un intercambio de OMF y se puede exportar en este formato a computadoras especializadas en la post producción de audio o Protools.

V.5.1 Compatibilidad con otros sistemas.

El audio no es solo un área en donde se es compatible con otros sistemas, en video se utilizan para Macintosh los formatos de QuickTime, los AVI para Windows y los OMF

para diferentes aplicaciones como en gráfico, títulos, animación, aplicaciones de audio y video para procesamiento de imágenes.

V.6 Opciones de Almacenamiento.

El almacenamiento de información es uno de los más significativos costos asociado con los sistemas de edición no lineales, particularmente por los proyectos on line. Al buscar un amplio rango de almacenamiento encontramos el almacenamiento fijo y almacenamiento removible, algunos sistemas de este tipo manejan discos duros removibles, esto permite mover proyectos de una manera sencilla entre los sistemas para obtener archivos importantes de material para su complementación. Se pueden cambiar incluso los drives sin la necesidad de abandonar las aplicaciones o apagar el sistema, para ella se tiene que estar seguro de que el almacenamiento del sistema soporte la reproducción de resoluciones de video más altas.

Algunos sistemas esta utilizando drives llamados Digital Linear Tape, discos DLT que trabajan a altas velocidades y con los mejores métodos de respaldo en video digital, estos discos son aproximadamente tres veces más rápidos y cuatro veces mayor capacidad que los discos DAT. Estos discos permiten crear registros de material o proyectos que permiten moverlos o crear nuevos.

Se puede almacenar en varios tipos de resoluciones, estas resoluciones tienen que ver directamente con la digitalización y el tipo de compresión que se le aplique a la imagen original.

AVR 25 (compresión 10:1) formatos VHS⁶⁷ (figura 5.13)

Maneja una imagen de 640x480 pixeles, con audio 2/4/2 con frecuencias de 44.1/48 Khz. 24 fps de soporte de cinta, edición tiempo línea o timeline y OMFI.



Figura 5.13

⁶⁷ <http://www.avid.com/digitize/>

Como se puede ver arriba la calidad de la imagen no es muy buena pero es utilizada para realizar pruebas sin que se ocupe gran cantidad de memoria para no saturar los discos.

AVR 27 (compresión 5:1) formatos Hi-8 (figura 5,14)

Maneja efectos en tiempo real, títulos en tiempo real, listas ópticas, Auto sincronía, soporte de enlace Photoshop de tipo Plug in, localizadores, acceso a red y detección de copias, sistema multicámara.



Figura 5.14

La calidad de la imagen mejora y es televisivamente aceptada para transmitirse, esta calidad es permitida en muchos noticiarios aunque tiende a desaparecer por los nuevos formatos digitales.

AVR 75 (compresión 3:1) formatos Betacam (figura 5.15).

Maneja imagen a 720x480 pixeles (estándar NTSC), audio de 4/4/4, codificador para Quick Time, Render de alta velocidad, efectos 3D en tiempo real, formatos gráficos PICT con opción de mate.

A continuación (figura 5.15) podemos ver que la calidad de la imagen es muy buena, es el tipo de imagen preferido en todas la televisoras ya que cumple con los estandares profesionales necesarios.

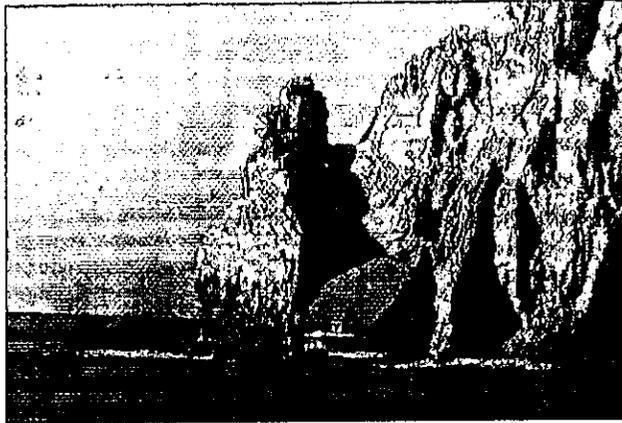


Figura 5.15

AVR 77 (compresión 2:1) formatos digitales (figura 5.16).

Audio de 4/8/4, mezcla de resoluciones, formatos nuevos de cinta, aplicación de tituladora, localizador de ventana, edición por ensamble, emulación de VTR, córner pinning, efectos 3D (tercera dimensión), puerto serial I/O, trim infiller.



Figura 5.16

Este tipo de compresión no tiene pérdida de calidad es la tendencia de las televisoras para futuros programas en Televisión de alta definición (HDTV)

V.7 Sistema de edición tipo DNG.⁶⁸

El sistema DNG Digital News Gathering, es un sistema basado en discos con un servidor central de producción diseñado para sistemas de noticias de televisión, éste maneja todo lo relacionado con la grabación, la edición y la reproducción sin necesidad de mover el Media. Es un sistema basado en un modelo cliente/servidor, en donde todo el video y el audio es almacenado en un servidor central con escalabilidad de almacenamiento en línea, utiliza lo último en video compresión, manejo de altos anchos de banda etc.

El sistema completo permite a las cadenas de T.V. trabajar con su Media al mismo tiempo que se trabaje con promos o historias manteniendo la más alta calidad.

V.7.1 Los componentes del sistema.

Estaciones de trabajo: Media Recorder, Media Editor, Airplay para estaciones de trabajo en las que se puede grabar, editar y reproducir el Media, estas estaciones de trabajo tienen bibliotecas de archivos Media almacenados en un Servidor Central.

Servidor de Media: Es el centro del sistema, recibe, almacena, organiza y devuelve bins y material Media según los requerimientos de sus clientes. Contiene material que se puede compartir con acceso simultaneo al Media.

Canal de fibra Raid 3, que maneja arreglos de discos para el almacenamiento central.

Tecnología de red: maneja las conexiones de las estaciones de trabajo al servidor.

Sistema de Base de Datos de Servidor de Media: es habilitado para establecer categorías y la organización del Media. Junto a este sistema se incluye un sistema del modelo Cliente/servidor, acostumbrado a manejar grandes cantidades de datos en almacenamiento y las velocidades requeridas para el manejo del material media digital.

El DNG maneja una configuración LAN Local Area Network para las conexiones de las estaciones de trabajo de Edición (Newscutter), airplay y el Media recorder con la biblioteca central. El Media es almacenado en el RAID 3 que son discos con arreglos conectados al servidor media. Las estaciones de trabajo comparten el acceso simultaneo al servidor central.

⁶⁸ <http://www.avid.com/avidnet/dng/peer-to-peer>

DIAGRAMA 1

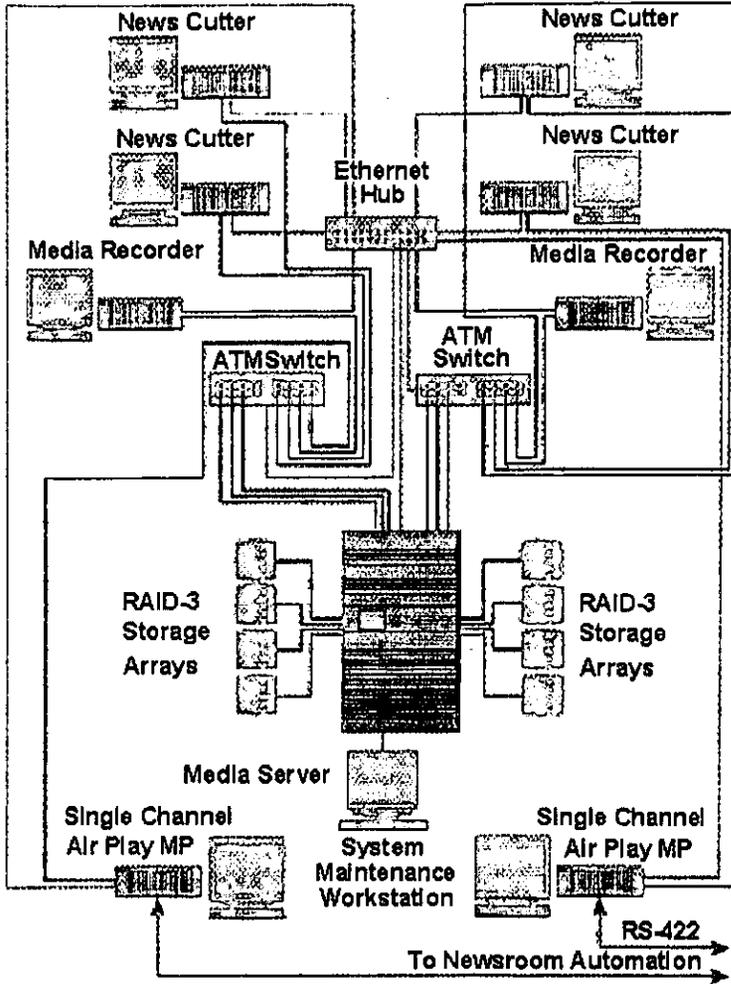


Figura 5.17

En la figura 5.17 se encuentra un diagrama donde se puede ver la implementación completa de estos sistemas. Además se puede observar el orden y la colocación de cada uno de sus elementos los cuales se mencionarán más adelante y así comprender mejor su función.

Los usuarios entran al servidor como clientes desde una estación de trabajo, la fibra óptica de una red ATM basada en un Modo de Transmisión Asíncrona que enlaza cada cliente con el servidor permitiendo acceso al audio, al video y a los gráficos. El material Media nunca se mueve del servidor hacia las clientes, en su lugar el Media es reproducido en tiempo real directamente desde la central de almacenamiento. Además la red ATM, un estándar de Ethernet pasa información de la base de datos entre el servidor de base de datos y los clientes.

Capacidades, se pueden conectar hasta doce usuarios al mismo tiempo en donde, diez pueden reproducir, editar y grabar, mientras los otros dos solo pueden reproducir. En una configuración normal, las estaciones de trabajo funcionan de la siguiente manera, diez Media recorders y Newscuter para grabación y edición, y las dos estaciones que sobran son para sistemas Airplay para transmisión. Cabe mencionar que este es un sistema modular y que depende de las necesidades para conformarlo.

V.7.2 Grabar y reproducir en tiempo real.⁶⁹

La recuperación de material media desde el servidor no necesita de tiempo adicional después de que se requiera reproducir para su salida. El material media se encuentra en forma digital lo que elimina la degradación de la calidad de la imagen.

Grabación directa al servidor. Los usuarios pueden grabar directamente desde las estaciones de trabajo a los discos duros del servidor en donde el Media esta disponible para otras estaciones que se encuentren conectadas a la red. Los archivos de Media están almacenados en formatos OMF Open Media Framework, esto elimina la búsqueda desde las cintas originales de material ahorrando tiempo.

Creación de acceso y compartiendo Bins. Los bins almacenan los nombres de los archivos usados para organizar los clips y las historias, estos incluyen información de la base de datos acerca de cada clip como su nombre y los tiempos código, los usuarios pueden agregar bins o solamente revisarlos en el servidor en cualquier momento. Los bins almacenados en el servidor son copiados a cualquier otra estación de trabajo y se pueden revisar en el servidor o almacenarlos localmente.

⁶⁹ <http://www.avid.com/avidnet/mediarecorder>

V.7.3 Consolidar.

Los usuarios puedan consolidar historias desde una estación de trabajo al servidor y viceversa. La consolidación permite a los usuarios trabajar de manera local y entonces enviar la historia completa y terminada al servidor. Los archivos individuales en los cuales se comprime una historia son copiados desde los discos locales al servidor sin necesidad de regrabar. Una vez que las historias están consolidadas en el servidor, están disponibles para todos los usuarios de la red.

V.7.4 La producción en las estaciones de trabajo.⁷⁰

Los clientes de las estaciones de trabajo como edición (newscutter), airplay y media recorder son los mismos sistemas que se están utilizando actualmente por cadenas como CNN, ITN y CBS. Todas las grabaciones, ediciones y la reproducción del material se lleva a cabo desde estos sistemas. Las estaciones de trabajo están conectadas a la red y al sistema ATN por medio de tarjetas de red llamadas NIC. Esta tarjeta permite mantener una interfase entre el cliente y la ATM permitiendo el mejor manejo de trafico de datos como lo indica la siguiente figura.(figura 5.18)

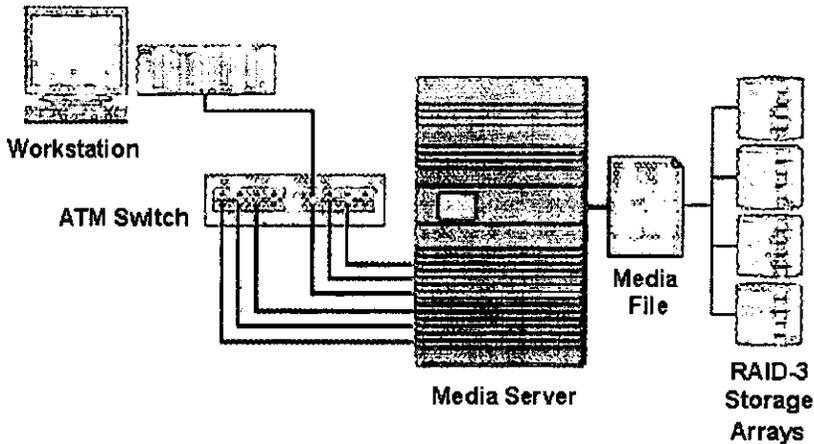


Figura 5.18

El Media recorder es una estación de grabación directa a discos que puede grabar Media desde una alimentación directa de Satélite, cinta u otras fuentes del servidor. Cabe señalar que en estaciones de trabajo de edición como Newscutter se puede grabar también, pero al usar la estación de grabación o Media recorder se puede liberar a la estación de edición (Newscutter) para que solamente realice el trabajo de edición. Las historias o bloque que

⁷⁰ <http://www.avid.com/avidnet/workstation>

fueron terminados en la estación de edición (Newscutter) son transmitidos al aire usando la estación Airplay como se muestra en la siguiente figura. (figura 5.19)



Figura 5.19

Las Televisoras pueden operar el Airplay a través de una interface gráfica para controlar el sistema, además se pueden hacer cambios del orden de las historias que serán transmitidas incluso segundos antes de salir al aire.

Cada estación de trabajo es un dispositivo independiente los cuales están conectados al servidor de Media a través de cableado de fibra óptica. El software de la red esta integrado con las aplicaciones de transmisión y aparece como una opción en pantalla. Este software maneja el trabajo en la red del cliente y permite integrar a la interfaz las características siguientes:

- El usuario puede seleccionar el volumen o espacio en un disco local o en el servidor para operaciones de grabación, render o consolidación. el volumen o espacio del servidor son listados a lo largo junto con los volúmenes locales.
- El menú de opciones permite crear bins en el servidor y desde el servidor, el proceso es similar a abrir, copiar y guardar la copia de un bin local. Los usuarios pueden tener acceso compartido a través de un volumen instalado al inicializar y puede visualizarlo como cualquier otro disco.

Un menú de aplicación de red permite al usuario conectarse o desconectarse del servidor.

V.7.5 Servidor de producción.⁷¹

Como es el centro del sistema, almacena y organiza el material Media usando los discos de arreglos externos RAID3 usando sus propios bins en su disco interno. Algunos servidores están basados en computadoras Silicon Graphics corriendo con sistemas operativos IRIX UNIX. Esto depende de la configuración.

El servidor es responsable de tres funciones principalmente.

- Grabación de Media a el canal de fibra de los discos de arreglos RAID3.

⁷¹ <http://www.avid.com/avidnet/server/>

- Reproducir el material Media desde el RAID3.
- Controlar el acceso de clientes.

V.7.6 Grabación a RAID3.⁷²

Durante la grabación un cliente pasa audio digital y video al servidor, el servidor pasa los datos a un administrador de archivos el cual escribe los datos de uno o más archivos en el volumen de memoria especificado por el cliente. Al mismo tiempo, el cliente crea un Master clip para el archivo en un bin local, el cliente puede usar el nuevo material media tan pronto como el Master clip aparece en el bin que se le asigno. Los otros usuarios pueden compartir el nuevo material tan pronto como el cliente revisa el nuevo clip en el bin dentro del servidor.

V.7.6.1 Reproducción de material desde el RAID3.

Cuando un usuario reproduce un clip o una historia desde una estación de trabajo, el servidor devuelve los archivos de audio y de video desde el almacenamiento central y le corre los datos al cliente. Al recibir los datos, el cliente los convierte en su información de audio y video.

V.7.7 Control de acceso.

Utiliza una combinación de asignación de anchos de banda y niveles de prioridad para determinar cuales clientes de la red pueden usar el servidor. Un usuario no especifica una conexión determinada, simplemente tiene funciones de log in y log out.

Los niveles de prioridad están basados en la importancia del cliente comparada con todos los demás clientes en el sistema. Por ejemplo, el sistema de transmisión al aire Airplay tiene una mayor prioridad que los demás clientes ya que la reproducción al aire no se puede interrumpir.

V.7.8 Base de datos del Servidor.

Es una base de datos poderosa que permite organizar material Media, asociando palabras o keywords con cualquier clip permitiendo al usuario seleccionar frecuentemente la información usada de las listas hechas por las librerías de la estación. El servidor de la base de datos soporta entrada de formas libres de texto y búsquedas en cualquier campo o

⁷² <http://www.avid.com/avidnet/raid/>

combinación de campos. Cada clip puede tener un catalogo múltiple de entradas, permitiendo al clip ser archivado en muchas localidades virtuales.

El administrador del sistema puede acostumbrar fácilmente a la base de datos según sea conveniente para el flujo de trabajo, por ejemplo, una vez conectados los periodistas pueden realizar búsquedas de nuevo material relacionado con alguna palabra clave como "Deportes", el administrador puede almacenar mas de 10,000 registros.

El servidor de Media utiliza un estándar Informix-online. La maquina Informix ofrece una combinación de funcionalidad y fuerza para las demandas de un ambiente de transmisión de T.V. y las especificaciones de manejo de video. El servidor tipo Informix online Dynamic esta diseñado para explotar las capacidades de ambas arquitecturas, multiprocesador y uniprocador para definir la escalabilidad de la base de datos, la manejabilidad y los procesos.

La base de datos de Servidor utiliza una base de datos relacional con tecnología RDBMS (Sistema Manejador de Bases de Datos Relacionales), que soporta Structured Query Language SQL.

Las estaciones de trabajo en la red tienen acceso simultaneo a todos los clips en la biblioteca central, la ventaja del sistema DNG es el manejo de herramientas para tener un manejo de material Media transparente para los usuarios y seguro para la integridad del propio Media dentro del sistema. El material Media es visible en tracks automáticamente una vez que se encuentra registrado dentro de un bin, por ejemplo, dos editores se encuentran trabajando con el mismo material Media para crear dos diferentes piezas. El primer editor termina y borra los clips que no uso, ignorando que el segundo editor sigue usándolo, la ventaja es que el sistema no borra el material mientras se sigue utilizando. El manejador de la estación puede borrar el material borrar los registros del mismo después de que este no haya sido accesado durante cierto periodo de tiempo.

V.7.9 Almacenamiento Central en línea.

Los Bins son almacenados como archivos de UNIX reflejados en discos pares en el servidor. Los usuarios accesan a los archivos del bin a través de las interfaces en sus estaciones de trabajo, El material Media esta almacenado en los arreglos del canal de fibra RAID3. RAID refiere a cualquier arreglo del disco en el cual o los cuales contengan información redundante acerca de los datos en los discos restantes. Cada controlador de disco del RAID maneja la actividad en el arreglo del disco y provee de una interface al servidor, así que cada arreglo del disco aparece como un disco individual.

Los arreglos de disco de RAID son ideales para almacenar Media ya que tienen un manejo independiente que reduce las búsquedas de entrada y salida en el servidor. Algunos sistemas manejan almacenamiento en disco con una técnica llamada striping, en lugar de escribir datos en archivos secuenciales en un solo volumen físico, el servidor parte los datos entre el número de arreglos de los discos que estén configurados como un volumen lógico.

A nivel de servidor, el striping agiliza encontrar las entradas y las salidas I/O desde el momento en que se escribe. Al nivel de drive, striping permite un rápido acceso ya que la mayoría de las cabezas de los discos leen la información en forma paralela, de otra forma un solo disco tendría que regresar a la posición inicial para cada lectura.

V.7.10 El Canal de fibra.

Esta arquitectura utiliza arreglos de tipo Cíprico 7000 los cuales proveen de 80 Mega bytes por segundo en ancho de banda. Cada arreglo de disco puede almacenar ocho horas de material Media y puede conectar discos a más de 10 kilómetros sin necesidad de repetidores. Además permite agregar discos estando en operación.

Los canales de fibra pueden transmitir 100MB por segundo por cada puerto dando un total de 200 MB/s para transmitir datos en cualquier dirección.

V.8 Red ATM.

Los clientes de una estación de trabajo están conectados a el servidor a través de dos redes separadas:

Una para la transferencia de material Media y otra para las transacciones de los bins y la base de datos.

La red de Media utiliza una transferencia asincrónica (Asynchronous Transfer Mode) llamado ATM que se trata de fibra óptica. Al combinar esto se permite tener una velocidad alta, menor problema de acceso de enlace y pérdida de información, ATM es utilizada para transferencias digitales de material de Media ya que su alta velocidad permite reproducir el video en tiempo real mientras otros enlaces o flujo de datos se están realizando sin interrumpirlos. La fibra elimina disturbios que pudieran afectar la calidad de la señal y se puede hacer transferencia a grandes distancias.

La red de base de datos maneja las transferencias de los bins entre los clientes y el servidor usando tecnología tradicional de Ethernet conectada a la base de datos.

V.9 Arquitectura de las estaciones de trabajo dentro de DNG.

La producción de las estaciones de trabajo comparten una arquitectura de hardware común. La imagen del DIAGRAMA 1 (figura 5.17) ilustra el diseño básico.

Cada estación de trabajo consiste en un CPU equipado con tarjetas de captura de video con entradas de video analógico y salidas de conversión Analógico/ Digital y Digital / Analógico, un componente de video digital con entrada y salida de puerto serial(que puede ser opcional al reemplazarlo por la tarjeta análoga de entrada y salida de conversión Analógico/digital), compresión, controlador de discos SCSI para discos locales e interface de red. Contiene también sub sistemas de audio externo que provee la conversión de audio A/D y D/A y entradas y salidas de audio seriales de tipo AES/EBU. El CPU no esta involucrado directamente en la compresión o la descompresión, estas funciones son realizadas por la tarjeta correspondiente.

V.10 Flujo de los datos.

Contando con un servidor central de producción, el sistema está diseñado para la producción múltiple en las estaciones de Trabajo con o sin almacenamiento local de material que están enlazadas al servidor de la biblioteca central de Media.

Las estaciones de trabajo tienen acceso en tiempo real simultaneo al material Media almacenado en los drives del servidor. No se necesita una transferencia de archivos ya que la arquitectura de este sistema hace que los drives del servidor tengan acceso a las estaciones de trabajo.

Las estaciones de trabajo graban directamente el material Media en los drives del servidor. El audio y el video que entran son convertidos a una forma digital es decir son digitalizados, y el video es comprimido por un subsistema de compresión JPEG. El material Media pasa a través del bus de datos del CPU a la tarjeta de ATM (la interfaz de la red). La velocidad de los datos promedia debajo de los 30 megabits por segundo para video con cuatro canales de audio, esto utiliza solo un fracción de 155 megabits por segundo que es una capacidad del ATM. El material Media viaja a través del cable de fibra óptica dentro del switch del ATM para rutearlo al servidor. Los datos entran al servidor a través de la tarjeta de ATM, pasan a través del bus de datos del servidor hacia el controlador de los drives hacia su almacenamiento en el subsistema RAID.

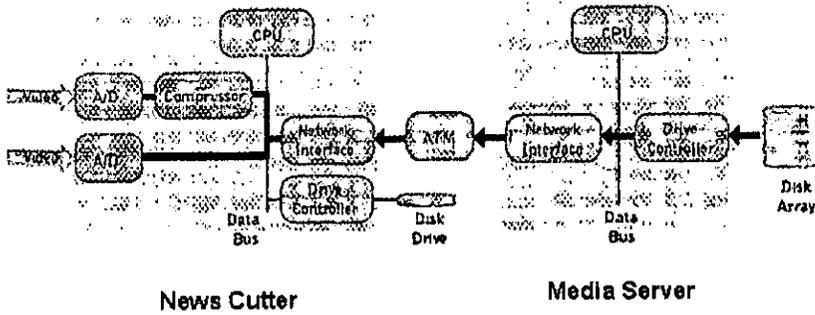


Figura 5.20

Durante la reproducción, el proceso se revierte. El flujo de audio no comprimido y el video comprimido desde el Subsistema RAID de almacenamiento del servidor va hacia el switch y después a la estación de trabajo, en el cual es descomprimido y convertido en una señal analógica de video y audio. Y otra vez la velocidad de los datos promedia por debajo de los 30 Mbps.

La mayoría de las estaciones de trabajo deben de incluir almacenamiento local que consiste en uno o más drives de tipo SCSI. El almacenamiento local es conveniente en dos casos, el primero, cuando un proyecto requiera de seguridad, los usuarios puedan editarlo localmente en un sistema de edición (Newscutter), y transferir la historia final al servidor. Segundo, dos usuarios adicionales pueden editar historias localmente, entonces consolidarlas y enviarlas al servidor cuando, la conexión este disponible.

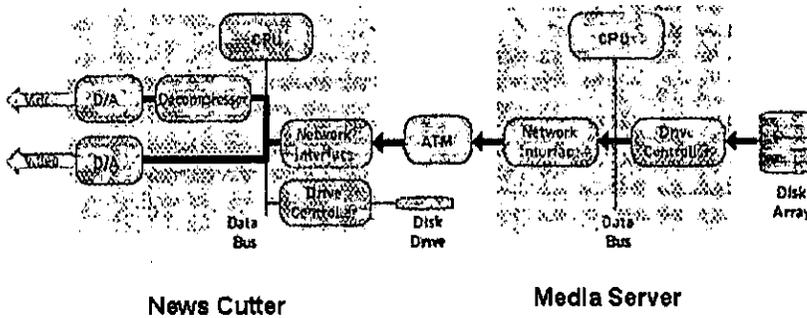


Figura 5.21

El flujo de datos empieza y termina en la estación es trabajo. Las funciones del servidor son de almacenamiento y manejo de archivos permitiendo que se compartan los accesos a las bibliotecas de materia Media. El servidor no juega un papel activo en la reproducción o en la grabación que no sea el de dar acceso al subsistema de almacenamiento.

El servidor no lleva a cabo funciones conversión A/D ni D/A, ni de procesos de compresión todos estos son llevados a cabo en las estaciones de trabajo.

La diferencia entre los dos diagramas anteriores consiste en que el segmento de Media Server o Servidor de Media solo es utilizado para la transferencia entre estaciones de trabajo, almacenamiento y manejo de archivos; se indica que las estaciones de trabajo tipo News Cutter realizan funciones de digitalización, compresión, edición, transmisión de datos entre muchas otras.

V.11 Video Compresión.

Una de las ventajas es la videocompresión, la JPEG (Joint Photographic Expets Group), es un tipo de videocompresión que permite que la velocidad de los datos del video que entra se reduzca aproximadamente 7:1 el cual resulta una velocidad aproximada de 30 Mbps incluyendo cuatro canales de audio.

Un estándar internacional aceptado, es el Intra frame de la compresión JPEG que permite trabajar con sistemas basados en discos para edición. La alta calidad de compresión de imágenes esta combinada con el tiempo de proceso en tiempo real.

V.12 Protección de almacenamiento RAID3.⁷³

El servidor soporta acceso simultaneo en una función multiusuario a el almacenamiento de material Media por el sistema de striping " ancho y profundo " (wide and deep). Est quiere decir, que cada uno de los subsistemas individuales de RAID3 esta partido internamente "profundo" en cuatro drives y con un quinto que es utilizado para paridad. El software controlador de los drives del servidor provee de una segunda dimensión de partición o striping para la extensión de datos " ancho" a través de los subsistemas individuales de almacenamiento RAID.

Este sistema de partición en "ancho y profundo" produce la acumulación de la velocidad de los datos que es buena y realizada por cualquier drive o arreglo. La combinación de alta velocidad de los datos y menor tiempo de acceso produce un sistema de almacenamiento capaz de soportar 12 usuarios con acceso sin restricciones a el material Media. Cualquier usuario puede tomar cualquier clip o cuadro en cualquier momento, es decir, que todos los usuarios pueden utilizar el mismo material o los mismos cuadros al mismo tiempo. Dos usuarios adicionales pueden trabajar con el almacenamiento local con drives de SCSI

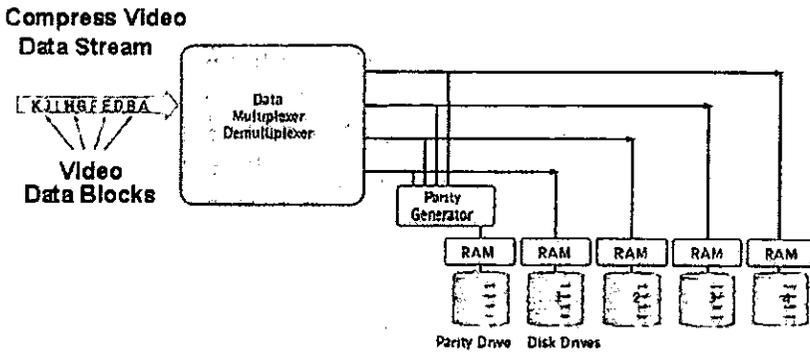
⁷³ <http://www.avid.com/avidnet/raid3>

ligados a el cliente, que se pueden conectar al servidor cuando alguien se desconecte. Ya que son 14 conexiones físicas las que se encuentran, y conectarse es un proceso simple.

El sistema de almacenamiento del servidor esta estructurado en bloques de cuatro sistemas RAID3 que proveen de la suficiente velocidad y poca acumulación de datos requerida para soportar un acceso multiusuario. Cada subsistema RAID3 recibe los datos a través de el canal de fibra que es un bus de datos.

Durante el proceso de grabado, el controlador del sistema RAID, que es un sofisticado multiplexor y demultiplexor, separa los datos que entran en cuatro partes las cuales están grabadas por separado en cuatro drives. Cada parte contiene determinado flujo de datos y es alimentado a una velocidad muy grande en un buffer de memoria RAM por su drive asociado. El buffer RAM permite acumular altas velocidades en el arreglo mientras permite a cada drive individual recibir los datos en su propia velocidad. El controlador puede entonces alimentar los datos al siguiente drive en el arreglo.

Una quinta parte del flujo de datos, los datos de paridad, es derivada por el generador de paridad de los datos contenidos en los cuatro segmentos principales y almacenada en el quinto disco. La paridad de los datos es utilizada para reconstruir cada uno de los dato si estos se llegaran a perder por alguna falla en alguno de los cuatro discos.



RAID-3 Storage Array System

Figura 5.22

De los cuatro segmentos de datos, es decir, de los acarreo de los números digitales representan muestras de el video o el audio, son sumados para crear un numero de paridad.

Arriba en la figura 5.22 se muestra el tipo de arquitectura entre los discos de almacenamiento en donde son almacenados por partes los clips digitalizados así como los segmentos o bloques finales ya editados.

Por ejemplo, si el valor de los datos en las cuatro partes fuera 7, 19, 5 y 12 la suma sería 43. Este dato es almacenado en el drive de paridad. Si el controlador detecta alguna pérdida de cualquiera de los cuatro drives de datos durante la reproducción, él utiliza el dato de paridad para reconstruir el valor faltante. Por ejemplo, si el drive tres se perdió el dato 5 estaría perdido. De cualquier forma el número de paridad es 43 y fue sacado de la suma de $7 + 19 + x + 12$. Despejando la ecuación se puede encontrar el valor $43 - 7 - 19 - 12 = x$. El valor faltante es 5 y se reconstruye fácilmente y la reproducción continúa sin interrupciones o errores de datos.

Estos sistemas RAID están equipados con características de reconstrucción automática, el controlador del drive puede reconstruir los datos del drive dañado por un repuesto en el background. Los tiempos de reconstrucción son menores a una hora.

Los RAID tienen su propio no brake y fuente de poder para mantener el sistema funcionando si algún otro componente falla.

El controlador suena una alarma en el momento de detectar cualquier falla en el sistema.

V.13 Interfaces del sistema.

Audio y video. Las estaciones de trabajo utilizan los estándares en los formatos de entrada y salida de audio y video. El video digital y analógico y las referencias de las señales están conectadas por conectores BNC y el audio por conectores de tres pins XLRs (conocidos como canon). El video I/O es swichable entre los componentes analógicos o componentes Y/R - Y/B - Y. Los componentes seriales tienen también adaptadores de salida analógica para propósitos de monitoreo. El audio digital utiliza AES/EBU y conexiones XLR.

Los OMF (Open Media Framework Interchange), son archivos que permiten intercambiar material de video y audio entre muchos sistemas ya que este formato puede ser leído por otros productos.

La manipulación de este tipo de información digital en cualquier tipo de formatos a través de este tipo de sistemas establece una propuesta muy interesante para toda la industria dedicada a la comercialización de video y la producción de programas de televisión, con este tipo de herramientas el tiempo y el costo de producción de este tipo de trabajos han bajado considerablemente en comparación con el antiguo método lineal de las máquinas analógicas que para realizar desde simples ediciones a corte directo hasta efectos complicados pudiesen acumular por lo menos entre 8 y 24 horas de trabajo con este tipo de sistemas la reducción de tiempo puede ser del 90% terminando el trabajo en un promedio de 45 minutos dependiendo de la complejidad de cada pieza de edición, a su vez la utilización de recursos materiales y humanos es aprovechada al máximo, considerando que

se puede dar un entrenamiento relativamente corto para poder obtener el mejor provecho a estos sistemas.

Para poder llegar a desarrollos tecnológicos como éste fue necesario comprender el origen de cada uno de los elementos que se han mencionado a través de estos capítulos. La representación clara de las ideas muchas veces se ve entorpecida al faltar algún concepto dentro de una explicación, por ese motivo es necesario obtener el conocimiento de los puntos más esenciales y básicos del que debe estar compuesto el sistema. Se puede deducir a simple vista que un sistema de este tipo pretende emular el funcionamiento del sistema visual humano al representar en la pantalla de televisión las imágenes que son captadas por el ojo humano y que el hombre en su afán por conservar esos recuerdos ha desarrollado su inteligencia para lograr que maquinas tales como la cámara de video logren mantener esas imágenes, aún más allá no basta con poder reproducirlas y difundirlas el hombre en un reto más ambicioso logra que su trabajo pueda representar esas imágenes como su creatividad se lo indique creando más tecnología que nos permite ahora verlas a través de una computadora haciendo conversiones analógicas – digitales y manipulándolas a su antojo con un esfuerzo cada vez menor.

CONCLUSIONES

CONCLUSIONES

Los sistemas de edición utilizados actualmente proveen de gran versatilidad a producciones diversas de televisión y publicidad. A través de la revisión de los procesos por los que atraviesa una imagen hasta el momento de ser visualizada en un monitor u pantalla de televisión se ha podido comprobar que no se trata de un sistema fácil de emular, ya que comprende desde los más sencillos elementos encontrados en la naturaleza como es el caso de la luz hasta el desarrollo de una nueva tecnología que aún se encuentra en desarrollo al intentar mejorar cada día.

La utilización de este tipo de sistemas en los últimos años por parte de la industria de Televisión ha crecido exponencialmente debido al ahorro de recursos, estos sistemas empezaron siendo muy caros, pero al demostrar su utilidad el costo se pagó por sí solo. La reducción del tiempo de trabajo y el aumento en la calidad de la imagen televisiva, disminuyeron los costos y aumentaron la productividad de las empresas. Actualmente no existe alguna televisora en el país que no cuente con un sistema con estas características, la maniobrabilidad de los mismos es tan grande que se puede tener desde un equipo de control remoto siendo operado desde una unidad móvil en los lugares más inhóspitos hasta una sala completa de post-producción en un set de televisión con la capacidad de elaborar piezas de noticias en unos cuantos minutos y transmitir las vía satélite desde la misma computadora.

Como un valor agregado el aprendizaje del uso de este tipo de sistemas es mucho más sencillo de entender ya que permite a los usuarios el manejo de materiales de producción desde el primer instante en que se encuentren frente a la computadora. En comparación con los sistemas analógicos y su gran extensión de tiempo para realizar cualquier edición aún siendo muy sencilla es ampliamente superada por cualquiera de estos sistemas teniendo hasta una reducción del 70% en tiempo, energía y desgaste humano.

Aún así con todas las ventajas que estos sistemas poseen, no debemos permitir que la tecnología analógica desaparezca, en su lugar se debe de combinar ambas. De esta manera aseguraremos la máxima calidad en la producción de este tipo de materiales.

Además se debe considerar que aunque ya existen en el mercado muchos sistemas aún es costoso mantenerlos, y los que pueden ser accedidos por personas que no se desarrollan profesionalmente en el medio son medianamente caros, por su puesto dependiendo el uso y la calidad que se pretenda obtener.

La importancia de no olvidar de donde se originan los procesos relacionados, el desarrollo de los mismos, los resultados obtenidos y la combinación de todos ellos es el impulsarnos a tratar cada día de mejorar aún más en la investigación y desarrollo de nuevos sistemas, ya que los límites nos los ponemos cada uno de nosotros.

De esta forma el presente trabajo pretendió dar a conocer el fondo de esta tecnología que muchas veces no tenemos la oportunidad de conocerla desde sus principios básicos, muchos de los cuales no los pudimos haber imaginado contenidos en el resultado final.

De la misma manera el compilar esta información derivada de muchas ciencias puede reducir el trabajo de investigación de diferentes áreas pudiendo tomar de un mismo texto diferentes perfiles de las mismas, así como facilitar el entendimiento de los procesos que se llevan a cabo de principio a fin, tanto como para personas relacionadas con términos de ingeniería como para personas relacionadas con la comunicación pudiendo representar un apoyo para todos ellos y quizá ser considerado como un texto alterno.

GLOSARIO DE TERMINOS.

GLOSARIO DE TERMINOS.

A

AC-3

Sistema de sonido envolvente digital, extrae 5.1 canales de una fuente láser disc o dvd, estos canales son surround derecho, frontal derecho, central, frontal izquierdo, surround izquierdo y subwoofer.

Acústica

Es la ciencia que estudia el sonido, su transmisión, efectos, producción, etc.

AES/EBU

Conexión digital balanceada, usa conectores del tipo XLR.

ADC

Convertor analógico a digital, convierte una señal analógica continua en un chorro digital de bits

AGC

Automatic Gain Control, control automático de ganancia, circuito que mantiene constante el nivel de la señal de luminancia.

ALC

Automatic Level Control, control automático de nivel, se usa en audio para mantener constante el nivel de volumen de sonido.

Alta fidelidad (HiFi)

Termino usado para describir la capacidad de un sistema de reproducir en sonido con un alto grado de realismo

Altavoz

Usado muchas veces para referirnos en realidad a un baffle. Transductor que convierte energía eléctrica en energía acústica, convierte la señal eléctrica proveniente de un sistema sonoro en sonidos.

Altavoz coaxial

Sistema que comprende dos altavoces generalmente montados sobre el mismo eje de radiación, suelen ser un altavoz de graves y otro de agudos.

Altavoz de agudos(Tweeter)

Altavoz especialmente diseñado para reproducir altas frecuencias.

Altavoz de graves(Woofers)

Altavoz especialmente diseñado para reproducir bajas frecuencias.

Altavoz electrodinámico

Es el altavoz más común hoy en día, formado por un imán y una bobina móvil montada generalmente en un cono móvil.

AM

Modulación de amplitud, tipo de transmisión usado en las bandas estándar de radiodifusión.

Amperio

Medida de la intensidad de la corriente eléctrica.

Amplificador (audio)

Circuito eléctrico para amplificar señales del espectro sonoro, se usa para aumentar el nivel de señal de una fuente y poder alimentar altavoces.

Analógico

Señal continua con muchos posibles valores aleatorios.

Ancho de banda

El margen de frecuencias que hay en una banda desde la más alta a la más baja.

Antena

Dispositivo usado para enviar y/o recibir ondas electromagnéticas.

Archivo Digital de Media

Es un archivo en gran escala capaz de almacenar video y audio pudiendo ser de diferentes formatos.

Armónicos

Las ondas que son múltiplos de una determinada frecuencia fundamental y cuyas intensidades suelen ser menores que la frecuencia fundamental.

Atenuador

Circuito que se usa para reducir los niveles de las señales.

ATM (Asynchronous Transfer Mode)

Protocolo de transmisión en red de alta velocidad de video y datos, capaz de conectar estaciones de trabajo y servidores.

Azimuth

Es la inclinación que tienen los cabezales de audio o video con respecto a la dirección de desplazamiento de la cinta.

A/V

Audio/Video, abreviatura usada para referirse a sistemas de audio y video relacionados con el home theater o cine en casa.

B

Baffle (Altoparlante)

Caja donde suelen estar alojados varios altavoces así como su filtro correspondiente.

Balance (audio)

Control usado para equilibrar los niveles entre canales musicales

Banda

Rango de frecuencias entre dos límites definidos

Beta

Formato de grabación/reproducción de video inventado por Sony y actualmente en desuso.

Biamplicación

Utilización de varios amplificadores en equipos de audio para cada una de la banda (graves, agudos).

Bi-cableado

Conexión a los altavoces mediante cables independiente para cada tipo de altavoz.

Bias

Corriente de polarización, para que en una cintase puedan grabar pasajes débiles y fuertes hay que tener en cuenta el ciclo de magnetización de los materiales que hace que no se magnetice el material hasta que sobrepasa un umbral mínimo, por tanto, en las grabadoras analógicas actuales se graba un tono de muy alta frecuencia (suele ser 100Khz) no audible pero lo suficiente para que los tonos débiles se graben ya que con este tono se sobrepasa el umbral de magnetización.

Binario

Sistema de representación numérico que usa el 0 y el 1 como elementos.

Bit

Unidad elemental de información digital.

Blanking Horizontal

El tiempo que tarda un haz de electrones en un dispositivo de salida en apagarse y colocarse en el filo izquierdo de la siguiente línea de escaneo.

Blindaje magnético

Sistema que tienen los altavoces de A/V para poder situarse cerca de pantallas de televisión sin producir distorsiones de imagen.

Bobina (audio)

Cableado especial en espiral, usado generalmente en el interior de los altavoces y en circuitos de radio.

BNMC (Broadcast News Machine Control).

Prepara los recursos de la maquina para transmisión o alimentación de las secuencias en orden, lleva el reloj master de transmisión.

C

Cabeza (o cabezal)

Componente electrónico destinado a grabar o leer en cintas o material magnético, convierte magnetismo en energía eléctrica y viceversa.

Cadena HIFI

Conjunto de elementos destinado a la grabación y reproducción de sonido, generalmente de uso domestico y con mando a distancia.

Campo magnetico

Area que se ve afectada por perturbaciones magnéticas, generalmente el creado por imanes y en las proximidades a estos.

Cassette

Cinta magnética para grabar y reproducir audio.

CATV

Televisión por cable, a veces también televisión por antena comunitaria.

Codificador/Decodificador

Es un dispositivo que convierte los componentes de color a bits y viceversa.

Componente de Color

Sistema de video que combina tres señales independientemente como RGB, YCbCr

Campo

La mitad de líneas horizontales necesarias para representar un cuadro de televisión. El primer campo de un cuadro contiene todos las líneas pares y la segunda mitad contiene las nones.

Cuadro

Es una imagen completa. En formato de televisión un cuadro se compone de dos campos.

CD

Compact disc, termino usado para designar tanto a los discos compactos digitales como a sus reproductores.

CD-R

Compact disc gravable

Cinta magnética

Cinta plástica de varias capas una de las cuales esta impregnada con partículas magnetizables, su usa generalmente en audio para grabación de música.

Circuito Integrado

Circuito electrónico en miniatura construido sobre un soporte de silicio y que viene generalmente en un encapsulado negro con patillas de metal

Clipping

Distorsión que consiste en que una onda es recortada por la parte superior y por la parte inferior, de tal modo que si fuese una señal sinusoidal, las partes superiores e inferiores de la onda serian planas en vez de curvas

Coherencia

Termino que se suele usar para referirse a lo bien que suena un equipo

Coloración

Termino que se suele usar para expresar el efecto audible de ciertas frecuencias cuyos niveles se ven aumentados artificialmente.

Conexión S-Video (video separado)

Conector que une cámaras, televisores y equipos de video manteniendo separadas las señales de luminancia (brillo) y crominancia (color).

Cono

Es la forma que tienen generalmente los altavoces de medios y graves.

Crominancia

Es la señal que lleva la información del color para poder formar una imagen de video

Crossover

Dispositivo que divide las frecuencias en diferentes márgenes, puede ser pasivo (usa resistencias, condensadores y bobinas) en cuyo caso suele ir conectado a altavoces, o activo (usa circuitos integrados, transistores, etc.) en el caso que divida frecuencias para ser amplificadas por separado.

Conteo lineal

Es el número total de líneas horizontales en una imagen.

D

DAB

Digital Audio Broadcast, radiodifusión digital

DAC

Convertor digital/analógico, a veces esta separado en otro elemento de la cadena de sonido entre el transporte de CD y el preamplificador.

DAT

Digital Audio Tape, cinta digital de audio de alta calidad.

DBX

Sistema de reducción de sonido para cintas magnéticas por compresión/expansión.

DCC

Sistema de grabación/reproducción con cassette digital, desarrollado por Philips

Decibelio (db)

La décima parte de un bel, expresa siempre una relación de potencias, intensidades y se suele usar para ver la amplificación o atenuación.

- 15-25 dB Un susurro
- 40-60 dB Ruido de fondo en casa o la oficina
- 65-70 dB Nivel normal de la voz
- 100 dB Algunos picos en una orquesta
- 120 dB Y mas es lo que sueles oír en un concierto rock
- 130 dB es el umbral del dolor
- 140 dB En adelante es el ruido de los aviones a reacción.

Decodificador

Sistema para convertir una señal cifrada en un código determinado a otro determinado.

Digital

Sistema que usa muestras digitales (valores discretos codificados en binario) para representar señales analógicas.

Digitalización

Proceso de conversión del campo analógico al digital

Disco Láser(LD)

Disco (y su reproductor de discos láser) que contiene información de video y audio

Disipador

Pieza que va asociada a un componente para liberar el calor producido por este.

Distorsión

Adiciones de armónicos no deseados a la señal original. Cualquier cosa que altere la señal original.

DSP

Digital Sound Processor, procesador digital de sonido, sistema que recrea digitalmente las características acústicas de diferentes recintos, efectos, etc.

DVD

Digital Versátil Disc (o Digital Video Disc), sistema basado en CD con capacidad para almacenar sonido y video.

DV

Digital Video, formato de video digital.

E

ECC

Sistema de corrección de errores durante las transferencias de audio digital.

Ecuallizador

Dispositivo electrónico para variar los niveles de diversos márgenes de frecuencias aumentándolas o disminuyéndolas

Espectro

Se suele denominar a un margen de frecuencias determinado.

Estéreo

Sistema de reproducción de sonido mediante dos altavoces con informaciones diferentes.

E

Fibra óptica

Cable de fibra transparente por donde se transmite información en forma de luz

Filtro

Circuito electrónico o eléctrico usado para limitar ciertas frecuencias en una señal. Filtro paso-bajo (Low-Pass) deja pasar las bajas y atenúa las altas frecuencias, el inverso en el filtro paso-alto (High-Pass) que deja pasar las altas y atenúa las bajas.

FM

Frecuencia modulada, modulación de frecuencia, usada para transmitir sonido, para registro de video, etc.

Frecuencia

Numero de ciclos por unidad de tiempo de una onda sonora. Se mide en Hz (Herzios). Un Herzio es un ciclo por segundo). La respuesta en frecuencia en las personas suele ir de 20 a 20.000 Hz.

G

Ganancia

Amplificación de señal.

GHz

Gigahertzio, equivale a un billón de ciclos por segundo.

H

HD CD

Compact Disc de Alta Definición. Requiere una codificación especial en el proceso de grabación. Alguna gente dice que tienen mejor sonido, aunque se requiere un lector CD especial.

HDTV

Televisión de alta definición.

Hertzio

Unidad de medida para la frecuencia que mide el numero de oscilaciones por segundo de una onda.

Hi 8

Versión de mayor resolución del formato de video 8mm

Hi-Fi

Alta fidelidad

I

Impedancia

Resistencia que ofrece un elemento al paso de la corriente. Los altavoces suelen ser de 4 a 8 ohmios (aunque es variable con la frecuencia y podría ir de 4 a 60 Ohm)

Infrarrojo

Radiación electromagnética que se suele usar para transmitir información.

IRE.

Una unidad IRE equivale a 0.714V (IRE= Instituto de Ingenieros de la Radio, hoy llamada IEEE).

J

K

KHz Kilo hertzio

Equivale a mil oscilaciones por segundo.

L

Láser

Dispositivo que emite un haz luminoso.

Led

Diodo que emite radiación luminosa.

Longitud de onda

Distancia entre picos y valles consecutivos en ondas periódicas.

LP

Larga duración, se suele usar en los videos y se refiere a la característica de duplicar la duración de una cinta al usar una velocidad mas lenta en grabación/reproducción.

Luminancia

Señal que lleva que lleva la intensidad de la luz y que hace que se vea la TV en blanco y negro o en color si se combina con la crominancia

M

Micrófono

Dispositivo que convierte señales acústicas en eléctricas

MHz

Megahertzio, equivale a un millón de ciclos por segundo

Mono

Opuesto a estéreo, información de audio a través de un único canal.

MPEG

Motion Picture Experts Group, Grupo de creadores de normas para desarrollar estándares de compresión de video.

N

NTSC

Norma de televisión americana, con 526 líneas de barrido.

O

8 mm

Formato de grabación de video en un casete con una cinta de 8mm

Ohmio

Unidad de medida de la resistencia de un circuito electrónico al paso de la corriente eléctrica.

P

PAL

Sistema de transmisión de señal de televisión en color, significa alternancia de fase por línea.

Pico

Valor máximo en una señal

Potenciometro

Resistencia variable que se usa para graduar intensidad de corriente.

Q

Quadrafónico

Sistema estéreo de cuatro canales

R

Rango de frecuencias

Margen de frecuencias que es capaz de reproducir con respuesta lineal plana un sistema.

RCA

En audio, se suelen denominar a un tipo estándar de conectores. Son por ejemplo los que van de un lector CD al amplificador, etc.

Resolución de video

Se usa para definir el máximo detalle que se puede obtener en una imagen, por ejemplo, en VHS se usan 240 líneas mientras que SuperVHS y láser disc se consiguen 400 y más.

RGB

Red, green, blue, se refiere a entradas, cables, conectores que llevan la información de cada uno de los colores, rojo, verde y azul.

Ruido

Interferencias o señales no deseadas que existen en señales de audio, video, etc.

S

SCART

Euroconector) Conector de múltiples patillas con información de Audio y Video. Literalmente es sindicato de constructores de aparatos radioreceptores y televisores.

SECAM

Sistema de transmisión de señal de televisión, usado en Francia y en algunos otros países.

Señal

La información de audio o video generada por una fuente que puede ser una emisión de radio o televisión, o una cinta o un CD, etc.

Servidor de Vista (Browse Server)

Es una computadora que revisa el video en resoluciones inferiores cuyo propósito es seleccionar las imágenes que serán utilizadas en la edición.

Sistema Aditivo de Color.

Sistema de especificación de color donde los colores primarios se unen para crear un color determinado. Por ejemplo, RGB (red/green/blue). Estos sistemas están generalmente asociados con la luz emitida por los CRT.

Scanning Standard

Es el parámetro asociado con el Raster de escaneo para un monitor, cámara o grabadora de video, denotado por el conteo lineal, frecuencia de campo y el entrelazado.

SMPTE

Society of Motion Picture and Television Engineers.

Sistema Substractivo de Color

Sistema donde los colores primarios son sustraídos de una referencia de color para obtener un color específico. Por ejemplo, cyan/magenta/yellow (CMY) y luminance/red - luminance/blue - luminance (Y,R-Y,B-Y).

S-VHS(Super VHS)

Sistema VHS de mayor resolución que el VHS

I

True Color

Esquema de relación de color referido a 16-bit or 24-bits

THX

Estándar de sonido de alta calidad para cines y para entorno domestico, bajo licencia de George Lucas (Star Wars)

Timbre

Cualidad especifica del sonido y sus armónicos que define cada instrumento, voz, etc.

Tono

Cualidad especifica del sonido que distingue entre agudos y graves.

Transductor

Dispositivo que convierte información en una energía determinada a otra energía diferente.

Transistor

Componente electrónico activo que actúa amplificando una corriente entre dos electrodos mediante otro electrodo de control.

Transporte

Sistema de arrastre, puede ser para una cinta o también para extraer el chorro de bits de un Compact disc

U

UHF

Ultra High Frequency, banda de frecuencias entre 300 y 3000 Mhz

V

VCR

Reproductor y/o grabador de video

Videocasete

Cinta de video alojada en una caja característica.

VHS

Formato de grabación/reproducción de video

Video In

Entrada de señal de video

Video Out

Salida de señal de video

Voltio

Unidad de medida de la diferencia de potencial

Vertical Blanking

Tiempo que tarda el haz de electrons en un dispositivo de salida en posicionarse en el filo superior izquierdo de la pantalla.

W

Watto

Unidad de medida de potencia.

Woofers

Altavoz de graves, diseñado para reproducir bajas frecuencias

X

Y

Y-Adapter

Adaptador en Y que divide una señal en dos, se usa para conectar dos conectores a una única salida.

BIBLIOGRAFIA.

BIBLIOGRAFIA

1. Tratamiento Digital de Imágenes, Alberto Domingo Ajenjo, Anaya Multimedia 1994, Serie Guía Monográfica.
2. Practical Image Processing Inc., Craig A. Liindley, Ed. Jhon Wiley & Sons 1991, Estados Unidos.
3. Fundamentos de Televisión, Otto Cimann, Ed. Marcombo Boixareu 1989, España.
4. La Televisión una Metodología para su aprendizaje, Llorenc Suter, Ed. G. Gili S.A.1999.
5. HDTV Advanced Television for the 90's, K. Blair Benson and Donald G. Fink, ED Mc Graw Hill 1991. Estados Unidos.
6. HDTV High Definition Television, Stann Pretiss, Ed. Tab Brooks 1994. Estados Unidos.
7. Fundamentals of Digital Image Processing, Anil K.Jain, Ed. Pentice Hall 1990.
8. Image Tecnology, Jorge L. Sans, Ed. Springer 1996. Estados Unidos.
9. Fundamentos Tecnológicos de Video y Televisión, Vicente Llorens, Ed. Paidos 1995. España.
10. La imagen, comunicación funcional, Abraham A Moles, Ed Trillas 1991.
11. Televisión, teoría y practica, Jorge E. González, Ed. Treviño 1989.
12. Manual de servicio Sony, Sony México 1995.
13. Anatomía Cromodinámica, Wynn Kapit, Ed. Fernández Editores1990.
14. Anatomía Descriptiva, Dr. J. A. Fort, Ed. Gustavo Gili 2ª edición Francesa 1995.
15. El Cuerpo Humano, sus partes y funcionamiento, Mitchell Wilson, Organización Editorial Novaro 1982.
16. Introducción a la Psicología, Linda Davidoff, Ed. Mc Graw Hill 1990.

Direcciones de Internet.

A. Sistema de edición no Lineal AVID
<http://www.avid.com/avidnet/raid3>

B. Fundamentos de Luz y Color
http://wwwetsi2.urg.es/depar/ccia/public_html/robotica/indice_apuntes.html

C. Almacenamiento de Información.
<http://titan.pue.udlap.mx/15097380/jair/SyQuest>

D. Televisión y Video.
<http://www.udec.cl.investigación/apuntes/video>