

92
2 es.



UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE INGENIERIA

DISEÑO E IMPLANTACION DE UNA TARJETA PARA EL PROCESAMIENTO DE VIDEO DIGITAL

T E S I S

QUE PARA OBTENER EL TITULO DE:

**INGENIERO MECANICO ELECTRICISTA
AREA ELECTRICA ELECTRONICA**

P R E S E N T A N :

**ANGEL GREGORIO LABASTIDA ALVARADO
JUAN RENE LOPEZ ESQUEDA**

DIRECTOR DE TESIS: DR. VICTOR GARCIA GARDUÑO



MEXICO, D. F.

1998

**TESIS CON
FALLA DE ORIGEN**

263764



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.

**DISEÑO E IMPLANTACIÓN DE
UNA TARJETA PARA EL
PROCESAMIENTO DE VÍDEO
DIGITAL**

AGRADECIMIENTOS

A la Universidad Nacional Autónoma de México

A la Facultad de Ingeniería

Al Departamento de Ingeniería en Telecomunicaciones

Al Laboratorio de Multimedia del Departamento de Ingeniería en Computación

y en particular al Dr. Víctor García Garduño

por su apoyo incondicional durante nuestra formación como universitarios.

Angel Labastida
Rene López

ÍNDICE	i
INTRODUCCIÓN GENERAL	v
PARTE I ESTADO DEL ARTE	
CAPÍTULO 1 VÍDEO ANALÓGICO	
1.1 Generalidades del video analógico	3
1.2 Señales que componen el vídeo analógico	5
1.3 Compatibilidad en los sistemas de vídeo analógico	6
1.4 Codificación en color	7
1.5 Sistemas de vídeo en color	10
1.5.1 Formato NTSC	10
1.5.1.1 Corrección Gamma	13
1.5.1.2 Manejo de la información en color	14
1.5.2 Formato PAL	15
1.5.3 Formato SECAM	17
1.5.4 Diferencias y similitudes entre las tres normas	17
1.5.5 Sistemas HDTV	19
CAPÍTULO 2 VÍDEO DIGITAL	
2.1 Introducción	22
2.2 Aspectos de la digitalización	22
2.2.1 Los pixeles	24
2.2.2 Exclusión de sincronía y blanking	24
2.2.3 Bits por pixel	24
2.3 Métodos de codificación en color	24
2.4 Bitmap	25
2.4.1 Lanzamiento de línea o Pitch	26
2.4.2 Mapeo del color	26
2.5 Proceso de digitalización de las imágenes	27
2.5.1 Muestreo y cuantización	27
2.5.1.1 Limitaciones del muestreo	29
2.5.1.2 Limitaciones de cuantización	31
2.5.2 Conversión de vídeo A/D	31
2.5.2.1 El convertidor Analógico-Digital	31
2.5.2.2 Conversión D/A	32
2.5.2.3 Efectos de vídeo digital	32
2.5.2.3.1 Pixelización	32
2.5.2.3.2 Contorneo	33
2.5.2.3.3 Aliasing	34

2.6 ¿ Por qué hardware?	35
2.6.1 Requerimientos de hardware para el procesamiento digital de vídeo	35
CAPÍTULO 3 CÁMARAS DE VÍDEO	
3.1 Introducción	38
3.2 Sensores de las cámaras o visualizadores	38
3.3 Dispositivos de carga acoplada	40
3.4 Arquitecturas CCD	42
3.4.1 Arquitectura de transferencia de cuadro	42
3.4.2 Arquitectura de transferencia interlínea	43
3.4.3 Arquitectura de transferencia cuadro-interlínea	43
3.5 Escaneo entrelazado en las CCD's	44
3.6 Offset espacial	45
3.7 Cámaras de sensor sencillo	45
3.8 Colorimetría de las cámaras	46
3.9 Relación Señal a Ruido CCD	46
3.10 Sistema óptico de las cámaras	47
3.10.1 Lentes	48
3.10.2 Apertura y número f	49
3.10.3 Profundidad del campo	49
3.10.4 Aberraciones	50
3.10.5 Filtros ópticos	51
3.10.6 Separación del color	51
3.11 Sistemas de vídeo básicos	52
3.11.1 Definiciones	52
3.11.2 Formato CCD	52
3.11.3 Sensibilidad de una cámara CCD	52
3.11.4 Formatos de señales	53
3.11.5 Control de Ganancia Automático (AGC)	53
CAPÍTULO 4 PROCESAMIENTO DIGITAL DE IMÁGENES	
4.1 Revisión	55
4.1.1 Introducción	55
4.1.2 Gráficos e imágenes animadas por computadora	55
4.1.3 Comunicación por imágenes	55
4.1.4 Tratamiento de las imágenes	56
4.2 Morfología Matemática	57
4.2.1 Dilatación y Erosión	57
4.2.1.1 Dilatación	58
4.2.1.2 Erosión	58
4.2.2 Opening y Closing	59
4.2.3 Transformada de Serra	61
4.2.4 Algunos algoritmos morfológico básicos	62
4.2.4.1 Extracción de límites o contornos	62

4.2.4.2 Llenado de regiones	63
4.2.4.3 Extracción de componentes	64
4.2.4.4 Corteza convexa	64
4.2.4.5 Adelgazamiento	66
4.2.4.6 Engrosado o espesado	67
4.2.4.7 Esqueletos	67
4.2.4.8 Recortado o podado	68
4.3 Compresión de imágenes	70
4.3.1 Técnicas de compresión de imágenes	70
4.3.1.1 Píxeles independientes y dependientes	70
4.3.1.2 La norma JPEG	71
4.3.1.3 La recomendación H.261	73
4.3.1.4 MPEG-I	74
4.3.1.5 MPEG-II	75
PARTE II DISEÑO DE LA TARJETA	
CAPÍTULO 1 INTRODUCCIÓN	
1.1 Introducción	78
CAPÍTULO 2 HARDWARE	
2.1 Generalidades	80
2.2 Procesamiento analógico	81
2.2.1 Amplificador de entrada diferencial	81
2.2.2 El separador de sincronía, LM1881	84
2.2.3 Referencia de negro a tierra	87
2.2.3.1 Circuito de retención y muestreo, LF398	88
2.2.3.2 Filtro paso bajas	88
2.2.3.3 Circuito completo	89
2.2.4 Rectificador de precisión	89
2.3 Digitalización	91
2.3.1 El convertidor Analógico-Digital	91
2.3.2 Oscilador de cristal	94
2.3.3 Búfer de memoria RAM estática	96
2.3.4 Contadores de direccionamiento	97
2.3.4.1 El diferenciador	98
2.4 Adquisición	100
2.4.1 El bus ISA	100
2.4.2 El controlador de puertos 8255	102
2.4.2.1 Manejo de puertos	104
2.4.3 El DMA	108
2.4.3.1 El 8237	108
2.4.3.2 Funcionamiento del ciclo DMA	110
2.4.3.3 Programación del DMA	113
2.4.3.3.1 Canal de transferencia	113

2.4.3.3.2 Modo de transferencia	113
2.4.3.3.3 El flip-flop DMA	114
2.4.3.3.4 Dirección del bloque de datos	114
2.4.3.3.5 Tamaño de la transferencia	114
2.4.3.3.6 Liberación del canal	115
2.5 Funcionamiento general de la tarjeta	115
CAPÍTULO 3 SOFTWARE DE OPERACIÓN	
3.1 Generalidades	119
3.2 Captura	122
3.3 Adquisición	122
3.4 Visualización	123
PARTE III RESULTADOS Y CONCLUSIONES	
CAPÍTULO 1 RESULTADOS	
1.1 Procesamiento implementado para la presentación de la tarjeta	126
1.2 Procesamiento sobre cuadros sencillos	126
1.2.1 Compresión de imágenes fijas	126
1.2.1.1 Filtro de mediana	131
1.3 Procesamiento sobre cuadros en secuencia	133
1.3.1 Erosión y Dilatación	133
1.3.2 Programa del contorno	135
1.3.3 Programa de la flecha	137
1.3.4 Consideraciones	140
CAPÍTULO 2 CONCLUSIONES	
2.1 Conclusiones	142
APÉNDICE A SISTEMAS COMERCIALES	
A.1 VFG 512-8	147
A.2 CG-7	148
A.3 MOTION PICTURE	149
A.4 AV2001	151
A.5 ULTRA II	152
A.6 RT300	154
BIBLIOGRAFÍA	158

Diseño e implantación de una tarjeta para el procesamiento de vídeo digital

Introducción general

El procesamiento digital de imágenes es un campo que ha tomado mucho interés en los últimos años, en particular el procesamiento de imágenes en movimiento, es decir, el vídeo digital. La correcta representación digital de la señal de vídeo analógico, así como su transmisión son de los aspectos más importantes a considerar en el campo de vídeo digital.

Este proyecto es un trabajo de investigación, diseño e implementación que tiene como propósito ser una herramienta base para nuevos diseños y futuras aplicaciones en el campo del procesamiento digital de imágenes. Se pretende hacer en hardware, una tarjeta que realice la adquisición, almacenamiento y transmisión de vídeo digital hacia una computadora personal.

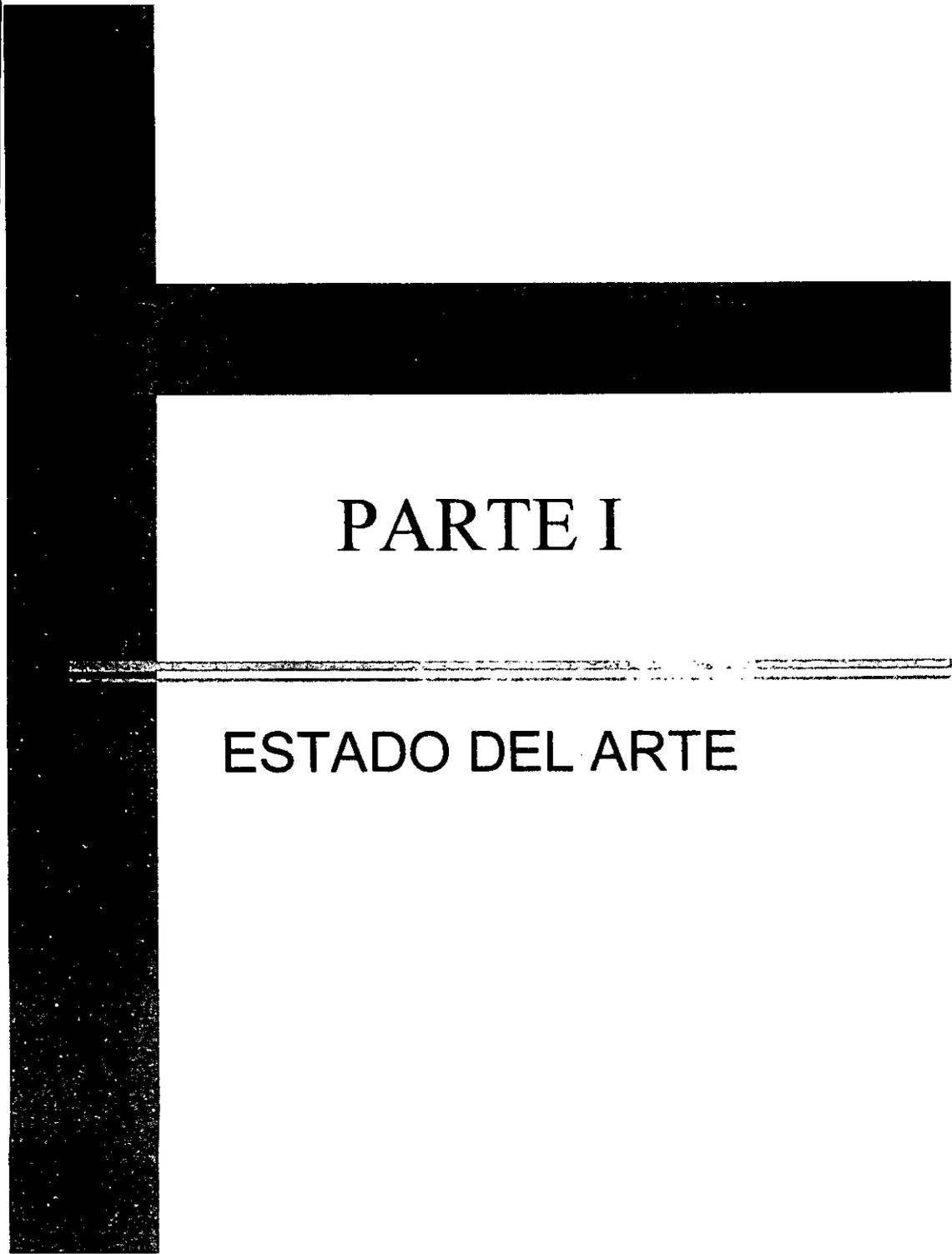
En su inicio, el trabajo fue desarrollado conjuntamente por el Departamento de Ingeniería en Computación y el Departamento de Ingeniería en Telecomunicaciones de la Facultad de Ingeniería, ya que el campo de acción del procesamiento digital de imágenes no se limita a una sola área de trabajo. La terminación del proyecto se hizo en el Departamento de Ingeniería en Telecomunicaciones, en donde se pretende mejorar el diseño, adicionando nuevos componentes, incrementando con esto la productividad de la tarjeta.

La realización de dicho trabajo fue con fines de investigación, sin tratar de competir comercialmente con los dispositivos similares existentes en el mercado, ya que éstos llevan muchos más años y la tecnología con que están diseñados es muy avanzada comparados con nuestro diseño que se hizo de manera económica y con circuitos muy elementales.

En lo que respecta al trabajo escrito, lo hemos dividido en tres partes principales, en donde se describe detalladamente el método empleado de acuerdo al enfoque establecido en los párrafos previos. La primera parte comprende lo que sería el estado del arte. Aquí damos una breve descripción del entorno necesario para poder entender el trabajo del dispositivo, tal entorno está formado por la descripción del vídeo analógico y el vídeo digital. En el capítulo de vídeo analógico se pueden encontrar las principales características de una señal de vídeo, la diferencia entre la señal a color y la de blanco y negro, así como los estándares existentes en el mundo. En el capítulo de vídeo digital se hace un recorrido por el camino que sufre una señal analógica hasta que se convierte en digital, estableciendo algunos conceptos y respetando algunos principios elementales. Por otro lado se presentan conceptos básicos de las cámaras, con esto queremos situar al lector en el hecho de que la cámara es una parte fundamental en la captura de imágenes. Por último se da una pequeña introducción de lo que es el campo del procesamiento digital de imágenes, describiendo algunos algoritmos y estableciendo una serie de ejemplos.

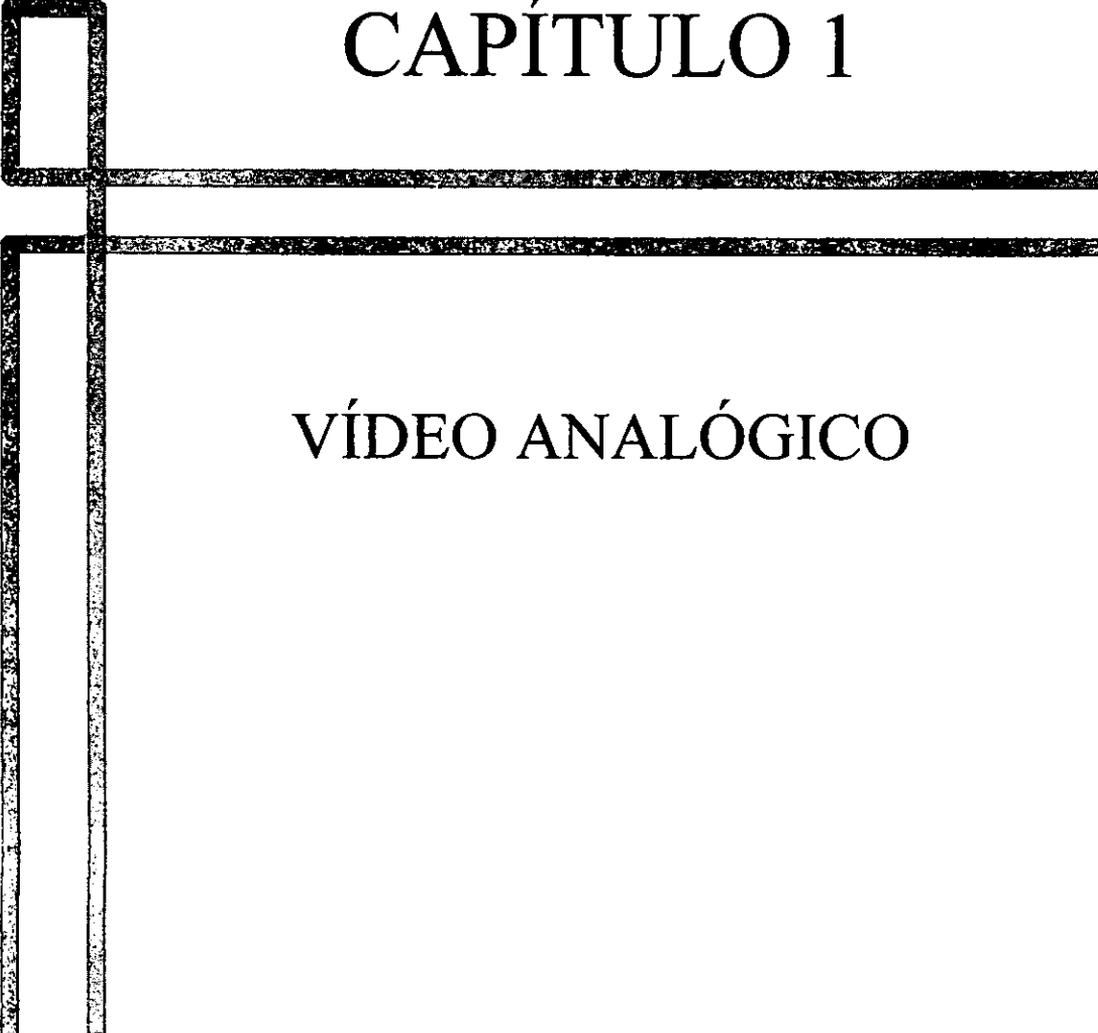
La segunda parte es donde mostramos el desarrollo práctico del circuito, damos una breve descripción individualizada de los componentes que forman la tarjeta, estableciendo su importancia dentro del sistema. Además, se describe a grandes rasgos el funcionamiento general de la tarjeta.

En la tercera parte presentamos una muestra de las pruebas que realizamos para ver el rendimiento del circuito ya como una aplicación práctica, dejando abierta cualquier posibilidad de mejoramiento del diseño. Al final de esta parte presentamos las conclusiones de nuestro trabajo.



PARTE I

ESTADO DEL ARTE



CAPÍTULO 1

VÍDEO ANALÓGICO

1.1 Generalidades del vídeo analógico

Una de las señales más interesantes desde el punto de vista de la electrónica es la señal compuesta de vídeo analógico, ya que es una señal formada a su vez de varias señales que se reúnen mediante un método de multiplexaje. Dichas señales trabajan en conjunto para obtener al final en el monitor una imagen idéntica de una escena tal y como si el ser humano la estuviera viendo directamente. Evidentemente, esta señal tiene que ver mucho con la vida cotidiana del ser humano, ya que es la que nos permite que nuestro receptor de televisión casero nos de la oportunidad de ver acciones o escenas que ocurren lejos de nosotros. Dicha señal se basa en el hecho de que se puede reproducir el movimiento de una escena si se presenta una secuencia de cuadros fijos con un espaciado de tiempo mínimo entre sí para que el ojo humano no detecte discontinuidades en el cambio de posición de los objetos entre un cuadro y otro.

Para el caso del cine, los cuadros son fotografías secuenciales de la escena; pero para la televisión es necesario hacer un "truco" que permita convertir la fotografía de dos dimensiones, en una señal de una dimensión (sin contar con la variable de tiempo), lo cual facilita su manejo y transmisión; esto se puede lograr descomponiendo la imagen en pequeñas componentes o puntos. A este método se le conoce como *exploración, barrido o escaneo de línea (scanning¹)*. De esta forma, los puntos pueden formar líneas horizontales y dichas líneas a su vez forman cuadros y los cuadros en secuencia forman imágenes animadas.

La característica principal de la señal de vídeo es la complejidad en cuanto a su formación, lo cual se debe al empleo de varias señales y a que el ancho de banda de la señal compuesta debe de ser de 4.2MHz, porque la cantidad de información que contiene reclama una gran cantidad de espectro, comparado con el ancho de banda de la señal de audio que es del orden de 20 kHz. En la figura I-1.1 podemos observar la señal de vídeo analógico vista sobre un osciloscopio.

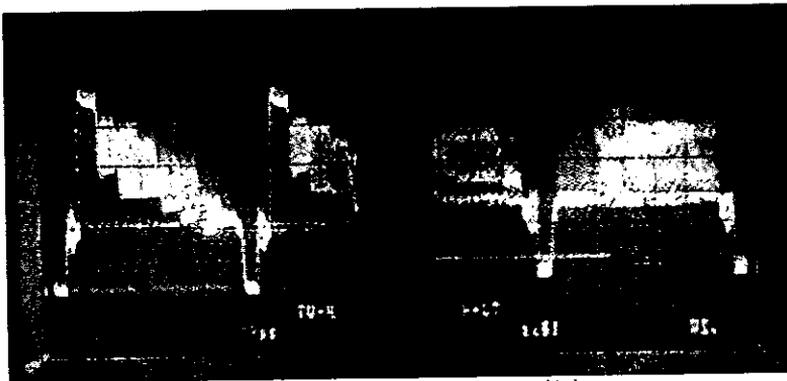


Fig. I-1.1 Señal de vídeo compuesto analógico.

¹ Andrew F. Inglis - Arch C. Luther; Video Engineering; McGraw-Hill, EUA, 1996; cap 1, pag. 4-6.

Cada señal que compone la señal final de video tiene su función específica con el objeto de dar al receptor los elementos necesarios para que no tenga problema en reproducir la imagen tal y como se captó por la cámara. Para el caso de una señal en colores se sabe que un color puede formarse a través de la combinación en proporciones determinadas de los tres colores primarios, según la colorimetría aditiva, rojo, verde y azul, con sus siglas en inglés RGB.

Existen varias consideraciones que limitan el manejo de dichas señales, las cuales se abordarán más adelante en este capítulo. Entre las limitaciones que se tienen, se encuentran los efectos espurios que pudieran ocurrir a la hora de desplegar la imagen, efectos ajenos a la escena que se incorporan a la señal por el mal manejo de las señales para su codificación y transmisión. Uno de estos detalles a considerar puede ser el efecto de persistencia de la visión, este efecto es una característica del ojo humano que consiste en que el mismo no reacciona a los cambios rápidos de iluminación, sin embargo, si la iluminación varía cíclicamente en una frecuencia relativamente baja, el ojo percibirá un parpadeo molesto de la imagen conocido en el idioma Inglés como "Flicker". Por esta razón, es necesario elegir la frecuencia a la que se presentarán los cuadros fijos lo suficientemente alta para evitar este efecto. La frecuencia de refresco en la que no se percibe el efecto es proporcional al grado de iluminación que se tenga en el medio, la frecuencia en la que el ojo humano no detecta en lo más mínimo el parpadeo está alrededor de 50 Hz para el caso de una iluminación media, que sería el caso de la televisión. A continuación se muestra una tabla más formal acerca de las frecuencias de refresco.

Sistema	Frecuencia de barrido (Hz)	Ambiente
Televisión NTSC	60 (59.94)	Sala
Televisión PAL	50	Sala
Monitor de computadora	72	Oficina
Cine	48	Cuarto oscuro

Tabla I-1.1 Frecuencias de refresco de pantalla.

En el caso específico de la señal de video compuesto, empleada en televisión, se verá más adelante que la frecuencia de refresco de cuadro varía desde 25 cuadros por segundo hasta 30 cuadros por segundo, dependiendo del estándar del que se hable. Sin embargo, para evitar el molesto parpadeo, fue necesario aumentar la frecuencia de refresco a 50 y 60 veces por segundo respectivamente, sin afectar para nada la resolución manejada ni el ancho de banda y esto se logró partiendo el cuadro en dos partes llamadas campos; lo que se conoce como video entrelazado. Para un entrelazado 2:1 se quiere decir que para producir un cuadro completo de video se hace un doble barrido vertical, en el primer barrido se pintan las líneas impares y en el siguiente se pintan las líneas pares.

La frecuencia de barrido en sentido vertical está estrechamente relacionada con la frecuencia de barrido horizontal. Esto porque para tener un patrón de N líneas horizontales, la frecuencia de barrido horizontal deberá tener la frecuencia $f_H = f_V \times N$. Pero para el video entrelazado la frecuencia de barrido vertical se incrementa al doble, por lo que la relación para el video entrelazado queda $f_H = f_V \times N/2$. Tomando en cuenta esta relación se obtiene un patrón de escaneo como el que se muestra en la figura I-1.2. Aquí es importante notar

que la última línea del campo impar es finalizada a la mitad, mientras que la primera línea del cuadro par inicia a la mitad, esto respeta el hecho de que cada campo tiene 262.5 líneas para el caso del estándar NTSC y ocasiona que las líneas de cada campo se entrelacen entre sí.

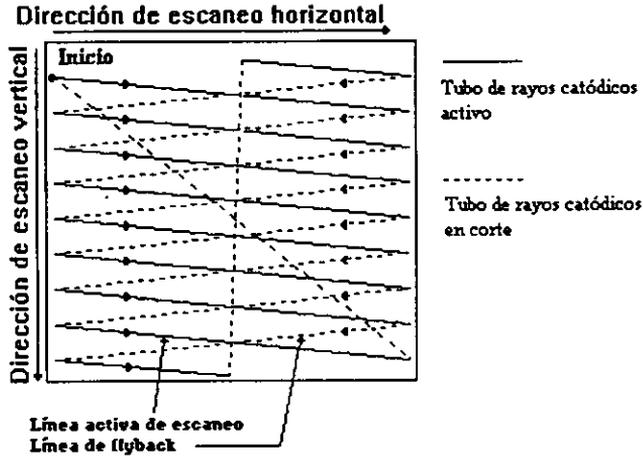


Fig. I-1.2 Modo de barrido entrelazado.

Los valores de dichas frecuencias dependen del estándar empleado. Los tres estándares más importantes en el mundo son: el NTSC, el PAL y el SECAM. El primero es empleado en Estados Unidos y Canadá, y también es el más importante para nuestro estudio, ya que es el que se maneja en nuestro país y en otros países de Latinoamérica.

Cada estándar se abordará más a detalle en los apartados siguientes, pero por el momento se puede decir que en general todos los formatos emplean el principio de Luminancia/Crominancia para el caso de imágenes a color. Este principio establece que la señal en color puede ser separada en dos partes, una es la **Luminancia**, la cual es una señal monocromática que contiene la información referente a la brillantez de la imagen; la segunda es la **Crominancia**, ésta es la señal que contiene la información de coloración de la misma imagen.

Como se mencionó antes, la información de color se puede descomponer en tres señales RGB, pero como la señal de vídeo compuesto analógico las representa de distinta manera, es necesario que la Crominancia se divida a su vez en dos componentes para tener así, junto con la Luminancia, las tres señales necesarias para hacer una transformación matricial y recuperar las tres señales RGB con las que se captó inicialmente la imagen en la cámara.

1.2 Señales que componen el vídeo analógico

A continuación se hace una pequeña descripción de las señales antes mencionadas, además de otras esenciales para la correcta codificación del vídeo:

La Luminancia.- Ésta es la señal que lleva la información de la intensidad o el brillo detectado en cada punto de la imagen tomada por la cámara.

La Crominancia.- Ésta contiene la información del matiz o tipo de color y el grado de saturación del mismo correspondiente a cada punto explorado de la imagen.

Sincronía del color o *burst*.- Ésta es una señal de control para mantener la señal de color recibida con la que se mandó de la emisora.

El blanking o los borrados.- Es la señal encargada de poner en estado de corte el tubo de rayos catódicos (encargado de emitir electrones para colisiones con la pantalla fluorescente y pintar la imagen) para que no emita electrones durante el regreso del cañón a su posición de inicio de pintado de la siguiente línea, es decir, durante los tiempos de retardo horizontal y vertical.

Las señales de sincronía.- Estas señales se encargarán de controlar la señal de barrido en el momento en que es "pintada" la imagen sobre la pantalla fluorescente, o en otras palabras para sincronizar las etapas de deflexión horizontal y vertical, y así tener la misma imagen obtenida con el haz del tubo de la cámara transmisora.

El sonido.- Contiene la señal de audio captada por el micrófono transmisor.

Como se puede observar, la cantidad de información que se necesita transmitir en cada cuadro de imagen es demasiada, por lo que se emplean métodos de multiplexaje en frecuencia, en fase y en amplitud.

Es de destacarse que para el caso del video compuesto en blanco y negro, la señal lleva todas las señales antes mencionadas, menos las que corresponden al color, las cuales son la Crominancia y la señal de sincronía del color o "*Burst*".

1.3 Compatibilidad en los sistemas de video analógico

Según se mencionó antes, existen distintos estándares para la construcción de la señal de video compuesto a color. En los inicios de la televisión, el tipo de imágenes que se transmitían eran monocromáticas, es decir, lo que se conoce comúnmente como blanco y negro, con niveles de grises intermedios.

Durante todo el tiempo que se empleó este modo de transmisión se fijaron algunos estándares propios para este propósito, como el ancho de banda asignado para cada canal, que es aproximadamente de 4 MHz. Se tuvieron problemas con la llegada de los primeros sistemas a color, el cambiar toda la estructura ya existente era complicado, por lo que la FCC (Federal Communications Commission) y la NTSC (National Television Standard Committee); que son los organismos encargados de administrar el uso del espectro en frecuencia en los Estados Unidos, diseñaron una serie de reglas que debían de cumplir los nuevos sistemas a color de tal forma que fueran compatibles con la estructura y los receptores en blanco y negro, es decir, que un receptor en blanco y negro pudiera recibir la señal de video compuesto a color con una calidad aceptable sin la necesidad de modificar en nada los parámetros del receptor. Esta situación arrojó algunos criterios para el diseño de los estándares de video a color.

El sistema monocromático consiste de una señal proporcional a los niveles de luminosidad existentes en cada punto de la escena explorada en tonos de grises, que como se mencionó antes se conoce como Luminancia. Se tiene la señal de blanking donde se colocan las señales de sincronía vertical y horizontal que controlan el despliegue de la imagen y se colocan en el nivel infranegro para poner el cañón de electrones encargado de pintar la imagen en corte, de modo de que no ocurra ningún error de pintado durante estos periodos de retraso, es decir, los tiempos en los cuales el cañón de electrones regresa del punto final de una línea al punto inicial de la siguiente línea.

1.4 Codificación en color

Existen algunos principios que hay que respetar en cuanto a la codificación del vídeo a color para tener la compatibilidad deseada con los sistemas monocromáticos, además de que se debe tener la capacidad de concentrar toda la información para transmitir la señal en color alrededor de 6 MHz para tener buenos resultados, dentro del ancho de banda establecido previamente para el sistema monocromático que es del orden de 4 MHz. Por estas razones fue necesario que los diseñadores tomaran en cuenta el comportamiento del ojo humano, aprovechando sus limitaciones y capacidades al máximo.

Uno de estos principios es la compartición de bandas empleando subportadoras de baja visibilidad. Evidentemente, para la determinación de este principio fueron necesarios algunos estudios y pruebas que llevaran a determinar cuál es la mejor condición. De acuerdo a estas investigaciones, se pudo aprovechar una característica del espectro de la señal de vídeo debida al tipo de escaneo que se efectúa en la imagen.

Se pudo ver, según se muestra en la figura I-1.3, que la energía de la señal se concentra principalmente alrededor de las armónicas de la frecuencia de escaneo o exploración de línea; además, se cuentan con bandas laterales de cada armónica espaciadas 60 Hz. Este comportamiento se debe a que existen pocos cambios entre líneas de escaneo consecutivas, y las bandas laterales son las que muestran estos pequeños cambios en la dirección vertical.

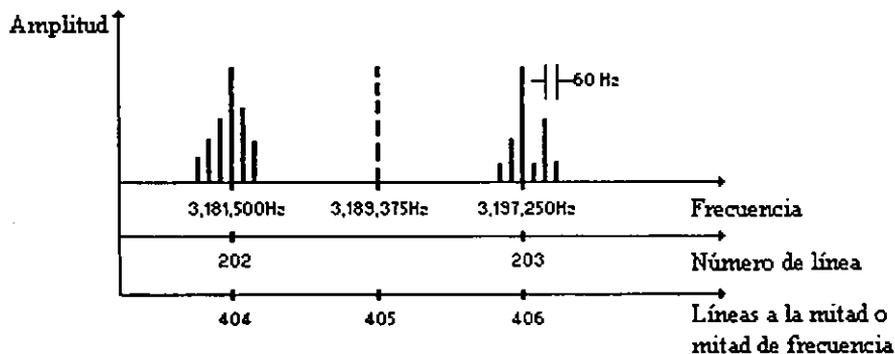


Fig. I-1.3 Espectro de la distribución de información de la señal de vídeo.

Como se puede ver en la figura queda espacio sin información entre cada armónica y se vio que era posible utilizarlo para mandar otro tipo de información sin interferir la señal de vídeo. En términos prácticos se puede ver que estas zonas libres corresponderían a las armónicas impares de una frecuencia de la mitad de la frecuencia de escaneo de línea. Para recuperar esta información intercalada se hace uso de un filtro de forma tipo peine (*comb-shaped filter*²), el cual es relativamente fácil de implementar. Si se aprovecha esta característica, la señal adicional será de muy baja visibilidad a la hora de reproducir la imagen.

Otra consideración es de que la visibilidad de esta información adicional para color es directamente proporcional a la frecuencia empleada para colocarla, esto porque a altas frecuencias la visibilidad será menor, debido a que los puntos generados por estas señales serán más finos en la imagen. Es por eso que se elige la subportadora de color lo más alta posible, sin embargo, si se aumenta esta frecuencia demasiado, se puede llegar a interferir con la frecuencia de la información de la señal de audio, lo que puede llegar a causar efectos molestos en la reproducción de la imagen. La subportadora de audio se encuentra a 4.5 MHz según el estándar NTSC, por lo que es necesario elegir la frecuencia de la información adicional sin interferir otras señales.

La codificación con mínima visibilidad de la subportadora de color fue establecida porque existen algunos efectos causados por interferencias entre las señales de color y la señal monocromática en el sistema de bandas compartidas empleado. Estos efectos son causados debido a la mala cancelación entre ambas señales mezcladas. Esta situación causa que la energía de la subportadora afecte la compatibilidad entre los sistemas monocromáticos y en color. Entonces, fue necesario establecer el requerimiento de que la energía de la subportadora y de sus componentes de banda lateral desaparecieran cuando se estén transmitiendo imágenes en blanco y negro. Por esta razón se emplean las señales llamadas diferencias en color con portadoras suprimidas, de esta manera, se tendría energía de la subportadora cuando las señales de diferencia de color difieran una de otra y desaparecería cuando las señales rojo, verde y azul fueran iguales, por lo cual se podría decir que la energía de subportadora sería proporcional a las señales de Crominancia, siendo definida ésta como la diferencia entre un área de color y un área de tonos de grises de la misma Luminancia.

Por otro lado, como ya se mencionó antes, para transmitir una señal a color se requiere que la señal cromática se componga a su vez de dos señales. Inicialmente se podría pensar que estas dos componentes serían transmitidas por modulación de dos subportadoras de baja visibilidad con frecuencias distintas, pero esto causaría deformación entre las subportadoras, lo cual podría causar un golpeteo de alta visibilidad, además, debido al espaciamiento necesario en ambas subportadoras, una de ellas debería tener una frecuencia relativamente baja y podría incrementarse la visibilidad de la misma. Es por esto que se ha pensado en transmitir las dos componentes en la misma frecuencia de subportadora al modular dos componentes de subportadora en cuadratura o como una subportadora, la cual es modulada tanto en amplitud como en fase y se considera lo que se mencionó anteriormente, que una reducción adicional en la visibilidad de la energía de la

² Andrew F. Inglis - Arch C. Luther; *Video Engineering*; McGraw-Hill, EUA, 1996; cap 5, pag. 155-156.

subportadora en un canal monocromático puede ser obtenida si se emplea una frecuencia de subportadora tan alta como sea posible.

También hay que tomar en cuenta que la información de Crominancia es mejor analizada en términos de una componente de banda ancha (I) y una componente de banda angosta (Q); la frecuencia de subportadora alta puede ser adoptada empleando una modulación conocida como *modulación de banda lateral residual* en la componente de banda ancha, la cual consiste en remover parcialmente una de las bandas laterales obtenidas de la modulación, de esta forma la frecuencia de subportadora es fijada apenas lo bastante alta para transmitir la doble banda lateral de la componente de banda angosta. Para el caso de la componente de banda ancha, se transmite su sección de frecuencia de modulación baja por modulación de doble banda lateral y su componente de frecuencia de modulación alta por modulación de banda lateral sencilla para el lado de baja frecuencia de la subportadora. De esta forma se obtiene un espectro parecido al que se muestra en la figura I-1.4, donde se puede observar como quedaría la compartición de las bandas así como las componentes de color.

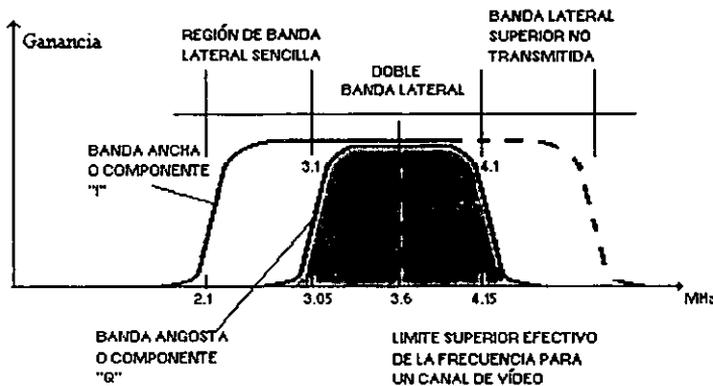


Fig. I-1.4 Espectro de la codificación del color en NTSC.

Finalmente, tomando en cuenta estas consideraciones, se pueden fijar las siguientes reglas en el manejo de la señal de vídeo a color compatible con blanco y negro:

- Diseñar el sistema de tal forma que la subportadora y todas sus bandas laterales se desvanezcan cuando la transmisión es en blanco y negro.
- Emplear solamente una subportadora para transmitir ambas componentes de la señal de Crominancia.
- Emplear transmisión de banda lateral residual de la componente de Crominancia de banda ancha para que una frecuencia de subportadora alta pueda ser empleada.

Adicionalmente, la codificación de color debería de tener una relación señal a ruido (S/N) máxima, dada por el canal de la subportadora de banda compartida, de esta forma se dará protección máxima contra ruido térmico y señales de interferencia externas. Además

de que la interferencia interseñal debida a la banda compartida podría ser visualmente minimizada. Es en esta parte donde se aprovecha el comportamiento del ojo humano ya que, según se ha investigado, es considerablemente más sensible a las fluctuaciones en la Luminancia que en las fluctuaciones de Crominancia. Por esta razón se estableció el *principio de Luminancia constante*³, el cual recomienda: "para permitir la codificación optima en color, las señales en el canal de la subportadora de color de banda compartida no deberían afectar la Luminancia de la imagen reproducida".

1.5 Sistemas de video en color

Una vez que se fijaron los principios anteriormente mencionados, se comenzó el diseño del estándar para la codificación del video a color. El primer estándar creado fue el formato NTSC que puso en práctica las consideraciones de diseño para que dicho sistema fuera compatible por completo con el estándar monocromático.

Posteriormente, entre las décadas de 50's y 60's, surgieron otros sistemas que hacen pequeñas modificaciones al formato NTSC y que son actualmente empleados en algunas partes del mundo. Estamos hablando concretamente del estándar PAL que es empleado en Europa, principalmente en Alemania; y del estándar SECAM, el cual es empleado en Francia. Cabe destacar que en la actualidad se tiene la introducción de un nuevo estándar con grandes adelantos tecnológicos conocido como el estándar HDTV.

1.5.1 Formato NTSC

En los inicios de la televisión a color, cuando se buscó cambiar las estaciones emisoras de televisión en blanco y negro para transmitir en color, se tuvieron algunos problemas en el acondicionamiento de los circuitos por errores en la fase del color.

Fue necesario modificar los primeros sistemas incorporando un control de matiz para modificar la fase de referencia de los decodificadores de color y era necesario ajustarlo para lograr la fase del tono de piel. Tenía algunos problemas al modificar las fases de otros colores, pero si los errores eran constantes, era capaz de corregirlos. Los problemas se presentaban cuando los errores eran diferenciales; esto porque al sumarse la portadora de color a la Luminancia, se producía un error de fase dependiente de la amplitud. A este sistema se le llamó NTSC (**National Television Standard Committee**). El NTSC rige las transmisiones de televisión en nuestro país, y por lo tanto es al que responde nuestra tarjeta.

Se desarrolló en los Estados Unidos en la década de los 50's; con el paso del tiempo, el sistema fue adoptado por otros países tales como Japón, México y gran parte de Latinoamérica. La manera en que está estructurado es para presentar 30 cuadros por segundo con cada cuadro compuesto por 525 líneas divididas en dos campos, uno de líneas impares y otro de líneas pares. Es un sistema basado en los conceptos de Luminancia y Crominancia tal y como se bosqueja al principio de este capítulo.

³ K. Blair Benson - Jerry Whitaker; Television Engineering Handbook; McGraw-Hill, New York, 1992; cap 4, pag. 4.31-4.32.

Para conservar la compatibilidad de este sistema con respecto al sistema monocromático, fue necesario acomodar las señales de tal forma que la señal final fuera idéntica a la señal del video monocromático. Como ya se mencionó antes, la Luminancia es la señal que lleva la información de los tonos de grises en una imagen, por lo que fue necesario usar esta señal como base para la transmisión. De esta forma, fue necesario "ocultar" lo más posible las señales propias del color, es decir, la Crominancia.

Los principios de codificación en color tomados para este estándar son el principio de Luminancia constante y una operación de banda compartida. La codificación de color fue diseñada para tener mínima visibilidad de la subportadora usando transmisión de subportadora suprimida con una frecuencia de subportadora sencilla para transmitir las dos componentes de Crominancia.

En lo que respecta al número de líneas que componen un cuadro, se realizaron estudios donde se vio que un espectador promedio se coloca aproximadamente a 5 veces la magnitud de la diagonal de la pantalla y se concluyó que se necesitaban entre 350 y 400 líneas horizontales para que el observador no pudiera distinguir las líneas y que las tomara como una imagen continua. Si se multiplica el número de cuadros presentados en un segundo por el número de frecuencias por cuadro tendremos: $30 \times 525 = 15.75 \text{ kHz}$, es decir, es la frecuencia de barrido horizontal, por lo tanto, sacando el recíproco se verá que una línea horizontal dura aproximadamente 63.5 microsegundos.

La sincronía horizontal se coloca en el intervalo de blanking horizontal, justo después de finalizada una línea para señalar al cañón que tiene que colocarse al inicio de la siguiente línea, dicha señal es un pulso hacia abajo, es decir, es sincronía negativa. Para el caso de la sincronía vertical, lo que se hace es que cada vez que se completa un campo, se introducen en la señal de video una serie de pulsos más anchos que los pulsos de sincronía horizontal rodeados por una serie de pulsos horizontales, separados entre sí tan sólo la mitad de tiempo de lo normal (ver fig. I-1.5). De esta forma se señala qué es el pulso de sincronía vertical y son fácilmente detectables con el uso de filtros.

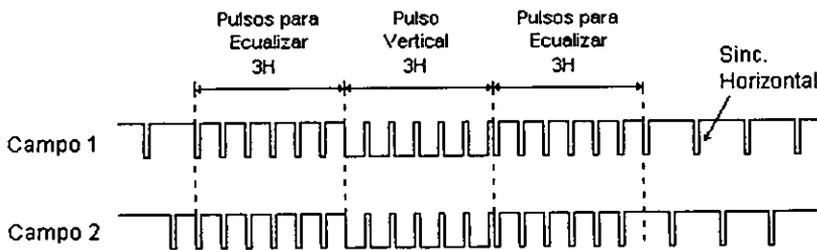


Fig. I-1.5 Sincronía vertical.

Una característica adicional para el valor de la frecuencia de escaneo horizontal es que ésta debe de ser $2/455$ veces la frecuencia de la subportadora de color, dando así un patrón de baja visibilidad para la subportadora y produciendo el menor golpeteo de baja visibilidad debido a la relación existente entre la subportadora de color y la de sonido.

En NTSC se simboliza la señal de Luminancia con la letra *Y*, y las dos señales de Crominancia con *I* y *Q*. *I* y *Q* están en fase y cuadratura, porque fueron moduladas en amplitud en dos fases sobre la subportadora de color (una en 0 grados, y la otra a 90 grados, en cuadratura). Esta forma de modulación doble y en cuadratura permite obtener fácilmente la modulación de los valores de matiz en fase de la subportadora y los valores de saturación para la amplitud de una manera sencilla a partir del empleo directo de las señales de diferencia de color *B-Y* y *R-Y* descritas más adelante. Para esto se alimenta la subportadora al modulador en amplitud de la señal *B-Y* y se coloca en una red de defasamiento de 90° y después al modulador en amplitud de *R-Y*. Finalmente, las salidas de cada modulador se suman, produciendo la modulación de la subportadora en fase y en amplitud.

La frecuencia de la subportadora de color es 3.579545Mhz; esta frecuencia debe ser respetada ya que también se basa en los principios antes discutidos para evitar que se puedan tener incompatibilidades con respecto a la señal monocromática. Al igual se emplearon dichos principios en las señales de diferencia de color de *I* y *Q*, en vez de *R-Y* y *B-Y*, que tienen reducido el ancho de banda. La Luminancia puede utilizar el ancho de banda total de 4.5 MHz asignados para un canal de televisión, el ancho de banda de *I* es de 1.5 MHz y la señal *Q* tiene el ancho de banda más reducido de 0.5 MHz.

Como se dijo antes, las señales *I* y *Q* son moduladas sobre la subportadora de color del sistema NTSC, donde la amplitud de la misma representa los valores de saturación mientras que la fase representa los valores de la matiz. Aquí es necesario señalar que el matiz es el color al que se está refiriendo la fase, mientras que la saturación es la intensidad de dicho color, es decir, el verde oscuro tiene el mismo matiz que un verde claro, sin embargo, lo que los diferencia es el nivel de saturación que tienen, siendo el verde oscuro el de mayor nivel de saturación. De esta forma se pueden entender los controles de los receptores de televisión a color: el control del matiz es el tinte y el control de la saturación se le conoce como color.

Desde el punto de vista matemático la estandarización en color tiene la siguiente composición:

$$E_M = E'Y + [E'Q \operatorname{sen}(\omega t + 33^\circ) + E'I \operatorname{cos}(\omega t + 33^\circ)] \quad (I-1.1)$$

$$E'Q = 0.41(E'B - E'Y) + 0.48(E'R - E'Y) \quad (I-1.2)$$

$$E'I = -0.27(E'B - E'Y) + 0.74(E'R - E'Y) \quad (I-1.3)$$

$$E'Y = 0.30 E'R + 0.59 E'G + 0.11 E'B \quad (I-1.4)$$

Como se puede ver de las ecuaciones, se pueden definir algunas componentes importantes. Primero, la expresión entre corchetes es la señal de Crominancia.

E_M es el voltaje de vídeo total correspondiente al escaneo de un elemento en particular.

$E'Y$ es el voltaje gama corregido de la porción monocromática de la señal de imagen a color correspondiente al elemento en particular.

$E'R$, $E'G$ y $E'B$ son los voltajes gama corregidos correspondientes a las señales rojo, verde y azul correspondientes al escaneo de ese elemento.

$E'Q$ y $E'I$ son las amplitudes de las componentes ortogonales de la señal de Crominancia correspondientes respectivamente a los ejes de banda angosta y de banda ancha.

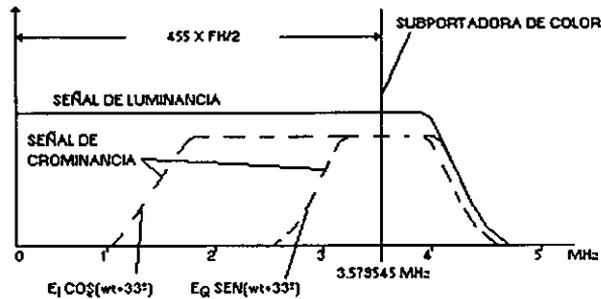


Fig. 1-1.6 Espectro de la señal de vídeo compuesto.

De las mismas expresiones anteriores se pueden ver las señales de diferencia de color $E'_R - E'_Y$ y $E'_B - E'_Y$, las cuales representan la otra alternativa para representar la señal de Crominancia. Las frecuencias de las señales de diferencia de color son bajas (abajo de 500 kHz). Empleando esta alternativa se puede representar la señal de vídeo compuesto a color por la siguiente ecuación:

$$E_M = E'_Y + \{1/1.14[1/1.78(E'_B - E'_Y)\text{sen}(\omega t) + (E'_R - E'_Y)\text{cos}(\omega t)]\} \quad (1-1.5)$$

De esta ecuación se puede concluir que la reproducción correcta de la imagen requiere que la luz de salida en el monitor sea proporcional a la intensidad de luz reflejada de la escena original en cada punto escaneado. Esto significaría que la señal de vídeo debe de ser una función lineal de la brillantez de la escena. Sin embargo, por las características propias de los dispositivos (por ejemplo, el TRC), se tiene una intensidad no lineal contra la señal de la función, es por eso que estas señales deben emplear el concepto conocido como *corrección gamma*, que ya se mencionó en la definición de las componentes de las ecuaciones anteriores.

1.5.1.1 Corrección Gamma

Muchos sistemas estándar han incluido la corrección en la producción de la señal de vídeo, ya que si se colocara dicho sistema en los receptores el remedio sería caro. Considerando que la mayoría de los sensores CCD de las cámaras tienen una respuesta lineal, entonces es necesario incorporarles la corrección gamma para introducir la amplitud no lineal característica de la señal. La gamma se refiere a la aproximación exponencial para la característica de transferencia; la gamma típica de un CRT está cerca de 2.2 lo que quiere decir que una cámara debe introducir una corrección gamma con una exponencial de $1/2.2$ o 0.45. Esta característica se muestra en la figura I-1.7. La corrección gamma está incluida en cada uno de las componentes a color RGB.

⁴ K. Blair Benson - Jerry Whitaker; Television Engineering Handbook; McGraw-Hill, New York, 1992; cap 4, pag. 4.37-4.38.

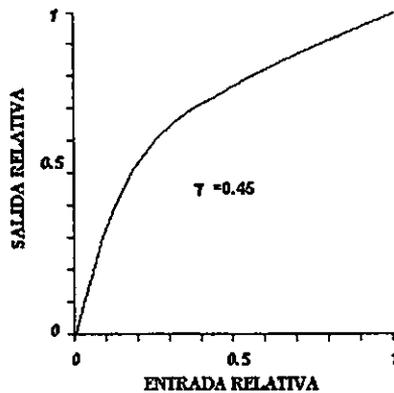


Fig. I-1.7 Corrección gamma.

1.5.1.2 Manejo de la información en color

Para el desarrollo del sistema en color, se dotó a las cámaras con tres tubos de rayos catódicos, uno para cada color (rojo, verde y azul). Pero en la actualidad sólo se emplea un tubo que maneja las tres señales y mediante un circuito electrónico se obtienen las distintas señales de vídeo de forma individual para cada color.

Estas señales son procesadas en forma matricial mediante un conjunto de circuitos sumadores, amplificadores inversores y no inversores; dicho arreglo combina en proporciones fijas estas señales para obtener la señal de Luminancia Y, así como las señales de diferencia de color que son función de la Luminancia Y. Evidentemente esta combinación está basada en la ecuación (I-1.5).

Otro punto importante, es que de las tres señales de diferencia de color R-Y, G-Y y B-Y sólo se transmiten las correspondientes al color rojo y azul, la señal correspondiente al color verde se reconstruirá a partir de las otras dos en conjunto con la señal de Luminancia Y en un "sistema matricial".

La subportadora de color se obtiene usando un cristal de precisión con una frecuencia de referencia de 3.57 MHz, ésta será la señal de referencia de 0°. Dicha señal se alimenta tanto al generador de sincronía como a los moduladores, los cuales emplean modulación de dos portadoras en cuadratura, mencionada anteriormente, mediante la cual se obtiene de forma sencilla una modulación en fase y cuadratura de la subportadora de color. La señal obtenida de la modulación se suma a la Luminancia para obtener la señal final.

Por otra parte, la subportadora de color se alimenta a los circuitos generadores de sincronía. Aquí se obtienen las señales de sincronía vertical y horizontal, y se coloca además la señal de sincronía de color. Luego de esto, la señal obtenida se suma a la Luminancia. La señal de sincronía de color consiste en una porción de la subportadora de color. Estos ciclos de la portadora llevan la fase de referencia necesaria para poder decodificar el matiz. Es necesario que en el receptor se encuentre un circuito capaz de tomar

estos ciclos intermitentes para obtener una señal continua que represente la recuperación de la subportadora suprimida.

La fase de referencia empleada para modular y demodular la componente de Crominancia no es una fase cualquiera. Según estudios hechos a este respecto, el ojo humano tiene mucho mayor sensibilidad a los errores que se pudieran producir en el rango de tonos correspondientes al color de la piel, es decir, es mucho más fácil ver cuando una piel está enrojecida o amarillenta, que cuando un coche tiene un poco quemada su pintura verde con respecto a otro del mismo color. Lo anterior ha llevado a establecer que la fase empleada como referencia sea precisamente la correspondiente al rango de los colores de la piel, es decir, entre los colores rojos y anaranjados.

El sonido se procesa en un modulador en frecuencia de la subportadora de sonido cuya frecuencia típica es de 4.5 MHz. Después se suma a las señales de vídeo y son finalmente transmitidas.

1.5.2 Formato PAL

Como se señaló anteriormente, existen otros estándares de codificación de vídeo a color empleados en el mundo. Estos son los sistemas PAL y SECAM, los cuales se originan en Europa. Dichos estándares son muy similares entre sí y con respecto al estándar NTSC, las diferencias fundamentales se presenta en el manejo de la señal de Crominancia.

El sistema PAL trabaja desplegando 25 cuadros por segundo con una resolución de 625 líneas por cuadro.

La historia del sistema PAL comienza en Alemania cuando se efectuó una modificación a la norma NTSC. La modificación consistió en modular la subportadora de color en dos fases, en vez de una como se hace en el sistema NTSC. Este proceso se conoció como *Phase Alternating Line* (PAL). Esta modificación da la ventaja de que ambas señales de Crominancia pueden tener el mismo ancho de banda, que es de 1.5 MHz.

Por lógica, las componentes de Crominancia obtenidas con esta modificación son diferentes a las obtenidas en el estándar NTSC por lo que se les conoce con otro nombre, es decir, las componentes de Crominancia en el sistema PAL son llamadas U y V en vez de I y Q.

Otra ventaja del sistema PAL gracias a la inversión de fase es de que sus señales son más tolerantes a ciertas distorsiones que pudiesen ocurrir en la transmisión afectando la calidad de la reproducción a color. Esto se consigue en la inversión de fase línea a línea en la subportadora de color del modulador de R-Y. En esta norma, a los circuitos moduladores se les agrega en el modulador en amplitud de la señal R-Y un circuito que da la portadora con fase alternada entre 90° y 270° . Las líneas impares tendrán 90° , mientras que las pares tendrán 270° . De esta forma los errores se van alternando en signo, y al momento de pintar las líneas con error alternado, el efecto de las líneas juntas producirá que el error se corrija.

Mediante la alternación de fase el promedio de errores de color tiene buenos resultados, aunque cuando los errores son muy grandes pueden ser obtenidos efectos espurios. En PAL el efecto conocido como barras de Hannover resulta de la promediación visual, lo cual puede afectar a la Luminancia constante cuando la fase no es perfecta y esto puede ser visible cuando el error es grande.

POLARIDAD

MATIZ	R-Y	A-Y
PÚRPURA	+	+
ROJO NAR. AMAR.	+	-
AMAR/VERD. VERDE	-	-
CIAN AZUL	-	+

FASE

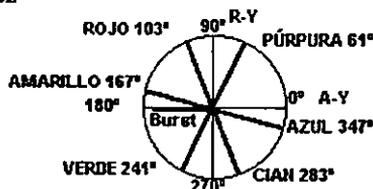


Fig. I-1.8 Manejo de colores.

Por otro lado, ya que las secuencias de fase opuesta se llevan acabo línea por línea, una línea de retardo igual a una línea de escaneo puede ser empleada para dar un promedio electrónico de los errores de fase, esto elimina substancialmente los patrones de línea espurios a cambio de reducir la resolución de Crominancia vertical.

Otro detalle importante se puede encontrar en el sistema PAL cuando se analiza el espectro en frecuencia de la Crominancia tal y como se muestra en la figura I-1.9. Según se puede ver, para la componente B-Y la energía se concentra mayormente en la frecuencia de subportadora de color y en las armónicas de la frecuencia de línea; pero para el caso de la señal R-Y la energía se concentra en frecuencias laterales, las cuales son armónicas impares de la mitad de la frecuencia escaneo de línea. Entonces, si se empleará un offset de un medio de línea como se hace en NTSC se tendrían problemas, ya que la energía R-Y se concentraría en la frecuencia de línea de escaneo, por lo que la señal B-Y sería de baja visibilidad, pero la señal R-Y producirá un patrón de alta visibilidad. Para evitar este problema, en PAL sólo se emplea un offset de un cuarto de línea para bajar la visibilidad de las dos componentes.

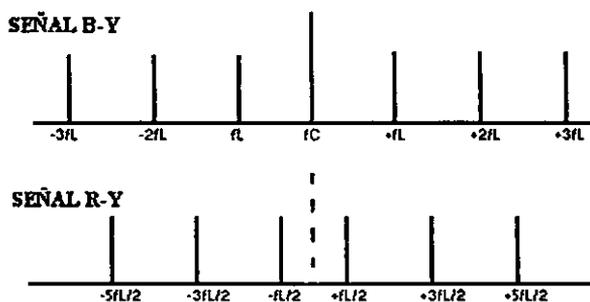


Fig. I-1.9 Espectro de distribución del sistema PAL.

Se pueden tener dos tipos distintos de receptores: el receptor PAL simple, el cual cuenta sólo con la promediación visual, y es muy parecido a un receptor NTSC, excepto por los circuitos para trabajar con esta propiedad.

El otro tipo es el receptor PAL de retardo de línea, que cuenta además con el soporte para emplear la promediación eléctrica.

Otra ventaja es que este sistema no necesita encimar tanto los espectros de Crominancia y Luminancia, por lo que se puede elegir una frecuencia de subportadora más alta ayudando a la compatibilidad con los sistemas monocromáticos. Aunque por otro lado, PAL sacrifica la resolución de Crominancia a cambio de la promediación de línea de retardo.

1.5.3 Formato SECAM

El sistema SECAM (*Sequentiel Couleur Avec Memoire*), desarrollado en Francia, usa una subportadora a color modulada en FM para la señal de Crominancia. Como el sistema PAL, el SECAM es también más tolerante con las distorsiones de la transmisión.

Emplea la transmisión secuencial de las dos señales de diferencia de color más memoria (es decir línea de retardo) en el receptor para tener ambas señales presentes para la reconstrucción de las señales R, G y B. Esto evita los problemas existentes en NTSC por la modulación de las dos señales en cuadratura.

El problema es que la subportadora de color sin modular se encuentra presente en la ausencia de color; además, cuando hay color, la subportadora es modulada en frecuencia, por lo que el grado de visibilidad cambia; este problema es más significativo que en NTSC o PAL. Por esta razón, SECAM tiene compatibilidad de menor calidad con los receptores monocromáticos.

Para contrarrestar esto un poco, en SECAM se eleva la frecuencia de dicha subportadora y se agrega un filtro con respuesta en forma de campana invertida.

1.5.4 Diferencias y similitudes entre las tres normas

Haciendo una comparación entre los tres sistemas antes descritos, se puede tener una mejor idea de las ventajas y desventajas de cada sistema. De inicio, como ya se mencionó antes que las diferencias principales son los procesos específicos de modulación usados para codificar la información de Crominancia. Por un lado, el estándar NTSC usa la modulación simultánea en amplitud y en fase de la subportadora de color, mientras que PAL suma la alternación de una de las componentes de modulación y SECAM usa la modulación de frecuencia de la subportadora de color de línea secuencial.

En lo que respecta a lo demás son bastante similares, es decir, todos los sistemas usan el principio de colorimetría aditiva, cuentan con tecnología de despliegue y cámara de vacío similar y la Luminancia tiene banda ancha mientras que para la Crominancia se emplea banda angosta. Además, los tres sistemas son compatibles con los sistemas monocromáticos.

Por otro lado sus diferencias, en cuanto al manejo de la información, son con respecto a la frecuencia de línea y campos, al ancho de banda específico de sus componentes y a la distribución de la frecuencia. La señal de Luminancia es obtenida de la

misma manera en los tres sistemas, es decir, sumar linealmente las proporciones específicas de las señales corregidas en gamma R, G y B. Además, las componentes de Crominancia (matiz y saturación) derivan en las diferencias entre las señal de Luminancia y las señales básicas R, G y B, son obtenidas por una matriz lineal.

En lo que respecta a las ecuaciones, se puede decir que la ecuación de la señal de Luminancia es común a los tres sistemas, pero en el caso de la Crominancia existen algunas diferencias:

Luminancia:

$$E'Y = 0.299 E'R + 0.587 E'G + 0.114 E'B \quad (I-1.6)$$

Crominancia:

NTSC:

$$E'I = -0.274 E'G + 0.596 E'R - 0.322 E'B \quad (I-1.7)$$

$$E'Q = -0.522 E'G + 0.211 E'R + 0.311 E'B \quad (I-1.8)$$

$$B-Y = 0.493(E'B - E'Y) \quad (I-1.9)$$

$$R-Y = 0.877(E'R - E'Y) \quad (I-1.10)$$

$$G-Y = 1.413(E'G - E'Y) \quad (I-1.11)$$

PAL:

$$E'U = 0.493(E'B - E'Y) \quad (I-1.12)$$

$$E'V = +0.877(E'R - E'Y) \quad (I-1.13)$$

SECAM:

$$D'R = -1.9(E'R - E'Y) \quad (I-1.14)$$

$$D'B = 1.5(E'B - E'Y) \quad (I-1.15)$$

En los sistemas NTSC, la expresión de la señal de croma total está dada por:

$$C_{NTSC} = ((B-Y)/2.03) \text{sen } \omega_{sc}t + ((R-Y)/1.14) \text{cos } \omega_{sc}t \quad (I-1.16)$$

La expresión de la señal de croma PAL es:

$$C_{PAL} = (U/2.03) \text{sen } \omega_{sc}t + (V/1.14) \text{cos } \omega_{sc}t \quad (I-1.17)$$

donde U y V son substituidos por $B-Y$ y $R-Y$, respectivamente, y la componente V es alternada 180° sobre una base línea a línea.

Los anchos de banda de las señales I y Q del sistema NTSC difieren de acuerdo con el eje de agudeza de color preferido, mientras que tanto U como V en PAL son del mismo ancho de banda y aproximadamente igual al ancho de banda de la señal I de NTSC.

1.5.5 Sistemas HDTV

Desde 1968, la Corporación de Transmisiones Japonesa (NHK) ha estado llevando a cabo investigaciones dentro del sistema de televisión de pantalla ancha de alta definición. En 1978, se realizaron pruebas de transmisión vía satélite de este sistema HDTV.

Se puede decir que existen dos sistemas principales de HDTV que han surgido en el mercado: el sistema 1125/60 (hecho por NHK y SMPTE) y el sistema Eureka (desarrollado por un consorcio de fabricantes europeos y otros interesados). En los Estados Unidos, el sistema 1125/60 ha sido aceptado como un estándar recomendado. Los parámetros técnicos de este sistema son los siguientes:

Escaneo:

Estructura de escaneo: entrelazado 2:1
1125 líneas totales por cuadro
1035 líneas totales activas por cuadro
Rango de aspecto 16:9
Duración del periodo de la imagen horizontal activo = 29.63 μ s
Duración del periodo de blanking horizontal = 3.77 μ s
Frecuencia de campo de 60 Hz

Ancho de Banda y Resolución

Luminancia = 30 MHz
Dos señales de diferencia de color = 15 MHz
Resolución de Luminancia horizontal = 786 TVL/ph
Resolución de Luminancia vertical = 750 TVL/ph

Representación Digital (tentativa)

Frecuencia de muestreo de la Luminancia = 74.25 MHz
Número total de muestras horizontales = 2,200
Número total de muestras horizontales activas = 1,920
Estructura de codificación digital = 22:11:11
Relación con CCIR-601 (4:2:2) = X5.5

La Comunidad Económica Europea de Naciones ha establecido conjuntamente una organización de investigación y desarrollo conocida como Eureka. Uno de los proyectos de Eureka, EU-95, es el de estabilizar los estándares para un sistema HDTV Europeo que sea compatible con los sistemas convencionales DBS D-MAC y D2-MAC/Packet. Esto porque el proyecto es uno de la familia MAC (Multiplexed Analog Components) de los sistemas DBS, conocido como HD-MAC.

HD-MAC es una señal híbrida digital/analógica. La porción del formato es la señal visual principal, la cual está basada sobre conceptos MAC; la porción digital cuenta con canales de sonido, teletext, datos de acceso convencional y datos de televisión asistidos digitalmente (DATV). Para mantener compatibilidad con señales convencionales MAC, el Burst principal de datos es de 10.125 Mbit/seg. Los datos DATV están a 20.25 Mbits/seg., acarreados en el intervalo de blanking vertical. La señal visual analógica tiene un ancho de

banda de 10.152 MHz, comparado con 8.4 MHz para el MAC convencional. El amplio ancho de banda significa que HD-MAC está operando en los límites de que es alcanzable ya sea por un satélite o un canal de cable en términos de ruido e interferencia. Medidas especiales, tales como el pre-énfasis no lineal, han sido adoptadas para incrementar la relación S/N.

CAPÍTULO 2

VÍDEO DIGITAL

2.1 Introducción

Al hablar del vídeo digital, se hace referencia a que una imagen es expuesta sobre una pantalla en forma de píxeles, los cuales fueron previamente almacenados en memoria en forma de bits. Esto se hace a una frecuencia tal que los cambios de una imagen a otra no son perceptibles al ojo humano durante la visualización (50 a 60 cuadros por segundo).

Uno de los sentidos más apreciados que posee el ser humano es el de la vista. El nervio óptico transmite información captada por el ojo hacia el cerebro a una velocidad tal, que si se hiciera una analogía a un sistema de computo podría decirse que transmite información a varios millones de bits por segundo.

Dentro de los principales campos de estudio en los que se ubica el vídeo digital se encuentran: la robótica y automatización, así como las telecomunicaciones. Durante muchos años se ha intentado dotar de una visión artificial a las maquinas, con el fin de poder hacer el reconocimiento de formas. En tanto que para las telecomunicaciones lo que se pretende es mejorar la calidad de las imágenes que son transmitidas.

El estudio del procesamiento en forma digital tiene como principal objetivo el substituir los ya tradicionales sistemas analógicos, al tener como ventaja principal que una vez digitalizada una imagen, ésta no presenta variaciones al copiarse o procesarse. Esto ha permitido el desarrollo de nuevas áreas de investigación, como por ejemplo, el reconocimiento de patrones, los gráficos y la animación por computadora, así como la comunicación visual.

2.2 Aspectos de la digitalización

Una señal analógica es una función continua que generalmente depende del tiempo, si se desea hacer un procesamiento digital sobre ésta, se deben tomar en cuenta los tres pasos básicos: convertir la señal analógica en digital, procesar la señal digital y convertir la señal digital nuevamente a su forma analógica. Cuando se hace la digitalización de la señal de vídeo analógico, es importante mantener en un valor constante la señal al momento de entrar al convertidor A/D, lo cual se logra con un circuito de retención y muestreo, esto porque la señal es continua tanto en amplitud como en tiempo. Con este circuito lo que se hace es convertir la forma de la señal en una especie de escalera y así obtener su equivalente en número binario a la salida del convertidor, estos valores representan la amplitud de la señal al momento de la digitalización. A la salida del convertidor A/D, se tienen los datos binarios que puede ser procesados. Al restaurar la señal, lo que se hace es someter los datos binarios en un proceso inverso (conversión digital- analógico), lo que a la salida mostrará una forma de escalera muy parecida a la señal original. Para hacer un suavizado de la señal obtenida se hace un filtrado para quitar las componentes de alta frecuencia. Todo este proceso es contemplado en la figura 1-2.1.



Fig. 1-2.1 Proceso de digitalización.

Una pregunta que tal vez muchas personas se hagan es: ¿Por qué trabajar con imágenes digitales, si un procedimiento analógico es mucho más simple?. Para responder a esto veamos las ventajas y desventajas que presenta el video digital frente a el analógico.

Los circuitos digitales no necesitan de valores exactos de las señales, esto implica que un circuito digital es menos sensible a los cambios de temperatura o bien al envejecimiento. Otro punto a favor es que los circuitos digitales pueden producirse en grandes cantidades, son pequeños y fáciles de transportar y actualmente, con el desarrollo de la tecnología, se han logrado colocar sofisticados y complejos sistemas en un chip.

En un sistema digital, los datos binarios que representan a la señal pueden corregirse con el hecho de cambiar el valor de estos datos. Varias señales digitales pueden trabajarse simultáneamente, aplicando métodos de multiplexaje.

Es posible efectuar un procesamiento multifrecuencia con implementación digital, mientras que con un sistema analógico no. Los circuitos digitales pueden conectarse en cascada sin problemas de carga, lo que no sucede con los circuitos analógicos. Además, el tiempo de almacenamiento digital puede ser casi indefinido sin pérdida de información.

Dentro de las desventajas que podemos encontrar en los sistemas digitales comparados con los analógicos están: se incrementa la complejidad del sistema, por eso es necesario un pre y un post-proceso, se procesa en forma digital pero inicia y termina en forma analógica. Se presentan problemas con el rango de operación de las frecuencias, debido al circuito de retención y muestreo y a la tecnológica del convertidor, pues la frecuencia de muestreo debe ser al menos el doble de la frecuencia de la señal analógica, según se verá más adelante. Otro problema es que los circuitos digitales están hechos con componentes activos, lo que produce un consumo de energía mayor que en los sistemas analógicos.

Como se ve, las ventajas superan a las desventajas, además la gran demanda que tiene el hardware para procesamiento digital indica que este tipo de aplicaciones crecen rápidamente.

Para darnos una idea aproximada del camino que sigue el procesamiento digital de imágenes, desde la captura, pasando por el manejo de las mismas hasta la obtención de resultados, presentamos los siguientes pasos:

1. Adquisición de la imagen. La fuente puede ser cualquier reproductor de vídeo analógico (cámara de vídeo, reproductora de vídeo, etc.).
2. Detección. Fijar el tamaño de la imagen en una matriz conservando la proporción de la señal original.
3. Extracción. Es aquí donde la señal analógica se transforma en digital.
4. Depuración. El objetivo aquí es el de eliminar ruido a la imagen quitando impurezas.
5. Extracción de características. En esta etapa la imagen pasa por una serie de algoritmos, por ejemplo: la extracción de contornos, la extracción de fronteras o la segmentación.
6. Interpretación. De acuerdo a los resultados de la etapa anterior se toman las decisiones.

2.2.1 Los píxeles

Cuando una imagen analógica es convertida a digital, se pretende que el resultado de esta digitalización permita hacer una representación de la imagen original sin distorsión, esto se hace con un apropiado manejo del muestreo y la cuantización. El resultado de la conversión A/D es una serie de números digitales, los cuales deben conservar la secuencia en la cual fueron obtenidos durante el barrido de la imagen, cuando se reconstruye la imagen estos datos serán representados en el monitor como unos puntos rellenos con algún color, los cuales se conocen como *píxeles*, el color dependerá del valor obtenido durante la digitalización. La imagen es reconstruida al colocar en un arreglo matricial de dos dimensiones la totalidad de los píxeles. La calidad de la imagen reconstruida depende del color y el número de píxeles usados en la digitalización.

El reloj de muestreo debe ser tal que obtenga el número de píxeles deseados (muestras), así como cumplir con la frecuencia de muestreo o de lo contrario la foto digital tendrá una incorrecta proporción.

2.2.2 Exclusión de sincronía y blanking

Como vimos en el capítulo 1, de vídeo analógico, la señal de vídeo cuenta con las señales de sincronía horizontal y vertical, además de las señales de blanking, que para el caso del vídeo digital no son utilizadas, por ello, si se hace una digitalización con un reloj constante se corre el riesgo de digitalizar estas señales y deformar los resultados esperados. Para evitar este problema debemos dejar de almacenar información durante este periodo y sólo digitalizar la parte concerniente a los datos de la imagen, dejando el control del despliegue al software.

2.2.3 Bits por pixel

El indicador de bits por pixel determina la cantidad de colores a la cual fue digitalizada la imagen, pudiendo tenerse desde 1 hasta 24 bits por pixel. Cuando sólo se tiene un bit por pixel, se entiende que la imagen está formada por dos colores, blanco y negro habitualmente. Cuando se hace con 24 bits se pueden tener sobre la imagen más de 16 millones de colores.

2.3 Métodos de codificación en color

La digitalización de una imagen en color se puede hacer de dos formas, la primera consiste en separar las tres componentes de color, empleando en paralelo un convertidor A/D para cada una de las tres componentes de color (RGB, YIQ, YUV, etc.), como se ilustra en la figura 1-2.2. Los resultados obtenidos de cada digitalización se unen para formar un solo grupo de bits que formará el total de cada muestra. Por ejemplo en un sistema RGB se pueden utilizar 8 bits para cada color primario, logrando al final un grupo de 24 bits por pixel, con los cuales podremos representar más de 16 millones de colores ($2^{24}=16,777,216$).

Un segundo método de codificación en color es digitalizar la señal compuesta con 8 bits por pixel. Para este caso la digitalización debe hacerse con un reloj de al menos 3 veces la frecuencia de la subportadora de color, con lo cual la cantidad de espacio requerido para el almacenamiento sería menor que con el sistema de 24 bpp que se mencionó anteriormente.

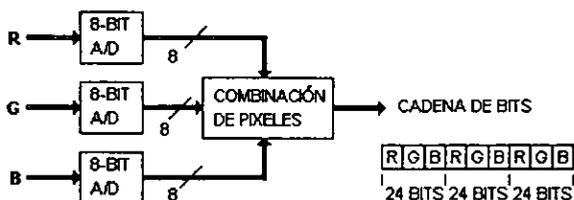


Fig. I-2.2 Digitalización en color con 24 bits.

Cuando se pretende hacer algún procesamiento a la imagen, no es conveniente hacerlo con resultados obtenidos de la digitalización de la señal compuesta dada su complejidad y por tanto no representa una ventaja importante sobre el sistema de 24 bpp. El uso de esta forma de digitalización es aplicable para sistemas de transmisión de televisión.

Con el fin de reducir el espacio requerido con las imágenes de 24bpp, se puede hacer un truncamiento a 5bpp para cada color, agregando un bit a la izquierda, se logra tener imágenes representadas en 16bpp, lo cual da una imagen de calidad aceptable teniendo 32,768 colores disponibles, suficientes para representar imágenes coloridas. La desventaja que esto implica es que puede presentarse el efecto de contorneo en imágenes con un sutil sombreado en grandes áreas.

Existe la posibilidad de desplegar buenas imágenes con un menor número de bits, y se logra empleando las técnicas de compresión de vídeo. Existe una gran variedad de técnicas, pero requieren un procesamiento tanto para la compresión a la entrada como para la descompresión en la salida (véase sección de técnicas de compresión).

2.4 Bitmap

Se conoce como bitmap al arreglo de bytes alojados dentro de una zona de memoria, los cuales formarán una imagen cuando se presenten en un monitor. Este nombre proviene del concepto de que se tiene mapeada la imagen física en forma de píxeles dentro de las direcciones vecinas de memoria.

No puede hablarse de un bitmap cuando se hace referencia a información en memoria con respecto a texto, puesto que no se manejan directamente los píxeles.

Para que un algoritmo identifique que la información en memoria representa a un bitmap, es necesario que éste cuente con la siguiente información:

- Dirección de inicio en memoria.
- Longitud de una línea en píxeles.
- Número de bits por píxel.

- Lanzamiento de línea (usualmente llamado pitch).
- El total de líneas verticales.
- Paleta de colores.

2.4.1 Lanzamiento de línea o Pitch

El pitch es un indicador que se usa para establecer la distancia en forma de bytes o pixeles que hay entre el inicio de una línea hasta el comienzo de la siguiente. Si durante la digitalización los intervalos de blanking fueron perfectamente excluidos, entonces el pitch será igual a el número de pixeles por línea.

Como se mencionó en el capítulo 1, la señal de vídeo se forma de manera entrelazada, es decir, se forma después de dos barridas del cañón, pintando el campo par y el campo impar. Para el caso del almacenamiento en memoria, es preferible no hacerlo entrelazado. Usando un doble pitch durante la digitalización y escritura en memoria se pueden ir almacenando las líneas de cada campo alternadamente y al final de las dos escaneadas de la imagen tendremos la representación digital no entrelazada.

2.4.2 Mapeo del color

Existe otra forma de usar el bitmap para representar una imagen en donde los valores de pixel no son precisamente los obtenidos del muestreo, es decir, teniéndose una base de datos de colores llamada "*CLUT*", la cual puede estar formada por tantos colores como se desee, por ejemplo con 24 bpp se tendría una base de datos de más de 16 millones de colores. Si una imagen se digitaliza con 8bpp se obtendrá una imagen con 256 colores, empleando los datos de la digitalización como apuntadores a un color dentro del CLUT, los pixeles que formarían a la imagen tendrían 256 colores de una paleta de 16 777 216 con 24 bpp. Los 256 colores a desplegar pueden cambiar dependiendo del tipo de imagen procesada. A este procedimiento de elegir los colores de una base de datos existente se le conoce como "*mapeo del color*".

Los resultados obtenidos con esta técnica dan una imagen poco real comparada con la escena original, pero si no se tiene la fuente original para hacer la comparación, el resultado es bastante aceptable. Esto no es bueno si se manejan imágenes en las cuales estén presentes tonos de la piel ya que, como se mencionó en el capítulo anterior, el ojo es muy sensible a cambios en estos colores, por lo que se requieren muchos colores del CLUT para tener un buen suavizado.

El mapeo del color da buenos resultados cuando la escena fuente a digitalizar presenta poco colorido, y se usa principalmente cuando no se desea hacer un procesamiento muy elaborado, permitiendo con ello un procesamiento en tiempo real.

La capacidad de manejo de colores de nuestra computadora va a depender de las características de nuestro equipo, principalmente en la parte concerniente al vídeo. En la tabla I-2.1 se listan algunos adaptadores de despliegue de vídeo para IBM PC AT y computadoras compatibles, con algunas de sus características. Podemos ver que sólo las tarjetas Targa son capaces de representar imágenes realistas. Además, debemos darnos

⁵ Andrew F. Inglis - Arch C. Luther, *Video Engineering*; McGraw-Hill, EUA, 1996; cap 7, pag. 188.

cuenta de la gran diferencia en cuanto a requerimiento de memoria entre las diferentes tarjetas. Algunas de estas tarjetas tienen la capacidad de cambiar la resolución y el número de colores.

TARJETA TÍPICA	RESOLUCIÓN DE PÍXELES	NÚMERO DE COLORES	BYTES EN MEMORIA PARA LA IMAGEN
EGA	320 X 200	4	16,000
VGA	640 X 350	16	112,000
MGA	640 X 480	16	153,600
TARJETA 16	512 X 480	32,768	491,000
TARJETA 32	512 X 480	16,777,216	983,000

Tabla I-2.1 Adaptadores de despliegue de vídeo para PC AT.

2.5 Proceso de digitalización de las imágenes

2.5.1 Muestreo y Cuantización

En el proceso de digitalización, la representación en forma digital, normalmente se hace con valores enteros en potencias de 2, así por ejemplo, una palabra digital de 16 bits puede representar valores de 0 hasta 65535, con esto se elimina el tedioso manejo de fracciones, que se puede hacer si se empleara notación numérica de punto flotante, lo que sólo causaría un trabajo más complejo para la computadora, consumiendo tiempo, lo cual no es deseable para trabajar con imágenes en tiempo real.

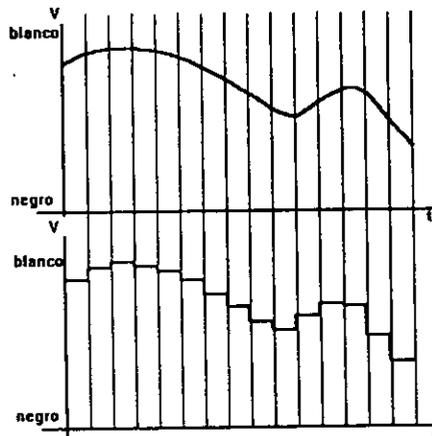


Fig. I-2.3 Conversión Analógica/Digital.

Una señal analógica es continua tanto en amplitud como en tiempo. Para convertirla a una forma digital es necesario cambiar ambas dimensiones a valores no continuos. Al proceso de conversión en tiempo a su forma digital se le conoce como *muestreo*, y el proceso para convertir amplitud a valores discretos se llama *cuantización*. A estos procesos

en conjunto se les conoce como *conversión analógico a digital*, o *digitalización*. La figura I-2.4 muestra el proceso de muestreo.

En un sistema de digitalización podemos hablar de un reloj, el cual tiene una frecuencia F_c (frecuencia de muestreo), éste produce pulsos en un periodo $P_c=1/F_c$. Durante el muestreo, se toman lecturas de los valores instantáneos de la señal analógica en el tiempo de cada pulso de reloj, como se ejemplifica en la parte inferior de la figura I-2.4. Las líneas punteadas representan el estado estable de la señal durante la lectura. El circuito que hace esto es llamado circuito de retención y muestreo. Si la frecuencia de muestreo es muy alta la representación digital es bastante parecida a la forma de onda analógica, pero ésta sigue siendo en parte analógica pues es discreta en tiempo y analógica en amplitud.

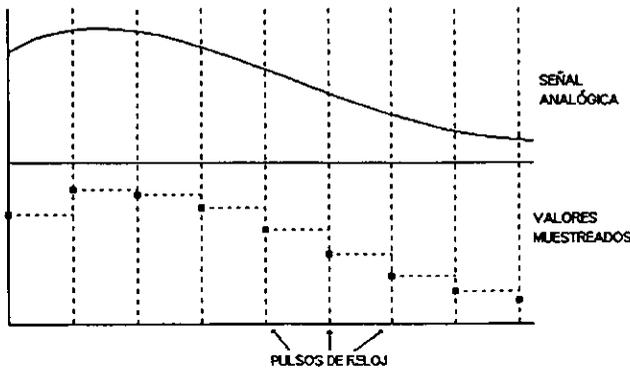


Fig. I-2.4 Muestreo.

Para transformar las muestras en una señal totalmente digital tiene que hacerse una cuantización (figura I-2.5). En una cuantización uniforme se fijan umbrales equidistantes en amplitud (niveles de cuantización). Así cada muestra es analizada individualmente, se fija al valor que corresponde y se le asigna un valor numérico apropiado. Este número es representado digitalmente para cada muestra (proceso conocido como codificación).

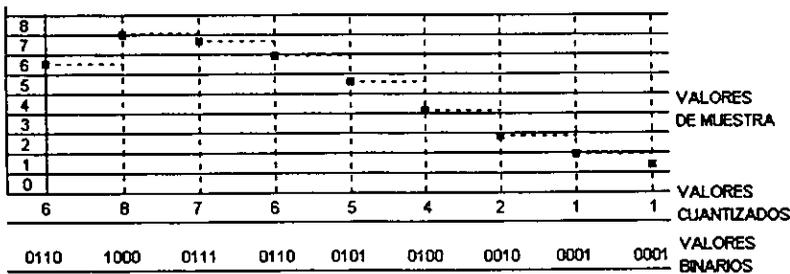


Fig. I-2.5 Cuantización.

En el ejemplo se tuvieron 16 niveles de cuantización, así el número para los niveles puede ser representado con cuatro bits, de esta forma se cuantiza en cuatro bits por muestra.

2.5.1.1 Limitaciones del muestreo

Los principales parámetros para que exista una buena conversión A/D son: tener una frecuencia de muestreo bastante alta y tener suficientes niveles de cuantización.

La figura I-2.6 muestra como la frecuencia de muestreo afecta los resultados. En esta figura la forma de onda analógica es una onda senoidal que es muestreada con tres diferentes relojes. En el superior, usamos una frecuencia muy alta, así que estas muestras tomadas para cada ciclo de la onda pueden representar satisfactoriamente la onda senoidal. En el segundo caso, se usa un valor de muestreo que es más bajo que la frecuencia de la senoidal, así que se tiene una muestra tomada para cada dos ciclos de la onda senoidal. Aquí se puede apreciar que la salida de este muestreador no lleva la onda senoidal en su totalidad. En la última figura se tiene una circunstancia especial, donde tomamos exactamente dos muestras por cada ciclo de la onda senoidal. En este caso la frecuencia senoidal está presente en la salida, pero la forma de onda es más cuadrada que senoidal.

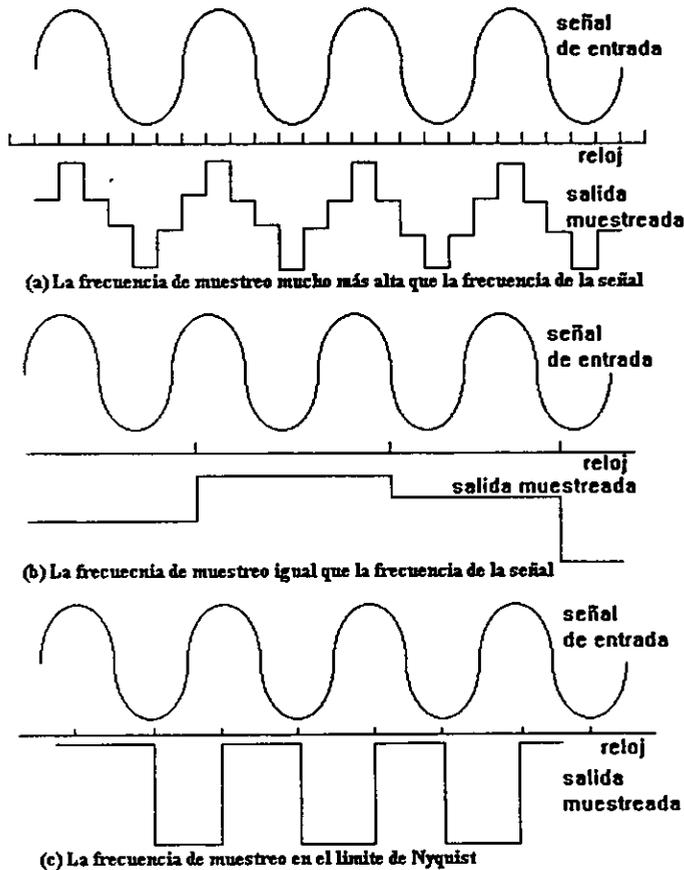


Fig. I-2.6 Muestreo a diferentes frecuencias.

Este caso es llamado el límite de Nyquist, la onda senoidal de entrada está en la frecuencia más alta que puede ser reproducida con el reloj de muestreo de este ejemplo. Para hacer una correcta digitalización, el rango de muestreo para un proceso de digitalización debe ser al menos el doble del ancho de banda de las señal a ser digitalizada. Si probamos digitalizar en un rango más bajo que éste, la salida de muestras puede ser incorrecta para cualquier componente de alta frecuencia en la entrada. Se puede observar que exactamente en el límite de Nyquist, la frecuencia es reproducida correctamente, pero la amplitud es confusa.

Si la frecuencia de muestreo es más baja que el límite de Nyquist, se producen salidas falsas, conocidas como "aliasing".

Para evitar que accidentalmente se supere el límite de Nyquist y se produzcan salidas falsas, se coloca un filtro paso bajas sobre la señal de entrada al frente del muestreador con el fin de remover aquellas componentes de entrada que pasen de la mitad del rango de muestreo.

El criterio de Nyquist puede ser también visto en el dominio de la frecuencia. Una serie de muestras puede ser considerada como una función de impulso periódica en la frecuencia de muestreo, cuya amplitud es modulada por los valores muestreados. El espectro de un impulso periódico consiste de una componente de frecuencia cero más la componente de amplitud igual en la frecuencia de repetición y sus armónicas. Cada componente tiene bandas laterales representando modulación por las frecuencias de la señal. Esto se muestra sobre el espectro en las tres condiciones correspondientes a los tres casos de la figura I-2.7. El aliasing ocurre cuando las bandas laterales de las componentes de frecuencia de muestreo enciman las bandas laterales de la componente de frecuencia cero. Si ocurre traslape, no hay forma de arreglarlo después del muestreo, esto debe ser prevenido por un filtrado de entrada.

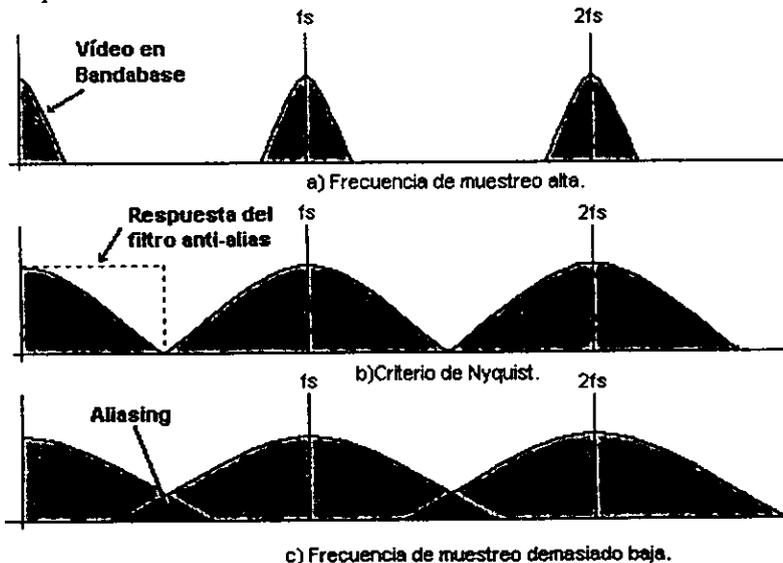


Fig. I-2.7 Límite de Nyquist sobre el espectro en frecuencia.

Quando es necesario operar con frecuencias de señal cerca del límite de Nyquist, el correcto filtrado es importante, como se puede ver en la figura I-2.7 b), la frecuencia de corte del filtro debe de ser precisa para evitar distorsión en la señal deseada mientras se remueven la frecuencias que podrían causar el aliasing.

2.5.1.2 Limitaciones de cuantización

Las limitaciones del proceso de cuantización dependen del número de niveles de cuantización empleados. Sin embargo, el fijar cuantos niveles de cuantización se usarán no es tan sencillo. Siempre hay que considerar que la señal de ruido siempre está presente y así elegir los niveles que permitan producir variaciones en la señal que sean más grandes que el ruido.

2.5.2 Conversión de vídeo A/D

En la figura I-2.8 podemos ver como quedaría digitalizada una imagen que entra en forma analógica. Cuando la imagen es digitalizada en tonos de grises, a cada muestra se le asigna un valor numérico, el cual indica el grado de iluminación o brillantez que corresponde al punto explorado. Cada valor numérico corresponde a un color proporcional de la paleta de colores.

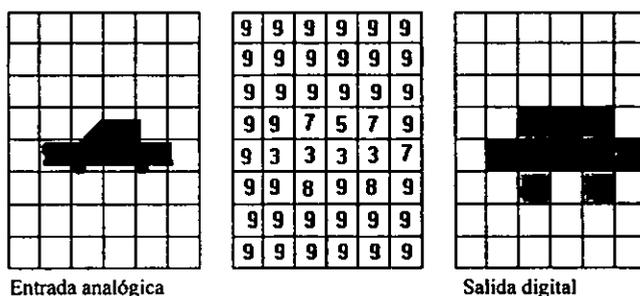


Fig. I-2.8 Digitalización de una imagen.

2.5.2.1 El convertidor Analógico-Digital

El convertidor A/D está basado en el circuito comparador, éste compara el valor de la señal de la entrada con el voltaje de referencia. La salida digital es cero cuando la señal es menor que el valor de referencia o umbral y 1 cuando la señal está arriba de la referencia. Los comparadores deben de ser muy estables y precisos, deben de ser de respuesta rápida para cambios muy pequeños de la señal de entrada.

La figura I-2.9 muestra una forma de como los comparadores son usados en la cuantización, hay un comparador por cada nivel de cuantización, la señal de entrada es aplicada a todos los comparadores y una cadena de resistencias es colocada para obtener los valores de referencia para cada comparador a niveles sucesivamente más altos. Como

requiere de 2^a comparadores, esto sólo es práctico para $n=8$ o 10. Este diseño es conocido como convertidor tipo *Flash*.

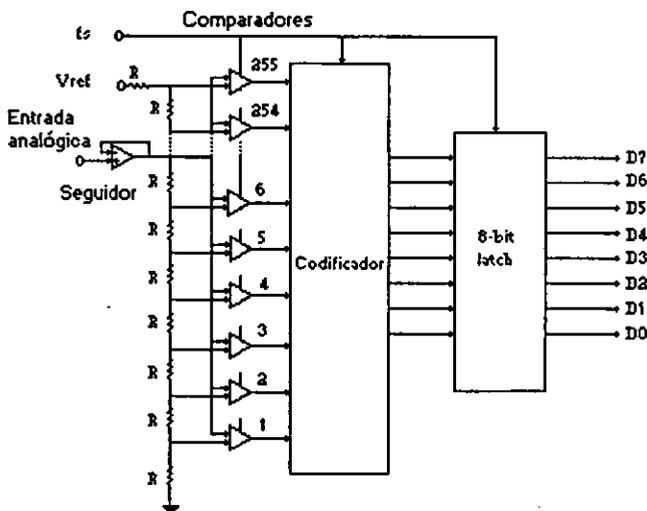


Fig. 1-2.9 El convertidor A/D tipo flash.

2.5.2.2 Conversión D/A

Para poder ver una imagen digital en un monitor analógico RGB, debemos de convertir la señal digital en tiempo real a una señal analógica. Esto se hace a través de un convertidor D/A que típicamente está implementado en un chip (tres para el caso de RGB).

Este convertidor cuenta con algunos switches que son controlados por los mismos bits de la señal digital para obtener valores analógicos iguales a los representados por cada bit. Comúnmente el D/A puede incluir circuitos de retención para el muestreo, los cuales aseguran que la salida cambie suavemente de un pixel a otro.

Si se pone un filtro analógico a la salida del convertidor D/A, puede reducirse la visibilidad de pixelización. Dicho filtro deberá tener un ancho de banda igual al límite de Nyquist, para la piel puede no cambiar mucho la resolución, pero puede suavizar algunas variaciones bruscas entre pixeles adyacentes (sólo en la dirección horizontal).

2.5.2.3 Efectos de vídeo digital

2.5.2.3.1 Pixelización

El número de pixeles en una imagen de un sistema digital es análogo a las resoluciones horizontal y vertical de un sistema analógico, excepto en que la apariencia de pixeles es muy diferente de los tipos de resolución analógica. En un sistema analógico con baja resolución, la foto puede llegar a ser difusa, pero en un sistema digital obtenemos un

efecto conocido como *pixelización*. La figura 1-2.10 muestra una imagen que ha sido digitalizada con un número diferente de píxeles. La apariencia de bloques de la imagen digitalizada con un pequeño número de píxeles es el efecto de *pixelización*.

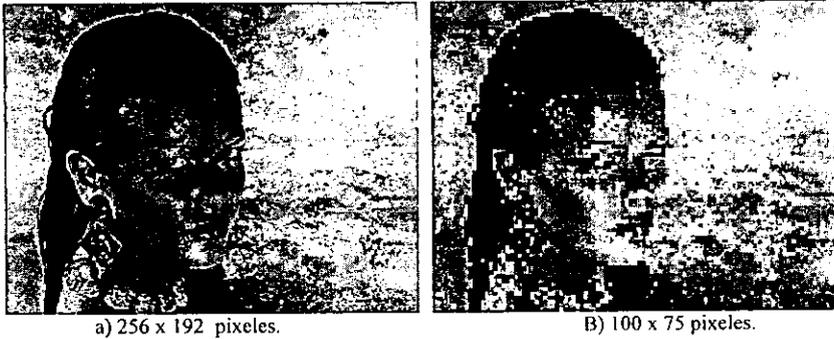


Fig. 1-2.10 Pixelización.

El grado de visibilidad de pixelización depende de la distancia entre el observador y la imagen. La razón de visibilidad es también una justificación válida para hablar acerca de los problemas en los sistemas de vídeo digital. La figura 1-2.11 muestra el número de píxeles por la altura de la foto necesarios para que la pixelización no sea visible por un observador común para diferentes razones de visibilidad. La figura asume que los píxeles son cuadrados, lo cual implica que estos tienen las mismas dimensiones sobre la pantalla horizontalmente y verticalmente.

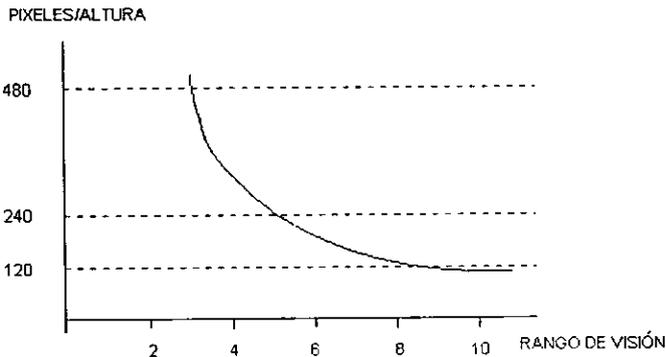


Fig. 1-2.11 Umbral de visibilidad de píxeles.

2.5.2.3.2 Contorneo

El *contorneo* se presenta cuando distintas partes de la señal analógica caen dentro del mismo nivel de cuantización, esto provoca que el valor de salida sea el mismo para estas distintas partes. Cuando la señal analógica tiene cambios suaves, la señal digital de salida es en forma de escalera debido a la cuantización. Los niveles de escalera son

altamente visibles, aun cuando son menores que el ruido aleatorio que está presente en la señal.

El contorneo produce un cambio brusco en los límites donde varía el color, ocasionando que la imagen pierda calidad, es por ello que a mayor número de bits por píxel se tengan en la cuantización, la calidad de la imagen crecerá. La figura I-2.12 muestra el grado de contorneo ocasionado al digitalizar una imagen con 2, 4 y 8 bits por píxel.



Fig. I-2.12 Ejemplo de contorneo.

2.5.2.3.3 Aliasing

El efecto de *aliasing* se produce al muestrear una imagen que cuenta con frecuencias arriba del límite de Nyquist. Por ejemplo, en la reproducción de texto sobre un monitor, un carácter de texto cuenta con una resolución infinita. Sin embargo, para desplegar el carácter sobre la pantalla de la computadora, la cual tiene un número finito de píxeles, se debe desarrollar un proceso de muestreo para crear la imagen del carácter en píxeles. La figura I-2.13 muestra esto para el carácter texto "V". El efecto de dentado sobre las partes diagonales de la "V" es el resultado del aliasing. Si el sistema tiene la capacidad de desplegar en escala de grises, entonces un proceso anti-aliasing puede hacerse sobre el carácter para llenar en algunos de los dientes con niveles de gris intermedio. Como se muestra en la figura I-2.13 c), el anti-aliasing es una forma efectiva para reproducir texto cuando se tienen bits múltiples por píxel.



Fig I-2.13 Aliasing y Anti-aliasing.

2.6 ¿Por qué hardware?

Para realizar algún tipo de procesamiento sencillo, generalmente no necesitamos de hardware muy complejo, nos basta con emplear los recursos existentes en una computadora de propósito general. Un ejemplo de esto es el diseño de filtros digitales o algún tipo de simulaciones. Se puede decir que la característica común en estas aplicaciones es que el procesamiento se puede llevar a cabo en tiempo no real, es decir, la velocidad de procesamiento puede no depender de la velocidad con que se reciben los datos. Sin embargo, cuando las señales que se manejan son señales de alguna forma dinámicas, como en el caso del vídeo, y es deseable analizar los datos de entrada en tiempo real, nuestro sistema debe de ser capaz de procesar toda la información; de no ser así, se tendrán problemas de almacenamiento y al tener pérdida de información el resultado esperado será erróneo.

En el caso de los sistemas de propósito general, el flujo máximo de datos es del orden de algunos miles de muestras por segundo. Arriba de estos rangos el sistema será incapaz de responder. Con el fin de seguir empleando la flexibilidad de las computadoras de propósito general, es necesario aumentar su capacidad de procesamiento mediante arreglos de procesadores adicionales dedicados para la aplicación. Este tipo de hardware puede procesar desde unos cuantos millones de instrucciones hasta miles de millones de instrucciones por segundo.

Cuando no se necesita la flexibilidad de una computadora de propósito general, ya se tienen sistemas completos de procesamiento digital con un subsistema de adquisición de datos y controladores de salida para interfaces a una computadora personal y/o a dispositivos de visualización.

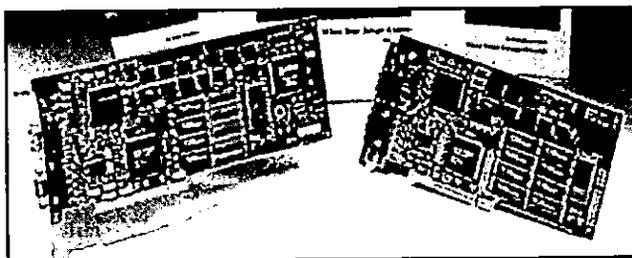


Fig I-2.14 Sistema de procesamiento digital.

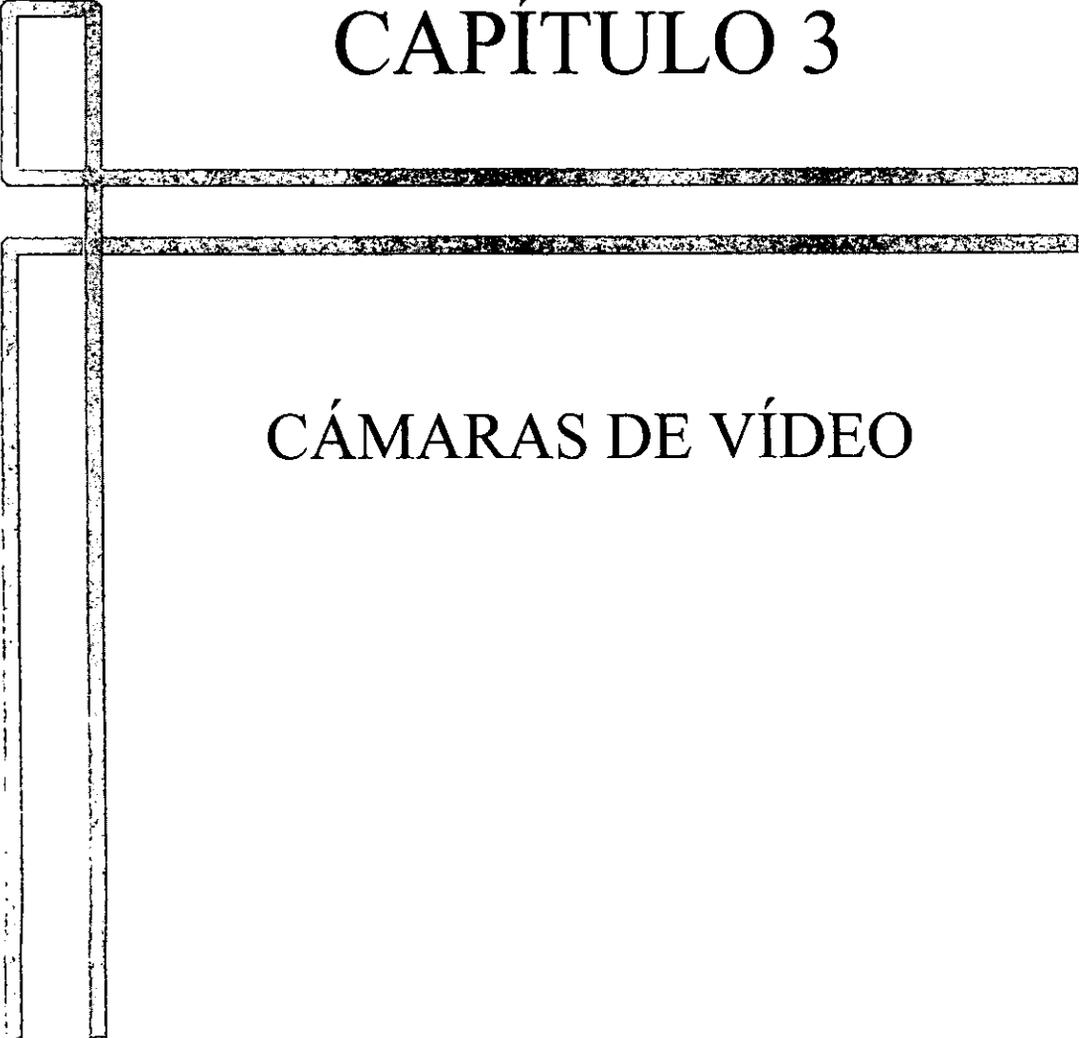
2.6.1 Requerimientos de hardware para el procesamiento digital de vídeo

En aplicaciones de vídeo es común encontrarse con algoritmos específicos para lograr algún efecto determinado. Por ejemplo, donde una imagen de televisión se alarga, encoge, rota o se distorsiona en tiempo real (30 o 25 cuadros por segundo). Si el sistema es incapaz de trabajar a estas frecuencias, incorrectos filtrados producirán efectos, como el aliasing o contorno.

Al tener la señal de vídeo analógico un ancho de banda grande, el procesamiento es un poco complicado en tiempo real. El hardware debe ser capaz de procesar por lo menos uno de los campos entrelazados a 60 Hz. Para el caso en donde se emplea el estándar para

muestreo SMPTE-EBU (Society of Motion Picture and Television Engineers - European Broadcasting Union) en 13,5 Ms/s, y considerando que la línea está activa durante 53.5 microseg, aproximadamente se recogen 722 muestras por línea. Con 480 líneas activas, 10.4 millones de muestras son recolectadas por segundo. Según la norma digital, cada muestra consta de una muestra de luminancia y una muestra de las dos señales de crominancia (color). Así el proceso acaba en 20 Ms/ s.

Aunque la norma de 13,5 Ms/s es para los varios estándares de televisión existentes, ésta no satisface el muestreo para señales compuestas como la señal NTSC. La velocidad de muestreo empleada para la señal NTSC es cuatro veces la frecuencia de subportadora de color, esto es, 14.3 Ms/s. A esta velocidad se simplifica la decodificación de la señal de color. Un filtro digital es entonces necesario para convertir la velocidad de 14,3 Ms/ s a una de 13,5 Ms/ s.



CAPÍTULO 3

CÁMARAS DE VÍDEO

3.1 Introducción

Una parte indispensable en la digitalización de una imagen es la cámara, mediante la cual se convierte la imagen óptica a señal eléctrica. La característica principal en el diseño de las cámaras es la forma en que se capta la imagen óptica para transformarla en una señal eléctrica, ya sea empleando tubos fotoconductores o dispositivos de estado sólido. A grandes rasgos, un dispositivo de este tipo puede dividirse en diferentes áreas funcionales, las que podemos encontrar en todas las cámaras (figura I-3.1). Una cámara "ve" una escena natural y debe enfocarla mediante elementos ópticos.

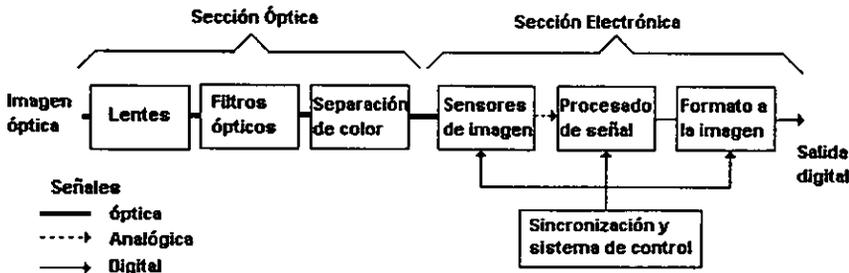


Fig. I-3.1 Diagrama de funcionamiento de una vídeo cámara.

La sección óptica se encarga de acondicionar la imagen para que llegue a la parte electrónica con mayor calidad. Esta sección requiere un análisis especial, por lo que se abordará más adelante. Por principio de cuentas iremos conociendo los dispositivos que forman una cámara.

3.2 Sensores de las cámaras o visualizadores

El dispositivo capaz de captar la imagen óptica y convertirla a una señal eléctrica es el visualizador.

La historia de los visualizadores ha pasado varios puntos cruciales en su desarrollo. El primer visualizador fue conocido como *escáner instantáneo*, éste realizaba un patrón de barrido a través de la pantalla con un punto brillante de luz, luego un fototubo recolectaba la luz reflejada para generar una señal eléctrica. Una imagen óptica puede ser escaneada mecánicamente o electrónicamente sin necesidad de crear una imagen eléctrica y un método electrónico que hace esto es el llamado *escáner de punto flotante*.

En un escáner de punto flotante para aplicaciones de lectura de películas, el haz de escaneo del TRC se enfoca sobre la película a ser escaneada, un fotodetector de alta sensibilidad es requerido para captar la luz que pasa a través de la película. Cuando el color es recolectado, la división del color se hace mediante tres fotodetectores (todo el proceso es instantáneo). La intensidad debe ser lo suficientemente alta para obtener una buena SNR, esta técnica sólo es aplicable para escanear áreas pequeñas, tal como un cuadro de película.

El proceso puede ser mejorado cuando se forma una imagen electrónica antes de escanearla, esto se puede lograr de la siguiente forma: cuando llega la luz de la imagen

óptica, se convierte a una carga eléctrica en un período completo de un cuadro escaneado. A esta acción se le conoce como "*imagen almacenada*" (storage imaging). El proceso no es instantáneo, la imagen electrónica se construye durante el tiempo de cuadro, y se va obteniendo la señal de vídeo conforme el escaneo pasa sobre cada parte de la imagen electrónica.

Basándose en lo anterior, se presentó un cambio en la manera de obtener la imagen, con la introducción de los tubos de almacenamiento fotoemisoro. En este tubo, la luz incidente de una imagen causa que la superficie fotoemisiva emita electrones en un patrón de acuerdo a la brillantez de la imagen. Además, tienen la capacidad de almacenar la energía de la luz en una carga eléctrica para un cuadro de imagen completo hasta el momento en que el proceso de escaneo o exploración se efectúa, por lo que cuentan con una gran sensibilidad.

A pesar de lo antes mencionado, estos dispositivos tienen problemas con el ruido, por lo que se reemplazaron con los tubos fotoconductores, en estos se utilizan los cambios en la resistencia eléctrica del fotoconductor cuando la luz incide. El primero y el más representativo de estos dispositivos es el "Vidicón" de la década de los 50's, el cual tenía una baja sensibilidad.

Los tubos de almacenamiento fotoconductivos físicamente son según se ve en la figura 1-3.2. La superficie frontal es un disco, donde se tienen las capas de materiales fotoconductivos. En la circunferencia de este disco se tiene un anillo conductor formando una estructura conocida como "*target*", donde se tiene un electrodo conectado al anillo. El cañón de electrones se encuentra en la parte posterior del tubo y genera los rayos de electrones para escanear el target. La imagen se forma sobre la superficie fotosensible en función de las variaciones de la luminosidad de la imagen. El encargado de transformar esto en señal eléctrica es el haz de escaneo.

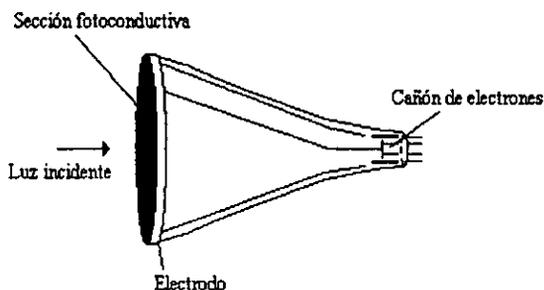


Fig. 1-3.2 Construcción de un tubo fotoconductor.

En la actualidad nos encontramos en una época de cambio al respecto, ya que los tubos se están reemplazando por dispositivos de estado sólido, como son los *dispositivos de carga acoplada* (*Charge-Coupled Devices, CCD's*). Estos dispositivos tienen ventajas superiores comparadas con los tubos al vacío, por lo que su uso se ha hecho más frecuente; un ejemplo muy claro, es el tipo de cámara que empleamos para respaldar el diseño y operación de la tarjeta. A continuación se muestran las ventajas y desventajas de los CCD's ante los tubos al vacío con base a datos anteriores a 1996:

- Los tubos tienen mayor resolución que los diseños CCD, excepto para los de HDTV. En los CCD's, la resolución es importante tanto en dirección vertical como horizontal porque se realiza muestreo en ambas direcciones. La resolución es obtenida al multiplicar el número de líneas de altura por el *factor Kell* (*kell factor*⁶). Al emplear el offset espacial, la resolución puede ser incrementada.
- Los CCD's tienen mejor definición.
- Los CCD's tienen menos retrasos que los tubos. La razón es porque en los tubos los retrasos se originan por defectos en el escaneo al no eliminar completamente la carga almacenada, mientras que en los CCD's esto no se presenta.
- Los CCD's manchan más la imagen, pero existen métodos para evitarlo. Se mancha en el proceso de transferencia de las cargas del área de recolección al área de salida.
- Los CCD' tienen mejor resolución dinámica.
- El ruido en los CCD's es generado internamente por el CCD, mientras que en los tubos la relación depende del preamplificador
- Los tubos producen menos efectos en la imagen, pero los CCD's tienen más sensibilidad, menos carga de fluorescencia (Blooming), mejor colorimetría, es menos sensible a vibraciones, son más fáciles de operar, tienen larga vida, consumen menos potencia y son de menor tamaño.

3.3 Dispositivos de carga acoplada

A grandes rasgos, un CCD es un transductor entre una imagen óptica y una señal de video escaneada. Es un dispositivo que desarrolla los dos pasos del sensado de una imagen: almacenamiento y lectura (escaneo).

Las primeras apariciones de estos dispositivos fueron en 1964 y posteriormente, fueron publicados en 1967. Un sensor de estado sólido constaba de un arreglo de diodos fotosensibles, donde cada diodo era un elemento de la imagen. La carga que se tenía presente en cada elemento era muestreada en forma vertical y horizontal (x,y). Estos circuitos demostraban la viabilidad del proyecto, aunque no solucionaban el problema en aplicaciones comerciales, ya que los primeros modelos experimentales apenas alcanzaban los 180 x 180 pixeles de resolución, lo cual es aproximadamente un décimo de lo requerido para aplicaciones de transmisión de televisión.

A principios de los 70's, se hicieron importantes avances en el desarrollo de sistemas CCD. En estos nuevos dispositivos la acción de los fotodiodos se aprovechó incorporando un sistema de transferencia y almacenamiento de carga, haciendo uso de la característica de almacenamiento de los capacitores. Cuando la carga es fotogenerada, ésta se transfiere a un capacitor de semiconductor de metal óxido (MOS) en el CCD y es almacenada para su subsecuente lectura como pixel. Con la acción de almacenamiento se evita la necesidad de tener que muestrear las cargas instantáneamente

La estructura de los fotosensores es compleja por las pequeñas dimensiones que manejan, un lado de un sensor mide típicamente menos de 0.001 pulgadas y un CCD tiene

⁶ Andrew F. Inglis - Arch C. Luther; *Video Engineering*; McGraw-Hill, EUA, 1996; cap 4, pag. 105.

500,000 o más. Como se mencionó antes, las distribución de los sensores es un arreglo en celdas de dos dimensiones, las cuales pueden ser directamente los pixeles. Cada arreglo de celdas CCD tiene dos modos: un modo de almacenamiento, donde la carga eléctrica se acumula en la celda de acuerdo a la luz incidente sobre el área sensitiva de la celda; y un modo de transferencia, donde la carga en una celda es movida a una celda adyacente en forma horizontal o vertical dependiendo de la arquitectura que se emplee .

Cuando se captura la imagen, el área de imagen en un CCD funciona en modo de almacenamiento, mientras que el área de lectura usa el modo de transferencia para mover parte de la imagen a un puerto de salida de vídeo. Una vez capturada la imagen, para moverla del área de imagen al área de lectura, se emplea el modo de transferencia, aprovechando el intervalo de blanking vertical de la señal de vídeo, es decir, la sección de la señal donde no se tiene información visible.

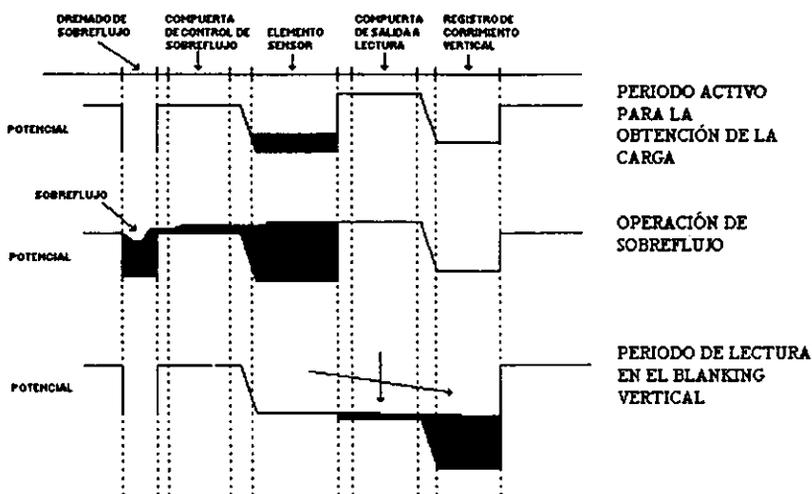


Fig. I-3.3 Operación del CCD.

El conteo de pixeles (*pixel count*), es decir, el número de renglones y columnas, es una especificación muy importante en los CCD's. Las dimensiones de cada sensor tienen que negociar entre la sensibilidad y la respuesta de apertura. Una gran sensibilidad es obtenida cuando el sensor se acerca al pixel completo, pero por otro lado, la respuesta de apertura se incrementa al reducir las dimensiones del sensor.

Para explicar el funcionamiento de un sensor haremos uso de la figura I-3.3, la cual es un mapa de los potenciales manejados por los componentes. En la primera parte, se pueden ver los electrones liberados por el fotosensor al incidir la luz sobre él. Si la sección de la imagen es demasiado brillante entonces sobrepasará el nivel de sobreflujo, lo que hará que el exceso de electrones se drene a través de la compuerta de sobreflujo, tal como se muestra en la segunda parte de la imagen. Al final del campo, el potencial de la compuerta de lectura, que es ligeramente más alto que la compuerta de sobreflujo cuando está cerrada,

es bajado y los electrones pueden pasar al registro de corrimiento como se muestra en la última parte.

3.4 Arquitecturas CCD

3.4.1 Arquitectura de transferencia de cuadro

La figura I-3.4 muestra el modo de transferencia de cuadro (FT). El área de transferencia es colocada sobre un área idéntica de lectura, pero el área de lectura es enmascarada para que no incida luz sobre ella. La imagen es capturada y almacenada en el área de imagen durante el tiempo que la imagen está activa. Cuando se presenta el intervalo de blanking vertical, la imagen se transfiere verticalmente hacia el área de lectura. Las celdas de cada línea son transferidas abajo simultáneamente una línea y este proceso es repetido para cada línea de la imagen.

Una vez que se tiene presente la imagen en el área de lectura, los intervalos de blanking horizontal se aprovechan también para transferir la línea más baja de la imagen a un registro horizontal, el cual va obteniendo la señal de video al ir leyendo en secuencia a la derecha de acuerdo a un reloj durante el tiempo de línea activa.

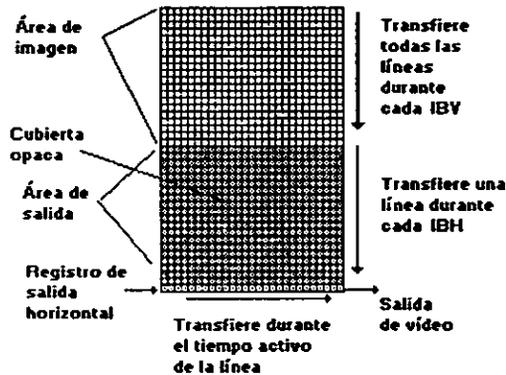


Fig. I-3.4 Arquitectura de Transferencia de Cuadro.

Sin embargo, se tiene un problema porque la acumulación de carga en el área de imagen no puede ser bloqueada cuando se lleva a cabo la transferencia hacia el área de lectura, esto produce un efecto de transferencia manchada, por lo que la imagen presenta una especie de arrastre sobre las áreas luminosas. Este problema puede ser corregido con el uso de una persiana mecánica en el sistema óptico del sensor. Se emplea de modo que pueda controlar la cantidad de luz en el sistema para prevenir sobre exposición y *blooming* (que son defectos de propagación de áreas blancas en zonas de la imagen donde las luces son muy brillantes). Generalmente el sistema de transferencia de cuadro se usa en todas las cámaras profesionales de calidad.

3.4.2 Arquitectura de transferencia interlínea

Los primeros visualizadores empleaban el concepto de transmisión interlínea. La figura I-3.5 muestra la arquitectura de transferencia interlínea (IT). Esta arquitectura alterna las columnas de los arreglos de imagen y de lectura, poniendo una columna de la zona de imagen junto a una de lectura, lo que ocasiona que el chip de esta arquitectura sea más pequeño que el de la arquitectura FT. Como la imagen óptica es enfocada sobre toda el área, las celdas de lectura deben ser cubiertas ópticamente. Obviamente se tendrá pérdida de luz por lo que la sensibilidad es reducida. Pero ya también para este problema se tiene la solución, ya que se emplea un arreglo de microlentes colocado sobre las celdas CCD, con el fin de que toda la luz incidente sea dirigida a el área sensitiva.

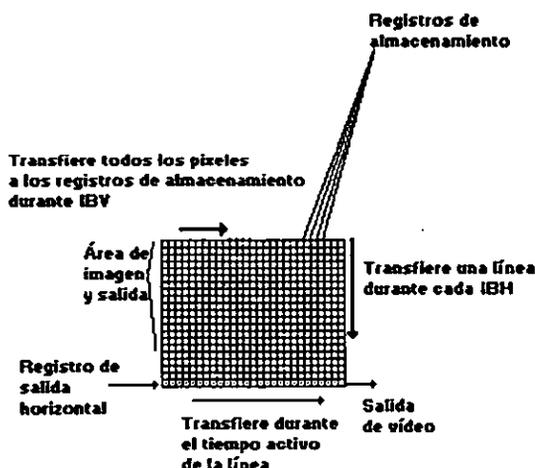


Fig. I-3.5 Arquitectura de Transferencia Interlínea.

En el periodo del blanking vertical, se transfiere la carga de todas las celdas de almacenamiento a las de lectura horizontalmente. De manera similar a la arquitectura de transferencia de cuadro, durante los intervalos de blanking horizontal, las celdas de lectura recorren su carga una línea abajo para entregarla al arreglo de lectura horizontal, en donde a través de un reloj durante el periodo activo se entrega la señal de video.

El problema existente en esta arquitectura es que el bloqueo de las celdas de lectura entrelazadas no puede ser perfecto. Esto causa efectos de línea vertical que aparecen en las zonas luminosas.

3.4.3 Arquitectura de transferencia cuadro-interlínea

Esta arquitectura combina las dos anteriores y se muestra en la figura I-3.6. Durante el blanking vertical, la imagen entera es transferida hacia los registros de transferencia intercalados y entonces transferida abajo al arreglo de lectura. Durante el tiempo activo, la lectura ocurre de la misma forma que en la arquitectura FT. La característica distintiva de esta arquitectura es que cuenta con un renglón de compuertas de selección de las áreas de

imagen y de lectura. Estas compuertas son manejadas para vaciar el exceso de carga antes de que se efectúe la transferencia, evitando así el uso de la persiana mecánica.

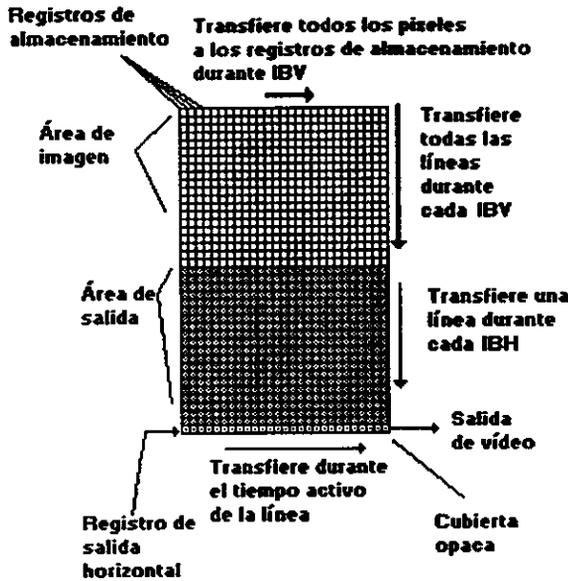


Fig 1-3.6 Arquitectura de Transferencia Cuadro-Interlínea.

3.5 Escaneo entrelazado en las CCD's

Existen varios métodos posibles para obtener la imagen entrelazada. Por un lado, la transferencia cuadro interlínea tiene la ventaja de que puede ser manejada de tal forma de que sólo transfiera líneas alternadas hacia el registro de salida horizontal, logrando así el entrelazado de campos. Esto da una alta sensibilidad y ocasiona que sólo sean necesarias en el área de lectura la mitad de las líneas totales. El problema que se tiene con esta técnica es, que se producen efectos entrelazados normales sobre movimientos rápidos y el aliasing vertical se verá sobre formas de niveles horizontales, esto se puede corregir empleando un filtro óptico.

Para obtener el entrelazado en otras formas es necesario que la frecuencia de escaneo sea duplicada para obtener la frecuencia de campo (1/60 de segundo). Se pueden enviar líneas saltadas a la salida del CCD y expandir las líneas restantes en el tiempo correspondiente al siguiente campo usando una línea de memoria. Ya que se está desperdiciando una mitad de la señal, se presenta una pérdida de la relación de señal a ruido. Otra forma es almacenar en memoria dos líneas del CCD y combinarlas para obtener una señal entrelazada. Esto mejora la SNR y el aliasing, pero por otro lado produce pérdida de resolución vertical. Ambos métodos de escaneo progresivo en el CCD requieren que los registros trabajen al doble de la velocidad requerida por el entrelazado directo.

3.6 Offset espacial

La resolución máxima del dispositivo está limitada por la estructura de las celdas de muestreo. Para el mejoramiento de la resolución, la construcción de las cámaras a color ha establecido una relación entre las posiciones de las columnas R, G y B con visualizadores individuales.

Como se muestra en la figura I-3.7, el offset espacial es obtenido al correr los dispositivos de rojo y azul horizontalmente por un medio del ancho de pixel comparado con el dispositivo verde. Ya que la señal de Luminancia combina el 59% de verde, el 41% de rojo y azul, el efecto es tal que el muestreo es el doble que la frecuencia espacial de un CCD. Se requiere que el ancho de los sensores sea menor de un medio del ancho de un pixel para evitar conflictos de traslape al unirlos.

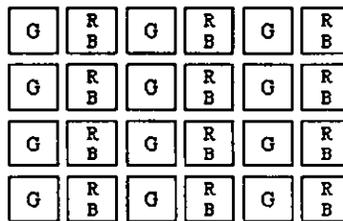
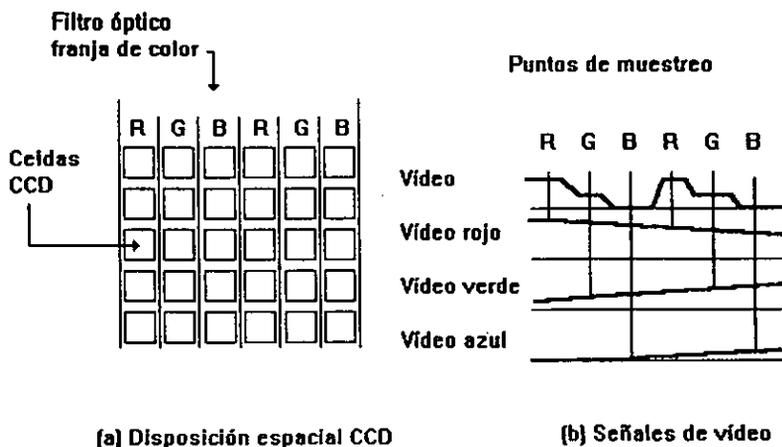


Fig. I-3.7 Offset Espacial en un CCD.

3.7 Cámara de sensor sencillo

Este tipo de cámara es más económico que las cámaras con tres sensores y se emplea en aplicaciones que no requieren alto desempeño. Su distribución se muestra en la figura I-3.8.



(a) Disposición espacial CCD

(b) Señales de vídeo

Fig I-3.8 Cámara de color CCD.

Es necesario emplear un filtro óptico para separar los colores, porque las celdas adyacentes corresponden a diferentes colores. En el procesamiento electrónico de señales, los colores RGB son recuperados al muestrear la señal de vídeo en tres fases como se muestra en la figura.

Ya que se emplea un tercio de los píxeles para cada color primario se reduce la señal por canal en un factor de tres. Una forma de mejorar esta situación, es dar la mitad de los píxeles al verde y asignar alternadamente al rojo y el azul. Esto complica la recuperación de la señal pero mejora la resolución verde a un medio de la resolución CCD desnuda.

3.8 Colorimetría de las cámaras

Para obtener una buena calidad en el manejo de las componentes RGB, es necesario conservar la colorimetría, lo cual consiste en cumplir con un grupo de gráficas estándar de la respuesta espectral para cada color, tal y como se muestra en la figura I-3.9 en NTSC.

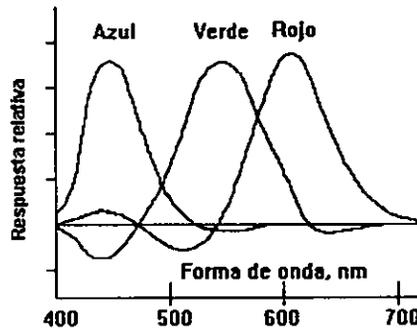


Fig. I-3.9 Espectro de las componentes RGB en NTSC.

Según se puede ver en las formas de onda, cada curva correspondiente a cada color tiene un pico y tiene una parte en la zona negativa de la gráfica, estas segundas zonas sirven para obtener la colorimetría correctamente, pero no pueden ser producidas por los sensores, incapaces de producir salidas negativas. Los lóbulos negativos son producidos por un proceso electrónico llamado "masking", que consiste en una combinación lineal de los tres colores primarios.

El masking puede ser usado para ayudar en el ajuste de la colorimetría en las cámaras, pero es un proceso complicado que no es práctico para desarrollarse manualmente, porque no existe relación entre los valores de los coeficientes y el comportamiento visible de la imagen. Por esta razón, las cámaras tienen un masking fijo.

3.9 Relación Señal a Ruido CCD

Para el caso de los visualizadores de tubo fotodetector, la relación señal a ruido depende principalmente del preamplificador. Sin embargo, para el caso de los

⁷ Arch C. Luther; Principles of Digital Audio and Video; Artch House, Inc., EUA, 1997; cap 6, pag. 135.

visualizadores CCD depende directamente del funcionamiento interno del dispositivo. Por el hecho de manejar los electrones, es decir, al almacenarlos y transferirlos, en el dispositivo se genera ruido térmico. Por otro lado existe ruido también debido a defectos en el proceso de fabricación. Los niveles de la señal de salida son bastante altos de modo que superan el ruido térmico que pudiese existir, así que éste, en circuitos amplificadores, no es un problema grave.

El desempeño de SNR de un CCD depende del área del pixel, mientras más grandes mejores. Sin embargo, pixeles grandes requieren un chip grande, óptica grande y más luz lo cual incrementa el costo.

El rango dinámico, es decir, la relación entre la salida más alta posible del CCD a el nivel de ruido, es también de consideración, por lo que los CCD's son calibrados de modo de que no lleguen a su nivel de saturación, para proteger de malos efectos de operación en zonas brillantes.

3.10 Sistema óptico de las cámaras

Como se dijo al principio de este capítulo, la sección óptica juega un papel importante dentro del desempeño de una cámara.

La cámara necesita que el objeto a visualizar se encuentre iluminado, misma situación que requiere el ojo humano. Si se tiene una fuente de luz que incida sobre el objeto en cuestión, éste refleja parte de esa luz y el reflejo es captado por el ojo y procesado por el cerebro. Es por eso que nosotros vemos los objetos de acuerdo a la luz que reflejan, entonces, la forma de un objeto es una función compleja de la iluminación, la reflectancia del objeto y de las propiedades físicas y psicofísicas de la visión humana. A este respecto, la cámara trata de emular este funcionamiento de la visión humana.

Como se ha bosquejado antes, para poder visualizar un objeto, una parte importante es la iluminación, ésta se puede obtener de varias maneras. La manera más fácil de obtenerla, es de la naturaleza, con ayuda del Sol, el cual irradia un amplio espectro de energía, no sólo el espectro visible, sino también el rango de infrarrojo y ultravioleta, aunque para el caso de la imitación del ojo, sólo es necesario poner atención en el rango visible, es decir, de 400 a 700 nm de longitud de onda. Otra forma de obtener la iluminación deseada es empleando fuentes de iluminación artificial, como las luces incandescentes o las fluorescentes. Para clasificar estas fuentes se hace uso de la densidad de energía en un escena, dada en watts por centímetro cuadrado. Una forma más práctica es hacer uso de la intensidad luminosa, es decir, la frecuencia del flujo de energía luminosa por segundo; el flujo de energía luminosa se mide en *lumens*, por lo que la intensidad luminosa se mide en lumens por unidad de ángulo sólido. Pero en este caso, la intensidad luminosa en el ojo varía con el ángulo en el cual se encuentra la fuente con respecto a la escena.

Por otro lado, ya que la cantidad de energía que incide sobre una escena cambia con el cuadrado inverso de la distancia entre la fuente y la escena, se expresa la luz en un escena en términos de lumens por metro cuadrado. La unidad "lux" ha sido definida para este propósito.

$$1 \text{ lux} = 1 \text{ lumen} / \text{m}^2$$

Y su unidad equivalente, el pie-candela;

1 pie-candela = 1 lumen/ pie²

Cabe destacar que la sensibilidad de una cámara de vídeo está usualmente expresada en términos de lux.

La brillantez de una superficie es el resultado de toda la iluminación incidente y la reflectancia de la superficie. La unidad que se emplea para medir la brillantez es el nit.

1 nit = 1 candela / m²

Una característica importante del ojo humano es de que éste es capaz de adaptarse a una gran variedad de distribuciones de iluminación, pero en el caso de una cámara, no tiene la capacidad de adaptarse como el ojo, es por ello que debe hacerse electrónicamente, esto se conoce como *balanceo de blanco* (*balancing white*⁸). Para un sistema RGB el blanco se obtiene a una cantidad igual de las tres señales rojo, verde y azul, en un sistema Luminancia-Crominancia, el blanco corresponde al cero en ambas diferencias de color.

En la figura I-3.10 se muestra un sistema óptico para una cámara a color con tres CCD's.

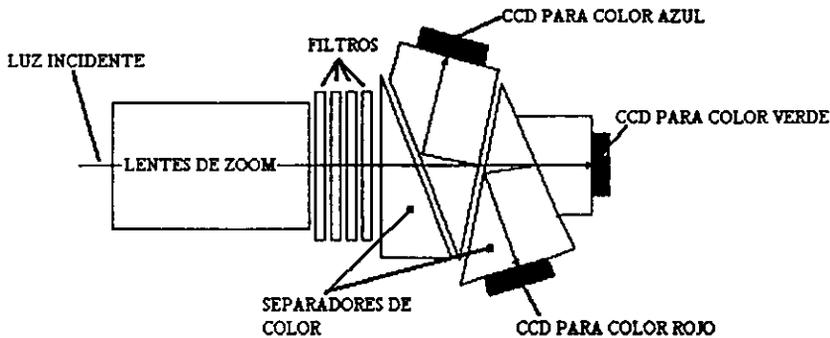


Fig. I-3.10 Sistema Óptico de las cámaras.

3.10.1 Lentes

Los primeros elementos ópticos en este sistema son los lentes. Con ayuda de los lentes se puede acondicionar la imagen para que pueda ser mejor captada por la cámara en la superficie fotosensible. Partiendo de que la luz es una emisión de rayos, podremos entender mejor su funcionamiento empleando óptica geométrica.

En el diseño de lentes para cámaras se emplean sus características de reflexión y refracción, que se presentan cuando un haz de luz incide a través de distintos materiales, para nuestro caso el vidrio. Anteriormente, las cámaras tenían mecanismos sencillos de lentes manuales de longitud focal fija, pero actualmente, las cámaras de vídeo modernas tienen lentes de longitud variable (zoom), esto es, un sistema de varios lentes, que pueden

⁸ Arch C. Luther; Principles of Digital Audio and Video; Artch House, Inc., EUA, 1997; cap 6, pag. 125.

llegar a ser de 10 a 20 elementos de lentes, los cuales son ajustados mecánicamente para obtener una longitud focal variable.

Cabe destacar que la longitud focal relaciona la distancia del lente al objeto y la distancia del lente a la imagen:

$$1/S_0 + 1/S_1 = 1/F \quad (I-3.1)$$

donde:

S_0 es la distancia objeto.

S_1 es la distancia imagen.

F es la longitud focal del lente.

En la práctica se puede considerar la distancia objeto infinita entonces, la ecuación se puede reducir a:

$$S_1 = F \quad (I-3.2)$$

Para el tamaño del plano de la imagen, como para el tamaño de la superficie sensible CCD, la longitud focal determinará el ángulo de vista de la cámara. Longitudes focales largas van a dar pequeños ángulos de vista.

Entre los parámetros más importantes de unos lentes de zoom, está su relación de zoom más alta de la longitud focal más baja. La mayoría de las cámaras de vídeo tienen relación de zoom óptico de 6x a 15x, aunque algunos llegan a ser de 40x o más. Relaciones de zoom más altas llegan a ser obtenidas electrónicamente en algunas cámaras, aunque es usual tener un umbral con respecto a la resolución. La función de zoom es controlada electrónicamente donde la longitud focal puede ser ajustada por el operador o remotamente.

3.10.2 Apertura y número f

El diámetro efectivo del lente controla la capacidad de recolección de luz y se conoce como apertura; las cámaras tienen controles de apertura. La apertura se representa mediante el uso del número f, el cual es la relación entre la longitud focal con el diámetro de apertura. Pero en la práctica se emplea una razón del número f de la raíz de 2 por 1.4 y representa el cambio 2:1 en la intensidad de luz por medio de los lentes.

3.10.3 Profundidad del campo

En un lente colocado a cierta distancia de la escena se tiene un plano donde la imagen es enfocada detrás del lente de acuerdo con la ecuación de los lentes. Pero, para puntos de la escena más lejanos o más cercanos que la distancia del plano enfocado al lente, estarán progresivamente fuera de foco en el plano de la imagen, tal y como se muestra en la figura I-3.11. El diámetro del flujo de rayos de luz desde este punto fuera del plano, cuando pasan a través del plano enfocado, es conocido como *círculo de confusión* (c-of-c). Se puede ver que el tamaño del c-of-c es una función de la apertura del lente A , la distancia objeto S_0 y la distancia D_1 . En el caso de una cámara CCD la profundidad de campo

representa la distancia en la escena dentro de la cual el c-of-c permanece menor que el tamaño de un pixel proyectado sobre la escena.

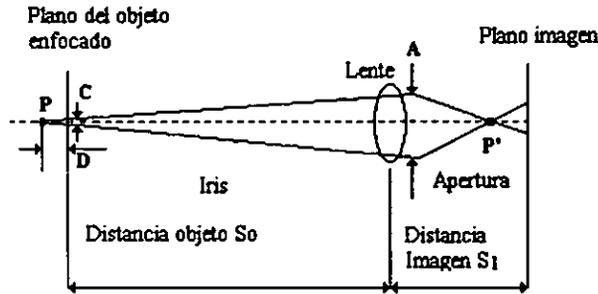


Fig. I-3.11 Profundidad de campo.

3.10.4 Aberraciones

En un lente se pueden presentar defectos al enfocar un objeto, que pueden ser molestos al presentar la imagen y se incrementan conforme los rayos de luz incidentes en el lente se alejan del centro del mismo. Estos defectos en la imagen se conocen como *aberraciones*, existiendo cuatro tipos:

- Geométricas.
- De foco.
- Cromáticas.
- De intensidad.

En general, las aberraciones se incrementan con el tamaño de la apertura y con el campo angular de vista.

La aberración geométrica depende del ángulo de vista, lo cual causa en las esquinas de la imagen un alargamiento o compresión geométricamente.

Las aberraciones de foco son:

- Aberración esférica, causada porque los rayos que pasan a través de la lente no se reúnen en el mismo punto, lo que causa un efecto borroso.
- El astigmatismo, donde el lente no puede enfocar a un punto fino, por lo que va de una elipse en una dirección a una elipse perpendicular cuando el foco es ajustado.
- La coma, similar a la aberración esférica en que los rayos fuera de los ejes no enfocan en el mismo lugar como los rayos sobre los ejes.
- La curvatura de campo, donde el plano de mejor foco no es un plano sino una superficie curvada.

Aberraciones de color producen corrimientos del color en la imagen y se presentan cuando el comportamiento del lente varía con la longitud de onda.

La aberración de intensidad es la variación de la intensidad de la imagen con el ángulo de vista. Esto es algunas veces llamado "*vignetting*".

3.10.5 Filtros ópticos

En el sistema óptico de la cámara también se incluyen filtros ópticos, los cuales tienen funciones distintas, se emplean filtros separadores de color que dejan pasar un cierto rango de longitudes de onda, permiten obtener los colores primarios de una manera más pura, ayudando a las superficies dicróicas (descritas más adelante).

Se cuenta con filtros que se ajustan a la temperatura característica de la fuente de luz, ya sea de luz de día o incandescente. Esta propiedad es necesaria para conservar la colorimetría.

Los objetivos del filtrado óptico son: primero obtener la mejor aproximación para los lóbulos positivos de las curvas mostradas para la colorimetría y segundo obtener la más alta eficiencia óptica y un aceptable nivel de luz en los sensores.

Se tienen filtros de densidad neutral, los cuales sirven para reducir la necesidad de grandes variaciones del número f y de la profundidad de foco en escenas con cambios grandes en la brillantez. Otros filtros necesarios para limitar un poco el efecto de aliasing en la imagen son los filtros paso bajas, los cuales dejan pasar la luz abajo de una cierta frecuencia.

3.10.6 Separación del color

La parte final en el sistema óptico de una cámara a color es el sistema de separación de color. Anteriormente, esto requería de un complicado mecanismo de espejos y lentes.

Ahora, aprovechando la refracción y reflexión de la luz se puede hacer la separación de colores para el caso de los sistemas CCD a color. Las superficies dicróicas son un recubrimiento de dos capas que tiene un espesor de $\frac{1}{4}$ de la longitud de onda de la separación de color deseada. Estos filtros pueden ser diseñados para reflejar longitudes de onda arriba o abajo de la frecuencia típica, pero no ambas al mismo tiempo, en la figura 1-3.12 se muestra un filtro dicróico.

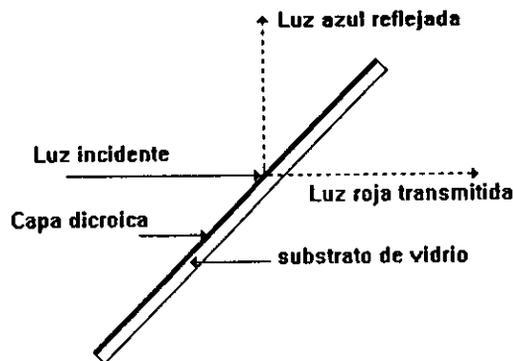


Fig 1-3.12 Operación de un filtro Dicróico.

3.11 Sistemas de vídeo básicos

3.11.1 Definiciones

Definiciones:

1. Campo de vista (FOV):

Es el área del objeto que va a ser visualizada en el monitor.

2. Resolución:

La característica más pequeña capaz de ser resuelta.

3. Distancia de trabajo mínima y máxima:

Medida del lente objetivo a el objeto.

Resolución de línea de TV. Las resoluciones del monitor y los CCD's están dadas frecuentemente en términos de "la resolución de línea de TV".

Resolución de cuenta de pixel. (número de pixeles). La mayoría de los CCD's no son columnas y renglones perfectos de pixeles. Frecuentemente cuentan con espacios muertos entre ellos. La resolución de los pixeles es una característica importante cuando se quiere aprovechar las capacidades de los dispositivos de captura de vídeo, aquí es necesario comparar dicha resolución con respecto a la cámara que se va a usar.

3.11.2 Formato CCD

La relación de aspecto para Dispositivos de Carga Acoplada (CCD) es 4:3. El tamaño del sensor no es equivalente para el área activa del sensor (figura I-3.13). Los lentes de vídeo pueden ser usados con cualquier cámara CCD, así como el formato de diseño de los lentes es más largo o igual al de la cámara.

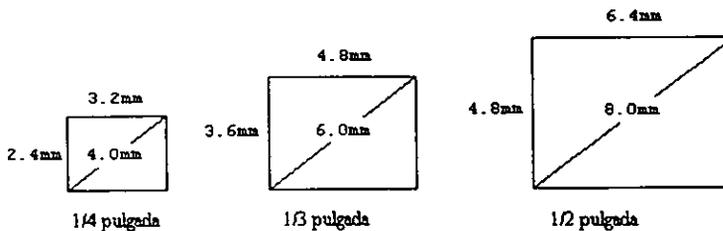


Fig. I-3.13 Formato CCD.

3.11.3 Sensibilidad de una cámara CCD

Abajo encontraremos una curva de respuesta típica para CCD's de silicon en blanco y negro (figura I-3.14). Las cámaras de color típicamente vienen con un filtro de bloqueo IR que corta en 700 nm.

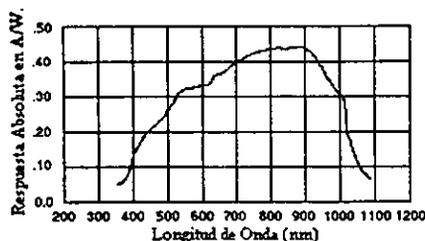


Fig. I-3.14 Respuesta típica de una cámara CCD.

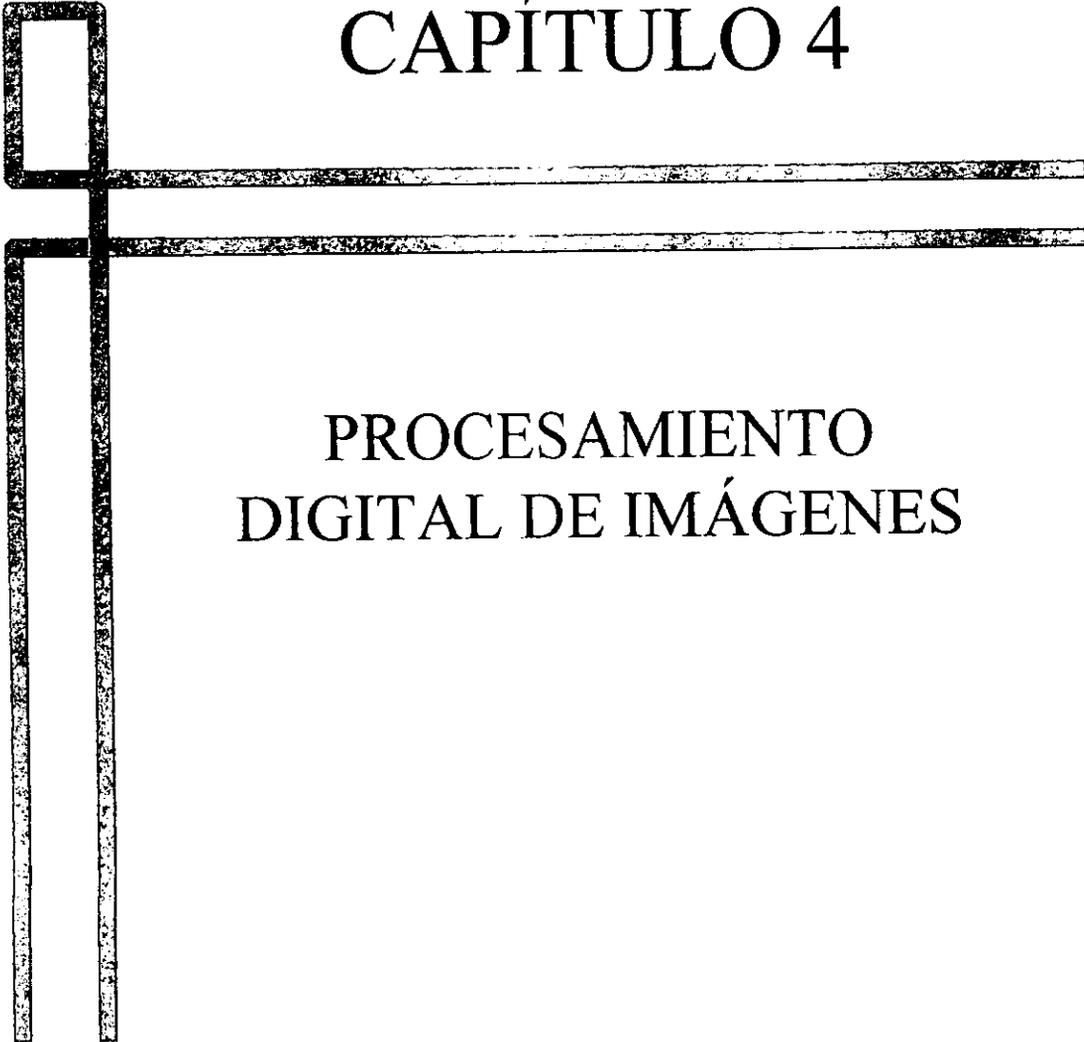
3.11.4 Formatos de señales

Existen cuatro tipos de señales básicas usados en cámaras CCD, la Compuesta (NTSC, EIA), Y-C(S-video), RGB y Digital (RS-422). Las señales NTSC (RS-170^o/Color) y EIA (RS-170/Monocromo) son las más comunes y se acomodan a la mayoría de las aplicaciones. La Y-C y RGB separan las imágenes en componentes y en consecuencia dan una superior calidad de la imagen para reproducción del vídeo y análisis de la imagen. Las cámaras digitales dan un nivel de desempeño que las hacen únicas. Usadas en conjunción con tarjetas de captura de imágenes, las cámaras digitales no sufren de los estándares impuestos para formato de vídeo. El resultado es una gran flexibilidad en la adquisición de las imágenes y calidad. En cualquier sistema electrónico, el formato de la señal podría ser constante.

Cada formato de señal de vídeo corresponde a un cable específico de conectores, las señales compuestas pueden usar ya sea conectores de tipo BNC o RCA. La Y-C usa un tipo Pin-DIN y la RGB usa cuatro conectores BNC.

3.11.5 Control de Ganancia Automático (AGC)

Ésta es una opción que automáticamente compensa para niveles de luz bajos. El Control de Ganancia Manual es una opción similar que es usada para ajustar la cámara a cambios en el ambiente del sistema. Incrementando la ganancia puede incrementarse la calidad de la imagen, al incrementar el contraste y la brillantez.



CAPÍTULO 4

PROCESAMIENTO DIGITAL DE IMÁGENES

4.1 Revisión

4.1.1 Introducción

Algunas de las aplicaciones que hacen uso de imágenes, ya sean fijas o en movimiento, requieren de algún tipo de tratamiento. Por ejemplo, para almacenamiento y/o transmisión, es necesario hacer una compresión de la información, con el fin de reducir el espacio en disco y el tiempo de transmisión, conservando la información relevante de las imágenes.

En el procesamiento de imágenes se cuenta con varias técnicas para conseguir un fin determinado. Obviamente, la elección del algoritmo va a depender de las necesidades del sistema, en cuanto a velocidad de procesamiento, capacidad de almacenamiento, etc. Todo esto involucra el uso, desde algoritmos sencillos y fáciles de implementar en estaciones de trabajo genéricas, hasta algoritmos complejos que necesitan de un hardware extra capaz de mejorar el rendimiento.

Enseguida a manera de ejemplo se muestran algunos usos del procesamiento de imágenes y se muestra un ejemplo sencillo de procesamiento.

4.1.2 Gráficos e imágenes animadas por computadora

La máquina puede ser capaz de manipular imágenes ya existentes o sintetizar otras completamente nuevas.

La animación por computadora da vida a las imágenes. Muchas de sus aplicaciones son obtener una representación dinámica que se parezca en lo posible al movimiento natural de los objetos. Un punto importante es que, una animación real se obtiene imitando el proceso físico que produce el efecto de las imágenes en movimiento. Es por esto que las reproducciones más complejas toman en cuenta, entre otras, la distribución tridimensional de las masas y las propiedades materiales de los objetos.

4.1.3 Comunicación por imágenes

Éste es uno de los campos de aplicación más importantes en la actualidad. Por el momento, se tiene la visión de que las futuras redes digitales de banda ancha permitirán la retransmisión de TV y la comunicación visual bidireccional con imágenes de alta calidad. La velocidad de transmisión de datos de una señal TV digital de calidad equivalente a las normas actuales es de 216 Mbits/s. La velocidad de transmisión de una HDTV (televisión de alta definición) de estudio será superior a 1000 Mbits/s.

Un punto importante es que se puede reducir la velocidad de transmisión necesaria para imágenes en movimiento sin perder necesariamente calidad, aprovechando algunas características comunes entre imágenes. Esto se logra porque las imágenes animadas se forman según las leyes de la Física, por lo que es posible encontrar cierta redundancia en las imágenes; además, se aprovechan las características del ojo humano.

En los últimos años se ha buscado la compresión de las señales de imágenes animadas, a fin de permitir una transmisión por la RDSI (red digital de servicios integrados) (2^{64} kbits/s). Con el desarrollo de procedimientos de compresión de alta

eficiencia, será posible ofrecer también amplios servicios de videotelefonía y de videoconferencia, sin inversiones adicionales en la red existente. Para lograr esto, es necesario un factor de compresión superior a 3000 para transmitir por ejemplo, una señal de TV digital por la RDSI. Este factor sólo puede obtenerse hoy en día limitando la resolución en el tiempo y el espacio, lo que implica una pérdida perceptible de la calidad de imagen en las secuencias muy rápidas.

4.1.4 Tratamiento de las imágenes

Por el momento, el tratamiento de imágenes en tiempo real es difícil de obtener, debido a que se necesita una velocidad de procesamiento bastante considerable. Es por ello que en El Instituto de Investigación de Stanford, por ejemplo, se desarrollaron algunos algoritmos en la década de los 70's que permiten tratar una imagen en tiempo real. Estos algoritmos reciben el nombre de algoritmos SRI (Stanford Research Institute). Para el uso de estos algoritmos debe cumplirse que:

1. Los algoritmos se tratan sobre una imagen binaria.
2. Los objetos que se visualizan no han de estar en contacto. De ser así, pasarían a considerarse una sola pieza.
3. Las imágenes serán fijas.

Un ejemplo fácil y demostrativo es: partamos de una figura geométrica, queremos que nuestro sistema pueda reconocerla. La imagen se trata en su forma más simple, por lo que eliminaremos toda aquella información que no sea útil y dejaremos solamente, aquella que pueda servir. El proceso que resalta sólo la información efectiva es la segmentación. En la mayoría de los casos el método de segmentación escogido es el de contorno. Para aplicarlo en su forma más sencilla transformaremos la imagen en binaria, es decir, sólo se admitirán dos tonalidades, blanco y negro. Para ello no tenemos más que aplicar un umbral al *histograma* (*histogram*⁹) referente a esa imagen. Así por ejemplo, si trabajamos dentro de un rango de diferentes tonalidades de gris pasaremos de 256 diferentes tonos a sólo dos, el 1 y el 0.

Para obtener el contorno de la imagen, realizaremos las siguientes operaciones:

1. Barreremos todas las filas, de arriba a abajo y de izquierda a derecha. Para cualquier pixel $P(x,y)$ tendremos que:

$$P(x,y)=0 \quad (I-4.1)$$

si $P(x,y-1)$ y $P(x,y)$ toman el mismo valor. De no ser así,

$$P(x,y)=1 \quad (I-4.2)$$

2. De forma semejante operaremos con las columnas y tendremos que:

$$P(x,y)=0 \quad (I-4.3)$$

si $P(x-1,y)$ y $P(x,y)$ toman idénticos valores. En caso contrario,

$$P(x,y)=1 \quad (I-4.4)$$

⁹ Theo Pavlidis; *Algorithms for Graphics and Image Processing*; Computer Science Press, EUA, 1982; pag. 50-51.

Después de aplicar estos dos barridos obtendremos una imagen que forma el contorno, aunque se produce además de un desplazamiento una deformación del objeto tenemos las características principales (figura I-4.1).

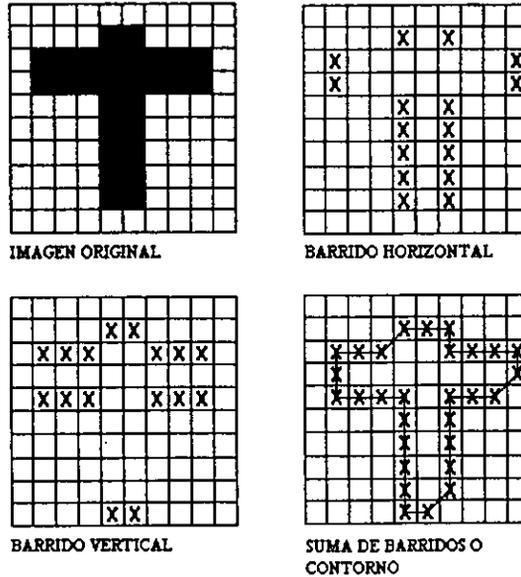


Fig.I-4.1 Detección de contorno de un objeto.

4.2 Morfología Matemática

La morfología matemática estudia las propiedades estructurales relacionadas con la forma de los objetos existentes en una imagen.

El análisis teórico de la morfología matemática se hace basándose en la teoría de conjuntos. La descripción hecha aquí está enfocada a trabajar con imágenes binarias (1=blanco, 0=negro), es decir, los conjuntos en cuestión pertenecen al espacio entero de dos dimensiones Z^2 , donde cada elemento de un conjunto está representado por un vector de dos dimensiones en las coordenadas (x,y) de un pixel negro o blanco, según sea el caso, sobre la imagen. Una imagen digital en escala de grises puede ser representada como un conjunto en Z^3 . En este caso, dos de las componentes del vector representan la posición y la tercera corresponde a la intensidad del tono. Sin embargo, el presente análisis puede extenderse a imágenes en escala de grises o en otros espacios más complejos.

4.2.1 Dilatación y Erosión

La morfología matemática cuenta con dos operaciones principales que son la base para generar las otras operaciones disponibles, tales operaciones básicas se conocen como: *la dilatación y la erosión*.

Como se mencionó antes, es necesario puntualizar algunos detalles de la teoría de conjuntos para poder comprender dichas operaciones.

Sean A y B conjuntos en Z^2 con componentes $a = (a_1, a_2)$ y $b = (b_1, b_2)$ respectivamente. Entonces, la traslación de A por $x = (x_1, x_2)$, denotado por $(A)_x$ está definida como:

$$(A)_x = \{c | c = a + x, \forall a \in A\} \quad (I-4.5)$$

la reflexión de B, denotada como \hat{B} , está definida de la siguiente forma:

$$\hat{B} = \{x | x = -b, \forall b \in B\} \quad (I-4.6)$$

El complemento del conjunto A está definido como:

$$A^c = \{x | x \notin A\} \quad (I-4.7)$$

Finalmente, la diferencia de dos conjuntos A y B, denotado por A - B, está definida como:

$$A - B = \{x | x \in A, x \notin B\} = A \cap B^c \quad (I-4.8)$$

4.2.1.1 Dilatación

En palabras sencillas de comprender, la dilatación consiste en expandir un objeto dentro de una imagen, pasando alrededor de su contorno un objeto que controlará la dilatación, tal objeto es conocido como *elemento estructurante*.

La dilatación de A por B, denotada por $A \oplus B$, está definida por:

$$A \oplus B = \{x | (\hat{B})_x \cap A \neq \emptyset\} \quad (I-4.9)$$

En la operación con conjuntos la dilatación puede entenderse como: siendo B el elemento estructurante, se hace la reflexión de éste sobre su origen (\hat{B}), la reflexión se traslada pixel a pixel por toda la imagen A, si en algún punto cualquier pixel de la reflexión coincide con un pixel de valor 1 dentro de A, se colocará en la posición que ocupa el origen de (\hat{B}) un 1, al final del recorrido la dilatación estará formada por todos los 1's que se encontraron, esto es:

$$A \oplus B = \{x | [(\hat{B})_x \cap A] \subseteq A\} \quad (I-4.10)$$

4.2.1.2 Erosión

De manera parecida a la dilatación, el objetivo de la erosión es el de contraer el contorno de los objetos dentro de una imagen en función de la forma de otro objeto conocido como elemento estructurante.

La erosión de A por B, denotado por $A \ominus B$, está definida como:

$$A \ominus B = \{x | (\hat{B})_x \subseteq A\} \quad (I-4.11)$$

Para obtener la erosión de A por B, se hace pasar el elemento estructurante por toda la superficie de A, para que un 1 sea puesto en la posición que ocupa el origen de B sobre A e ir formando así la erosión, es necesario que todos los elementos de B estén contenidos sobre valores 1 de A. Al final la erosión estará formada por todos los elementos de valor 1 que se hayan encontrado.

Cabe señalar que la dilatación y la erosión son duales entre sí con respecto al conjunto complementario y a la reflexión, esto es:

$$(A \ominus B)^c = A^c \oplus \hat{B} \quad (I-4.12)$$

4.2.2 Opening y Closing

Como se estableció anteriormente, la dilatación expande el conjunto, mientras que la erosión lo contrae. Las operaciones que describiremos a continuación están basadas en las anteriores definiciones, estas operaciones son conocidas como *opening* y *closing*.

El *opening* es una operación que tiene como efecto general un suavizado del contorno de una imagen, rompiendo los istmos delgados y eliminando las protuberancias delgadas. El *closing* también tiende a suavizar las secciones de contornos, pero en forma opuesta al *opening*, tiene un efecto de fusionar las rupturas angostas y los golfos delgados largos, elimina los pequeños agujeros, y rellena aberturas del contorno.

El *opening* de un conjunto A por un elemento estructurante B, denotado por $A \circ B$, está definido por:

$$A \circ B = (A \ominus B) \oplus B \quad (I-4.13)$$

Esto es, el *opening* de A por B hace primero la erosión de A por B y el resultado es dilatado por B.

El *closing* del conjunto A por el elemento estructurante B, denotado por $A \bullet B$, está definido por:

$$A \bullet B = (A \oplus B) \ominus B \quad (I-4.14)$$

en este caso se puede ver que de manera opuesta al *opening*, el *closing* es la dilatación de A por B y sobre el resultado se le aplica la erosión por B.

Los resultados de estas operaciones pueden verse ejemplificados en la figura I-4.2. El proceso de *opening* se ilustra de las figura I-4.2 a) a la I-4.2 e), en este procesamiento podemos darnos cuenta que las esquinas exteriores han sido redondeadas, mientras que las esquinas interiores no fueron afectadas. El *closing* en las figuras I-4.2 f) a I-4.2 i) muestra que las esquinas interiores fueron redondeadas, mientras que las esquinas exteriores ahora no fueron afectadas.

Como en el caso de la dilatación y la erosión, el *opening* y el *closing* son duales¹⁰ entre sí con respecto a la complementación y a la reflexión. Se puede demostrar que:

$$(A \bullet B)^c = (A^c \circ \hat{B}) \quad (I-4.15)$$

Estas propiedades fueron empleadas en algunos de los programas que se hicieron para ejemplificar el procesamiento digital que puede hacerse con ayuda de la tarjeta (véase capítulo de resultados). Las operaciones de *opening* y *closing* pueden ser usadas para construir filtros morfológicos.

¹⁰ Y. Mahdavi; Advances in Image Analysis; Spie Optical Engineering Press. USA. 1992; pag. 283-285

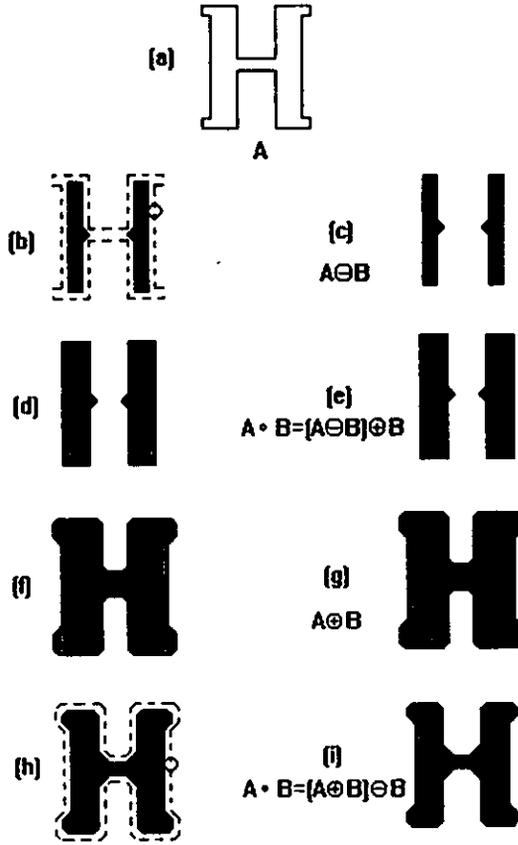


Fig. 1-4.2 Opening y Closing.

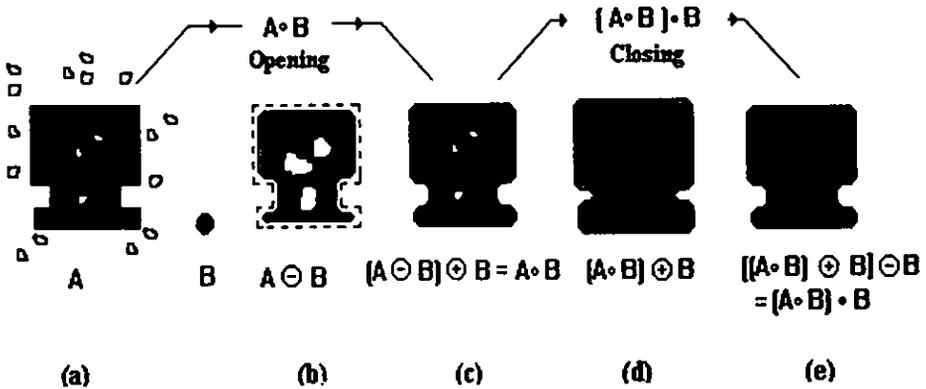


Fig. 1-4.3 Ejemplo de filtrado morfológico sobre una imagen ruidosa.

Por ejemplo, en la figura I-4.3 a) hay que ver que el conjunto A consiste de un objeto con ruido de fondo y ruido en su interior. El objetivo es eliminar el ruido sobre el objeto, tanto como sea posible, empleando un filtro morfológico $(A \circ B) \bullet B$. Al aplicar el opening sobre el conjunto A con un elemento estructurante B, el cual es más grande que todas las componentes del ruido, se puede ver que el ruido exterior ha desaparecido, pero no afecta los hoyos internos, figura I-4.3 c).

El resultado de la operación del closing cierra estos huecos, consiguiendo tener un objeto completamente sólido y muy parecido al que se tenía al inicio, figura I-4.3 e).

4.2.3 Transformada de Serra

Esta transformada es un instrumento básico para la detección de formas. Para analizar este concepto consideremos la figura I-4.4.

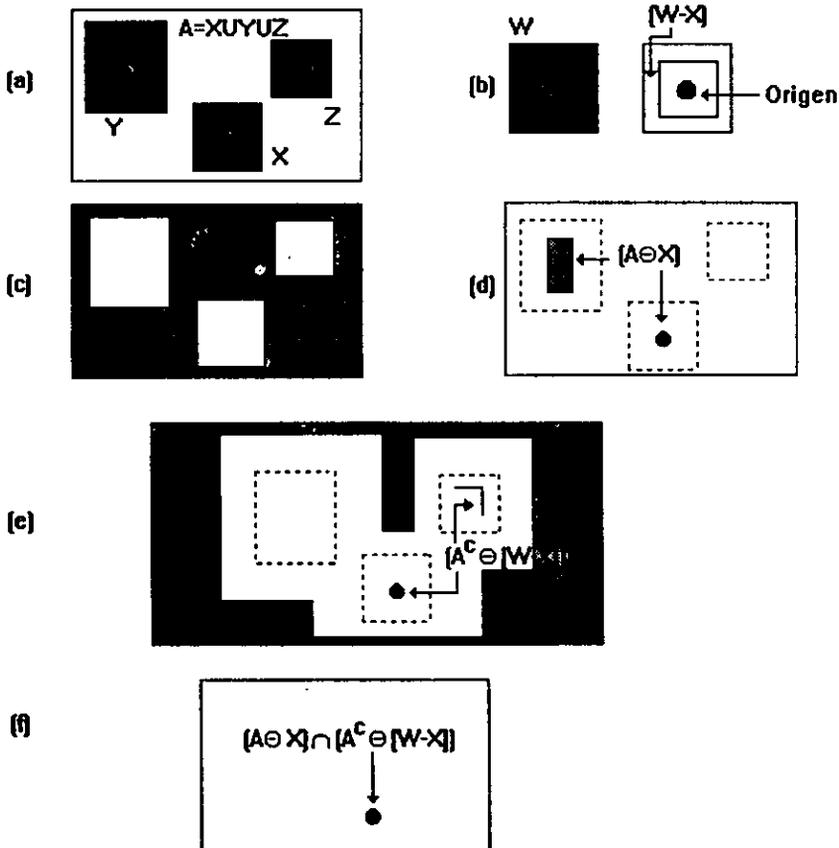


Fig. I-4.4 Transformada de Serra.

Se puede observar un conjunto A que se compone de tres formas denotadas por X, Y y Z. El objetivo es encontrar la ubicación de las formas, por ejemplo X.

La operación de la transformada de Serra es: se considerará que el origen de cada forma se encuentra en su centro de gravedad. Se extrae el fondo de X con una ventana W ligeramente mayor que X, quedando (W-X), además obtenemos el complemento de A (A^c). Una vez hecho esto, se realiza la erosión de A por X y la erosión de A^c por (W-X), el resultado de la transformada es la intersección de los puntos de las erosiones antes mencionadas.

Podemos denotar la transformada de Serra como:

$$A \ominus B = (A \ominus B) \cap (A^c \ominus [W-X]) \tag{I-4.16}$$

4.2.4 Algunos algoritmos morfológicos básicos

Esta parte comprende el análisis de algunos métodos de gran utilidad que emplean morfología matemática.

4.2.4.1 Extracción de límites o contornos

El límite o contorno de un conjunto A, denotado por β(A), puede obtenerse simplemente al restarle al conjunto A la primera erosión de A por B, donde B es un elemento estructurante apropiadamente elegido.

Es decir,

$$\beta(A) = A - (A \ominus B) \tag{I-4.17}$$

El mecanismo de extracción del contorno es ilustrado en la figura I-4.5.

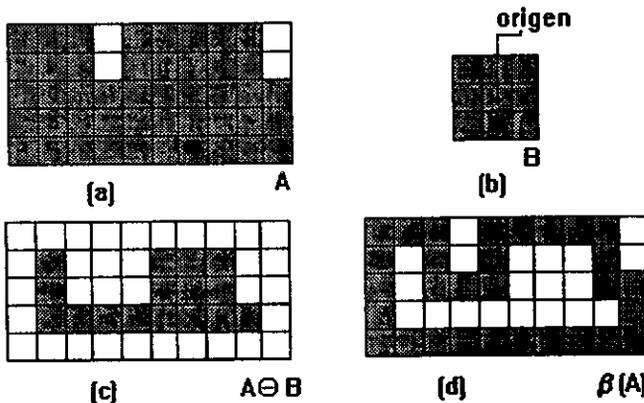


Fig. I-4.5 Extracción de límites.

4.2.4.2 Llenado de regiones

Un algoritmo simple, basado en el uso de la dilatación, el complemento y las intersecciones, puede ser empleado para llenar de regiones. La figura I-4.6 muestra un conjunto A, el cual contiene una conexión limitando una región. Dado un punto, p, dentro del los límites, el objetivo es llenar dicha región con unos.

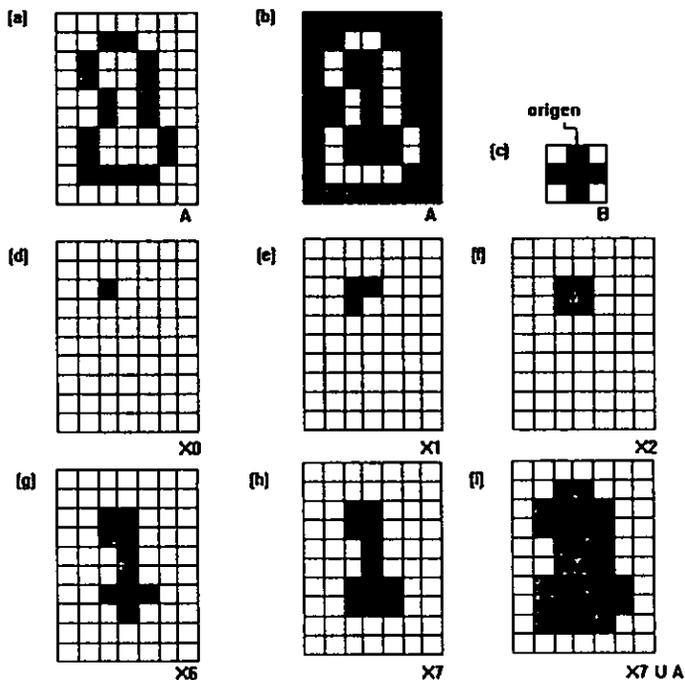


Fig. I-4.6 Llenado de regiones.

El llenado se hace partiendo del punto conocido p, el cual es un elemento ubicado dentro del contorno, éste se ubica en una imagen de iguales dimensiones a la de análisis en donde el único punto de valor 1 es el indicado por p, sobre esta nueva imagen se hace una dilatación por un elemento estructurante, como el mostrado en la figura I-4.6 c), el resultado se interseca con A^c, formando una nueva imagen a la cual se le hace una dilatación con el mismo elemento B, el resultado se interseca con A^c, formando una nueva imagen, el procedimiento se repite hasta que el resultado de las intersecciones en la iteración k y k-1 sean iguales. La región llena junto con el contorno se tiene en la iteración k, figura I-4.6 i).

Esto en conjuntos es:

$$X_k = (X_{k-1} \oplus B) \cap A^c \quad k = 1, 2, 3, \dots \tag{I-4.18}$$

4.2.4.3 Extracción de componentes

La extracción de componentes obtenidas de una imagen binaria es algo muy útil en aplicaciones de análisis sobre imágenes. Sea Y que representa una componente conectada en un conjunto A y se asume que nosotros conocemos un punto p de Y . Entonces, podemos emplear la siguiente expresión iterativa para encontrar todos los elementos de Y :

$$X_k = (X_{k-1} \ominus B) \cap A \quad k = 1, 2, 3, \dots \quad (1-4.19)$$

donde $X_0 = p$ y B es un elemento estructurante apropiado. Si $X_k = X_{k-1}$ se puede decir que el algoritmo es convergente y se puede decir que $Y = X_k$.

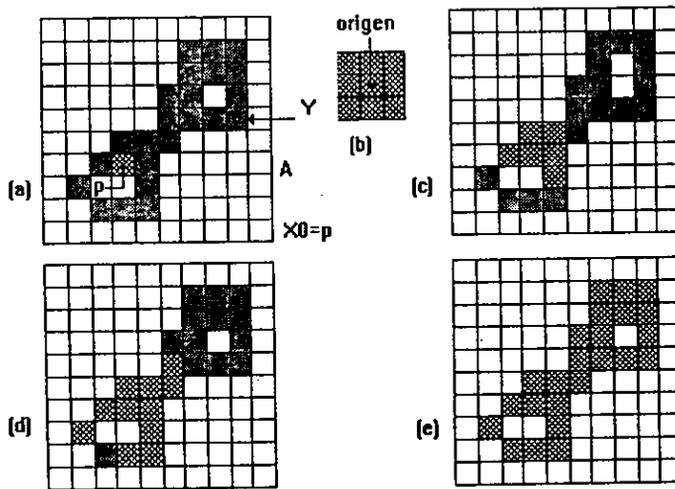


Fig. 1-4.7 Extracción de componentes.

El procedimiento es idéntico al de la aplicación anterior, lo único que cambia es el elemento estructurante empleado y que las intersecciones se hacen directamente con A .

4.2.4.4 Corteza convexa

Sean B^i , $i=1,2,3,4$, cuatro elementos estructurantes para ser empleados, este procedimiento consiste en aplicar la siguiente ecuación:

$$X_k^i = (X_{k-1}^i \ominus B^i) \cup A \quad i = 1, 2, 3, 4; k = 1, 2, 3, \dots \quad (1-4.20)$$

con $X_0^i = A$. Si dejamos $D^i = X_{conv}^i$, donde el subíndice "conv" indica convergencia en el sentido de que $X_k^i = X_{k-1}^i$, entonces podemos escribir la corteza convexa de A como:

$$C(A) = \bigcup_{i=1}^4 D^i \quad (1-4.21)$$

En otras palabras, el procedimiento consiste en aplicar la transformada de Serra de A con B^1 ; al resultado se le aplica la unión de conjuntos con A, llamando al resultado D^1 . Entonces iniciamos la misma operación, pero ahora usando B^2 , y así sucesivamente hasta terminar con B^4 . La unión de los cuatro resultados D^i 's constituye la corteza convexa de A. Este procedimiento se ilustra en la figura I-4.8.

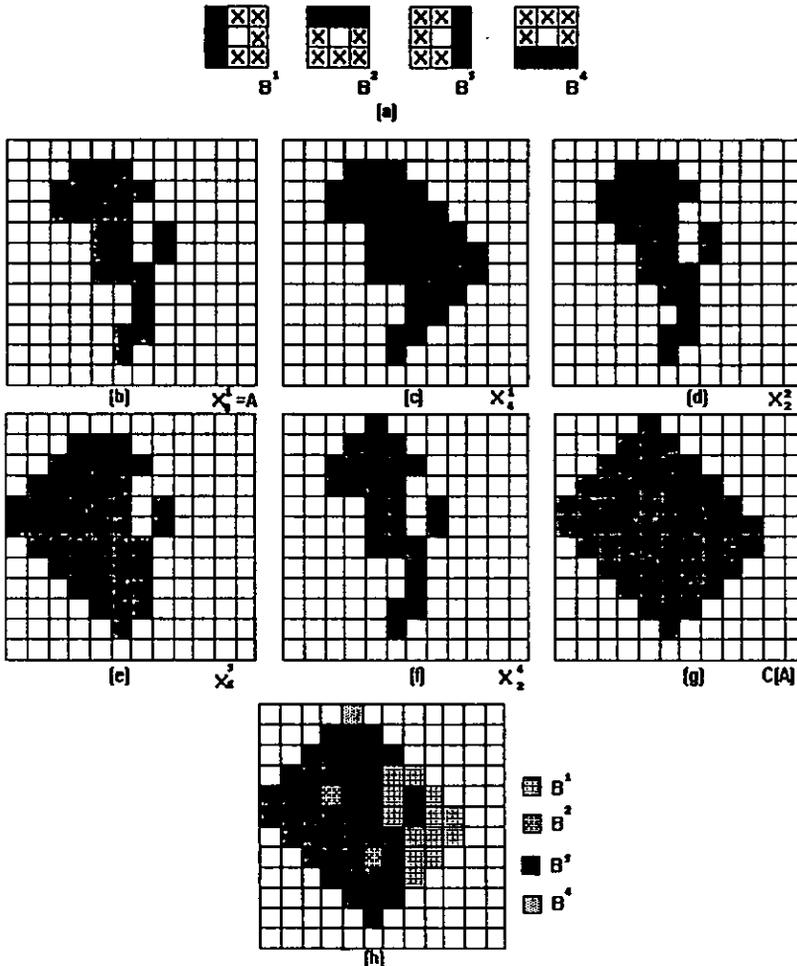


Fig. I-4.8 Corteza convexa.

La figura I-4.8 a) muestra los elementos estructurantes usados en la extracción de la corteza convexa. La figura I-4.8 b) muestra un conjunto para el cual nosotros deseamos encontrar la corteza convexa. De la figura I-4.8 c) a la I-4.8 f) se muestra el resultado D para cada elemento estructurante. La corteza convexa se ilustra en la figura I-4.8 g). La contribución de cada elemento estructurante se puede ver en la figura I-4.8 h).

4.2.4.5 Adelgazamiento

El adelgazamiento de un conjunto A puede estar definido en términos de la transformada Serra como:

$$A \otimes B = A - (A \otimes B) = A \cap (A \otimes B)^c \tag{I-4.22}$$

O en términos de elementos estructurantes secuenciales, definido como:

$$\{B\} = \{B^1, B^2, B^3, \dots, B^n\} \tag{I-4.23}$$

donde B^i es una versión rotada de B^{i-1} . De esta manera el adelgazamiento realizado por una secuencia de elementos estructurantes es:

$$A \otimes \{B\} = (((((A \otimes B^1) \otimes B^2) \dots) \otimes B^n) \tag{I-4.24}$$

En otras palabras, el adelgazamiento de un conjunto A se hace aplicando la transformada de Serra de A con B^1 , al resultado se le aplica la transformada de Serra, pero ahora con B^2 y así sucesivamente hasta terminar con B^n . El proceso completo es repetido hasta que no exista ningún cambio en la forma de la imagen.

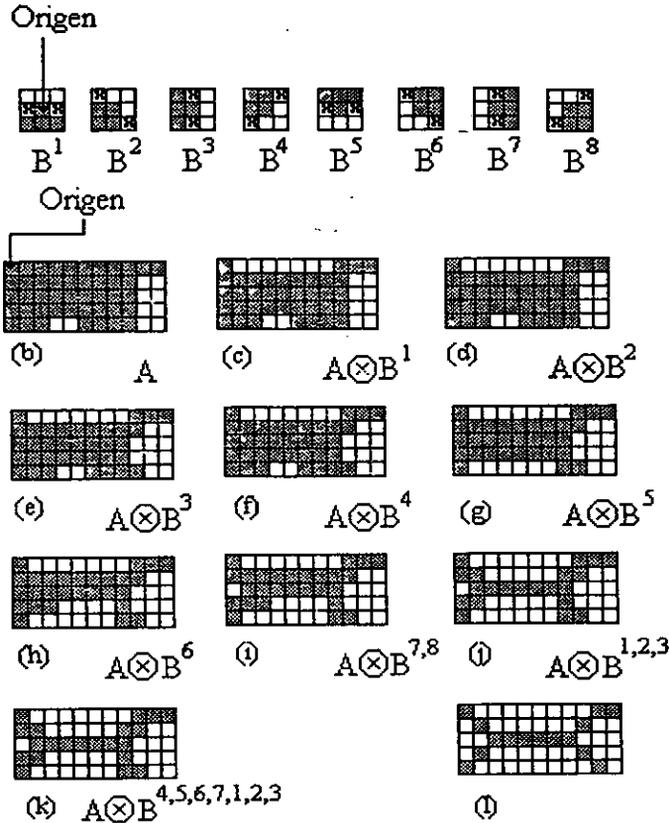


Fig. I-4.9 Adelgazamiento morfológico.

La figura I-4.9 a) muestra un conjunto de elementos estructurantes usados para adelgazamiento y la figura I-4.9 b) muestra el conjunto A que deseamos adelgazar. El resultado de adelgazamiento de A con B^1 es mostrado en la figura I-4.9 c) y el resultado de los pasos con los otros elementos estructurantes se muestran de I-4.9 d) a I-4.9 k). La convergencia se ha conseguido después de la segunda pasada con B^4 . El resultado del adelgazamiento se muestra en I-4.9 k). Finalmente, la figura I-4.9 l) muestra el conjunto adelgazado convertido a conectividad-m para eliminar múltiples caminos.

4.2.4.6 Engrosado o espesado

El engrosado está definido por la expresión

$$A \odot B = A \cup (A \otimes B) \quad (I-4.25)$$

donde B es un elemento estructurante elegido para el engrosado. Esta operación también se puede definir como una operación secuencial:

$$A \odot B = (\dots((A \otimes B^1) \otimes B^2) \dots) \otimes B^m \quad (I-4.26)$$

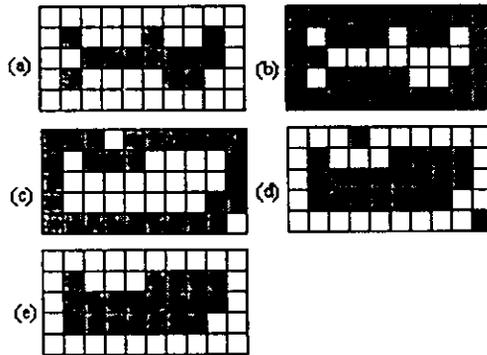


Fig. I-4.10 Ejemplo de engrosamiento.

El elemento estructurante usado para engrosar tiene la misma forma como se muestra en la figura I-4.9 a), pero con los unos y los ceros intercambiados. Sin embargo, este algoritmo para el engrosado es poco usado en la práctica. En vez de eso, el procedimiento usual es adelgazar el fondo del conjunto y entonces el resultado es complementado. En otras palabras, si nosotros deseamos engrosar el conjunto A, obtenemos $C = A^c$, adelgazar C, y entonces obtener C^c . El procedimiento se muestra en la figura I-4.10.

4.2.4.7 Esqueletos

El esqueleto de un conjunto A puede ser expresado en términos de erosiones y opening's. El esqueleto de A está definido como $S(A)$, donde:

$$S(A) = \bigcup_{k=0}^K S_k(A) \quad (I-4.27)$$

con

$$S_k(A) = \bigcup_{k=0}^K ((A \ominus kB) - [(A \ominus kB) \circ B]) \quad (I-4.28)$$

donde B es un elemento estructurante definido adecuadamente, $(A \ominus kB)$ indica k erosiones sucesivas de A: esto es $(A \ominus kB) = (((... (A \ominus B) \ominus B \ominus ...)) \ominus B)$, y k es el último paso de iteración antes de que desaparezca por completo el conjunto A.

La información dada en las ecuaciones precedentes establece que el esqueleto de A, puede ser obtenido como la unión de los esqueletos de los subconjuntos $S_k(A)$.

La reconstrucción de A, a partir de su esqueleto, se puede hacer usando la siguiente ecuación:

$$A = \bigcup_{k=0}^K (S_k(A) \oplus kB) \quad (I-4.29)$$

donde $(S_k(A) \oplus kB)$ denota k dilataciones sucesivas de $S_k(A)$; esto es:

$$(S_k(A) \oplus kB) = (((... (S_k(A) \oplus B) \oplus B) \oplus ...)) \oplus B \quad (I-4.30)$$

4.2.4.8 Recortado o podado

Éste es un método que complementa las operaciones de la obtención del esqueleto y el adelgazamiento, debido a que dichos procedimientos tienden a dejar componentes parásitas que necesitan ser limpiadas con un post-procesamiento.

Veamos sobre un ejemplo la aplicación de este algoritmo. Partiendo de que estas componentes parásitas no exceden los 3 pixeles.

En la figura I-4.11 podemos ver una letra "a", la cual fue escrita a mano, ésta presenta componentes que no corresponden al carácter, las cuales deben ser eliminadas.

Estableciendo A como el conjunto de entrada, podemos quitar las componentes parásitas adelgazando el conjunto A con una secuencia de elementos estructurantes, mostrados en la figura I-4.11 b) y I-4.11 c), diseñados para detectar sólo puntos finales. Con esto obtenemos:

$$X_1 = A \otimes \{B\} \quad (I-4.31)$$

Los dos elementos estructurantes mostrados en la imagen son girados 90° cada vez, hasta tener un total de 8.

Aplicando la ecuación anterior sobre A tres veces se obtiene el conjunto X_1 mostrado en la figura I-4.11 d). Como se puede ver de la imagen, además de los parásitos se eliminaron también algunos componentes del carácter.

El siguiente paso es restaurar el carácter a su forma original, pero con las ramas parásitas removidas. Para hacer esto, primero hay que obtener el conjunto X_2 a partir de X_1 de acuerdo a:

$$X_2 = \bigcup_{k=1}^8 (X_1 \oplus B^k) \quad (I-4.32)$$

donde B^k son los mismos elementos estructurantes empleados arriba. Haciendo la dilatación de X_2 tres veces, usando el conjunto A como delimitador, de la forma:

$$X_3 = (X_2 \oplus H) \cap A \quad (I-4.33)$$

donde H es un elemento estructurante de 3x3 de 1's. Este tipo de dilatación condicional previene valores de elementos unos fuera de la región de interés, este resultado se ilustra en la figura I-4.11 f). El resultado final se obtiene al hacer la unión de X_1 y X_3 como:

$$X_4 = X_1 \cup X_3 \quad (I-4.34)$$

lo cual es mostrado en la figura I-4.11 g).

Todo lo anterior muestra que con esta morfología matemática, utilizando las operaciones existentes, es posible resolver problemas prácticos, combinándolas de la manera correcta.

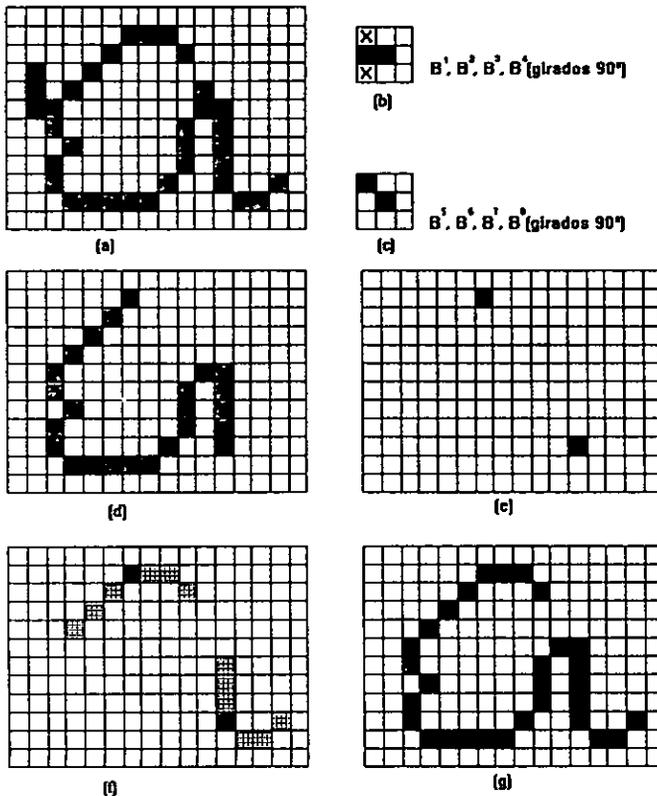


Fig. I-4.11 Ejemplo de recortado.

4.3 Compresión de imágenes

Cuando una imagen es digitalizada, la cantidad de datos resultante es tan grande como la calidad deseada al reproducir esta información, si se desea una reproducción con una calidad alta, el tamaño de memoria que se necesitará para almacenarla será enorme, es por ello que las técnicas de compresión debieron de surgir para tratar de reducir este problema. Por ejemplo, veamos la cantidad de espacio requerido en memoria para almacenar algunos tipos de imágenes:

- Calidad de TV, imagen de vídeo en color: 512x512 pixeles/color, 8 bits/pixel, y 3 colores. Requiere \approx 6 Mbits.
- Una fotografía de 24x36 mm (35mm) escaneada a 12 μ m: 3000x2000 pixeles/color, 8 bits/pixel, y 3 colores. Requiere \approx 144 Mbits.
- Una radiografía de 14x17 pulgadas escaneada a 70 μ m: 5000x6000 pixeles, 12 bits/pixel. Requiere \approx 360 Mbits.

Cualquier técnica de compresión intenta reducir la cantidad de datos que hay que transmitir y/o almacenar, para conservar los principales rasgos y calidad de la información, con la menor deformación posible.

4.3.1 Técnicas de compresión de imágenes

El gran auge que ha tomado el uso de vídeo digital en el mundo del procesamiento de señales ha repercutido en el desarrollo de estándares de compresión. Estas técnicas hacen más fácil el manejo de imágenes digitales, además de reducir el costo del hardware empleado para el procesamiento en tiempo real.

Uno de los principales puntos de apoyo para poder hacer compresión de imágenes es la redundancia que se presenta en una escena fija o en una secuencia de imágenes. Pueden presentarse tres tipos de redundancia en imágenes digitales, estas son:

1. Redundancia espacial, que se debe a la correlación entre los valores de pixeles vecinos.
2. Redundancia espectral, debida a la correlación entre diferentes planos de color en una imagen RGB.
3. Redundancia temporal, debida a la correlación entre diferentes cuadros en una secuencia de imágenes.

La tendencia en los avances de las técnicas de compresión, están encaminadas a desarrollar estándares genéricos para los diferentes tipos de imágenes que pueden presentarse:

- Imágenes bi-nivel.
- Imágenes fijas con tonos continuos en color o monocromas.
- Imágenes secuenciales en tonos continuos.

4.3.1.1 Pixeles independientes y dependientes

Una ligera compresión puede hacerse empleando esta técnica, la cual ubica ciertas localidades o pixeles como independientes, los cuales conservarán sus características y ayudarán a reconstruir la imagen, el resto de los pixeles serán dependientes de los primeros.

De esta manera, sólo los píxeles independientes serán transmitidos y los píxeles dependientes se calcularán de acuerdo a una lógica predeterminada. El elegir cuales de los píxeles serán los independientes y cuales los dependientes dependerá de la estructura que se quiera preservar sobre la imagen, líneas horizontales, líneas verticales, curvas, etc. Así, la calidad de la imagen resultante dependerá de esta elección, tanto como de la naturaleza de la misma imagen de entrada.

Esta técnica será empleada en el capítulo de resultados para hacer una demostración práctica de los alcances que pueden lograrse con esta tarjeta.

4.3.1.2 La norma JPEG

La necesidad de un estándar internacional para compresión de imágenes fijas en tono continuo, dio como resultado en 1986, la formación del JPEG (Joint Photographic Experts Group). Este grupo fue respaldado por ISO (International Standards Organization) y la CCITT (Consultive Committee on Telephone and Telegraph) para desarrollar un estándar de compresión de propósito general.

El algoritmo JPEG ofrece una excelente calidad para la mayoría de imágenes comprimidas con un rango de compresión 24:1. La figura I-4.12 muestra el diagrama base de codificación bajo la norma JPEG.

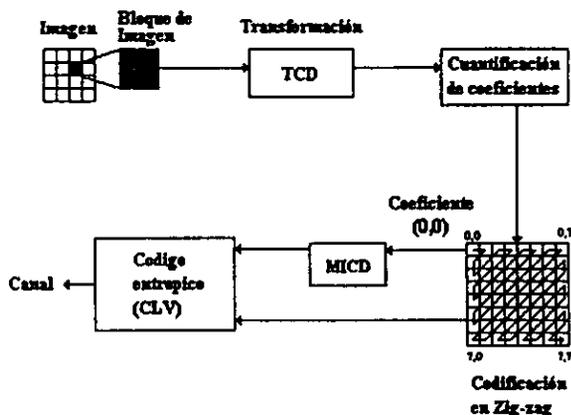


Fig. I-4.12 Codificación JPEG.

El algoritmo JPEG parte de la segmentación de la imagen en bloques de 8x8 píxeles, de los cuales se obtiene la frecuencia dominante usando TCD (Transformada Coseno Discreta, incorporación de JPEG a los algoritmos de compresión). La TCD concentra la mayoría de la energía del bloque en pocos componentes, usualmente en la esquina superior izquierda del bloque TCD. El coeficiente en la esquina superior izquierda es llamado coeficiente de DC, debido a que posee la intensidad de energía promedio del bloque al dividirse por 8. Los coeficientes de AC correspondientes a frecuencias más altas

se encuentran alejándose del coeficiente de DC. La ecuación para la transformada coseno discreta está dada por:

$$F(u, v) = \frac{1}{4} C(u) C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos\left(\frac{\pi u(2x+1)}{16}\right) \cos\left(\frac{\pi v(2y+1)}{16}\right) \quad (I-4.35)$$

donde

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & \text{para } u, v = 0 \\ 1 & \text{otro caso} \end{cases} \quad (I-4.36)$$

Después de haber pasado por la TCD, el bloque se somete a un proceso de *cuantización* empleando la ecuación:

$$Fq(u, v) = \text{redondear} \left[\frac{F(u, v)}{Q(u, v)} \right] \quad (I-4.37)$$

donde,

$F(u, v)$ es el bloque después de TCD

$Q(u, v)$ es un bloque de cuantización

Cuantizar significa dividir los valores reales entre una constante, para ello se utiliza una tabla de cuantización estándar, con la cual se logra reducir la amplitud de los coeficientes en el bloque. Después de que han sido cuantizados, los bloques de coeficientes usualmente cuentan con muchos coeficientes de AC de valor cero, si los coeficientes son reordenados, usando un escaneo en Zig Zag, ilustrado en la figura I-4.13, habrá una tendencia a tener cadenas de ceros. Sólo los coeficientes de AC no ceros son entonces codificados.

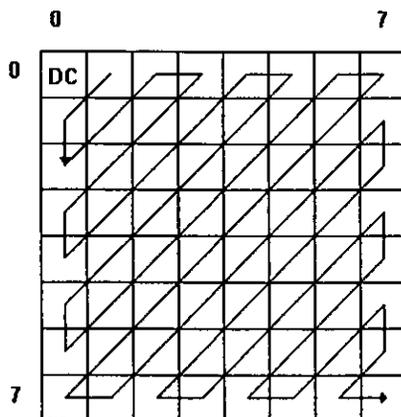


Fig. I-4.13 Escaneo en Zig-Zag.

Al hacer el escaneo en Zig Zag, se logra ubicar más rápidamente la información que se conserva en la imagen debido a que los valores son muy parecidos, esta información se codifica para su transmisión en forma entrópica usando códigos de longitud variable para modificar la tasa de compresión (a los coeficientes más probables se les asigna menos bits que a los menos probables), en JPEG generalmente se emplean los códigos de Huffman.

4.3.1.3 La recomendación H.261

En diciembre de 1984 se inició el estándar H.261, al comienzo el estándar contemplaba una frecuencia de transmisión de $m \times 384$ bits/seg (m entre 1 y 5). Después se contemplaron rangos de $n \times 64$ kbits/seg (n entre 1 y 5). Finalmente, el estándar H.261 quedó fijado en $p \times 64$ kbits/seg (situando p entre 1 y 30).

Al igual que JPEG, el estándar H.261 hace uso de la TCD (transformada coseno discreta) y codifica con CLV (códigos de longitud variable), la diferencia principal entre estos métodos es que JPEG sólo hace compresión sobre una sola imagen, mientras que H.261 lo hace para una imagen fija (intraframe) o para una secuencia de escenas en movimiento (interframe). En la figura I-4.14 se muestra el diagrama de bloques de el codificador $p \times 64$.

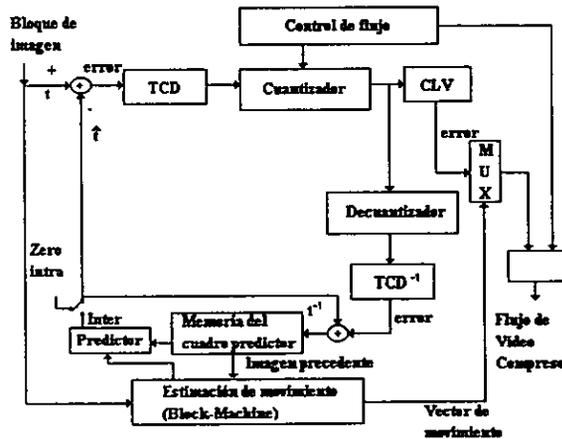


Fig. I-4.14 Diagrama de bloques del codificador $p \times 64$.

Acorde a la disponibilidad de la línea de transmisión, el codificador tiene la opción de cambiar de transmisión intraframe a interframe. Durante una transmisión intraframe, H.261 hace una compresión prácticamente igual a JPEG. La transmisión interframe fue la aportación más importante de esta recomendación, con ella se puede predecir la imagen actual a partir de la anterior. Partamos del supuesto de que una primer imagen se ha transmitido de manera intraframe, la cual fue reconstruida y almacenada en memoria (imagen precedente), esta imagen precedente conjuntamente a la imagen actual se someten a una técnica conocida como *Block Matching*, en la cual se propone un vector de desplazamiento para hacer una predicción de la imagen actual de la siguiente manera: el

Block Matching propone un vector de desplazamiento, el cual se ubica al centro de una ventana de búsqueda de 20x20 pixeles, esta ventana se desplaza pixel a pixel sobre las 400 posiciones posibles de la imagen precedente tratando de buscar la zona con la mayor aproximación de la imagen actual, formando una imagen predicha. Como se muestra en la figura I-4.15.

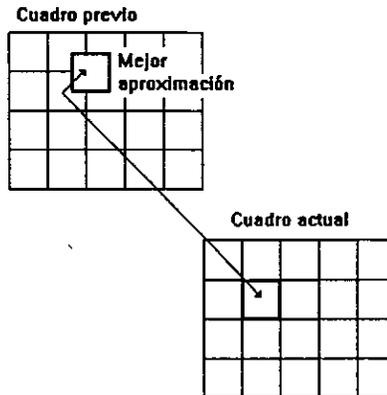


Fig. I-4.15 Compensación de movimiento.

Algunas formas de estimar el vector de desplazamiento son: por recursiva (un vector para cada pixel es obtenido), técnicas por bloques (Block Matching - se propone un vector de movimiento), por transformadas (en el dominio de la frecuencia) y Morfología Matemática entre otras. Una vez que se obtiene la imagen predicha, se compara con la imagen actual estableciendo una imagen de error, la cual sufre el mismo proceso de codificación antes descrito.

Para hacer la reconstrucción de la imagen actual en el receptor se requiere: de una imagen original precedente, de la imagen de error y del vector de movimiento, la imagen original precedente será la base principal para poder hacer una predicción de la imagen actual. Simplemente se hacen las modificaciones pertinentes de acuerdo a los datos que cambian usando la señal de error y el vector de movimiento sobre la imagen original precedente.

Cuando se transmite simplemente la señal de error y el vector de movimiento, la cantidad de datos es mucho menor que si se transmitiera la imagen actual directamente. Pero cuando la imagen actual ha cambiado considerablemente o es una escena totalmente distinta, se tiene que volver al inicio, es decir, hay que transmitir una imagen de forma intraframe, para después regresar a interframe y seguir con la codificación.

4.3.1.4 MPEG-I

MPEG estuvo respaldado por ISO para estandarizar una representación de código de audio y vídeo conveniente para medios de almacenamiento digital, como el disco compacto de memoria de sólo lectura (CD-ROM), la cinta de audio digital (DAT), etc. La meta del grupo, sin embargo, ha sido el desarrollo de un estándar genérico, uno que pueda ser usado en otras aplicaciones de vídeo digital, como es la telecomunicación.

Este estándar, dentro de su enfoque al vídeo, plantea la codificación de tres diferentes tipos de cuadros para formar una secuencia y lograr así la compresión de vídeo para su posterior reproducción, los tres tipos de cuadros que codifica la técnica tienen diferentes características e importancia. Esta técnica de compresión introduce los conceptos de imágenes predichas e imágenes interpoladas, para lograr lo que se conoce como "Predicción Temporal Bidireccional". La estructura de los cuadros en una compresión MPEG-I es como lo indica la figura I-4.16.

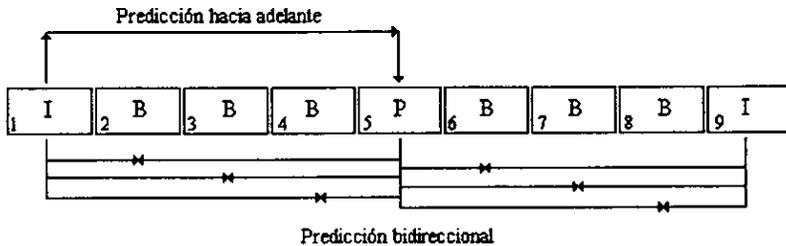


Fig. I-4.16 Predicción bidireccional.

De la figura anterior se puede ver que el primer cuadro está denotado por la letra I, esto significa que se codifica en intraframe similarmente a lo que hace JPEG, éste es el cuadro de mayor importancia, pues a partir de él se derivan los siguientes. El cuadro que sigue en importancia es el denotado por P, el cual es predicho y codificado teniéndose como referencia un cuadro I u otro P anterior a éste, algo muy parecido a lo que hace H.261 en interframe. Finalmente, tenemos los cuadros intermedios B, estos son cuadros que se obtienen al hacer una interpolación entre un cuadro pasado y uno futuro (I y P regularmente).

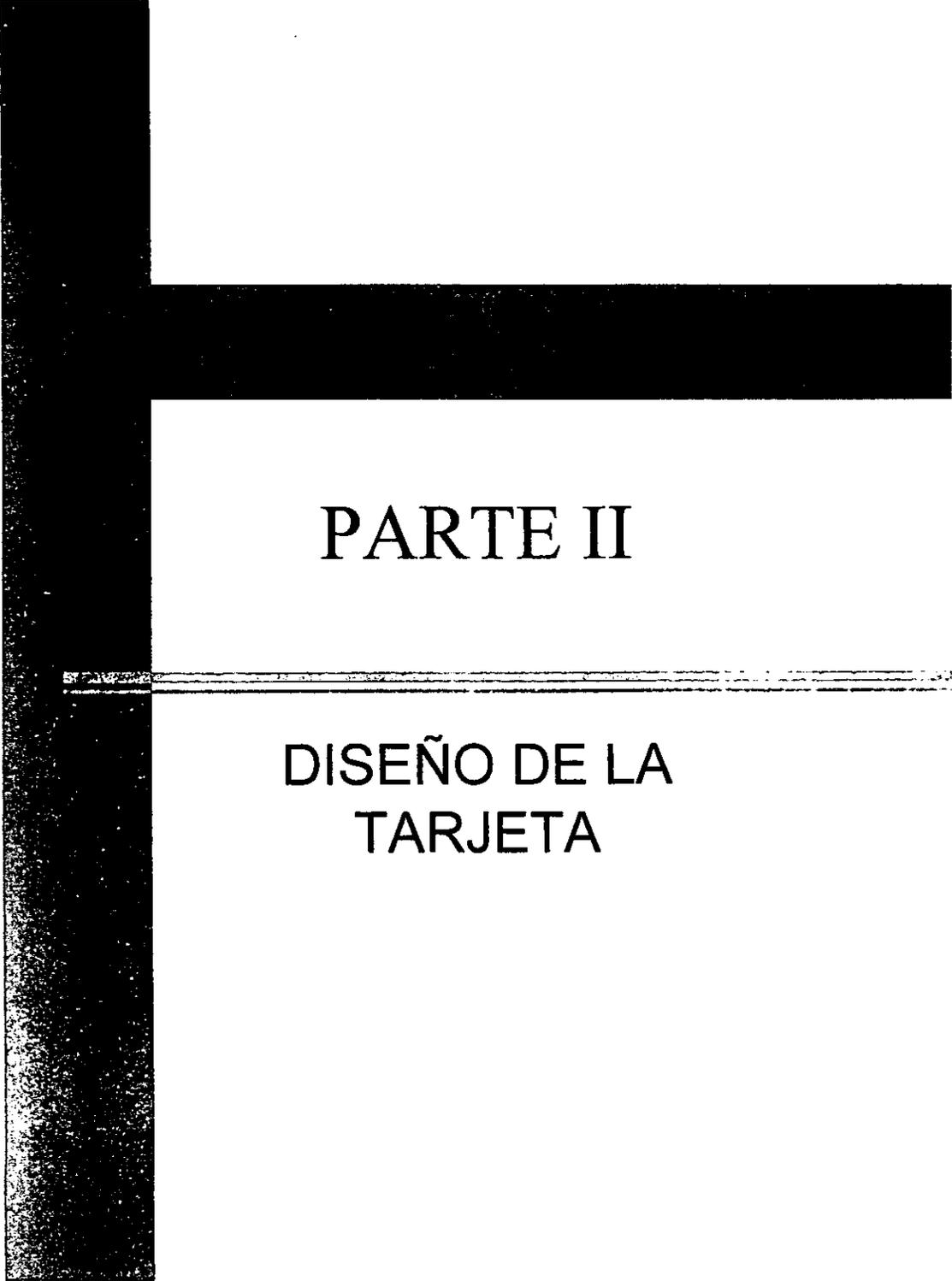
La transmisión de los cuadros no se hace secuencialmente, como uno pensaría, pues como mencionamos el orden de codificación de los cuadros es: imagen intra (cuadro 1), imagen predicha (cuadro 5), imagen interpolada (cuadro 2, 3 y 4), imagen intra (cuadro 9), imagen interpolada (cuadro 6, 7 y 8), así la transmisión de cuadros para este caso sería {1, 5, 2, 3, 4, 9, 6, 7, 8}.

4.3.1.5 MPEG-II

La segunda fase de MPEG (MPEG-II) fue aprobada a finales de 1994. Esta fase está orientada a codificación de las señales de vídeo con 720 píxeles, 480 líneas, 30 cuadros por segundo, entrelazado 2:1 en frecuencias de bits de 2 Mb/s, o superiores.

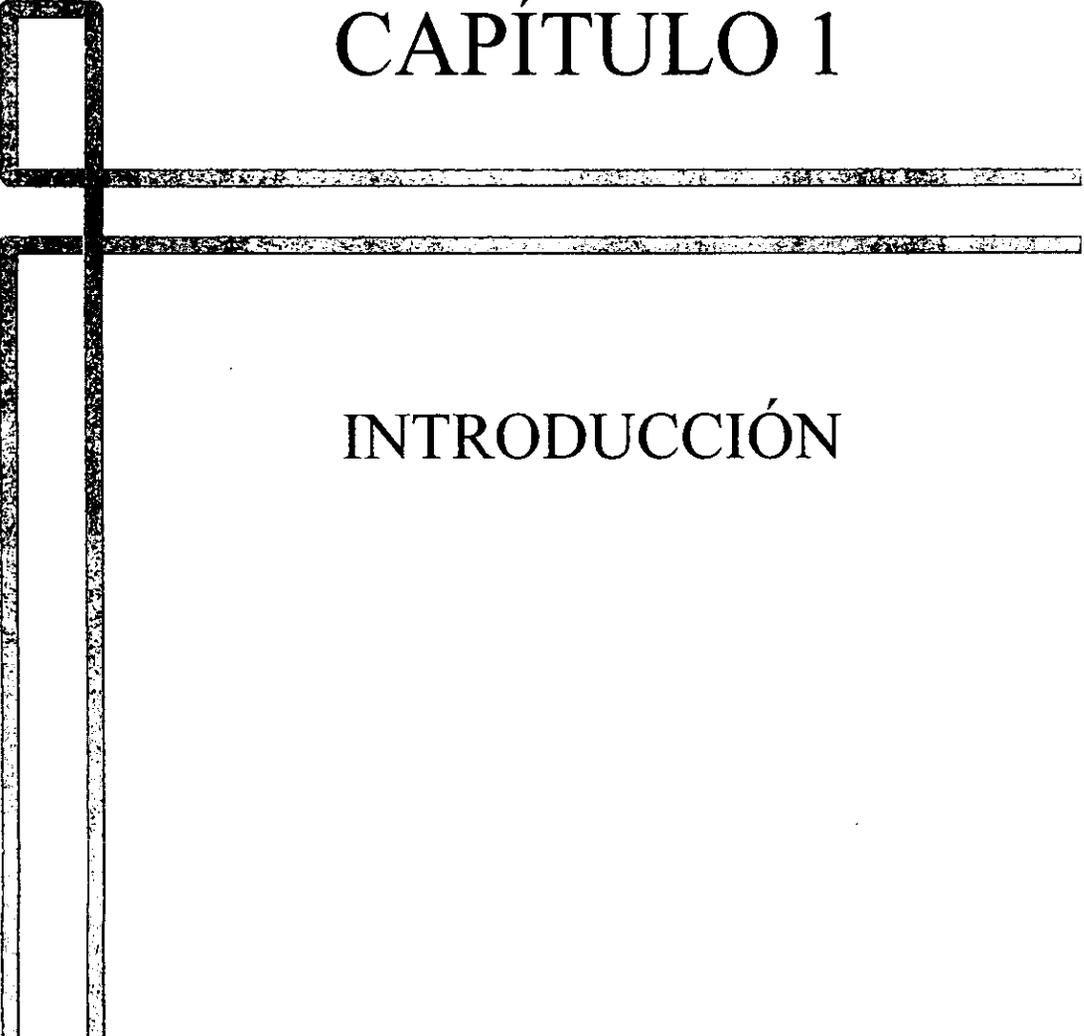
Las innovaciones importantes que presenta MPEG-II son que, además de lo que tenía MPEG-I, se agrega la posibilidad de transmitir a varios formatos y lo más importante, la posibilidad de transmitir en diferentes calidades de vídeo, utilizando los modos escalables¹¹: escalabilidad temporal, escalabilidad espacial y escalabilidad al ruido.

¹¹ D.G. Morrison, Video Compression Standars; Where we are and how we got there, In M.J.J. Holt and C.F.N. Cowan, editors, EUSIPCO94, volume 3, Edinburgh, Scotland U.K. September 1994.



PARTE II

DISEÑO DE LA TARJETA



CAPÍTULO 1

INTRODUCCIÓN

1.1 Introducción

El procesamiento digital de imágenes es un campo que ha tomado mucho interés en los últimos años, en particular el procesamiento de imágenes en movimiento, es decir, el vídeo digital. Un sistema de procesamiento de vídeo en tiempo real es aquel en el cual, a la hora de visualizar los resultados un observador no nota la diferencia en tiempo entre la imagen real y una procesada. Existen muchos trabajos que proponen distintos algoritmos que trabajan con la calidad, el tamaño, la redundancia espacial y temporal de la imagen para fines específicos como: el reconocimiento de patrones, la transmisión o simplemente la visualización.

Estos trabajos para su aplicación práctica presuponen la existencia de imágenes digitalizadas previamente. El objetivo de este estudio es el contar con una tarjeta electrónica capaz de obtener y procesar sus propias imágenes.

El procesamiento se lleva a cabo digitalmente sobre los elementos o puntos llamados píxeles que juntos componen una imagen, estos píxeles son muestras tomadas de una señal de vídeo analógica, la cual ha sido realizada por una cámara, que explora la imagen existente sobre una superficie fotosensible, aquí es donde se tiene la visualización de una imagen real. Esta imagen óptica es transformada a una señal analógica realizando un exploración secuencial, ya sea por medio de un cañón de electrones en el caso de una cámara de tubo de rayos catódicos, o de cada celda en el caso de un dispositivo de carga acoplada CDD. La unión de todas estas muestras secuenciales representarán las líneas de una imagen, las cuales serán coordinadas por señales de sincronía con el fin de no perder la imagen que originalmente se capturó.

El presente trabajo consiste en el diseño e implantación de una tarjeta de captura de vídeo para su aplicación en distintos experimentos de procesamiento digital de imágenes. Con el fin de contar con un cierto grado de continuidad en los movimientos de los objetos de la imagen, esta tarjeta tiene como objetivo el capturar el mayor número de cuadros de vídeo sobre una señal de vídeo compuesto con sincronía negativa, mediante una computadora tipo PC. La tarjeta hace la captura de una fuente de vídeo a cuadros de 256x192 píxeles en 256 tonos de grises. El desarrollo de esta tarjeta está hecho con componentes (circuitos integrados) de funcionamiento relativamente simple, como amplificadores operacionales y contadores. Sobre los cuadros adquiridos y en base a la velocidad de procesamiento disponible en la PC se intentan realizar diferentes tipos de procesamiento sencillo.

La fuente de vídeo compuesto a digitalizar puede provenir de una videoreproductora o cualquier cámara de vídeo, aunque es ideal una cámara CCD de blanco y negro muy básica ya que, como se verá mas adelante, es una parte importante para obtener un buen procesamiento de una imagen.

CAPÍTULO 2

HARDWARE

**ESTA TESIS NO DEBE
SALIR DE LA BIBLIOTECA**

2.1 Generalidades

Con respecto al hardware de la tarjeta, se tuvo que desarrollar un sistema que trabajará de acuerdo a las características de operación de la señal de vídeo y obtener así una imagen de buena calidad y de la manera más sencilla posible.

El funcionamiento de la tarjeta en cuanto a hardware, puede decirse que está dividido en tres etapas: procesamiento analógico, digitalización y adquisición, éstas trabajan en sincronía para obtener el resultado deseado. En la figura II-2.1 se muestra el diagrama de bloques simplificado de la tarjeta.

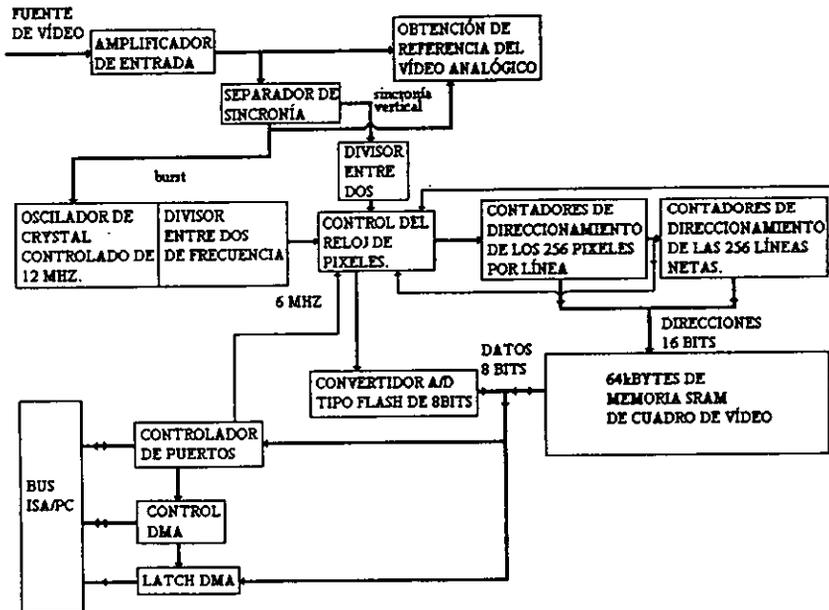


Fig. II-2.1. Diagrama de bloques de la tarjeta de captura.

Durante el procesamiento analógico, se trabaja la señal de vídeo de manera analógica con el objeto de acondicionarla para la digitalización, obteniendo algunas señales de control necesarias para este fin. Después se obtiene la digitalización mediante el uso de un convertidor A/D del tipo Flash, el cual es indispensable para obtener una buena digitalización del vídeo, ya que en el caso de este convertidor, el MC10319P, nos permite manejar frecuencias de muestreo de hasta 25 MHz. La digitalización es controlada por un conjunto de señales provenientes del controlador de puertos paralelos 8255, el cual es controlado por software. Las señales de sincronía propias de la señal de vídeo, obtenidas con ayuda de un circuito integrado separador de sincronía (el LM1881), ayudan a manejar el proceso de digitalización. Existe también una etapa de almacenamiento del cuadro capturado para su posterior lectura. Se cuenta con un reloj de cuarzo que nos da la frecuencia a la cual la señal de vídeo es digitalizada (6 MHz), por lo que éste se encuentra

conectado al convertidor A/D, además de conectarse a un sistema de cuatro contadores del tipo SN74HC193, de 4 bits; como estos se encuentran conectados en cascada, se tiene un contador conjunto de 16 bits. Dichos contadores sirven para direccionar un búfer de memoria propio de la tarjeta de 64 kbytes. Estos 64 kbytes son obtenidos mediante el arreglo de dos memorias de matrícula 43256, las cuales tienen una capacidad de 32 kbytes cada una. Estos circuitos fueron escogidos de manera que fueran lo bastante rápidos en su tiempo de acceso como para soportar la velocidad de captura del sistema. Puesto que se empleó un reloj de aproximadamente 6 MHz, significa que cada muestra del convertidor A/D es obtenida cada 167 nanosegundos, por ello fue conveniente seleccionar esta memoria, cuya velocidad de acceso es de 100 nanosegundos.

Posteriormente, se efectúa la lectura del búfer a la memoria RAM de la computadora mediante dos medios posibles, ya sea DMA (acceso directo a memoria) o como direccionamiento a uno de los puertos del controlador de puertos 8255; este búfer representará nuestro cuadro de vídeo en forma de píxeles cuantizados en 8 bits, por lo que se tendrá una resolución de 256 tonos de grises.

El reloj de 6MHz es obtenido a partir de un cristal de cuarzo resonante en paralelo de 12 MHz, esta frecuencia es dividida entre dos por un flip-flop divisor de frecuencia con el fin de obtener los 6 MHz deseados. Para tener controlado el inicio del reloj, se cambiaron los inversores comúnmente empleados como elementos activos del arreglo por compuertas NAND.

En los apartados siguientes se detallará el funcionamiento individual de cada elemento en la tarjeta y finalmente se hará un análisis global para ver como funcionan en conjunto.

2.2 Procesamiento analógico

La parte analógica se compone de un conjunto de amplificadores operacionales MC34082, cuyo ancho de banda sin compensar es de 16 MHz, y TL082, cuyo ancho de banda es de 4 MHz en ganancia unitaria. Contando con que la señal de vídeo tiene un ancho de banda aproximado de 5 MHz, la elección de este tipo de amplificadores es justificada.

2.2.1 Amplificador de entrada diferencial

El primer amplificador se configuró para que trabaje como un amplificador restador, mejor conocido como de entrada diferencial¹², el cual cuenta con un alto rechazo al ruido y da una señal aumentada en amplitud, lo que nos permitirá una mejor digitalización de la imagen; esto porque la señal de vídeo compuesto a la entrada del amplificador tiene una amplitud, en la información de vídeo, de 1 V_{pp} y se aumenta a 1.2 V_{pp}, que es un voltaje que da mayor campo de acción para la conversión con el circuito A/D MC10319, según sus especificaciones. La configuración típica de este circuito se muestra en la figura II-2.2.

Para el análisis de este circuito nos basaremos en el principio que dice: *“en amplificadores ideales, los voltajes en las entradas del amplificador operacional son*

¹² Luces M. Faulkenberry; Introducción a los Amplificadores Operacionales con aplicaciones ACI lineales; Editorial Limusa 2ª Edición, 1992.

aproximadamente iguales¹³”, es decir, que su diferencia es cero volts. En el caso de un amplificador real, esta diferencia es despreciable.

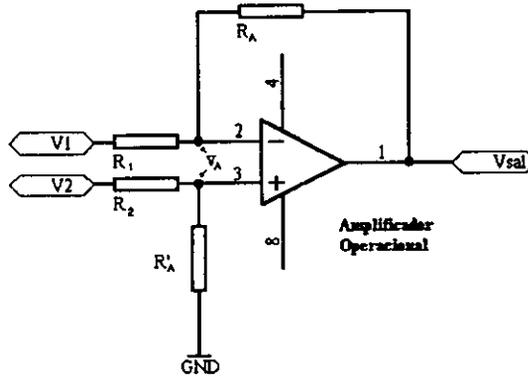


Fig. II-2.2 Amplificador de entrada diferencial.

Para realizar las operaciones llamaremos al voltaje en cada terminal de entrada V_A , que como se supuso en el párrafo anterior, es el mismo en ambas terminales. Se puede ver que si conectamos a tierra el voltaje de entrada V_2 , lo que se obtendrá es un amplificador en configuración inversora con entrada en V_1 . En esta situación, la corriente de la entrada no inversora del amplificador es cero, por lo tanto, también la corriente en R_2 y R'_A , y lo mismo ocurrirá con el voltaje V_A que será cero. En la situación contraria, cuando la entrada que tenemos a tierra es V_1 y alimentamos sólo la entrada V_2 , entonces tendremos una configuración de un amplificador no inversor, pero con un divisor de voltaje con las resistencias R_2 y R'_A que darán el voltaje de entrada V_A a la terminal no inversora.

En el caso en donde contamos con ambas entradas simultáneamente, la entrada inversora ocasiona que la señal tenga un voltaje de salida de tal modo de que el arreglo de las resistencias R_1 y R_A tengan en su unión el voltaje V_A , por lo que se tendrá la siguiente ecuación:

$$V_A = V_2[R'_A/(R_2 + R'_A)] \quad (II-2.1)$$

donde se tiene una diferencia con respecto a la operación de un amplificador inversor normal. Aplicando *ley de corrientes de Kirchoff*¹⁴ y considerando que las corrientes que fluyen hacia las entradas del amplificador operacional son despreciables, ya que las impedancias de entrada son muy grandes, se podrá tener lo siguiente:

$$I_{R_1} = I_{R_A} \text{ y } I_{R_2} = I_{R'_A} \quad (II-2.2)$$

¹³ Donald L. Schilling - Charles Belove; *Electronic Circuits*; McGraw-Hill, EUA, 1989; pag. 379-380.

¹⁴ Gabriel A. Jaramillo - Alfonso A. Alvarado; *Electricidad y Magnetismo*; Trillas, México, 1990; pag. 237-240.

es decir,

$$I_{R_1} = \frac{V_1 - V_A}{R_1} = I_{R_A} = \frac{V_A - V_{sal}}{R_A} \quad (II-2.3)$$

Resolviendo esta igualdad y despejando V_{sal} tenemos:

$$\frac{V_1 - V_A}{R_1} = \frac{V_A - V_{sal}}{R_A} \quad (II-2.4)$$

$$R_A V_1 - R_A V_A = R_1 V_A - R_1 V_{sal} \quad (II-2.5)$$

$$R_1 V_{sal} = R_1 V_A + R_A V_A - R_A V_1 \quad (II-2.6)$$

$$R_1 V_{sal} = (R_1 + R_A) V_A - R_A V_1 \quad (II-2.7)$$

$$V_{sal} = \left(\frac{R_1 + R_A}{R_1} \right) V_A - \frac{R_A}{R_1} V_1 \quad (II-2.8)$$

Cabe destacar que esta ecuación es válida para un amplificador no inversor con voltaje de entrada V_A , junto con la ecuación del voltaje de salida para un amplificador inversor con voltaje de entrada V_1 , pero lo que queremos es dejar esta ecuación en términos de sus voltajes de entrada V_1 y V_2 . Para esto sabemos que V_2 se relaciona con V_A por el divisor de tensión:

$$V_A = V_2 \frac{R_A'}{R_2 + R_A'} \quad (II-2.9)$$

Si sustituimos esta relación en la ecuación anterior se tiene:

$$V_{sal} = V_2 \left(\frac{R_A'}{R_2 + R_A'} \right) \left(\frac{R_1 + R_A}{R_1} \right) - \frac{R_A}{R_1} V_1 \quad (II-2.10)$$

Esta ecuación se puede simplificar más sabiendo que $R_1 = R_2$ y $R_A = R_A'$, entonces:

$$V_{sal} = V_2 \left(\frac{R_A}{R_1 + R_A} \right) \left(\frac{R_1 + R_A}{R_1} \right) - \frac{R_A}{R_1} V_1 \quad (II-2.11)$$

simplificando:

$$V_{sal} = \frac{V_2 R_A}{R_1} - \frac{V_1 R_A}{R_1} \quad (II-2.12)$$

es decir,

$$V_{sal} = \frac{R_A}{R_1} (V_2 - V_1) \quad (II-2.13)$$

Para el caso de la tarjeta, la señal de video se alimentó en el voltaje V_2 , mientras que el voltaje V_1 se conectó a tierra. Como se dijo antes, una de las funciones de este amplificador en la tarjeta fue la de amplificar un poco más la señal de video para acondicionarla de modo de aprovechar el convertidor A/D. La ganancia fijada finalmente

para este propósito fue de 1.2, por lo que se emplearon los siguientes valores de resistencias:

$$R_1 = 10 \text{ kohms.}$$

$$R_A = 12 \text{ kohms.}$$

Inicialmente se intentó con una ganancia mayor, pero esta ganancia distorsionaba un poco la señal de video, produciendo unas ondas de ruido en la imagen. Además, la amplitud obtenida con esta ganancia final, da la oportunidad de ajuste de la captura en lo que toca a la brillantez de la imagen por medio de un potenciómetro que da el voltaje de referencia del convertidor A/D.

Además, la señal de video se alimenta al separador de sincronía, el cual nos da la señal de sincronía compuesta, la señal de sincronía vertical, la señal de campo par/impar y la señal de burst.

2.2.2 El separador de sincronía, LM1881

El separador de sincronía de video extrae la información de tiempo de una señal de video compuesto con sincronía negativa NTSC, PAL o SECAM, con una amplitud posible de 0.5 a 2 Vpp, fig. II-2.3.

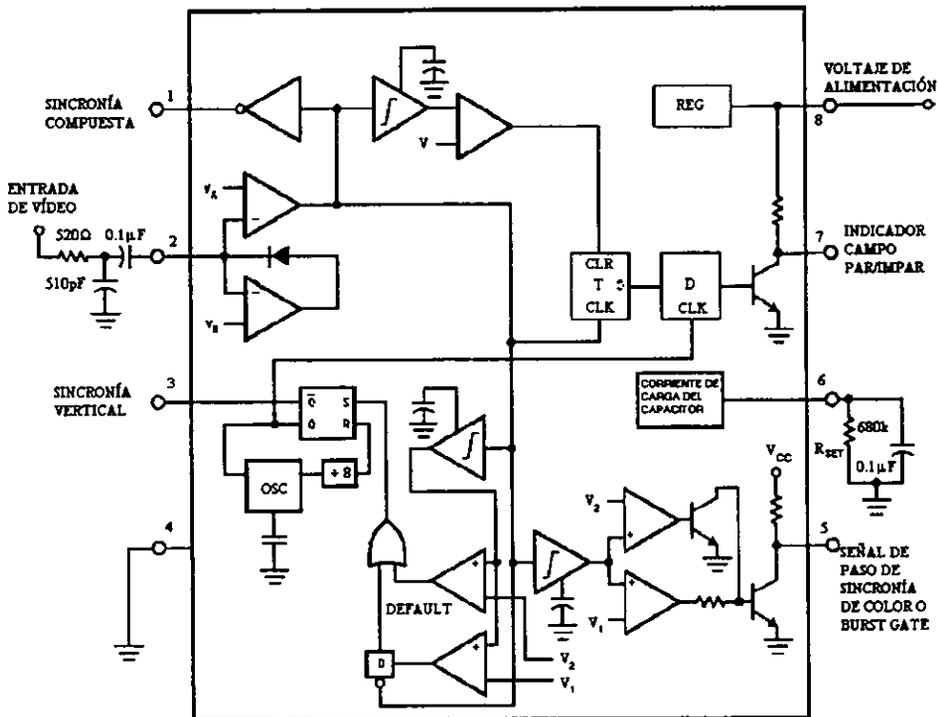


Fig. II-2.3 El separador de sincronía, LM1881.

Este circuito puede operar con voltaje de alimentación entre 5 y 12 volts, en nuestra tarjeta lo alimentamos a 5 volts para que las señales de salida fueran de 0 a 5 V, conservando compatibilidad con TTL. Para el funcionamiento del circuito se requieren pocos componentes externos. A parte del capacitor de acoplamiento, se requiere una resistencia que ajusta al LM1881 para poder funcionar con señales cuya frecuencia de línea es distinta de 15.734 kHz.

Como se señaló antes, este separador nos da cuatro señales de información de tiempo acerca de la señal de entrada: la señal de sincronía compuesta, que incluye la señal de sincronía vertical junto con la horizontal; la señal de sincronía vertical; la señal de campo y la señal de paso de burst. La señal de campo nos identifica cuál campo de una señal de video entrelazado se encuentra disponible.

La señal de sincronía compuesta es simplemente la reproducción de la señal de entrada recortada abajo del nivel de negro. Para esto las puntas inferiores de los pulsos de sincronía de la señal de video de entrada son elevados a 1.5 Volts DC, entonces se usa un comparador interno con un umbral de comparación de 70mV arriba del nuevo nivel de referencia, este umbral es independiente de la amplitud de entrada, por lo que para el caso de una entrada con amplitud de 0.5 Volts el umbral se encontrará casi a la mitad de la amplitud del pulso de sincronía. La corriente de entrada para el capacitor de acoplamiento de entrada es de 0.8 mA, mientras que la de descarga es de sólo 11 μ A, lo cual permite valores de capacitores relativamente bajos, el recomendado y el empleado para la tarjeta fue de 0.1 μ F.

Además, cuando no se cuenta con una señal completamente limpia de ruido que pudiera interferir con la separación, se puede poner un filtro paso bajas a la entrada que corte aproximadamente hasta 500 kHz, lo cual es suficiente para obtener la separación deseada. Esto se emplea en ocasiones cuando se trabaja con señales que tienen, por ejemplo, sus pulsos de burst para la cromina dentro de las líneas del blanking vertical.

Para obtener la señal de sincronía vertical es necesario integrar la señal anterior. Por eso se cuenta con un par de comparadores, un comparador tiene un voltaje generado internamente V_1 , mientras que el otro comparador tiene otro voltaje de referencia generado internamente V_2 , pero los dos comparadores tienen en su entrada no-invertida la señal común que viene del integrador. Dicho integrador recibe la señal de sincronía compuesta. El capacitor del integrador es también interno. La corriente de carga del capacitor es fijada por la resistencia externa R_{SET} . La salida del integrador estará en bajo voltaje cuando estén presentes las líneas normales de sincronía horizontal, esto porque el integrador tiene un tiempo muy corto de carga para el capacitor. En el momento en que se presenten los pulsos de equalización se mantendrá el voltaje en un nivel casi continuo, abajo del voltaje V_1 . Durante el periodo de sincronía vertical los pulsos angostos positivos son conocidos como pulsos dentados. En el inicio del pulso de sincronía vertical, antes de que el primer pulso dentado ocurra, el integrador carga el capacitor a un voltaje más alto. En el primer pulso dentado el voltaje del integrador se encontrará entre V_1 y V_2 . Esto va a dar un nivel alto a la salida del comparador con V_1 . Este nivel de alto hará que el flip-flop D pase a alto cuando se presente el flanco de bajada del primer pulso dentado. La salida del flip-flop D va a una compuerta OR y ésta se dirige a la entrada S del flip-flop R/S de salida, la salida Q de este

flip-flop habilita un oscilador, el cual es a su vez, dividido entre 8, va a la entrada R (reset) del flip-flop. La frecuencia del oscilador es fijada por un capacitor interno y la resistencia de ajuste R_{SET} . La salida /Q del flip-flop es la que va a la salida en el pin 3 del circuito. La duración del pulso de sincronía vertical obtenido dependerá de la frecuencia de oscilación elegida.

Con una R_{SET} demasiado grande, la corriente de carga del integrador va a ser demasiado pequeña para cargar el capacitor a V_1 , por lo que no se tendrá pulso de sincronía a la salida. Si el oscilador corre demasiado rápido en sus ocho ciclos, la salida va a ser demasiado corta, con esta condición se puede generar un pulso de sincronía doble por ser el primer pulso de salida más corto en relación al tiempo en que pasan los pulsos dentados.

El LM1881 también puede generar un pulso de sincronía vertical por default, cuando el periodo de sincronía es inusualmente largo y no cuenta con pulsos dentados. Esto porque al tener el integrador demasiado tiempo para la carga del capacitor, a la salida de éste se superará el voltaje V_2 , que está en uno de los comparadores. Como la salida de este comparador va al flip-flop D, entonces no habrá flanco de bajada, por lo que el único estado alto en la compuerta OR será el de este comparador e iniciará el pulso de sincronía al cambiar el estado del flip-flop R/S.

Para detectar el campo par/impar, se integra nuevamente la señal de sincronía compuesta. Un capacitor interno es cargado durante el periodo entre los pulsos de sincronía y descargado cuando el pulso de sincronía se presenta. El periodo normal entre los pulsos de sincronía horizontal es suficiente para permitir que el voltaje del capacitor alcance el nivel de umbral de un comparador que limpia el flip-flop, el cual tiene por reloj la forma de onda de sincronía. En el momento en que el intervalo vertical es alcanzado, el corto tiempo de integración entre los pulsos de equalización evita que el nivel de umbral sea alcanzado y la salida Q del flip-flop es cambiada por cada pulso de equalización. Ya que el periodo de la mitad de línea en el final del campo impar va a tener el mismo efecto como un periodo de los pulsos de equalización, la salida Q va a tener una polaridad diferente en campos sucesivos. De esta manera por comparar la Q con el pulso de salida vertical, un índice de campo par/impar es generado. El pin 7 permanece en bajo cuando está presente el campo par y será alto cuando pase el campo impar.

En una señal de vídeo compuesto, la señal de burst de la cromina está localizada sobre el "backporch", es decir, la sección de la señal compuesta de vídeo entre el nivel del pulso de sincronía horizontal y el final del periodo de blanking horizontal (donde comienza la información). Este periodo es aproximadamente de 4.8 μ s y es también el nivel de referencia negro para línea de escaneo subsecuente. El LM1881 genera un pulso en el pin 5, el cual puede ser usado ya sea para recuperar el burst de cromina de la señal compuesta de vídeo o como un control para tomar el nivel de restauración de CD de la señal de vídeo. Esta salida es obtenida simplemente por cargar un capacitor interno, iniciando sobre el nivel de retorno del pulso de sincronía horizontal. Simultáneamente, la salida del pin 5 es jalada a nivel bajo y mantenida hasta el tiempo de carga del capacitor, es decir, 4 μ s después.

Para el funcionamiento de la tarjeta nosotros requerimos de las señales de sincronía vertical y de paso de burst. Mediante la señal de sincronía vertical nosotros podemos discriminar un campo, ya sea el par o el impar, para solamente trabajar con un campo. Lo

anterior porque la parte de la señal que nosotros estamos capturando es de 256 x 256 para llenar el búfer de 64kb, y si se sabe que un campo se compone de 262.5 líneas, entonces sólo es necesario trabajar sobre uno solo. Como la señal de sincronía vertical se presenta en cada campo, es necesario dividir la frecuencia de esta sincronía con un flip-flop tipo D para obtener sólo los pulsos de sincronía de un campo. Inicialmente se trató de obtener esta señal de control directamente de la salida de campo par/impar, pero por cuestiones del manejo de umbrales y flip-flops internos en el separador de sincronía, se producía una cierta inestabilidad a la hora de capturar cuadros consecutivos, lo que ocasionaba que la imagen brincara de vez en cuando. La señal de paso del burst coincide con el comportamiento de la información visible del vídeo, por lo que es útil para separar esta parte de la señal para alimentarla al convertidor A/D.

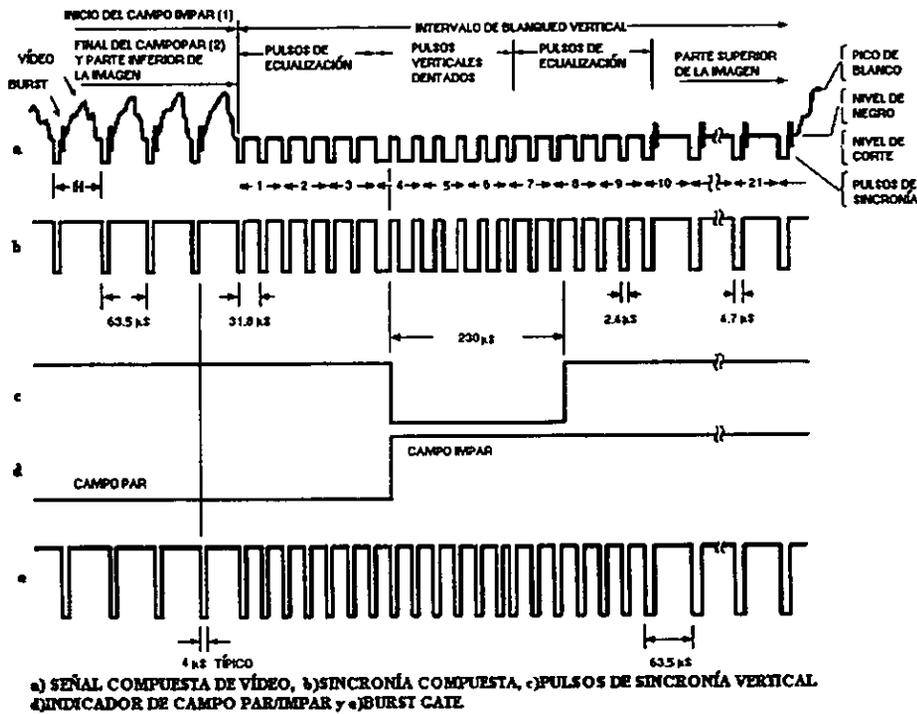


Fig. II-2.4. Señales obtenidas del separador de sincronía.

2.2.3 Referencia de negro a tierra

Contando con que a la salida del amplificador de entrada diferencial tenemos una señal de vídeo amplificada, es necesario separarla para dejar solamente la parte de información visible, lo cual se logra mediante los siguientes amplificadores operacionales. El amplificador inmediato se trata de otro amplificador de entrada diferencial, pero en esta

ocasión tiene ganancia unitaria, en la terminal de entrada negativa se alimenta la señal de vídeo amplificada proveniente del primer amplificador. La terminal de entrada positiva es alimentada con un nivel de referencia de CD, el cual corresponde al nivel en donde se encuentra el negro de nuestra imagen. Para obtener este nivel de referencia se emplea un circuito de muestreo y retención, el LF398, el cual en coordinación con la señal invertida de paso de burst y un filtro paso bajas nos da el nivel de negro de la señal de vídeo.

2.2.3.1 Circuito de retención y muestreo, LF398

Este es un circuito monolítico; una sección de entrada bipolar es usada para obtener un offset bajo y un amplió ancho de banda. Este circuito puede trabajar sin complicaciones en un circuito retroalimentado con amplificadores operacionales hasta 1 MHz sin tener problemas de estabilidad. Su impedancia de entrada es de 10^{10} ohms, lo cual permite tener fuentes de alta impedancia sin afectar el rendimiento. Dispositivos FET canal P están combinados con transistores bipolares en el amplificador de salida para dar relaciones de inclinación de hasta 5 mV/min. con un capacitor de toma de $1 \mu\text{f}$.

2.2.3.2 Filtro paso bajas

Para filtrar la señal de referencia de vídeo y dejarla limpia de cualquier rizo que interfiriera, se empleó un sencillo filtro de primer orden como el que se muestra en la figura II-2.5. Como se puede ver, se emplean componentes pasivos, como la resistencia y el capacitor, pero cuenta con un componente activo, como lo es el amplificador operacional. Este filtro sólo deja pasar las componentes de la señal que se encuentran abajo de la frecuencia de corte.

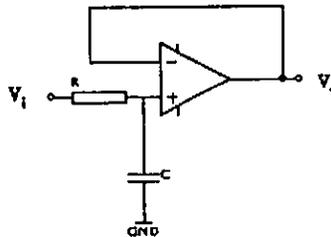


Fig. II-2.5 Filtro paso bajas de primer orden.

Un filtro paso bajas de primer orden emplea un capacitor y una resistencia, tiene una pendiente de 20 dB por década. Ya que el amplificador empleado aquí está en configuración seguidor común, la ganancia del filtro será unitaria; su frecuencia de corte es de

$$f_0 = \frac{1}{2RC\pi} \quad (\text{II-2.14})$$

Para este caso el capacitor es de 1 μf y la resistencia de 1 kohm, por lo la frecuencia de corte es de :

$$f_0 = \frac{1}{2(1k)(1\mu f)(3.1416)} = 159.15\text{Hz} \quad (\text{II-2.15})$$

2.2.3.3 Circuito completo

Recordando que la señal de video compuesto se encuentra en nivel de negro o infranegro en el momento en que se presenta la señal de burst o de sincronía de color, se guardará en el capacitor del muestreador este nivel, el cual será tomado en el momento en que la señal de control del LF398 se coloque en alto, es decir, durante el pulso alto de la señal de paso de burst invertida con la ayuda de un inversor "high-speed cmos" 74HC04 y dicha carga se retendrá hasta la siguiente activación del circuito por conducto de la señal de paso de burst nuevamente. Este nivel de negro no es totalmente recto, por lo que se emplea un filtro paso bajas para remover las señales de alta frecuencia sobre la señal de CD que corresponde al nivel negro. Una vez obtenido este nivel se resta, en el segundo amplificador de entrada diferencial, a la señal de video amplificada desde el amplificador de entrada y obtener así el video con el nivel de negro referenciado a tierra, tal y como se muestra en la figura II-2.6.

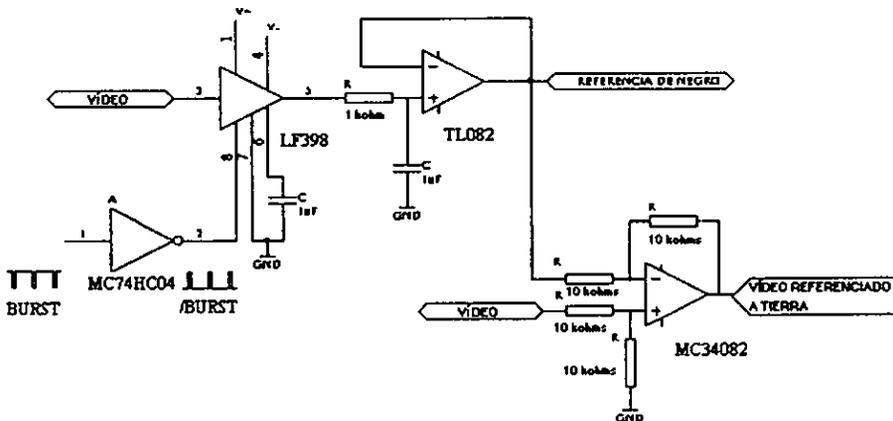


Fig. II-2.6 Circuito de referencia a tierra.

2.2.4 Rectificador de precisión

A la salida de este segundo amplificador de entrada diferencial se tendrá la señal de video amplificada y con un nivel de negro coincidente con el nivel de tierra. Esta señal deberá ser rectificadora para remover toda la información de sincronía y control, y así dejar sólo la información visible. Esto se alcanza usando un amplificador operacional en configuración rectificador de precisión de media onda.

Los amplificadores operacionales pueden operar muy cercanamente al diodo ideal. Para lograrlo, se emplea la retroalimentación negativa para hacer que el amplificador funcione como un rectificador con voltaje de barrido muy aproximado a cero y con una característica lineal de entrada-salida. Las curvas características se pueden apreciar en la figura II-2.7 para el diodo ideal y para el diodo de silicio.

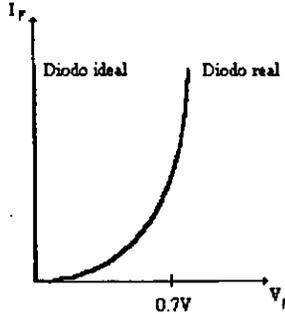


Fig. II-2.7 Curvas características del diodo.

El tipo de rectificador empleado para la tarjeta fue uno de media onda, esto porque sólo nos interesa la información arriba del voltaje de barrido, en este caso tierra. La operación de este circuito es explicada en dos partes. Como se puede ver en la figura II-2.8, se tienen marcados dos distintos sentidos de corriente, las flechas continuas corresponden a los semiciclos positivos de un señal senoidal de entrada, mientras que las flechas punteadas corresponden a los semiciclos negativos. Cuando la señal es positiva, la señal de salida del amplificador comienza a ser negativa, conduciendo D_1 . Ya que el diodo D_1 se encuentra en estado de conducción, la salida del amplificador (V_o) va alcanzar un nivel máximo cercano a $-0.7V$ sin importar como va el voltaje positivo de entrada (V_i). Esta situación no es suficiente para que D_2 conduzca, por lo que el voltaje V_o será cero volts.

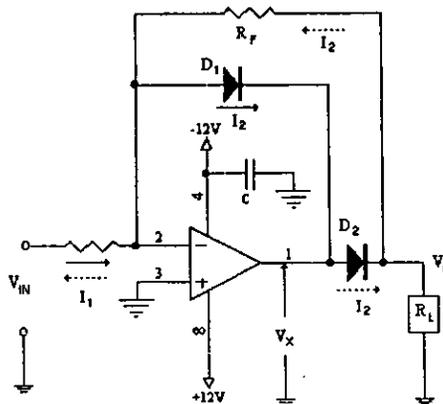


Fig. II-2.8 Rectificador de precisión de media onda.

Cuando se presentan los semiciclos negativos de entrada, el diodo D_1 permanece en corte, por lo que la retroalimentación negativa reduce efectivamente el voltaje de barrido de D_2 a $0.7 V/A$, lo cual como se puede ver para fines prácticos se puede considerar como cero, donde A es la ganancia en malla abierta del amplificador. De esta forma si se tendrá señal de salida invertida con respecto a la entrada. La ganancia en esta sección estará dada por la ecuación:

$$A = -R_F/R_I \quad (II-2.16)$$

Cabe señalar que el hecho de conectar el diodo D_2 dentro de la retroalimentación negativa reduce la no linealidad de la curva del rectificador. El diodo D_1 tiene como función la de mantener V_x a voltajes negativos pequeños.

Los diodos empleados fueron los 1N4148 los cuales son rápidos para conmutar, ya que se manejan altas frecuencias, evitando también altas capacitancias del diodo que pudieran causar errores.

2.3 Digitalización

El proceso de digitalización comienza después de que se ha obtenido la información de vídeo visible totalmente pura al final del rectificador de precisión, esta información pasará a través del convertidor A/D, en donde éste se encargará de muestrear y cuantizar el vídeo, se hará una cuantización a 8 bits, cada muestra será una parte diminuta de la representación digital de la imagen, dicha parte se conoce como *pixel*, obteniéndose con ello una señal que podrá ser visualizada en 256 tonos de grises.

2.3.1 El convertidor Analógico-Digital

El convertidor que se empleó en esta tarjeta fue un convertidor analógico digital del tipo flash. Este convertidor es de 8 bits y emplea una estructura interna de código *Grey*¹⁵, con el fin de evitar grandes errores en la salida cuando se tienen a la entrada señales muy rápidas. Este circuito es compatible con TTL. Requiere una fuente de alimentación de 5 V y una fuente negativa desde -3.0 a -6.0 V. Está provisto además de salidas tres estados. Puesto que es un convertidor de 8 bits, cuenta con 256 comparadores. Las salidas de los comparadores están conectadas a latches y de ahí se conectan a la red codificadora, la cual produce la salida de 8 bits y un bit de sobreflujo. El diagrama de bloques del circuito integrado se muestra en la figura II-2.9.

Los comparadores están conectados de manera de que una de las entradas de cada comparador tenga un voltaje de referencia espaciado, de acuerdo a una arreglo de resistencias. Por otro lado, las otras entradas de cada comparador están conectadas entre sí para recibir el voltaje de la señal de entrada. Cuando la señal de entrada está presente, algunos comparadores tendrán a la salida un uno lógico, mientras que otros tendrán un cero lógico, dependiendo de su posición relativa. Las salidas de estos comparadores son

¹⁵ Louis Nashelsky; *Fundamentos de Tecnología Digital*; Limusa, México, 1993; pag. 77-80.

mantenidas y convertidas a un código Grey por medio de un arreglo diferencial de latches. Al emplear el código Grey se evita que cualquier error producido por el sobrecruce cause problemas con la retroalimentación o disparidades de tiempo "glitches" en la salida con los bits menos significativos.

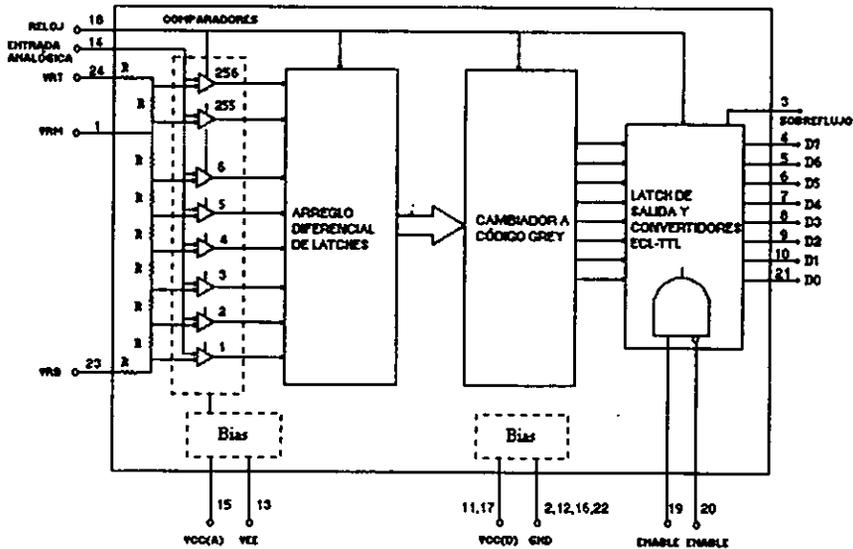


Fig. II-2.9 El convertidor A/D, MC10319P.

El código Grey es entonces transformado a un código binario de 8 bits y los niveles diferenciales son trasladados a niveles TTL antes de ser aplicados al latch de salida. Con ayuda de la terminal de habilitación (enable) se pueden colocar las salidas de este último latch en alta impedancia.

En la sección analógica, la salida depende de la posición relativa de cada comparador con respecto a su voltaje de referencia. Los comparadores tienen un ancho de banda de más de 50 MHz, por lo que este circuito es capaz de recibir entradas de hasta 12.5 MHz, respetando el teorema de Nyquist.

La corriente en el Pin 14 varía linealmente desde 0, en el momento en que el voltaje de entrada es igual al V_{RB} , hasta $60 \mu A$, en el momento en que es igual a V_{RT} . Para el caso en que el voltaje se encuentre abajo del voltaje V_{RB} o arriba del voltaje V_{RT} , la corriente permanecerá en el valor de V_{RB} o V_{RT} , respectivamente. Sin embargo, debe de ser mantenida dentro del rango absoluto de $\pm 2.5 V$ con respecto a tierra, de lo contrario habrá corrientes excesivas en el Pin 14 debido a problemas internos.

La capacitancia de entrada en este Pin es de aproximadamente 36 pf si la diferencia entre los voltaje de referencia es de 2.0 V, y se incrementa a 55 pf si la diferencia se reduce 1.0 V.

La impedancia de la fuente debe de ser mantenida debajo de 100 ohms para evitar errores de muestreo.

La escalera de voltajes de referencia está compuesta por resistencias del mismo valor lo que produce 256 voltajes igualmente espaciados. La diferencia entre voltajes adyacentes corresponde al voltaje de un bit menos significativo del rango de entrada. El comparador más cercano al voltaje V_{RB} está referenciado $\frac{1}{2}$ del voltaje de un bit menos significativos arriba de V_{RB} , mientras que el último comparador está colocado $\frac{1}{2}$ del voltaje de un bit menos significativos abajo de V_{RT} . El valor nominal de la cadena de resistencias completa es de 130 ohms con una tolerancia de $\pm 20\%$, por lo que requiere una corriente de 15.4 mA a 2.0 V y de 7.7 mA a 1 V.

La diferencia mínima entre los voltajes de referencia es recomendada para ser de 1 V. Una diferencia menor causará offsets y no linealidades que pueden llegar a ser significantes. La diferencia máxima recomendada es de 2.1 V por las limitaciones de potencia de la cadena de resistencias. La diferencia puede encontrarse en cualquier parte dentro del rango de -2.1 a +2.1 V con respecto a tierra, y V_{RB} debe de ser al menos 1.3 V más positivo que V_{EE} . Los voltajes de referencia deben de ser estables y libres de ruido y rizados, ya que la calidad de la conversión está relacionada con la calidad de la referencia.

El Pin 1 V_{RM} es el punto medio de la cadena de resistencia, excluyendo el comparador de sobreflujo. El voltaje en V_{RM} es : $(V_{RT} + V_{RB})/2.0 - \frac{1}{2} \text{ LSB}$

En la mayoría de las aplicaciones un capacitor de bypass a tierra de 1 μf es suficiente para mantener la calidad.

La fuente de voltaje $V_{CC(A)}$ es la alimentación positiva para los comparadores y $V_{CC(D)}$ es la alimentación positiva para la porción digital. Ambas deben de ser de +5.0 V $\pm 10\%$ con una diferencia de 100 mA entre sí. V_{EE} es la fuente de alimentación negativa para los comparadores y debe estar dentro del rango de -3.0 a -6.0 volts y debe ser al menos 1.3 V más negativa que V_{RB} .

La entrada del reloj es compatible con TTL con un rango de 0 a 30 MHz. No existen limitaciones del ciclo de trabajo. La secuencia de conversión se muestra en la figura II-2.10.

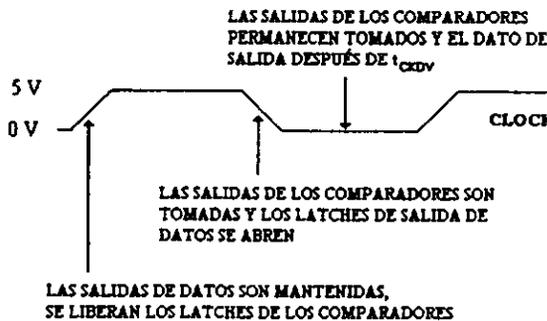


Fig. II-2.10 Secuencia de conversión.

Sobre el flanco de subida, los latches de salida de datos mantienen el dato anterior, y los latches de la salida de los comparadores están descargados para seguir la señal de entrada (V_{IN}). Durante el tiempo alto los comparadores rastrean la señal de entrada. El latch de salida de datos retiene el dato anterior.

En el flanco de bajada, las salidas de los comparadores son tomadas con el dato inmediatamente después de este nivel. La conversión a digital ocurre dentro del dispositivo, y el latch de salida de datos es liberado para indicar el nuevo dato dentro de 20ns.

Durante el tiempo de reloj bajo, las salidas de los comparadores permanecerán fijos, y el latch de salida permanecerá transparente.

Un sumario de las secuencias es que el presente dato en V_{in} justo después del flanco de bajada es digitalizado y disponible en la salida de datos inmediatamente después de ese mismo flanco de bajada.

Los latches de salida de los comparadores dan al circuito la función de mostrar y tomar, eliminando la necesidad de un circuito externo.

Las dos entradas de habilitación son compatibles con TTL y son usadas para cambiar la salida de datos (D7-D0) del estado activo a 3 estados. Para que las salidas estén activas, el Pin 19 debe de estar en uno lógico y el Pin 20 debe de estar en cero lógico. Cambiando cualquiera de las entradas, se pondrán las salidas en el estado de alta impedancia. Las entradas de habilitación afectan solamente el estado de las salidas, no anulan la conversión. El Pin 3 de sobreflujo no se afecta por las entradas de habilitación, por lo tanto, no cuenta con la característica tres estados.

La salida de sobre flujo va a alto cuando la entrada V_{in} es más positiva que $V_{RT} - \frac{1}{2}$ LSB. Esta salida siempre está activa.

2.3.2 Oscilador de cristal

Los osciladores son dispositivos muy confiables ya que dan una gran estabilidad y calidad necesarias en muchas aplicaciones, sobre todo para casos como el de nuestra tarjeta.

La potencia disipada en un cristal está referida al nivel que maneja y está especificada en mW. En bajos niveles de manejo, la resistencia resonante del cristal puede ser tan grande como para causar problemas de inicio. Para superar este problema, el elemento activo, como un inversor, debe de dar suficiente amplificación, pero no demasiada como para sobre empujar el cristal.

Un cristal puede ser modelado de la siguiente forma, fig. II-2.11:

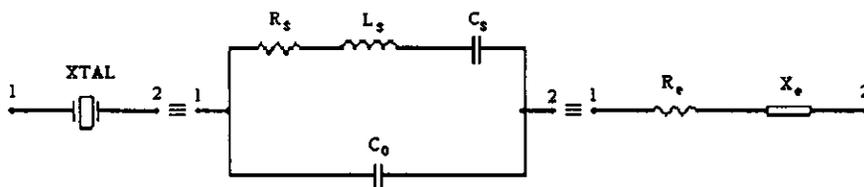


Fig. II-2.11. Circuito equivalente del cristal.

La figura II-2.12 muestra la configuración más popular en el uso del cristal. Para elegir R1, es necesario saber que la potencia es disipada en la resistencia serie efectiva del cristal. El nivel de manejo especificado por el fabricante es la máxima presión que el cristal puede aguantar sin daño o un excesivo corrimiento de frecuencia. R1 limita el nivel de manejo.

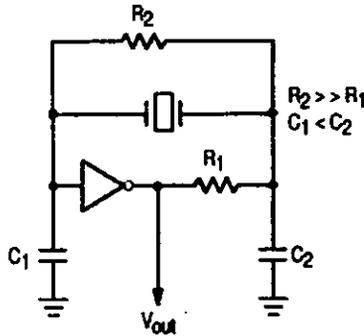


Fig. II-2.12 Configuración del cristal como reloj.

Para verificar que el máximo voltaje de alimentación de DC no sobre empuje el cristal, es necesario supervisar la frecuencia en la salida V_{out} . La frecuencia se puede incrementar muy ligeramente cuando el voltaje de alimentación de DC es incrementado. Un sobre empuje del cristal decrementa en frecuencia, o llega a ser inestable con un incremento en el voltaje de alimentación. El voltaje de alimentación de operación debe de ser reducido o $R1$ debe de ser incrementado en valor si la condición de sobre empuje existe. El tiempo de puesta en marcha depende del valor de $R1$.

La resistencia de retroalimentación ($R2$) típicamente es de hasta 20 Mohms. $R2$ determina la ganancia y el ancho de banda del amplificador. Un apropiado ancho de banda asegura una oscilación en la frecuencia correcta. La $R2$ debe de ser lo bastante grande como para no afectar la fase de la red de retroalimentación en una manera apreciable.

Para el caso del reloj de la tarjeta sólo se empleó la resistencia $R2$, la cual fue de un valor de 10 Mohms, los valores de los capacitores fueron de 22 pf y 47 pf, como se muestra en la figura II-2.13, y se empleó además un divisor para bajar la frecuencia obtenida del cristal a 6 MHz, un flip-flop sirvió para este propósito.

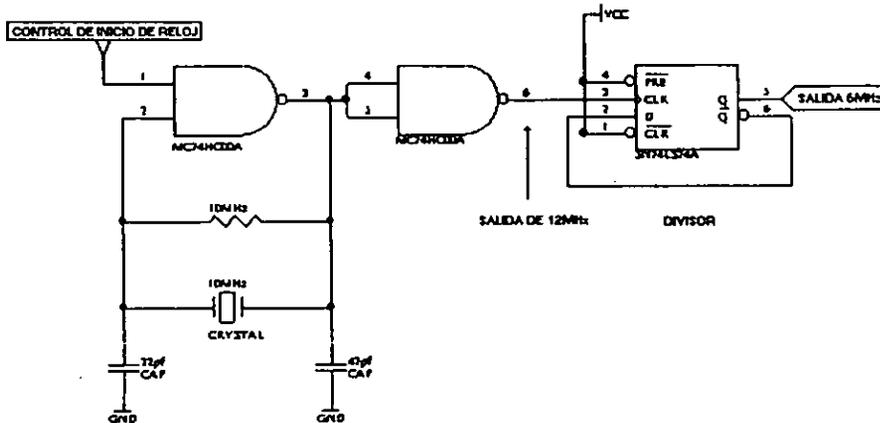


Fig. II-2.13. Reloj implementado.

Además, como se puede observar en la figura, en lugar de los tradicionales inversores se emplearon compuertas NAND. Conociendo el comportamiento de esta compuerta, se puede controlar el inicio de las oscilaciones del reloj colocando un uno lógico en una de la terminales del la compuerta principal o un cero cuando no se desee que el reloj actúe. Para ahorrar problemas de circuitos integrados adicionales se usa otra compuerta NAND como búfer.

2.3.3 Búfer de memoria RAM estática

La zona de almacenamiento empleada en este circuito fue la unión de dos memorias de 32 kb para obtener una capacidad de 64 kb, las memorias empleadas fueron los circuitos 43256-100, son memorias del tipo RAM estática, la elección de este tipo de memoria fue con el fin de no entrar en problemas con los dispositivos de refresco de un memoria RAM dinámica, además de que esta memoria cumplía con las especificaciones necesarias para poder trabajar a las velocidades de la tarjeta. A continuación se muestra un diagrama de bloques de la memoria, fig. II-2.14.

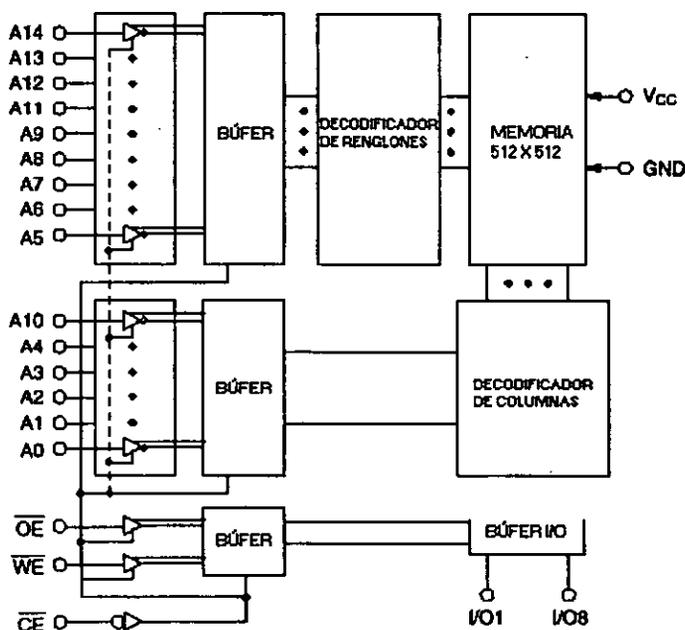


Fig. II-2.14. Memoria RAM estática.

Las características principales de este circuito son las siguientes:

- La capacidad de cada circuito es de 32,768 X 8 bits.

- Tiene un tiempo de acceso bajo comparado con otras memorias RAM estáticas, este tiempo es de 100 ns, el cual es suficiente para las frecuencias manejadas en el reloj de conversión.
- Además cuenta con una potencia de disipación muy baja, están hechas bajo la tecnología MOS. Cabe desatacar que, ya que se trata de una tarjeta de expansión, la potencia disipada por toda la tarjeta debe de ser la menor posible.
- Su voltaje de alimentación es de 5 V con una tolerancia del 10%.
- Su tiempo de acceso es el mismo que el tiempo de ciclo de lectura o escritura.
- El bus de datos es bidireccional y cuenta con una entrada /OE, la cual nos da la condición de tres estados.
- Es compatible con TTL.

2.3.4 Contadores de direccionamiento

Las muestras generadas por el convertidor A/D se irán almacenando secuencialmente en las memorias RAM, provistas en la tarjeta para hacer un almacenamiento temporal. La memoria será direccionada por medio de los contadores 74HC193. Este circuito es un contador síncrono binario de 4 bits hacia arriba o hacia abajo. La condición de sincronía está dada porque los flip-flops de conteo son cambiados de estado simultáneamente, con esto se asegura que todas las salidas funcionarán al mismo tiempo. El cambio de estado en los flip-flops se presenta con una transición de bajo a alto del reloj. La dirección de conteo está determinada por la entrada de reloj que es empleada, mientras la otra entrada debe de permanecer en alto. Este dispositivo tiene la ventaja de que se puede programar, esto gracias a unas entradas de carga a través de las cuales se fija el estado en el que queremos que cuente y colocando la entrada de "LOAD" a bajo. Se tiene también una entrada para limpiar los flip-flops, es decir, colocar todas las salidas a bajo.

Este circuito fue diseñado para poderse conectar en cascada sin necesidad de un circuito externo, esta función es posible gracias las salidas "BORROW" y "CARRY". La salida "BORROW" produce un pulso de ancho igual a la entrada de conteo hacia abajo cuando el contador se desborda hacia abajo, de manera similar, el "CARRY" da un pulso de ancho igual al reloj de conteo hacia arriba cuando el contador se desborda hacia arriba. Para realizar una conexión en cascada con estos contadores, sólo hay que conectar el "BORROW" y el "CARRY" con la entrada de reloj hacia abajo y reloj hacia arriba, respectivamente del siguiente contador.

Los contadores son controlados por la presencia del campo deseado a digitalizar (par o impar), es decir, los contadores funcionarán sólo cuando el campo deseado esté presente y se detendrán cuando este ausente. Cuando el campo se presenta, es decir, cuando se presenta un cambio de bajo a alto en la entrada del reloj del flip-flop de campo, aparece un uno a la salida del flip-flop, con lo que se autoriza el paso del reloj para la digitalización. El momento en que los dos primeros contadores llegan al límite de su conteo (256), quiere decir que una línea de la imagen ha sido digitalizada, se genera un pulso que pasa por un diferenciador que limpiará el flip-flop controlador de línea para que espere a que la siguiente línea a digitalizar se presente, durante todo este tiempo el reloj también se detiene.

El uso de la red diferenciadora es con el fin de obtener un pulso pequeño para resetear el flip-flop sin trabar el ciclo de los contadores.

2.3.4.1 El diferenciador

El circuito diferenciador empleado aquí es un circuito RC. Como se puede observar en la figura II-2.15, este arreglo está hecho en forma de filtro paso altas, esto porque la reactancia del capacitor en serie disminuye para altas frecuencias. Cuando la frecuencia es cero la reactancia es infinita. Es por esto que se puede deducir, que cuando la entrada es constante no habrá paso de la señal a la resistencia de salida.

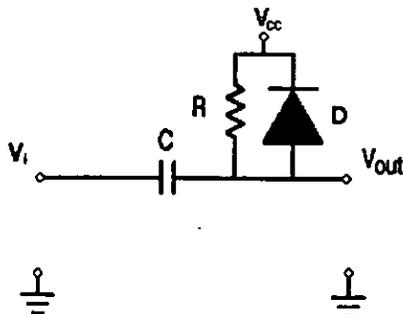


Fig. II-2.15. Circuito diferenciador.

Antes de que se produzca el pulso de fin de línea, el voltaje en la entrada es 5 volts y a la salida del circuito es también 5 volts, entonces la diferencia de voltaje en el capacitor es 0 volts, lo que significa que el capacitor estará descargado. En el momento que ocurre la transición de alto a bajo, el voltaje de entrada cambiará a 0 volts, lo que produce que el voltaje de salida caiga a cero volts, esto porque el capacitor no está cargado. A partir de este momento el capacitor se comenzará a cargar hasta 5 volts y la salida retornará a 5 volts como se muestra en la figura II-2.16.

Posteriormente, cuando ocurre el cambio de bajo a alto en la entrada, la suma de voltajes en la malla I nos dará 10 volts de salida (fig. II-2.17), esto porque además de los 5 volts de entrada, se tiene cargado el capacitor a 5 volts, con esto se tiene un diferencial de voltaje en el capacitor, por lo que éste se descargará hasta que en la salida se tenga un voltaje de 5 volts.

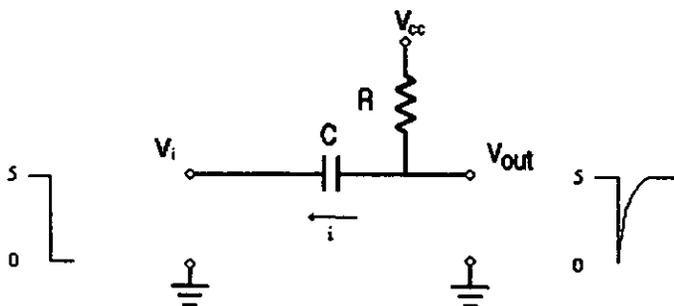


Fig. II-2.16. Salida para flanco de bajada.

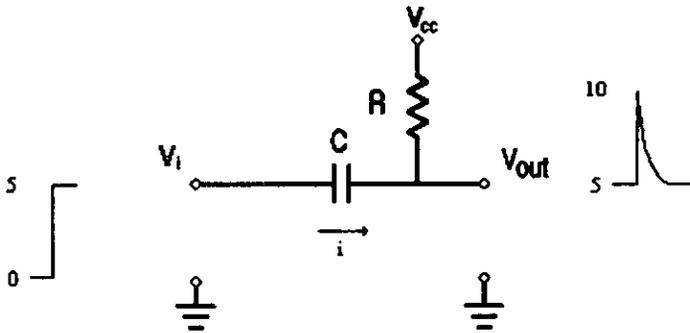


Fig. II-2.17. Salida para flanco de subida.

Juntando las dos secciones, se obtiene el resultado como ilustra la figura II-2.18. Como se puede observar, el pico positivo de salida excede los 5 volts de salida TTL, por lo que es necesario colocar un diodo volante para controlar el exceso de voltaje a sólo 0.7 V arriba de los 5 V. La parte que nos interesa de esta salida son los picos negativos, con los cuales se resetea el flip-flop controlador de línea, de manera de no trabar la captura.

Los últimos contadores serán los encargados de llevar el conteo de las líneas restantes para formar la imagen completa. Cuando se termina de digitalizar todo un cuadro, un pulso proveniente del contador cuatro resetea el flip-flop controlador de cuadro y el reloj se desactiva, porque al resetear el flip-flop, el reloj queda bloqueado, por lo que no llega a los contadores, entonces la tarjeta se quedará en este estado indefinidamente hasta que se desee efectuar la lectura hacia la PC.

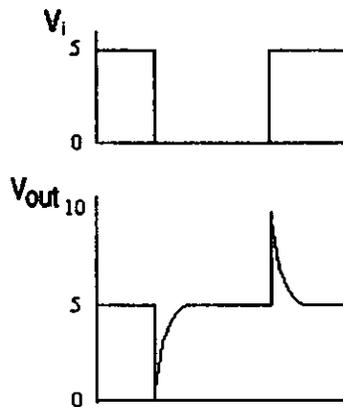


Fig. II-2.18. Respuesta a un pulso negativo.

Las memorias RAM que tenemos en la tarjeta son sólo para almacenamiento temporal, pues se tomarán cuadros continuamente. Para almacenar la información digitalizada por el convertidor A/D, la memoria es direccionada por los contadores en una forma secuencial, obteniéndose 65536 muestras que formarán una imagen. Cuando un cuadro completo ha sido capturado, se hace la lectura mediante el puerto, generando como reloj un tren de pulsos en uno de los pines del puerto de control del 8255 cuando la PC desea leer un byte. Este procedimiento se repite hasta finalizar totalmente de leer la memoria. La lectura puede hacerse también vía DMA, y se realizará hasta que el software de la autorización para que empleando la señal de READ de la computadora, se pueda obtener una señal de reloj para direccionar nuevamente, pero ahora para lectura de información de la RAM de la tarjeta hacia la PC. Con esto se asegura que la lectura será controlada de acuerdo a la velocidad de lectura de la PC.

2.4 Adquisición

Para la parte de la transmisión del cuadro de video del búfer de 64 kb a la memoria de la PC se probaron dos posibles caminos, uno de ellos fue la lectura de los bytes mediante un simple direccionamiento a puertos. Esto consiste en obtener un reloj de lectura mediante un tren de pulsos proveniente de uno de los pines del puerto de control del 8255. Este circuito, como ya se mencionó antes, sirve para obtener una interface de puerto paralelo y se logra con ayuda de un comparador de 8 bits, que es el 74HCT688. Cabe señalar, que el bus sobre el cual se trabajó la interface es el bus ISA. Dicho bus cuenta con señales de control, un bus de direcciones y un bus de datos. Una de las ventajas por las que se decidió trabajar directamente sobre el bus es de que éste nos da también la alimentación de +12, -12, +5 y -5 volts.

2.4.1 El bus ISA

Las ranuras de una computadora sirven para que el procesador de la computadora tenga la posibilidad de comunicarse con sus dispositivos externos. Prácticamente se puede conectar cualquier dispositivo a la ranura. Dichas ranuras cuentan con conexiones de alimentación, líneas de entrada y salida de datos, para el direccionamiento, para el manejo de interrupciones, el manejo del DMA, etc. Estas ranuras dan acceso directo a lo que se conoce como bus, a través del cual el procesador y cualquier dispositivo puede compartir conexiones e información.

El conector más popular en las computadoras personales es conocido como bus ISA (Industry Standard Architecture), figura II-2.19. También se conoce como bus AT, cuenta con un bus de datos de 16 bits, su conector tiene 98 patas colocadas en doble lado. Este conector está separado en dos secciones, una sección es la que corresponde al bus ISA inicial, conocido como XT, el cual cuenta con un bus de datos de 8 bits, la otra sección es la adición al conector para abrir el manejo de datos hasta 16 bits, aumentando más interrupciones, y el bus de direccionamiento. El bus original de la PC tenía solamente 20 líneas de dirección, que permitían alcanzar hasta 1 MB de memoria ($2^{20} = 1,048,576$ bytes).

La capacidad de direccionamiento en una computadora personal está en función directa del procesador, pero en general, el direccionamiento de puertos está limitado a 512,

de los cuales algunos ya están ocupados de fábrica en configuración de los periféricos comunes en la PC.

A continuación, se hace una breve descripción de la función de cada pata en dicho bus.

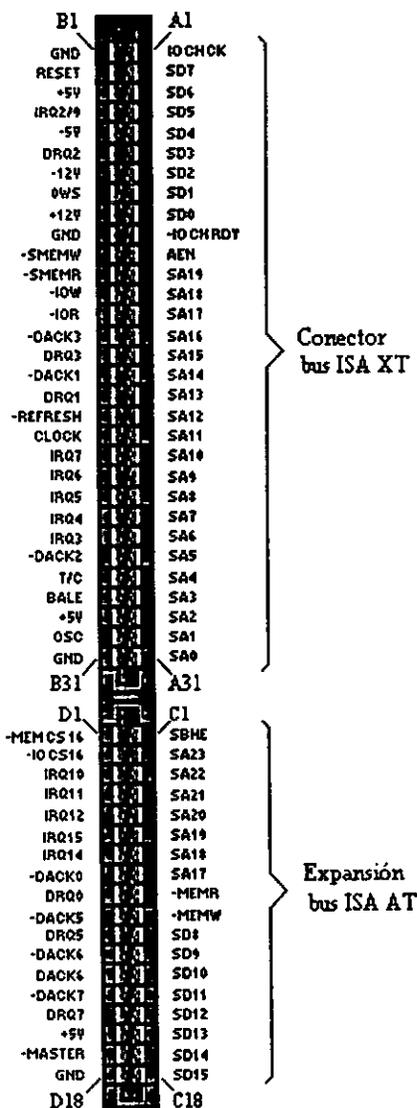


Fig. II-2.19. Conector ISA.

- SD0-SD15.** Primeramente se cuenta con el bus de datos. Estas líneas son bidireccionales. La dirección de datos depende de las señales -IOR , para leer, o -IOW, para escribir.
- SA0-SA19.** Estas líneas son de salida y sirven para efectuar el direccionamiento de las tarjetas o de la memoria.
- CLOCK.** Esta salida nos da la frecuencia a la que trabaja el micro.
- OSC.** Es una línea de salida con una señal de 14,318 MHz, la cual se emplea para las interfaces RS232 por ejemplo.
- RESET.** Es una salida proveniente del reset general de la computadora.
- BALE.** Es una salida empleada para validar las señales presentes en el bus de direcciones. Es empleada para sincronización de las direcciones con los datos, esta señal es dada por el controlador de bus (8288).
- I/O CH-CK.** Es una señal de entrada a la computadora que permite reportar al procesador que existe un error en el bus. Esta entrada produce una interrupción de alto nivel (NMI).
- IOW.** Esta salida señala que los datos han sido escritos, es decir, fueron puestos por el procesador.
- SMEMW y -SMEMR.** Estas señales son similares a las dos -IOW y -IOR respectivamente, pero sirven para direccionar memoria.
- DRQ1-DRQ7.** Estas son señales de entrada que sirven para pedir el inicio de un ciclo DMA y se dirigen a los dos controladores DMA conectados en cascada.
- DACK0-DACK7.** Son salidas que señalan el reconocimiento de un ciclo DMA.
- AEN.** Es una señal de salida, la cual sirve para mostrar que la dirección presente en el bus, es válida cuando está en bajo, y señala que se está llevando a cabo un ciclo DMA cuando está en alto.
- TC.** Esta señal de salida indica que se concluyó un ciclo DMA, en transferencia por bloques o por carácter.
- MASTER.** Esta salida es empleada por el procesador con DRQ para ganar el control del sistema.
- REFRESH.** Esta señal indica que el ciclo de refresco de la memoria está en proceso.
- NOWS.** Esta entrada sirve para acortar el ciclo de espera del procesador.

Además, se puede ver en la figura que se cuentan con patas que nos dan los voltajes alimentación para las tarjetas que se conectan a dichos bus. Los voltajes dados son: +12 V, -12 V, +5 V y -5 V.

2.4.2 El controlador de puertos 8255

Este circuito es una interface periférica programable. Un circuito que cuenta con 24 entradas/salidas, las cuales pueden ser programadas de forma individual en grupos de 12 y en tres modos de operación. En el modo 0, cada grupo de 12 puede ser programado en conjuntos de cuatro y 8, para ser entradas o salidas. En el modo 1, 8 líneas de entrada se eligen para ser de entrada ó de salida, de las cuatro líneas restantes, tres se emplean como

señales de control de interrupciones. En el modo 2, se tiene una configuración de bus bi-direccional.

El circuito cuenta con un bus de datos bi-direccional tres estados. Es el medio de comunicación del controlador con el CPU. Cuenta con lógica de control y las líneas /WR y /RD, todo lo anterior sirve para controlar las transferencia externas e internas. La forma de operación de cada puerto estará determinada vía software.

Como se puede ver en el diagrama del 8255, se cuenta con bloques de control , el grupo A y el B, el bloque de control del grupo A maneja el puerto A y las cuatro líneas superiores de C, mientras que el bloque de control B se encarga del puerto B y la parte baja de C.

Las 24 líneas pueden ser divididas en tres puertos de 8 líneas cada una, llamados puerto A, puerto B y puerto C. Esencialmente, tienen el mismo funcionamiento aunque hay diferencias en la forma de operación de sus latches. Una característica particular del puerto C es de que puede ser dividido a su vez en dos puertos de 4 bits.

El diagrama de bloques del circuito i8255A-5 se muestra en la figura II-2.20.

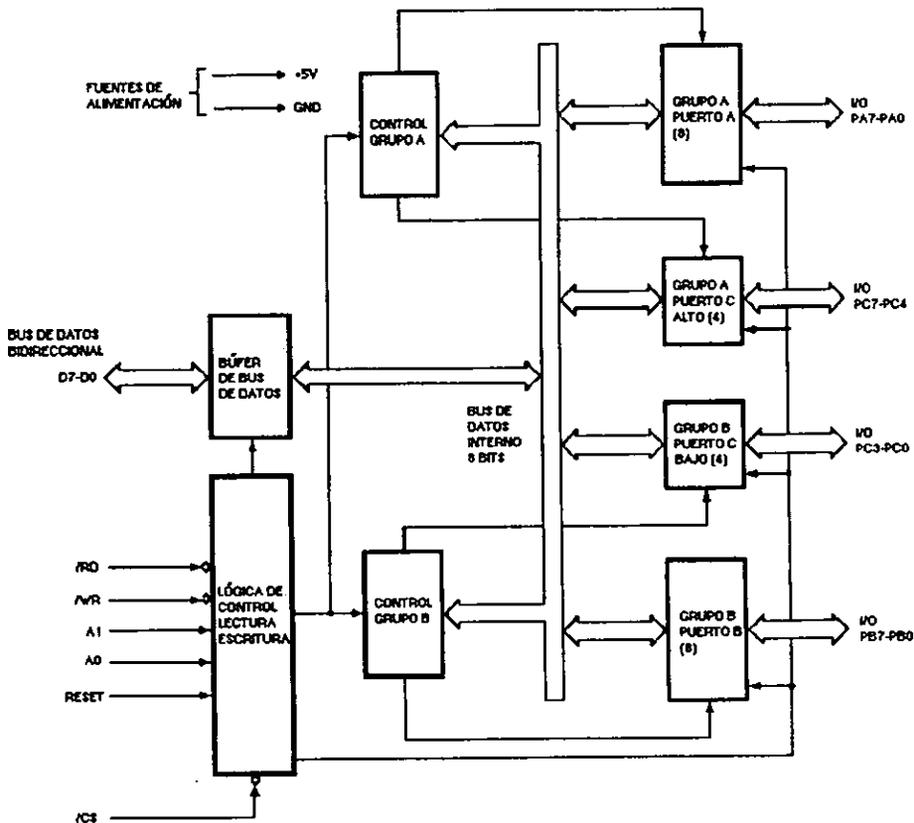


Fig. II-2.20. El 8255.

Como se mencionó anteriormente, el circuito tiene tres modos de operación.

- El modo 0 - básico entrada-salida.
- El modo 1 - entrada-salida con estrobo.
- El modo 2 - bi-direccional.

Después de un reset, las 24 líneas permanecerán en estado de entrada. En el modo 0, el manejo de los puertos se lleva a cabo sólo escribiendo a, ó leyendo desde, el puerto especificado. Se cuentan con las siguientes características funcionales.

- Dos puertos de 8 bits y dos de 4 bits.
- Cualquier puerto puede ser de entrada o salida.
- Las salidas son latcheadas.
- Las entradas no son latcheadas.
- Se cuentan con 16 distintas configuraciones de entrada/salida.

En el modo 1 requiere del uso de señales estroboscópicas junto con el manejo de los puertos. Los puertos A y B usan al puerto C para el manejo de las señales de protocolo. En este modo se tienen las siguientes características:

- Dos grupos el A y el B.
- Cada grupo contiene un puerto de datos de 8 bits y un puerto de 4 bits para control.
- El puerto de datos puede funcionar como entrada y salida, y en ambas situaciones es latcheado.
- El puerto de 4 bits es empleado para control y estatus del puerto de 8 bits.

En el modo 2, se tiene la comunicación con el periférico sobre un bus de datos bidireccional de 8 bits. En este modo, al igual que en el modo 1, se emplean líneas adicionales para la correcta comunicación. Las características de este modo se muestran a continuación:

- Se usa solamente el grupo A.
- El puerto A es el puerto bi-direccional y un puerto de control de 5 bits, que pertenece al puerto C.
- Tanto las entradas como las salidas son latcheadas.
- El puerto C será empleado para control y estatus del puerto de datos.

Se pueden tener varias combinaciones de modos de operación.

2.4.2.1 Manejo de puertos

Volviendo al asunto del puerto, se sabe que los puertos de una PC corresponden al siguiente mapa de memoria, figura II-2.21.

0200H		
0201H	NO USADO	1
0202H	JOYSTICK	1
0278H	NO USADO	118
0280H	LPT2	8
02F8H	NO USADO	120
0300H	SERIE 2	8
0300H	NO USADO	120
0378H	LPT1	8
0380H	NO USADO	48
03B0H	MONITOR MONOCROMO	16
03C0H	NO USADO	16
03D0H	MONITOR COLOR	16
03E0H	NO USADO	16
03F0H	FLOPPY	8
03F8H	SERIE 1	8
0400H		

Fig. II-2.21. Mapa de puertos.

Como se puede observar, para direccionar dichos puertos sólo es necesario emplear las primeras 10 líneas menos significativas del bus de direcciones, es por eso que el comparador de 8 bits, el 74HC688, fue colocado de tal manera que comparara los 8 bits más significativos de la señal presente en el bus de direcciones con 8 bits de referencia, que corresponden a los 8 bits más significativos de nuestra dirección base de la tarjeta (figura II-2.22).

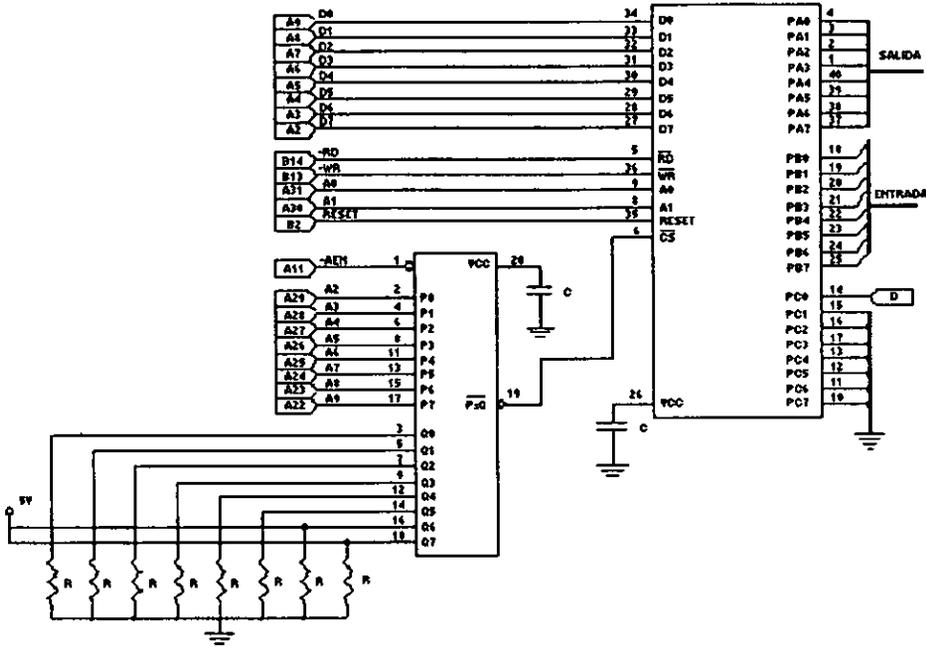


Fig. II-2.22. Implementación de la interface paralela.

Pero como se puede ver, los dos bits menos significativos aun quedan volando. Esto se aprovechó sabiendo que el 8255 tiene cuatro puertos, uno de control o configuración para los otros tres restantes. Como para acceder a cada puerto se requiere de una combinación binaria de dos bits, se decidió conectar los dos bits menos significativos a estos dos bits de direccionamiento del 8255, con esto se logró una interface paralela con cuatro puertos. Los cuales son direccionados a partir de una dirección base dada por la combinación de referencia presente en el comparador de 8 bits. El 8255 cuenta además con la señal de READ y la de WRITE, mediante las cuales el bus indica si se está realizando una operación de lectura o escritura de un dispositivo externo hacia la memoria de la computadora.

Para el empleo de este circuito en la tarjeta, se usó el modo básico o cero, eligiendo el puerto A como de salida, por el que se mandan los comandos de operación de la tarjeta; el puerto B se tomó como puerto de entrada, es decir, como el puerto de datos de la tarjeta; mientras que el puerto C se eligió completo como de entrada con el fin de sensar el término de la captura.

Las direcciones fijadas para la operación de la tarjeta fueron:

Puerto A = 300 hexadecimal.

Puerto B = 301 hexadecimal.

Puerto C = 302 hexadecimal.

Puerto Cont = 303 hexadecimal.

Este último puerto es el puerto de configuración interno al 8255, donde colocamos el modo de operación de los otros tres puertos, es el acceso a la palabra de control donde se reciben los comandos de operación. Para obtener la configuración buscada aquí se manda un 139 decimal. Esto porque la configuración para el modo de operación se fija en la palabra de control de acuerdo a la siguiente forma, figura II-2.23.

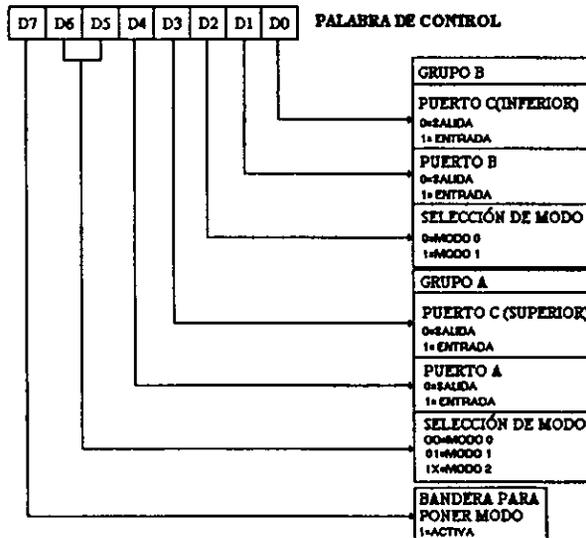


Fig. II-2.23 Formato para el modo de operación.

Y para obtener la configuración de los puertos deseada nos basamos en la tabla II-2.1.

MODO	A		B		GRUPO A		GRUPO B	
	D4	D3	D1	D0	Puerto A	Puerto C (alto)	Puerto B	Puerto C (bajo)
0	0	0	0	0	Salida	Salida	Salida	Salida
1	0	0	0	1	Salida	Salida	Salida	Entrada
2	0	0	1	0	Salida	Salida	Entrada	Salida
3	0	0	1	1	Salida	Salida	Entrada	Entrada
4	0	1	0	0	Salida	Entrada	Salida	Salida
5	0	1	0	1	Salida	Entrada	Salida	Entrada
6	0	1	1	0	Salida	Entrada	Entrada	Salida
7	0	1	1	1	Salida	Entrada	Entrada	Entrada
8	1	0	0	0	Entrada	Salida	Salida	Salida
9	1	0	0	1	Entrada	Salida	Salida	Entrada
10	1	0	1	0	Entrada	Salida	Entrada	Salida
11	1	0	1	1	Entrada	Salida	Entrada	Entrada
12	1	1	0	0	Entrada	Entrada	Salida	Salida
13	1	1	0	1	Entrada	Entrada	Salida	Entrada
14	1	1	1	0	Entrada	Entrada	Entrada	Salida
15	1	1	1	1	Entrada	Entrada	Entrada	Entrada

Tabla II-2.1 Configuraciones entrada-salida.

La palabra de control tienen los siguientes valores:

D ₇	D ₆	D ₅	D ₄	D ₃	D ₂	D ₁	D ₀	
1	0	0	0	1	0	1	1	= 139 DECIMAL

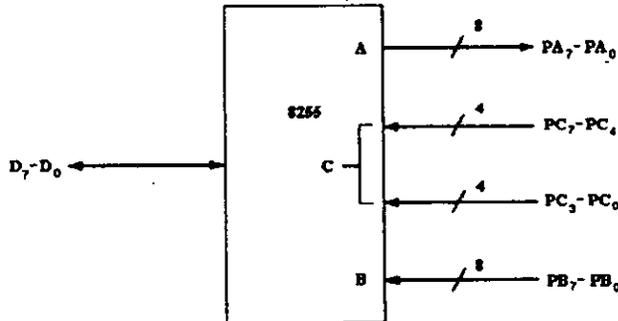


Fig. II-2.24 Modo de entrada salida elegido.

Para el caso de la adquisición de la imagen hacia la PC, se aprovechó el puerto B del 8255 para leer los datos presentes a la salida de la memoria. Para ir direccionado cada localidad de memoria de manera secuencial, con el fin de que estuviera disponible en la siguiente operación de lectura, se empleó un simple arreglo con una compuerta NAND que controla el paso del reloj hacia los contadores de direccionamiento, ya sea el de 6 MHz para captura o el tren de pulsos para la adquisición.

Mediante el arreglo anterior, se logró tener una velocidad de lectura de aproximadamente 300 kHz de señal del tren de pulsos en una computadora 486 SX a 25 MHz. Con esta velocidad el proceso completo de captura, adquisición y visualización fue de aproximadamente 3 cuadros por segundo. Lo cual no fue muy satisfactorio porque los movimientos en dichas escenas no tenía continuidad. Por esto se buscó otro camino para leer los datos de una manera más rápida, entonces se investigó y se encontró que mediante el acceso directo de memoria se podía tener mayor velocidad (en relación a esta computadora).

2.4.3 El DMA

El controlador DMA (Direct Memory Access) hace un acceso directo a memoria, lo cual permite un incremento en la velocidad de transferencia de datos, esto es gracias a que este circuito toma el control del bus de la computadora inhabilitando temporalmente al CPU, haciendo una transferencia a 8 bits de hasta 64 kbytes de tamaño. Desde que aparecieron las computadoras XT que trabajaban a 4.77 MHz, el DMA parecía ser el dispositivo de entrada/salida más veloz para transmitir y/o recibir información de los periféricos. Con los 4 canales con los que cuenta cada controlador DMA se tiene la capacidad de atender cuatro periféricos a la vez, por supuesto que la atención depende del orden de petición, es decir, los canales del controlador tienen prioridades, así el canal 0 tiene la mayor prioridad y el canal 3 la menor. En esta época el DMA tomó gran importancia pues trabajaba a 6 MHz, superando por mucho las transferencias comunes.

El desarrollo de las computadoras no fue de la mano del DMA, es decir, mientras que ahora podemos encontrar computadoras que trabajan por encima de los 200 MHz, el DMA desde las PC/AT disminuyó su velocidad a 3 MHz, fue este el momento en que el DMA dejó de ser tan popular. En las computadoras AT se incrementó a 2 el número de controladores teniéndose 8 canales, (siete en realidad) el canal 3 del primer controlador se conecta al segundo para tener acceso a los otros cuatro canales y transferencias a 16 bits de hasta 128 kbytes. El controlador DMA más popular es el 8237, es en base a éste que desarrollamos nuestro trabajo y del cual hablamos en este escrito.

2.4.3.1 El 8237

Este chip consta de 40 pines con las funciones siguientes:

Pin 1 (-IOR). I/O Read, es una señal activa baja, bidireccional, línea tres-estados, usada durante el ciclo inactivo cuando el micro está esperando por una petición DMA. Es usada para leer el registro de control. En el ciclo DMA, o ciclo activo, esta señal permite al 8237 el acceso de datos de dispositivos externos durante una transferencia de escritura DMA.

Pin 2 (-IOW). I/O Write, es activa baja, tres-estados. En el ciclo inactivo, ésta es usada por el micro para cargar información en el 8237. En ciclo activo, ésta es una señal de control de salida usada por el 8237 para cargar datos a dispositivos externos durante una transferencia DMA.

- Pin 3 (-MEMR).** Es una salida activa baja tres-estados, usada durante una operación de lectura DMA para acceder datos desde una locación de memoria seleccionada.
- Pin 4 (-MEMW).** Es una salida activa baja tres-estados, usada para escribir datos a una locación seleccionada durante una operación de escritura DMA.
- Pin 5 (Vcc).** Alimentación del circuito.
- Pin 6 (READY).** Es una señal activa alta, usada para acomodar las líneas de lectura y escritura con memorias bajas o periféricos de I/O.
- Pin 7 (HLDA).** Señal activa alta del micro, ésta indica que el micro ha dado el control de sus buses.
- Pin 8 (ADSTB).** Se utiliza esta señal para cargar el byte más significativo de una dirección de transferencia en el registro externo latch del DMA. El controlador activa esta línea tan pronto como deposita el byte de dirección en las terminales DB_0 - DB_7 .
- Pin 9 (AEN).** Señal habilitada, es activada como salida alta. Es usada para habilitar un latch de dirección que está conectado a las líneas de datos D_0 - D_7 . Recordando que estas líneas son multiplexadas para proveer las direcciones más significativas A_8 - A_{15} .
- Pin 10 (HRQ).** Pregunta al micro por el control del bus del sistema.
- Pin 11 (-CS).** La señal de chip select es una señal activa baja usada para seleccionar el 8237 como un dispositivo I/O. Ésta permite al microprocesador comunicarse sobre el bus de datos.
- Pin 12 (CLK).** Reloj de entrada que controla los intervalos de operación del 8237 y puede operar a frecuencias arriba de 10 MHz.
- Pin 13 (RESET).** Es una señal activa baja que limpia los comandos, estado, petición, y registros temporales.
- Pin 14,15,24,25 ($DACK_0$ - $DACK_3$).** Son usadas para notificar al dispositivo externo que el DMA ha sido otorgado a uno de los dispositivos.
- Pin 16-19 ($DREQ_0$ - $DREQ_3$).** Las líneas de petición de DMA son canales de entrada asincronos usados por los dispositivos externos para obtener operación DMA. Estas señales son programables para ser activas altas o activas bajas. El RESET inicializa las señales $DREQ_0$ - $DREQ_3$ como activas altas.
- Pin 20 (GND).**
- Pin 21-23, 26-30 (DB_0 - DB_7).** Las líneas del bus de datos son tres-estados, señales bidireccionales conectadas al sistema de bus de datos. Durante la operación DMA estas líneas son multiplexadas a salidas, las cuales serán las direcciones más significativas A_8 - A_{15} .
- Pin 31 (Vcc).**
- Pin 32-35 (A_0 - A_3).** Las cuatro direcciones menos significativas, bidireccionales, tres-estados. Durante el ciclo inactivo son usadas como entradas. Durante el DMA o ciclo activo son usadas como salidas.
- Pin 36 (-EOP).** Fin de proceso es una señal activa baja, bidireccional. Pone un bajo sobre éste cuando termina un servicio DMA. Cuando se usa como una salida, sirve para interrumpir al procesador señalando el final de una operación DMA.
- Pin 37-40 (A_4 - A_7).** Las cuatro líneas de direcciones de orden alto son salida tres-estados y proveen los cuatro bits de información. Ellas son habilitadas solamente durante el servicio DMA.

2.4.3.2 Funcionamiento del ciclo DMA

El controlador DMA INTEL 8237 es un dispositivo programable de 40 pines que es compatible con el microprocesador 8086/8088. El 8237 tiene cuatro canales DMA independientes (figura II-2.25). La transferencia de datos comienza con el requerimiento de DMA por las líneas de petición DREQ0-DREQ3. De la mano de las DREQ's están las líneas DACK0-DACK3. Estas líneas son usadas para confirmar una petición DMA. El 8237 usa la señal de bloqueo sobre el 8086/8088 para tomar el bus del sistema. Después es inicializado por el microprocesador, el 8237 toma control del bus para hacer una operación DMA.

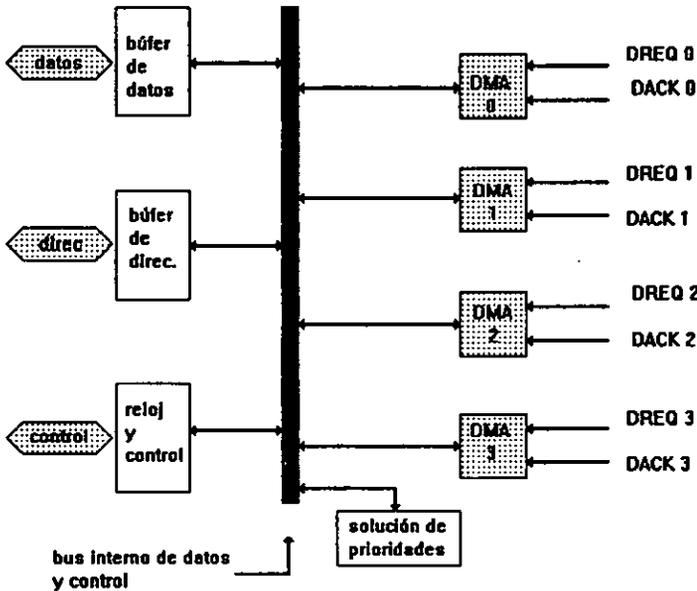


Fig. II-2.25. Diagrama de bloques del controlador DMA.

El controlador DMA está diseñado para operar en dos ciclos, el ciclo activo y el ciclo inactivo. Cuando el 8237 está en ciclo inactivo, ningún dispositivo externo está solicitando una transferencia DMA. En este ciclo el 8237 patea las líneas DREQ cada que termina un ciclo de reloj para determinar si alguno de los cuatro canales requiere el servicio. Cuando una de las líneas DREQ0-DREQ3 requiere DMA, el 8237 entra al ciclo activo. Cuando el 8237 está en el ciclo inactivo y un canal requiere transferencia DMA, las salidas HRQ (hold request-petición de bloque) del 8237 informan al microprocesador la entrada al ciclo activo, como muestra la figura II-2.26. Es en este ciclo que la transferencia toma lugar. Cuando está listo, el microprocesador responde a el 8237 con un HLDA (hold acknowledge- confirmar petición), indicando que éste tiene libre el control del bus. El 8237 entonces responde a el dispositivo externo con un DACK (DMA acknowledge), indicando el inicio del ciclo activo. La figura II-2.27 ilustra el diagrama de bloques básico de la

interface DMA. Cuando el controlador DMA completa la transferencia de datos, genera una señal EOP (end of process- fin de proceso), la cual indica que se ha completado la transferencia, regresando el control del bus al microprocesador.

Puesto que nuestra tarjeta adquiere muestras de 8 bits y nuestras memorias son de 8 bits, resultó conveniente emplear uno de los canales disponibles de 8 bits, el canal elegido fue el canal 3. Para poder efectuar el direccionamiento, es necesario realizar una programación previa del controlador para hacer la transferencia deseada. Además, fue necesario incorporar nuevos elementos a la tarjeta, como un latch de 8 bits para no conectar directamente el bus de datos a la memoria de la tarjeta, un flip-flop es el encargado de hacer el protocolo necesario para iniciar un ciclo DMA. Fue necesario incorporar un nuevo arreglo de compuertas para manejar la señal de READ, también presente en el ciclo DMA pero ahora no es manejada por el CPU sino por el controlador DMA, de manera de que se pudiera emplear para el avance de los contadores de direccionamiento.

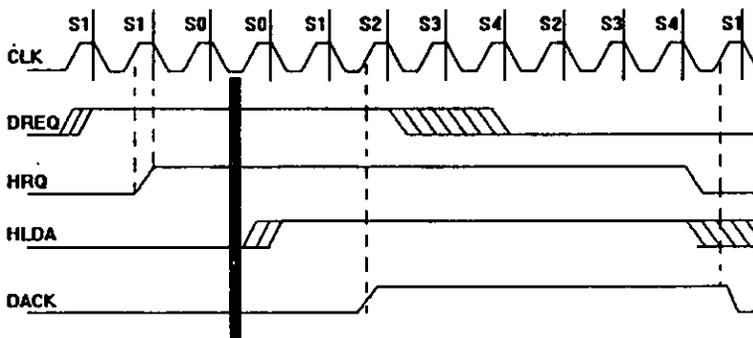


Fig. II-2.26. Diagrama de tiempos del DMA.

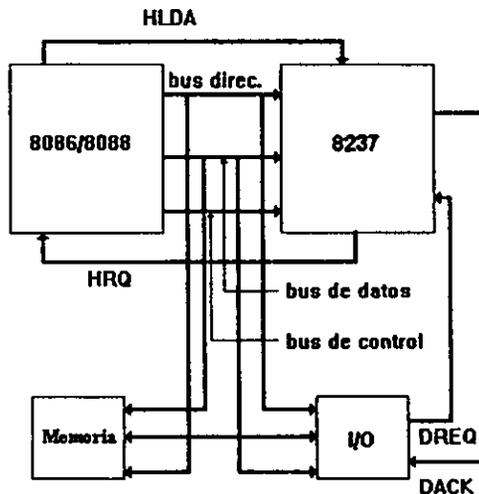


Fig. II-2.27 Diagrama de bloques básico de la interface DMA.

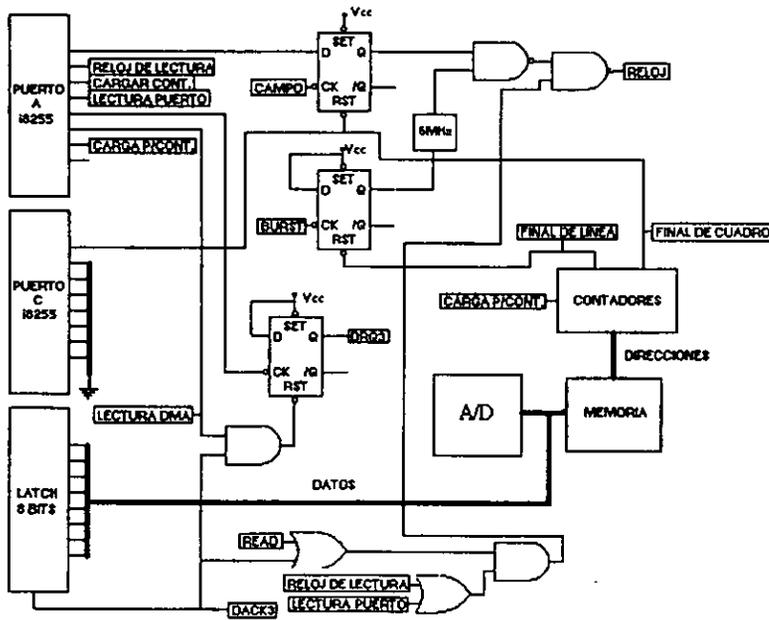


Fig. II-2.28. Sección DMA de la tarjeta.

Como se puede ver en la figura, el flip-flop que controla el inicio de la transferencia tiene su salida Q conectada directamente a la entrada DREQ3 del bus, y es cambiado de estado por una línea (la línea PA4) del puerto A del 8255. Además, para evitar conflictos con las peticiones es necesario limpiar la salida Q del flip-flop después de iniciar el ciclo, por lo que se aprovecha la salida DACK3 del bus para conectarla al CLEAR o RESET del flip-flop. Como se puede ver en el circuito, se tiene una compuerta AND controlada por la línea del puerto A (PA5), por lo que se mantiene el flip-flop en estado bajo en Q por la condición permanente de RESET en el flip-flop, evitando que se pudiera disparar la petición DMA innecesariamente.

Con esta nueva señal de READ se obtuvo una frecuencia de lectura de 1 MHz. Lo cual era significativamente más rápido con respecto al direccionamiento de puertos realizado en la misma computadora 486SX a 25 MHz.

El inconveniente del empleo del DMA, como ya se señaló antes, es de que éste ha permanecido prácticamente sin cambios desde que se empleó por primera vez en las computadoras XT. Lo cual hace suponer que esta velocidad de lectura permanecerá casi constante en cualquier computadora, mientras que en el direccionamiento de puertos depende del CPU empleado en la computadora. Se puede ver que en las nuevas computadoras Pentium a 200 MHz el DMA se convierte en un adorno más, comparado con la velocidad que se podrían alcanzar con el direccionamiento a puertos en estas computadoras.

Por otro lado, se tiene la ventaja que con el DMA el CPU no interfiere en modo alguno en la operación de manejo de memoria mientras que se captura el cuadro. Cosa que se presenta cuando se efectúa la lectura con direccionamiento a puertos.

2.4.3.3 Programación del DMA

Para poder controlar el DMA (8237) es necesario realizar los siguientes pasos vía software:

- Elegir el canal de transferencia.
- Modo de transferencia.
- Limpiar flip-flop interno.
- Fijar la dirección del bloque de datos.
- Establecer el tamaño de la transferencia.
- Liberar el canal empleado.

2.4.3.3.1 Canal de transferencia

Existen dos formas de establecer el canal a usar, esto es vía (des)enmascaramiento. La primer forma de (des)enmascarar un canal es mediante el Single-Mask-Register, con este procedimiento se puede hacer alusión a sólo un canal de los 8 posibles. La descripción de los canales se muestra en la tabla II-2.2.

El segundo método de (des)enmascaramiento es a través del All-Mask-Register, con este método se pueden modificar todos los canales al mismo tiempo, pero hay que tener cuidado si no se sabe manejar pues podría desconfigurarse el sistema, es por ello que se recomienda el uso del Single-Mask.

Canal	Controlador	Descripción
Canal 0	Controlador 1	RAM-refresh
Canal 1	Controlador 1	libre, habitualmente SB, GUS etc.
Canal 2	Controlador 1	Unidad de disquetes
Canal 3	Controlador 1	Unidad de disco duro
Canal 4	Controlador 2	Cascada Contr. 1 => Contr. 2
Canal 5	Controlador 2	libre
Canal 6	Controlador 2	libre
Canal 7	Controlador 2	libre

Tabla II-2.2. Canales DMA.

2.4.3.3.2 Modo de transferencia

El DMA maneja cuatro modos de transferencia, el que usará debe especificarse antes de cada transferencia.

- **Transferencia única.** En este modo el 8237 es programado para transferir un sólo byte cada vez que la petición es activada.

- **Transferencia por bloque.** En este modo el 8237, una vez activado por el DREQ, continua haciendo transferencias hasta que un bloque de datos es transferido. El 8237 es programado con la dirección de inicio y el número de bytes del bloque a transferir. La transferencia continua hasta que se presenta el final del conteo o un EOP es recibido.
- **Transferencia por demanda.** En este modo el 8237 es programado para continuar haciendo transferencias hasta que un EOP es recibido, o hasta se termina un DREQ desde un dispositivo externo. En este modo un byte es transferido cada vez que se presenta DREQ.
- **Transferencia en cascada.** Este modo es usado para montar un arreglo de 8237's para un sistema de expansión.

La información fluye de la PC al periférico, en cuyo caso se dice que hay una operación de lectura y de escritura, si la información va del periférico a la memoria principal, existe un tercero que se conoce como *Verify*, que es sólo un *test* de autocomprobación del controlador DMA.

El flujo de la información es hacia adelante (incremento) o hacia atrás (decremento) en la memoria, la transferencia termina cuando se genera la señal en el TC (Terminal Count-Fin de Conteo) o cuando el hardware externo lo detiene mediante una señal en el EOP. Puede producirse un autoinicialización de la configuración si se programa con anterioridad, así se puede estar haciendo una serie de transferencia en forma continua.

2.4.3.3.3 El flip-flop DMA

Antes de algunas aplicaciones de programación del controlador DMA, es necesario borrar el flip-flop. Con ello se reinicializa el flip-flop interno que muestra el estado de una transmisión de 16 bits. Es el encargado de controlar si se está accediendo a los 8 bits altos del registro de 16 bits o a los 8 bits bajos. Esto se hace mediante el registro Clear-flip-flop.

2.4.3.3.4 Dirección del bloque de datos

El controlador DMA necesita la página física del bloque de datos a transferir. Ésta se calcula según la fórmula, de acuerdo al direccionamiento en modo real:

$$\text{direcc} = 16 * \text{longint} (\text{Seg} (\text{Bloque}^{\wedge})) + \text{ofs} (\text{Bloque}^{\wedge}) \quad (\text{II-2.17})$$

Para escribir esta dirección existen los registros Lower-Page (8 bits bajos) y Upper-Page (8 bits altos), debe escribirse primero el registro Lower-page pues de lo contrario borra el contenido de Upper-Page.

2.4.3.3.5 Tamaño de la transferencia.

Antes de poder escribir el tamaño del bloque a transferir en el controlador DMA, primero hay que borrar el flip-flop. Después se transfiere el byte bajo y a continuación el byte alto al registro DMA-Count. Con ello se ha establecido el número de datos a transferir.

2.4.3.3.6 Liberación del canal

Finalmente el uso de una transferencia DMA termina liberando el canal que se había elegido al inicio de la configuración simplemente quitando la máscara que se había puesto previamente.

2.5 Funcionamiento general de la tarjeta

Una vez detallados todos los elementos empleados en el hardware de la tarjeta en forma individual, se procederá a hacer un bosquejo de cómo se hace la captura de los cuadros de vídeo.

Como ya se señaló antes, la señal de vídeo se alimenta a la entrada positiva del amplificador diferencial, mientras que su referencia se conecta a la entrada negativa, a la salida de este amplificador se tiene una señal de vídeo aproximadamente amplificada por 1.2. Esta señal es conectada al separador de sincronía, el cual será el encargado de controlar el reloj de 6 Mhz en la etapa de captura.

La señal de vídeo es también alimentada al circuito LF398, que será el encargado de obtener el nivel de negro de nuestra señal de vídeo, controlado por la señal invertida de paso de burst del separador de sincronía.

Finalmente, la señal de vídeo será conectada al amplificador diferencial en la terminal positiva, aquí será restado a dicha señal de vídeo el nivel de negro obtenido mediante el arreglo del LF398 y el filtro paso bajas, teniendo así a la salida de este amplificador diferencial la señal de vídeo con nivel de negro coincidente a tierra. Esta señal es rectificadora para sólo tener a la salida del rectificador de precisión la señal de vídeo visible, que es la señal de tierra a positivo.

Mientras esto sucede, se obtiene por otro lado el funcionamiento del reloj deseado para que trabajen correctamente el A/D y la memoria.

El reloj es controlado por dos flip-flop's, cada uno es controlado a su vez por la señal de paso de burst y la señal de sincronía vertical. En el caso del flip-flop de línea, es controlado por la señal de paso de burst conectada directamente del LM1881 al reloj del flip-flop D. Este flip-flop va conectado a la compuerta NAND que controla la activación o desactivación del reloj y es reseteado por un circuito diferenciador hecho con un arreglo de capacitor, resistencia y diodo.

El diferenciador nos dará un pulso negativo de 5 V a tierra que nos reseteará el flip-flop cuando se han tomado las 256 muestras correspondientes a la línea de vídeo. Las 256 muestras por línea son registradas por los dos contadores de direccionamiento menos significativos. La terminal de dato D está conectada directamente a 5 V con el objeto de tener disponible siempre el uno lógico a la entrada.

El otro flip-flop controla el reloj mediante una compuerta NAND para que pase sólo durante el campo de vídeo deseado y también es el encargado de activar la captura. Este flip-flop tiene en su dato D la señal de cero cuando no se quiere adquirir cuadro alguno y uno cuando se quiere adquirir el campo. Esto en coordinación con la señal de reloj pondrá en marcha la captura.

El reloj de cambio del flip-flop es obtenido a partir de la señal de sincronía vertical dividida entre dos mediante otro flip-flop para obtener sólo la sincronía o los flancos de

subida correspondientes a un sólo campo. Este flip-flop presentará un uno lógico a la salida cuando se tenga un uno en la terminal D y se tenga a la vez el flanco de subida de la señal de sincronía vertical del campo en cuestión.

A partir de este momento, el reloj en la compuerta pasará hacia los contadores de direccionamiento y hacia la terminal CLK del convertidor A/D, que es la encargada de controlar la conversión y a la terminal /WR de la memoria. El flip-flop se reseteará en el momento en que se hayan tomado 256 líneas con 256 muestras cada una, es decir, cuando los contadores de 16 bits se hayan vaciado completamente.

La salida Q de este flip-flop de campo también controla la activación del bus bidireccional de datos de la memoria, deshabilitando las salidas cuando se captura y activándolas con un cero lógico cuando se necesite leerlas, además es la encargada de habilitar el convertidor A/D.

En coordinación con el paso del reloj hacia los contadores de direccionamiento se tendrá la escritura simultánea a la memoria RAM estática de las conversiones hechas por el A/D en el momento en que la información visible se encuentra en la señal de vídeo.

En la señal que informa la finalización de un conteo completo de los cuatro contadores de direccionamiento no se cuenta con algún circuito diferenciador, sino que es alimentada directamente al flip-flop. El reloj permanecerá desactivado después de finalizar la captura de un campo, por lo tanto la tarjeta quedará "trabada". Esta señal será sensada continuamente mediante software con ayuda de el puerto C del 8255, usado como entrada para saber cuándo se ha finalizado la etapa de captura y se puede iniciar la etapa de adquisición.

La etapa de captura es controlada por el 8255 mediante la señal que pone en marcha dicho procedimiento, es decir, el dato D en el flip-flop de campo, así como también resetea y carga los contadores para que inicien la cuenta desde el valor deseado. También es el encargado de acondicionar el sistema para adquirir en forma de direccionamiento a puerto o en forma DMA.

La etapa de adquisición es puesta en marcha vía software, reseteando los contadores y cargándolos con el número binario 1110110100000000, de modo que la lectura de líneas se inicie exactamente 19 líneas después de la línea uno en que se inició la captura, esto porque hay que recordar que cada campo tiene 21 líneas que no tienen información visible correspondiente al blanking vertical, por lo que es conveniente saltar estas líneas.

Una vez reseteados los contadores se procede a realizar la lectura del búfer de memoria, si es mediante direccionamiento de puertos, se manda un cero lógico al bit PA3 del puerto A, que es el de control de la tarjeta, el cual está conectado a la compuerta OR que permite el paso del tren de pulsos de lectura hacia el arreglo de compuertas que distribuyen el reloj a los contadores, el convertidor y las memorias.

Ya que se ha hecho esto, la lectura del puerto B del 8255, de entrada de datos, puede ser efectuada una vez para cada byte de la memoria. Conforme se lee un byte, se manda un pulso con el cual durante su flanco de subida cambiará al siguiente estado los contadores para direccionar la siguiente localidad de memoria. Este proceso se lleva a cabo 49152 veces correspondientes al número de pixeles por trama capturada de 256 x 192.

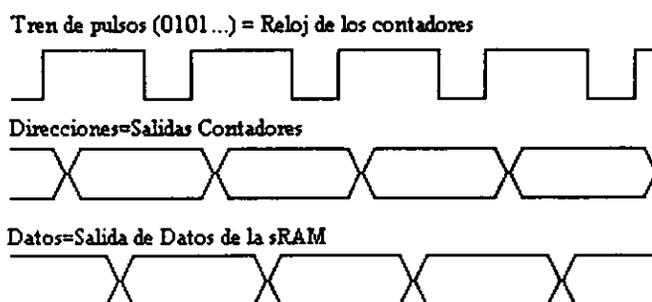
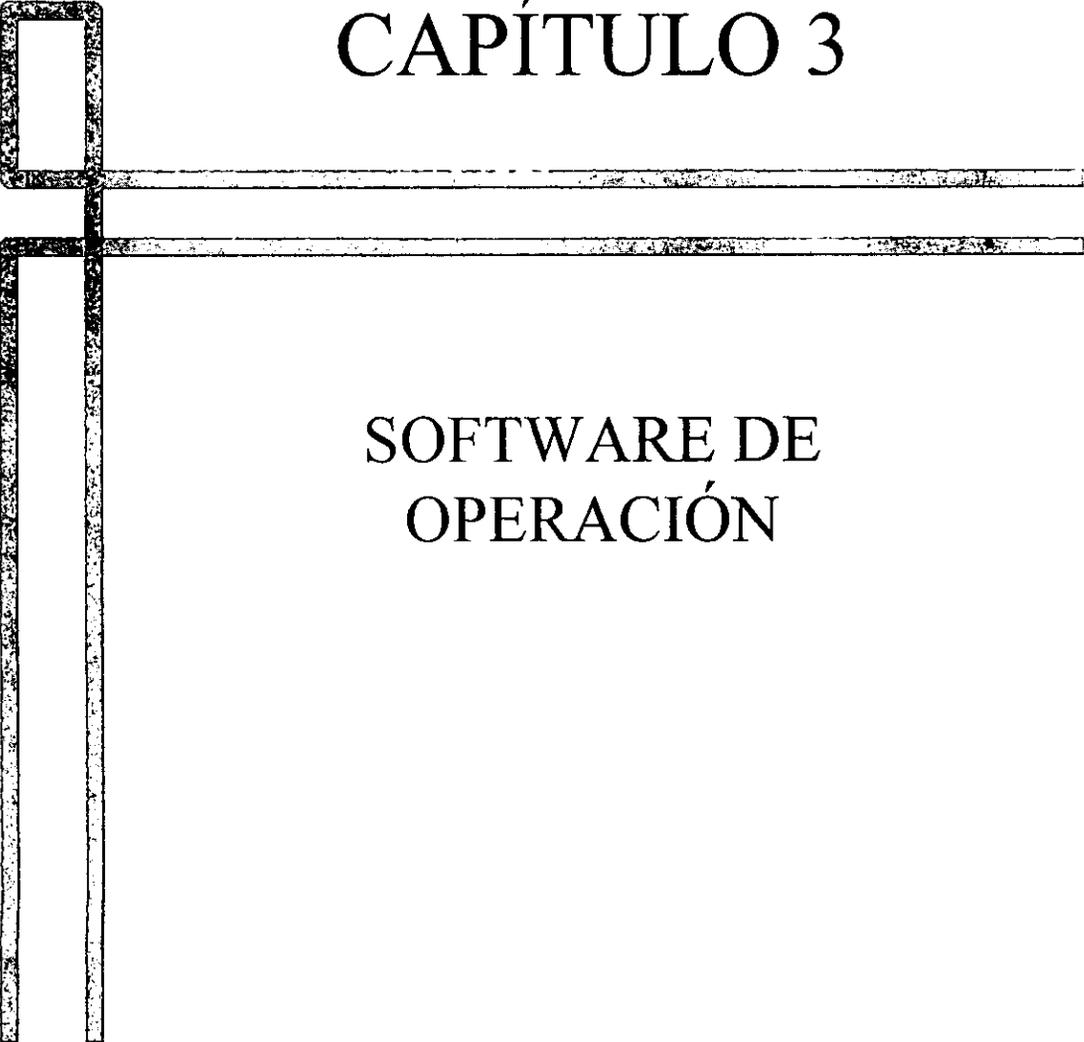


Fig. 11-2.29 Diagrama de tiempos de adquisición mediante lectura a puertos.

Si por otro lado, la lectura se lleva por DMA, una vez puestos los contadores como se desea, se cambia el estado del flip-flop DMA de modo de poner un uno lógico en su salida, lo cual significa que la tarjeta pide un ciclo DMA. Cabe destacar que el ciclo DMA ha sido programado al inicio de la secuencia de programación para que el controlador realice una lectura DMA de 49152 bytes en modo de transferencia de bloque en el canal 3 sin reinicialización (esto porque el EMM386 no soporta la autoinicialización). Al momento de recibir la petición DMA por parte del flip-flop en la señal DREQ3, el controlador pide el permiso al CPU y cuando éste lo da, el controlador pone en bajo la señal DACK3, limpiando el flip-flop DMA y dejando pasar la señal de READ hacia el sistema de direccionamiento. Existe el pin PA5 del 8255, que al igual que el pin PA3 del direccionamiento a puertos controla el paso del tren del pulsos a los contadores, en este caso el tren de pulsos es el READ al efectuar procesos continuos de lectura; el paso se logra poniendo un 1 en el pin PA5. Además, esta señal DACK3 activa el latch de 8 bits 74LS373 para que pueda trabajar con el bus de datos del bus ISA.



CAPÍTULO 3

SOFTWARE DE
OPERACIÓN

3.1 Generalidades

En este capítulo se describe la operación del programa que maneja la tarjeta de captura. Dicho programa controla la secuencia descrita en el capítulo anterior para la operación de la tarjeta. Es un programa relativamente sencillo, hay que recordar que el objetivo de éste es llevar la secuencia deseada con el menor tiempo posible, ya que la continuidad del movimiento de los cuadros depende de que tan rápido se efectúe todo el ciclo de captura, adquisición y visualización. Evidentemente, el tiempo de trabajo va a ser la suma de los tiempos de cada sección, es por eso que nosotros buscamos las mejores opciones para reducir el tiempo.

El programa fue realizado tanto para la versión de sistema DOS como para el ambiente WINDOWS. Esto porque en la primera versión se obtiene mejor velocidad de operación, pero en la segunda se cuenta con la estética que da la interface de WINDOWS para la presentación.

La versión de DOS fue programada en BORLAND C++, mientras que la de WINDOWS se realizó en VISUAL C++ de la empresa MICROSOFT. Ambos paquetes tienen como base el lenguaje de programación C++, sólo que en el caso de la versión WINDOWS se incorporaron los elementos indispensables para la programación en este ambiente (la clase, la creación y la visualización de la ventana, así como los mensajes a los que responderá el programa), además de que el manejo de gráficos es diferente en ambos casos, lo mismo que el direccionamiento de memoria.

En la versión de WINDOWS se hizo uso de las funciones SDK (Software Development Kit), conocidas más formalmente como API (Interfaz de Programa de Aplicación), las cuales permiten un funcionamiento que entiende el WINDOWS de antemano.

Una de las diferencias de WINDOWS con respecto a DOS, es de que en DOS el programa inicia la interacción con el sistema y trabaja con una sola aplicación a la vez, mientras que en WINDOWS, el sistema inicia la interacción y puede atender varias tareas a la vez. El proceso funciona así, un programa de WINDOWS espera hasta que WINDOWS le envía el mensaje. El mensaje es dirigido al programa por medio de una función especial manejada por WINDOWS. Ya que se recibió el mensaje, el programa puede ejecutar una función determinada. De manera general la interacción de los programas en WINDOWS está basada en mensajes con WINDOWS.

WINDOWS es un sistema multitareas y ejecuta sus tareas sin prioridades, es decir, un programa puede conservar el uso del procesador hasta que lo libera.

Para el caso de las funciones API, éstas son funciones que nos permiten entrar al entorno de WINDOWS. En estas funciones API se cuenta con unas funciones denominadas GDI (Graphics Device Interface), las cuales proporcionan el soporte para los gráficos de forma independiente del dispositivo, es por eso que éstas logran que un programa en WINDOWS pueda funcionar en una gran variedad de hardware.

En lo que respecta al esqueleto del programa, éste es prácticamente el mismo, por lo que se hará el análisis de dicho esqueleto y se señalarán las particularidades en las dos versiones. El diagrama de flujo del programa se muestra a continuación (figura II-3.1).

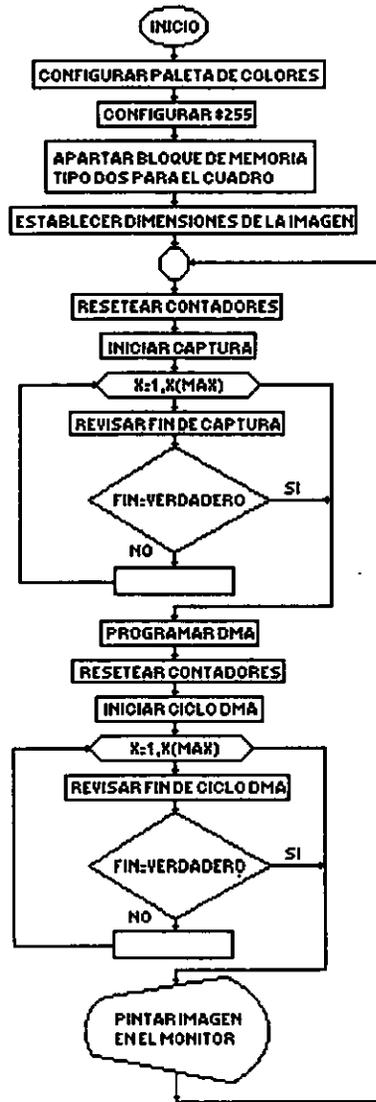


Fig. II-3.1. Diagrama de flujo de operación de la tarjeta

Nos basaremos en el diagrama de flujo anterior para hacer la descripción. Según se puede ver, primeramente se establece el manejo de colores que se necesita para que cada byte del cuadro de vídeo represente un tono específico de gris.

Para el caso de la versión de DOS, es necesario cambiar el modo de vídeo del monitor, para esto, se uso un controlador de vídeo del archivo BGI (Borland Graphic Interface) svga256.bgi, el cual nos pone el monitor en una resolución de 640 x 480 pixeles

con 256 colores. Estos archivos BGI deben de estar disponibles en el sistema y contienen los controladores gráficos.

El paso siguiente en esta sección es el establecer estos 256 colores como 256 tonos de grises, para esto hacemos nuestra propia paleta de colores. Para generar dicha paleta, sabemos que cada color está compuesto por una combinación de los colores Rojo, Verde y Azul (RGB), por lo que para generar una color en tono de gris sólo es necesario mezclar la misma porción de cada color. Como en este modo cada valor de Rojo, Verde o Azul es de 8 bits, nosotros tenemos 256 combinaciones de gris, suficientes para nuestra aplicación. Con la ayuda de la interrupción 10h nosotros hacemos válida nuestra paleta en el monitor. Esta interrupción de BIOS tiene a su cargo las funciones de vídeo, en este caso se empleó el servicio 12 hexadecimal el cual cambia toda la paleta de colores. Para el caso de la versión para Windows. El manejo de los colores es distinto, nosotros necesitamos crear nuestra paleta lógica para que sólo sea válida en la ventana donde veremos los cuadros de vídeo.

La paleta por default que se encuentra en uso actual en WINDOWS se conoce como paleta de sistema. Una paleta lógica es una que la aplicación crea, también puede especificar que colores quiere que tenga dicha paleta. Se emplean ciertas llamadas GDI (Graphics Device Interface), pudiendo así reemplazar temporalmente la mayoría de la paleta del sistema con la paleta lógica. WINDOWS guarda algunos índices de la paleta para si mismo.

Además, al crear nuestro cuadro de vídeo como un BITMAP independiente nosotros podemos tener 256 tonos que deseamos. Sino hacemos lo anterior, el Windows apartará 20 colores para uso, conocidos como de sistema, lo cual generará confeti en la imagen.

El paso siguiente es configurar el 8255 para que trabaje como deseamos. Como se vio en la parte de hardware, esto se hace con las operaciones de direccionamiento a puertos, debemos mandar al puerto de la palabra de control el número 139 decimal. Además de que se manda el número 8 al puerto A (el puerto de control de la tarjeta) con esto sólo tenemos un uno lógico en el bit 4 y cero en todos los demás, así tenemos un estado de reset de la tarjeta para iniciar la operación.

Además de estos arreglos iniciales, necesitamos apartar nuestros bloques de memoria que recibirán los cuadros digitalizados. Para este programa inicial sólo se apartó un bloque de 49152 bytes para un cuadro de vídeo de 256 x 192. Para el caso de la versión de DOS se empleó la función "malloc" de C, mientras que en WINDOWS se usó la función SDK "GlobalDosAlloc", ésta última reserva memoria global de la memoria de WINDOWS. Un detalle importante en WINDOWS es que éste trabaja en modo protegido y maneja un modo distinto de direccionamiento al del modo real, el cual maneja el direccionamiento segmento:offset, mientras que el protegido maneja selector:offset, esto para facilitar la paginación de la memoria. Es por eso que se emplea la anterior instrucción ya que, aunque trabaja bajo WINDOWS, nos reserva la memoria en forma continua, es decir, como si estuviera en modo real.

Ya teniendo nuestro búfer de memoria, podemos establecer la dimensiones del cuadro en el programa. El apartar la memoria antes de dimensionar el cuadro es un requisito en las dos versiones. Esto por que en el caso de DOS para la visualización se usaron las funciones "Getimage" y "Putimage", las cuales manejan los gráficos más rápidamente que si empleamos la operación de pixel por pixel, es decir, la función "Putpixel". En el caso de WINDOWS, para crear el BITMAP requerimos tener el búfer de datos.

3.2 Captura

Todos los pasos anteriores son de configuración inicial, por lo que se llevarán a cabo sólo una vez en el programa. Una vez cumplidas estas acciones, se entra al ciclo de captura, adquisición y visualización.

Primero, se manda un cambio en el bit que controla el reset de los contadores, con el objeto de asegurar que la captura se lleve a cabo correctamente. Esta acción sólo requiere de tres instrucciones consecutivas de direccionamiento a puertos para generar el pulso bajo para cargar los contadores. Es importante señalar un punto importante aquí, ya que la imagen capturada en el búfer de la tarjeta, es de 64kb, es decir, de 256 x 256, es necesario sólo hacer 49152 lecturas por el tamaño de la imagen que manejamos, pero como se señaló antes en la sección de hardware, se capturan las líneas de barrido donde se encuentran los pulsos de ecualización, por lo que esta zona no tendrá información visible, es necesario entonces, saltarse esta parte de la memoria, para sólo leer la información de la imagen. Con este fin, antes de resetear esta primera vez dentro del ciclo, nosotros cargamos todos los bits a alto, para que el contador apunte en la captura a la localidad de memoria más alta. Más adelante, después de la captura, se vuelven a resetear los contadores para la adquisición pero ahora se cargan los contadores de modo de que apunten a la sección donde se tiene sólo la información visible, saltándonos la parte no visible.

Con la acción anterior, se tiene todo listo para iniciar la captura, y se logra poniendo un uno lógico en el bit menos significativo del puerto A, el cual se alimenta al dato D del flip-flop que inicia la captura.

Como se explicó en la sección de hardware, la tarjeta mandará un cero en el puerto C cuando se haya terminado la captura, por lo que aquí se inicia un ciclo de lectura de este puerto, para esperar el momento en que la captura finalice y así seguir con la secuencia

3.3 Adquisición

Esta sección establece las diferencias en cuanto al funcionamiento del programa para DMA o sólo para direccionamiento de puertos. En el caso de sólo direccionar a puertos es necesario hacer un ciclo de lectura al puerto B (nuestro puerto de entrada de datos). Este ciclo se ejecuta 49152 veces y hay que recordar que se genera un pulso para avanzar el contador y tener el dato de la siguiente lectura. Este tren de pulsos es creado mandando 1's y 0's por uno de los pines del puerto de control.

En el caso del DMA, esta sección es más complicada. Ya que se emplea en ambas versiones el modo de transferencia en bloques sin reinicialización, nosotros necesitamos reprogramar el 8237 cada vez que se hace una transferencia, mediante los siguientes pasos:

- Enmascarar canal.
- Limpiar flip-flop.
- Fijar modo de transferencia.
- Poner dirección de inicio.
- Limpiar flip-flop.
- Fijar tamaño de transferencia.
- Liberar canal.

Estos datos le dicen al controlador que efectúe una transferencia en el canal 3, de 49152 bytes en la dirección del búfer de cuadro, y en transferencia en bloque. Para poner en acción el ciclo DMA, primero hay que habilitar el paso al DMA colocando un uno en la compuerta AND que pone a disposición el flip-flop que controla la transferencia DMA en la tarjeta, y luego es necesario poner un cambio de cero a uno en el reloj CLK del flip-flop para que cambie a alto la conexión con DREQ3, pidiendo así un ciclo DMA. Para detectar el final del ciclo, el programa tiene un ciclo para vigilar el cuarto bit del puerto 08 hexadecimal, el cual señala la finalización de un ciclo en el canal 3. Una vez cumplido esto, el programa puede continuar.

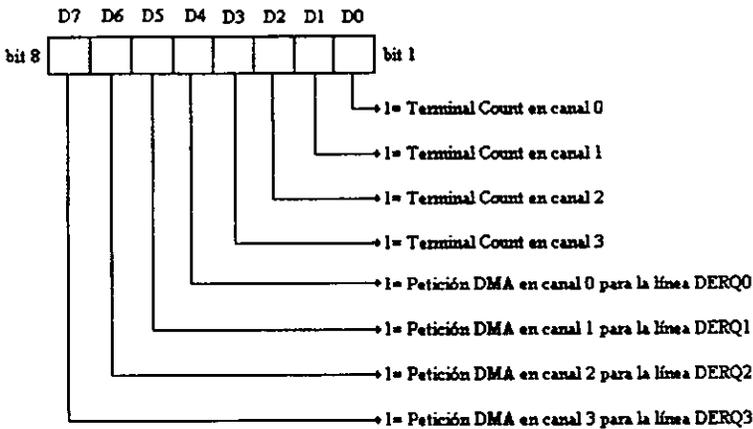


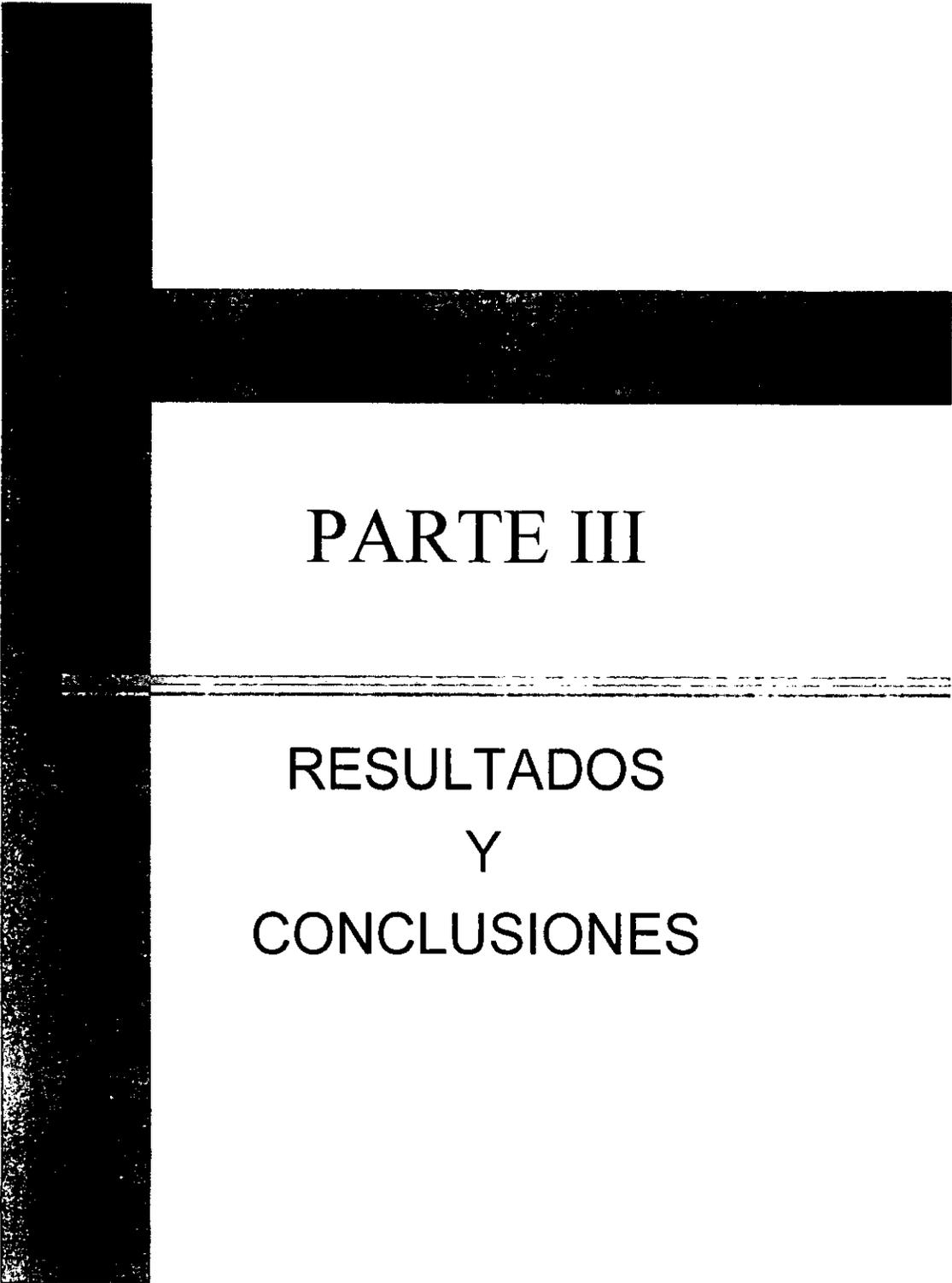
Fig. II-3.2. Registro de estado del controlador DMA en el puerto 08h y 0d0h.

3.4 Visualización

Al finalizar esta parte, la computadora ya tiene en su búfer el cuadro listo para visualizarse. Para la versión de DOS sólo es necesario ejecutar la instrucción "Putimage", mientras que en la versión de WINDOWS, se usa una función especial SDK para visualizar el BITMAP creado, la cual es "StretchDIBits". Esta última instrucción es ejecutada de manera que pinte la imagen en forma contraria a como se pintaría el Bitmap, esto porque el Bitmap tiene ordenadas sus líneas partiendo desde la línea más baja del cuadro, es decir, pinta la imagen de abajo hacia arriba.

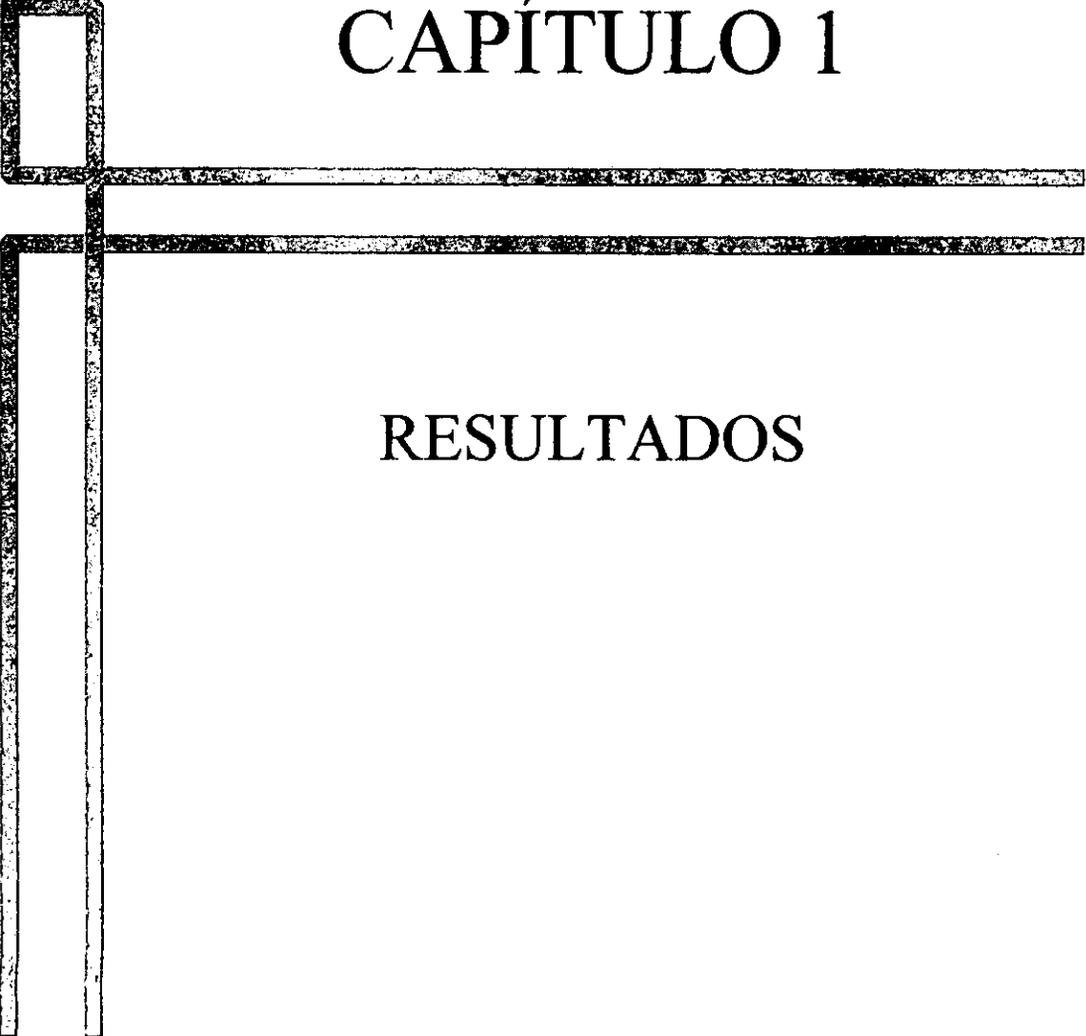
Este ciclo se efectúa continuamente tan rápido como sea posible y hasta que el usuario finaliza el programa.

Como dato adicional, se buscó la forma más rápida posible para visualizar la imagen, ya que al usar el simple "Putpixel" se generaba un efecto de barrido molesto debido a que la instrucción era relativamente lenta. Al tratar de mejorar esto se encontró que la imagen completa podía manejarse desde un bloque de memoria con la instrucción "Putimage" en el caso del sistema operativo; o el manejo como BITMAP en el caso de WINDOWS.



PARTE III

RESULTADOS
Y
CONCLUSIONES



CAPÍTULO 1

RESULTADOS

1.1 Procesamiento implementado para la presentación de la tarjeta

Los programas realizados en este trabajo tienen como objetivo el demostrar que se pueden hacer procesamientos sencillos, con ayuda de la tarjeta y con pocos recursos, como sería el caso que nos atañe. Cabe recordar que la tarjeta de digitalización sólo efectúa la transformación de la imagen analógica a una imagen digital, sin efectuar ningún tipo de compresión o procesamiento dentro de los dominios de la tarjeta. Por lo anterior, si se desea efectuar algún tipo de tratamiento con las imágenes de la tarjeta, es necesario realizarlo con los recursos existentes en la estación de trabajo tipo PC, dependiendo por completo de su velocidad de procesamiento. Tomando en cuenta esto, se realizaron algunos ejemplos demostrativos, algunos se pudieron hacer de modo que los resultados se presentarán en tiempo real, pero hubo otros para los cuales sólo se podían trabajar con imágenes capturadas previamente y guardadas como archivos de imagen.

El formato empleado para guardar los cuadros de vídeo deseados fue el formato RAW, el cual es la representación más burda de la imagen, ya que no requiere de ningún tipo de cabecera. En principio se trabajó con el formato BMP; pero para manejar este formato se requiere incorporar al archivo, aparte de la información visible o píxeles, ciertos datos que indican el número de bits por píxel, el número de planos de la imagen, el tamaño del archivo, las dimensiones de la imagen, la paleta de colores de la imagen, entre otras cosas. Esto es deseable para dar flexibilidad en el manejo de las imágenes, pero para nuestro caso, trabajamos sobre la información visible, por lo que fue conveniente dejar sólo la información directamente. Este es el caso del formato RAW, no requiere de ningún parámetro extra, como cabecera y al momento de la visualización, sólo requiere que se especifiquen las características indispensables como son la dimensiones de la imagen y el número de bits por píxel.

Los procesamientos que se efectuaron por cuadro aislado de la imagen fueron pequeñas pruebas de compresión sencillas, como transmitir ciertos píxeles de la imagen e interpolar los demás a partir de los recibidos. Se probó un pequeño filtro de ruido para mejorar los resultados de esta compresión.

Para el caso de los procesamientos en tiempo real, se buscó hacer ejemplos sencillos y que requirieran poco tiempo de procesamiento para no perder la poca continuidad de la imágenes capturadas. Tales ejemplos son: un programa que realiza la obtención del contorno de un objeto en movimiento dominante en la imagen con un cierto grado de contraste con respecto al fondo de la escena, el otro programa realiza la obtención de un vector que señala el movimiento realizado sobre la escena de un objeto dominante, esta señalización se realiza por medio de una línea que simula una flecha. Ambos programas están basados en las operaciones de morfología matemática, como son la dilatación, la erosión, el opening y el closing.

1.2 Procesamiento sobre cuadros sencillos

1.2.1 Compresión de imágenes fijas

Dentro de las aplicaciones que se realizaron como ejemplo básico de lo que se puede hacer dentro del campo de procesamiento digital de imágenes, tenemos la compresión. Un

ejemplo breve y sencillo, pero muy ilustrativo de lo que es la compresión, es lo que a continuación describiremos.

Basándonos en algo que se conoce como “reconstrucción de bitmap”, pudimos desarrollar lo que fue una ligera compresión de imágenes fijas. Existen diferentes formas de codificar la información, la que nosotros empleamos es la que se conoce como “*pixeles dependientes e independientes*” dicha técnica es usualmente empleada para reducir la cantidad de información que forma una imagen, es decir, no toda la información será transmitida, sino que sólo la parte de la imagen que uno desee y el resto se obtiene a partir de ésta, mediante una interpolación lineal.

Dentro de los ejemplos de compresión que mostraremos, está el de una imagen que se le practicarán dos distintas formas de compresión basadas en el mismo principio. Para ver las diferencias que hay entre un algoritmo y otro para un mismo objetivo (comprimir la imagen).

Como primer ejemplo, sea la imagen original como la que se ilustra en la figura III-1.1, esta imagen tiene un tamaño de 256 x 192 pixeles, la cual fue capturada por nuestra tarjeta.



Fig. III-1.1. Imagen sin procesamiento.

De la imagen se transmitirá sólo parte de la información y el resto se obtendrá interpolando linealmente con los pixeles vecinos. Esto es, de la imagen original transmitiremos un pixel si y uno no, tanto en columnas como en renglones, teniendo una imagen formada por los pixeles, como muestra la imagen (figura III-1.2).

X	X	X	X	X	X	X	X	X
X								X
X		X		X		X		X
X								X
X		X		X		X		X
X								X
X		X		X		X		X
X								X
X	X	X	X	X	X	X	X	X

Fig. III-1.2. Pixeles transmitidos.

donde X son los pixeles transmitidos.

Los pixeles faltantes se calcularon de la siguiente manera (fig. III-1.3).

X	X	X	X	X	X	X	X
X	O	Z	O	Z	O	Z	X
X	Y	X	Y	X	Y	X	X
X	O	Z	O	Z	O	Z	X
X	Y	X	Y	X	Y	X	X
X	O	Z	O	Z	O	Z	X
X	Y	X	Y	X	Y	X	X
X	X	X	X	X	X	X	X

Fig. III-1.3. Imagen reconstruida.

Los pixeles O son los primeros en calcularse de la forma siguiente:

X ₁		X ₂
	O	
X ₃		X ₄

Fig. III-1.4. Primera interpolación.

$$O = (X_1 + X_2 + X_3 + X_4) / 4 \tag{III-1.1}$$

en donde

- X₁ es el pixel en la esquina superior izquierda al pixel a calcular
- X₂ es el pixel en la esquina superior derecha al pixel a calcular
- X₃ es el pixel en la esquina inferior izquierda al pixel a calcular
- X₄ es el pixel en la esquina inferior derecha al pixel a calcular

Los pixeles Y se calculan como:

	O ₁	
X ₁	Y	X ₂
	O ₂	

Fig. III-1.5. Segunda interpolación.

$$Y = (X_1 + X_2 + O_1 + O_2) / 4 \tag{III-1.2}$$

En este cálculo se emplean pixeles transmitidos (X) y pixeles interpolados (O) en donde

- O₁ es el pixel superior al pixel a calcular
- O₂ es el pixel inferior pixel a calcular
- X₁ es el pixel a la izquierda del pixel a calcular
- X₂ es el pixel a la derecha del pixel a calcular

y finalmente los pixeles Z se calculan de acuerdo a:

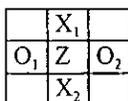


Fig. III-1.6. Tercera interpolación.

$$Z = (X_1 + X_2 + O_1 + O_2) / 4 \tag{III-1.3}$$

En este caso al igual que en el anterior se emplean pixeles transmitidos (X) y pixeles interpolados (O) en donde

- X₁ es el pixel superior al pixel a calcular
- X₂ es el pixel inferior pixel a calcular
- O₁ es el pixel a la izquierda del pixel a calcular
- O₂ es el pixel a la derecha del pixel a calcular

Dado que en el limite de la imagen se tienen problemas para la aplicación de este algoritmo, se decidió transmitir todo el contorno sin compresión. La tasa de compresión en este caso es de 4:1 (son necesarios 2 bits por pixel). Veamos el resultado de la imagen comprimida aplicando este método en la figura III-1.7.

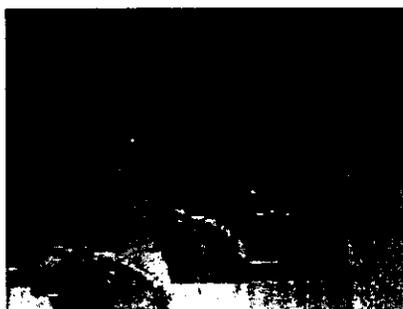


Fig. III-1.7. Resultado de la compresión.

Un segundo ejemplo será ahora el transmitir un renglón si y uno no y los renglones no transmitidos se interpolarán a partir de los otros. La parte de la imagen que se transmite es la concerniente a los renglones pares, iniciando en el renglón 0.

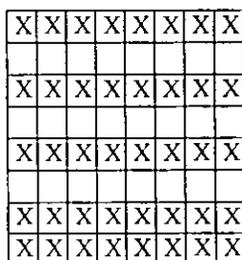


Fig. III-1.8. Renglones transmitidos.

Para poder aplicar el algoritmo sin tener problemas con el cálculo del último renglón, el cual es un renglón impar, se decidió transmitir éste íntegramente. Como se puede observar, la cantidad de información transmitida es mayor por lo que se espera que la calidad de la imagen mejore considerablemente.

Los pixeles interpolados se obtienen de la siguiente manera:

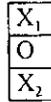


Fig. III-1.9. Interpolación de pixeles de líneas no transmitidas.

$$O = (X_1 + X_2) / 2 \tag{III-1.4}$$

en donde:

X_1 es el pixel del renglón superior a el pixel a calcular

X_2 es el pixel del renglón inferior a el pixel a calcular

Finalmente, la imagen quedará formada por los pixeles transmitidos y los pixeles interpolados de la forma:

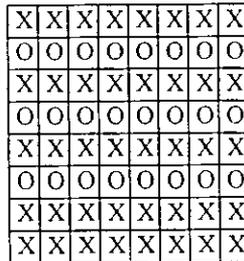


Fig. III-1.10. Imagen reconstruida.

Resultados prácticos pueden verse en la siguiente imagen en la figura III-1.11. La tasa de compresión es de 2:1 en este caso (son necesarios 4 bits por pixel).



Fig. III-1.11. Resultado de la compresión.

Como puede verse, esta imagen tiene mayor calidad que la imagen mostrada en la figura III-1.10. Sin embargo, la calidad repercute en mayor cantidad de píxeles a transmitir, por lo que dependerá del propósito de la compresión cuál de estos algoritmos será tomado en cuenta.

Pueden hacerse un sin fin de algoritmos como los aquí planteados de acuerdo a la aplicación que se desee realizar y de acuerdo a las características de la imagen, es decir, los rasgos que quieran ser identificados o conservados en una imagen.

Para mejorar los resultados obtenidos, se implementó un prefiltrado antes de la elección de los píxeles a transmitir, dicho filtrado nos permite distribuir la información de la imagen a lo largo y ancho de ésta, para que a la hora de hacer la discriminación de los píxeles, la información que pudiese servir no sea eliminada, como pueden ser líneas rectas, curvas o hasta los mismos contornos, el filtro que se empleó se conoce como "*filtro de mediana*" (*Median Filter*¹⁶).

1.2.1.1 Filtro de mediana

Este filtro es muy sencillo de implementar y da buenos resultados, funciona de la siguiente manera: como mencionamos algunos párrafos atrás, el contorno se transmite íntegro, por lo que el primer píxel a transmitir es el ubicado en la posición (1,1) considerando una numeración de 0 a 255 en los píxeles horizontales y 0 a 191 en los píxeles verticales.

Este píxel no es exactamente el que se transmitirá, sino lo que hace el filtro es: a partir de la ubicación del píxel en cuestión con los píxeles vecinos empleando una ventana de 3x3 píxeles, en donde el píxel central es el de estudio, se hace un ordenamiento en forma ascendente y el valor de píxel que quede situado en el centro del ordenamiento es el que tomará la posición (1,1), lo mismo sucede con el resto de los píxeles que forman la imagen excepto el contorno. Véase a forma de ejemplo esta parte de alguna imagen en la que el valor del centro es el píxel ubicado en la posición (1,1).

0	2	5
9	3	4
8	9	3

Fig. III-1.12. Región de la imagen en estudio para filtrar.

Los píxeles de esta ventana tienen los valores: 0, 2, 5, 9, 3, 4, 8, 9, 3, el filtro hace el ordenamiento de estos valores quedando en forma ascendente como 0, 2, 3, 3, 4, 5, 8, 9, 9. Como podemos ver el valor de píxel que queda al centro es el 4, éste será el valor que tome el píxel ubicado en la posición (1,1), en lugar del 3 que era el que correspondía antes del filtrado. Lo mismo se hace con el resto de las localidades o píxeles que forman el cuadro hasta culminar con el de la posición (254,190). Una vez que se ha realizado el filtrado podemos hacer la compresión que se desee, los ejemplos de compresión que realizamos previamente ahora se ven mejorados por el prefiltrado teniendo los resultados que se muestran en las imágenes (figuras III.1.13 a) y III-1.13 b)).

¹⁶ William K. Pratt; *Digital Image Processing*; John Wiley & sons, inc., EUA, 1991; pag. 294-296.

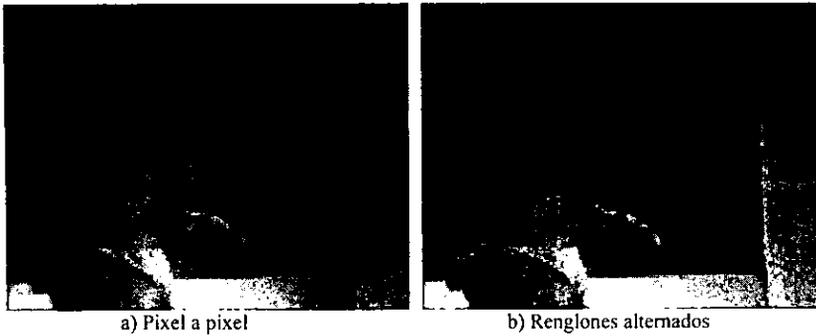


Fig. III-1.13. Algoritmos de compresión empleando el filtro antes de comprimir.

Veamos finalmente como queda la imagen desde que es capturada (figura III-1.14 a), hasta que sufre el proceso de compresión y reconstrucción sin filtro (figura III-1.14 b) y el proceso de filtrado compresión y reconstrucción (figura III-1.14 c). Para que esto fuera más notorio, usamos la imagen que sufrió mayor compresión, es decir, la que se transmitió un pixel sí uno no. Como se puede observar el prefiltrado ocasionaría que la imagen reconstruida sea más lisa, esto es, sus contornos están más borrosos (fig. III-1.14 c). Este tipo de filtrado simple requiere de mayor tiempo de procesamiento, lo que provoca un consumo de tiempo valioso si se quisiera trabajar con este tipo de imágenes en tiempo real.

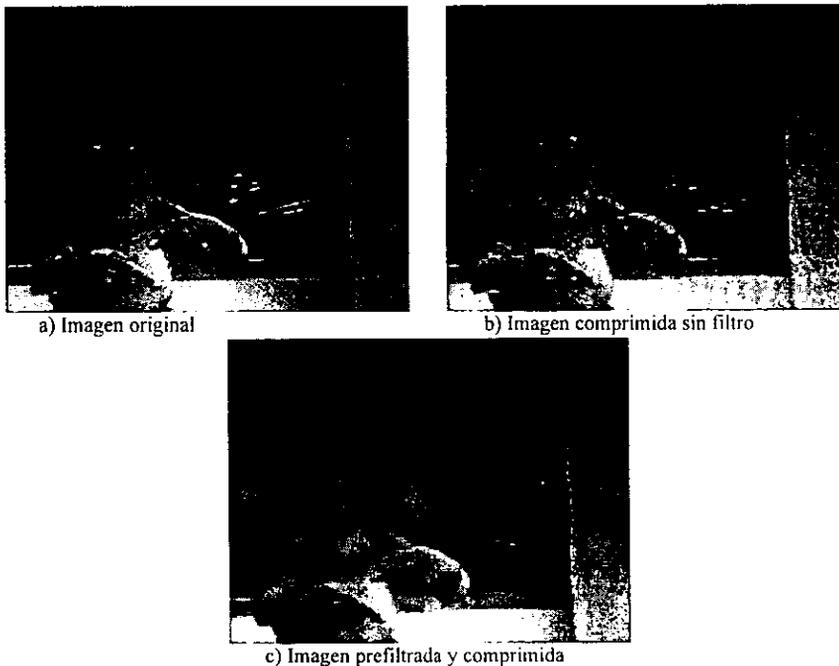


Fig. III.1.14 Comparación de los resultados de compresión sin y con filtro de mediana.

1.3 Procesamiento sobre cuadros en secuencia

Estos programas fueron realizados usando el paquete VISUAL C++, el cual está basado en el lenguaje de programación C++ orientado a Windows, por lo que dichos programas corren usando la interface gráfica de Windows.

1.3.1 Erosión y dilatación

Para el desarrollo de estos programas se empleó la morfología matemática, en especial las operaciones de erosión y dilatación. Se usó un algoritmo mediante el cual el proceso se hacia más rápido y más sencillo; se logró sabiendo que la erosión se puede obtener haciendo la operación lógica AND de todas las proyecciones de la imagen posibles, una por cada punto que forma el elemento estructurante. Dichas proyecciones consisten en: por cada punto del elemento estructurante, se realiza un desplazamiento de la imagen con respecto al origen del elemento estructurante. Es decir, para el caso de un elemento de tres puntos en línea horizontal con el origen del elemento estructurante en el pixel central, la erosión se puede obtener de la AND de la imagen recorrida un pixel hacia la izquierda (por el elemento de la izquierda), con la imagen sin mover (por el pixel de centro) y con la imagen movida un pixel a la derecha (por el pixel de la derecha), tal como se muestra en la figura III- 1.15.

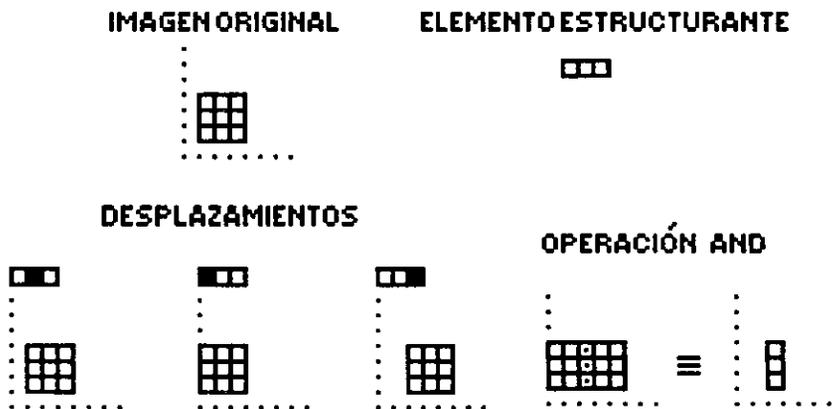


Fig. III-1.15 Erosión.

Además, se sabía que la dilatación también se puede obtener realizando la operación lógica OR de las proyecciones de la imagen de acuerdo al elemento estructurante, de manera muy similar a la erosión, como se muestra en la figura III-1.16.

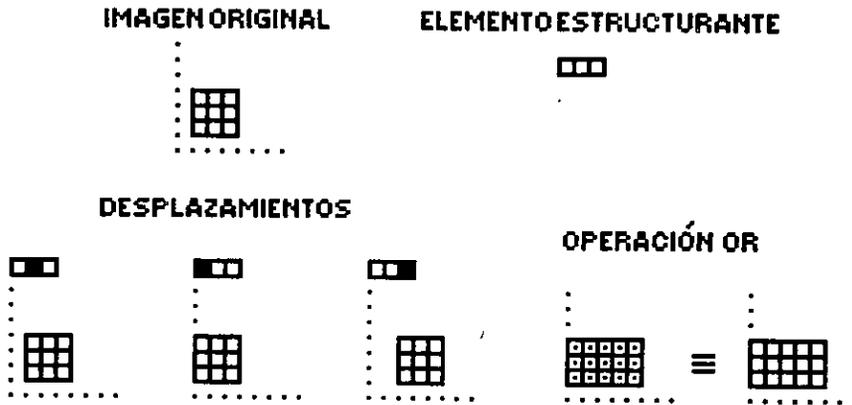


Fig. III-1.16 Dilatación.

Para el caso de la programación, esto se reduce a sólo efectuar operaciones OR o AND con diferentes apuntadores desplazados de acuerdo al elemento estructurante, reduciendo así el tiempo de procesamiento. Esta operación se realiza para cada elemento de la imagen, por lo que los diagramas de flujo son:

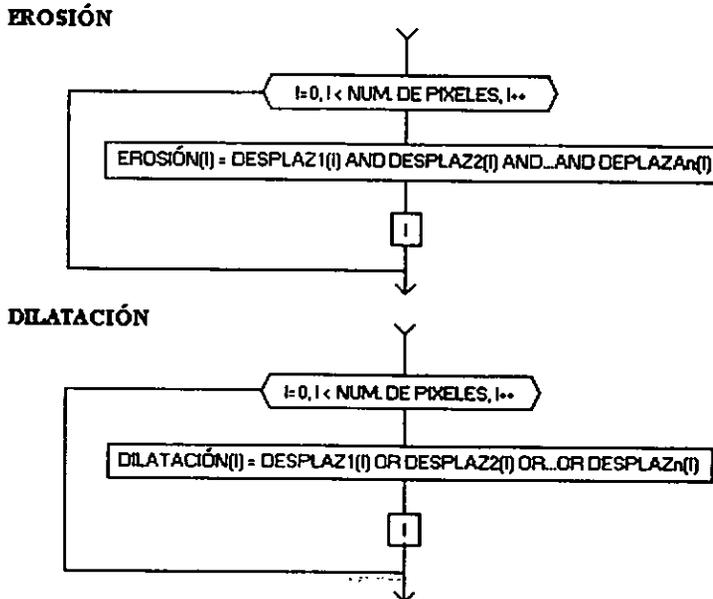


Fig. III-1.17 Diagrama de flujo de erosión y dilatación.

donde n es el número de puntos del elemento estructurante.

1.3.2 Programa del contorno

El primer programa realizado fue la obtención del contorno de un objeto dominante en movimiento sobre una escena. Para esto era necesario obtener dicho objeto separando cualquier otro elemento que pudiera interferir con la aplicación.

Para apartar los objetos en movimiento de una imagen de los que no lo están, es decir, de los objetos de fondo, se realizó una substracción de la imagen actual con respecto a una imagen previa. Con esto, en la resta sólo aparecen las secciones que se encuentran en movimiento, es decir, que hayan cambiado de posición o que son nuevas en la escena, con respecto a la imagen de referencia. Sobre esta misma substracción se realizó la binarización de la imagen resultante, es decir, colocarla en únicamente dos tonos, para poder aplicar la morfología matemática. Para lograr la binarización se utilizó un umbral, el cual determinaba si el pixel comparado se representaría como uno o como cero. Una vez que se contó con la imagen binaria, se realizó un proceso de opening-closing con el objeto de limpiar la imagen binaria de pequeños objetos espurios que podían interferir en el resultado. Una vez limpiada la imagen a un punto aceptable, se tenían dos caminos para obtener el contorno de los objetos restantes. Un camino era el barrer la imagen en sentido vertical y horizontal, de modo de colocar un pixel de contorno en aquellos pixeles donde se encontrara un cambio de cero a uno o de uno a cero en los pixeles consecutivos, lo cual indicaba que se trataba de un límite de un objeto. El otro camino posible era el de efectuar una erosión sobre la imagen, para adelgazar los objetos, de modo de que al efectuar una resta de la figura obtenida con respecto a la original, se tendrían las secciones perdidas durante la erosión que serían las correspondientes al contorno de la imagen. Por razones meramente estéticas, se quería que la representación de contorno se visualizará con un color distinto de los usados en la imagen, como rojo por ejemplo. Finalmente, se optó por la primera opción, la cual permitía la manipulación individual de los puntos de contorno de manera sencilla y rápida, para emplear la instrucción *SetPixel* o *PutPixel*, dependiendo del paquete empleado, la cual permite poner un pixel en una posición y color deseados sin interferir con los colores de la paleta. A continuación se muestra un diagrama de flujo del programa en general, figura III-1.18.

Como se puede ver es muy similar al diagrama de flujo que se describió en la operación de la tarjeta. Las diferencias principales se pueden ver en que para estas operaciones se utilizan bloques de memoria adicionales para el manejo de los cuadros de vídeo. En este caso se apartaron dos secciones de memoria en modo real, mientras que también se apartó un bloque de memoria en modo protegido. Esto se hizo, porque en uno de los dos bloques de memoria tipo DOS, se guarda el cuadro de referencia o antiguo, mientras que el otro recibe el cuadro actual y posteriormente, junto con el tercer bloque de memoria, sirve para realizar la operación de opening y closing, para limpiar un poco la imagen. Por el uso en estas operaciones de los apuntadores, se apartó memoria adicional al tamaño del cuadro para que los apuntadores se puedan mover libremente sin causar problemas. También se puede ver la sección en la que se hace la resta de las dos imágenes y se efectúa directamente pixel a pixel la binarización. Después de la binarización, sobre la imagen resultante, se realizan las operaciones de erosión y dilatación de acuerdo al procesamiento de opening y closing, es decir, erosión-dilatación dilatación-erosión.

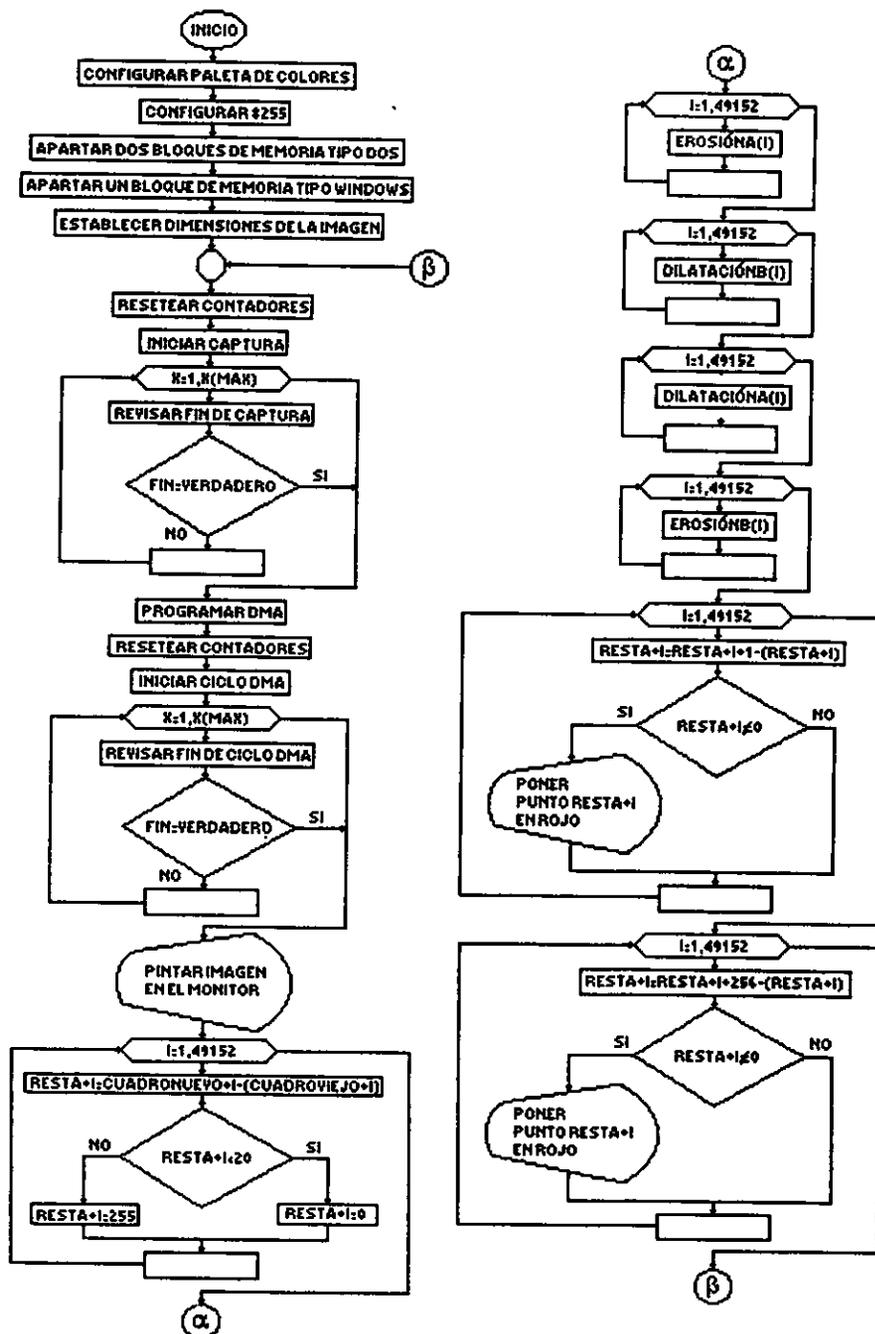


Fig. III-1.18. Diagrama de flujo del programa de contorno.

Después de esto se hace el barrido de la imagen de izquierda a derecha, colocando punto de color rojo sobre la imagen actual pintada previamente, de acuerdo a donde se tuviera cambio de tono en la señal binaria procesada.

Posteriormente, se hace algo similar pero barriendo de arriba hacia abajo. Al final se tiene el cuadro actual con contornos rojos sobre los objetos que tuvieron un cambio substancial de tono, lo cual en ocasiones significa movimiento. El proceso se lleva a cabo indefinidamente y en forma continua para dar la sensación de movimiento. Pero por los pasos adicionales, esto decreta la velocidad de despliegue de los cuadros con respecto al programa de operación sencillo.

En los primeros intentos del programa se trabajó como cuadro de referencia el cuadro inmediato anterior, pero esto ocasionaba que se tuviera, en la binarización de una imagen con un objeto en movimiento, la zona que había ocupado el objeto por su movimiento y también la zona que había dejado libre con respecto al cuadro anterior, lo cual ocasionaba que se obtuvieran los contornos para las dos situaciones. Por esta razón, se decidió tomar un cuadro de referencia cualquiera elegido por el usuario del programa mediante una tecla, con el fin de que se pudiera capturar una imagen de fondo, para que al pasar un objeto en movimiento sólo se tuvieran las zonas que dicho objeto ocupa en el cuadro actual con respecto al fondo, esto se obtuvo vigilando la pulsación de una tecla, entonces sólo se intercambian los apuntadores del cuadro actual y del de referencia.

En la figura III-1.19 se muestra el trabajo de este programa, el segundo cuadro es capturado en memoria como cuadro de referencia, lo que para nosotros sería el fondo, mientras que el primer cuadro es el cuadro actual que se tiene presente en el momento en el que se presenta la entrada de un objeto ajeno al fondo, entonces este objeto en movimiento causa que el algoritmo detecte el cambio, separe el objeto y coloque el contorno en la imagen actual del objeto en movimiento.

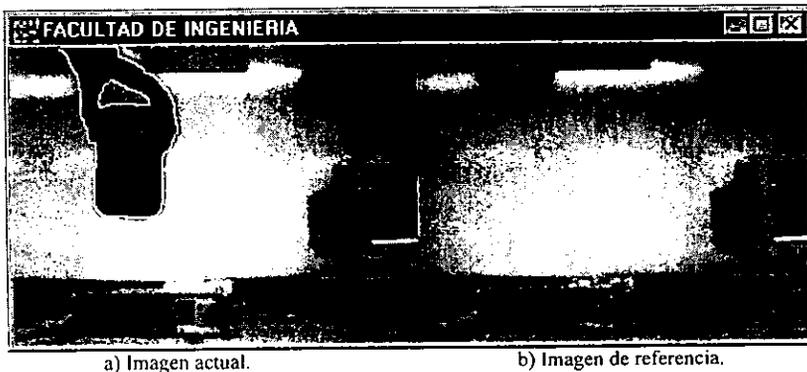


Fig. III-1.19. Ejemplo de la detección de contorno.

1.3.3 Programa de la flecha

El segundo programa realizado fue la obtención de un vector de movimiento del objeto más grande o dominante en una escena. Para hacer esto, nuevamente era necesario

apartar sólo los objetos en movimiento en una escena en un cuadro actual con respecto a un cuadro de referencia, por lo que se procedió de manera similar al programa anterior.

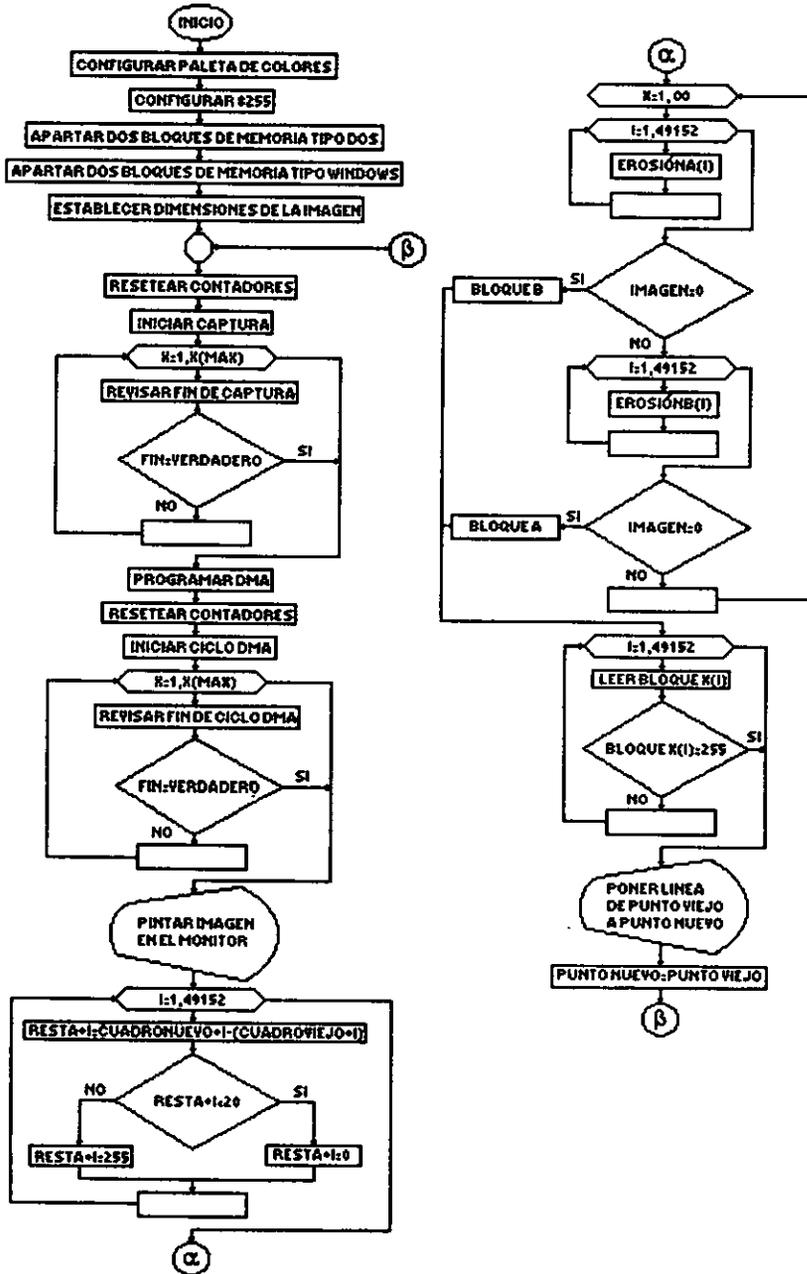


Fig. III-1.20. Diagrama de flujo del programa de la flecha.

Como se muestra en el diagrama de flujo de la figura III-1.20, también se efectúa la substracción de dichos cuadros y se obtiene la binarización de la misma. Con esto se tiene una representación de los objetos en movimiento, pero era necesario obtener algo parecido al centroide del objeto de mayor dimensión. Como se ve en el mismo diagrama, el proceso hecho para esto es efectuar una erosión iterativa sobre la binarización hasta que no se tuviera ningún elemento o pixel de tono blanco en la imagen final, lo cual indicaba que la penúltima imagen contenía los últimos elementos del objeto más grande, para esto, se hacen ciclos indefinidos de erosiones con condicionales de salto para detectar el momento en el que se tiene la imagen binaria a cero (última erosión). Finalmente, sólo es necesario detectar alguno de estos pixeles con el fin de tomarlo como una aproximación del centroide, con esta posición, comparada con la posición del procesamiento del cuadro anterior, se obtiene una línea que señala el cambio de posición del objeto entre los cuadros consecutivos.

Como en el otro programa, se probó la operación del algoritmo tomando el cuadro de referencia como el cuadro anterior inmediato, pero se tuvieron problemas similares, por lo que se eligió también la opción de tomar el cuadro de referencia en un tiempo elegido por el usuario.

Evidentemente, por la cuestión del adelgazamiento a cero, se puede deducir fácilmente que para objetos que ocupen mayor área en el cuadro de la cámara los resultados tardarán más en producirse, es por eso que para tener una mayor sensación de continuidad se emplean áreas pequeñas, como la de una pluma por ejemplo. A continuación se muestran cuadros del programa con una pluma, figura III-1.21.

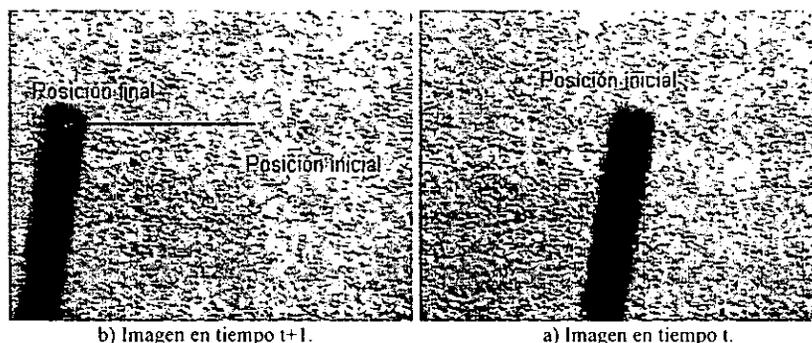


Fig. III-1.21. Ejemplo de la detección de movimiento.

La imagen de fondo es guardada en memoria como cuadro de referencia, tal y como se hizo en el programa del contorno, de acuerdo a lo anterior, el primer cuadro ha detectado la presencia de la pluma en el tiempo t , por lo que ha desarrollado el algoritmo con respecto al tiempo $t - 1$, por lo que al final de esta iteración ha guardado el punto de localización de la punta de la flecha del cuadro t como punto anterior. Al cambiar al siguiente cuadro, la pluma ha cambiado de posición, por lo que el algoritmo del cuadro $t + 1$ con respecto al cuadro t arroja una nueva coordenada como punto actual, pintando así una flecha desde el punto anterior al punto actual, tal y como se muestra en la figura del cuadro $t + 1$.

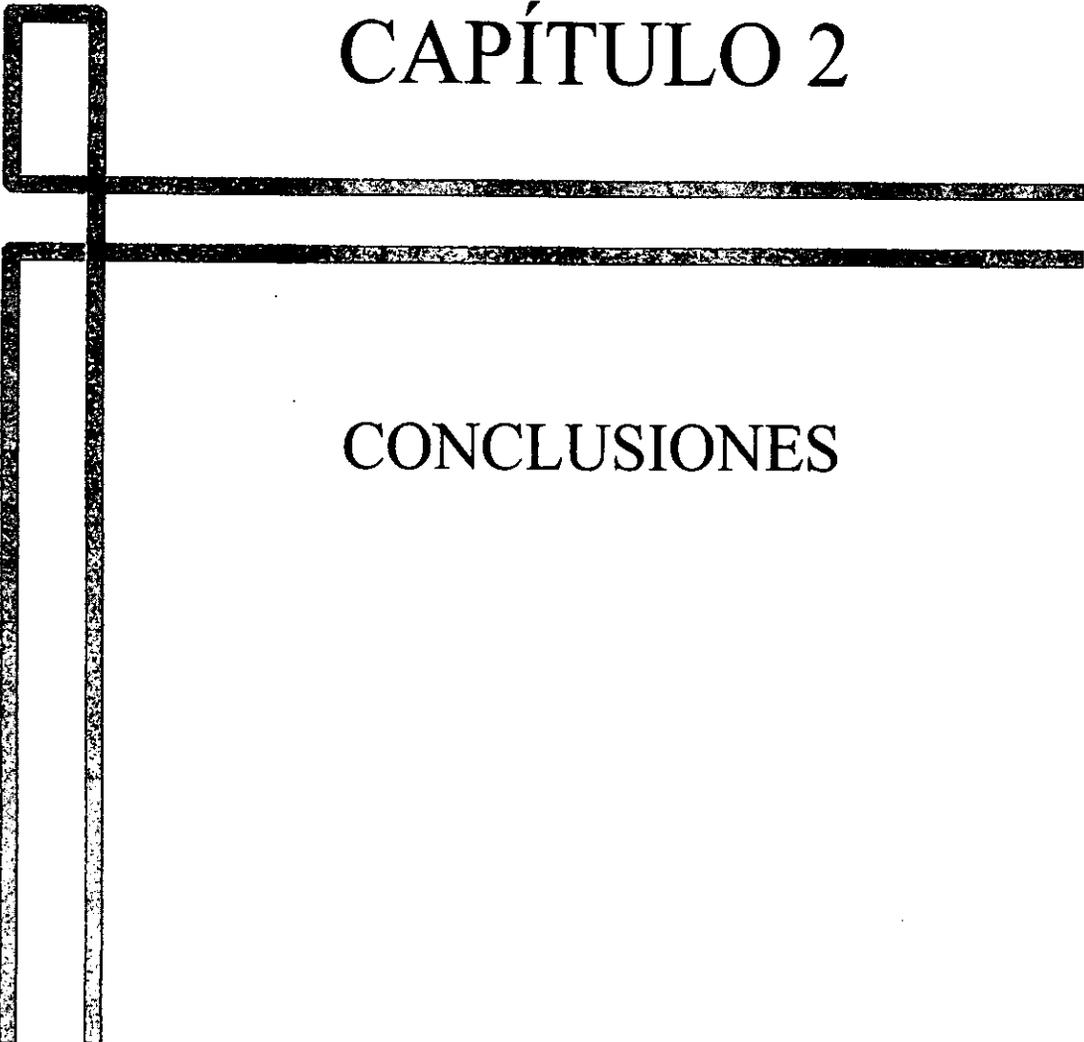
1.3.4 Consideraciones

Estos programas demandaron tiempo de procesamiento de la computadora, por lo que el proceso de captura, adquisición y visualización se hizo lento por la adición de este tipo de procesamiento. El promedio de cuadros por segundo en ambos programas fue de 1 cuadro por segundo aproximadamente, lo cual permite ver cierto grado de continuidad en la imagen.

Se puede ver que los programas tienen muchas limitaciones en cuanto a velocidad de la máquina empleada, es muy sensible a cambios de luz o a tonos similares entre el fondo y el objeto en movimiento.

Además, se intentó poner elementos adicionales que mejoraran la calidad de los resultados, como fue la adición de filtros o la detección del umbral de binarización mediante el histograma de la imagen, pero estos elementos hacían mucho más lento el proceso por lo que se optó por descartarlos del algoritmo.

Estos programas son muy básicos y se les pueden incorporar muchas opciones más que mejorarían substancialmente la calidad de los resultados, sin embargo, se debe negociar con la velocidad de procesamiento existente en la PC en cuestión.



CAPÍTULO 2

CONCLUSIONES

2.1 Conclusiones

Angel Gregorio Labastida Alvarado

Como se puede observar, ésta es una tarjeta que puede tener distintas aplicaciones posibles. Es un circuito relativamente sencillo y con componentes muy simples. Cabe destacar que es posible reducir de manera substancial el circuito usando circuitos GAL. Evidentemente, el corazón de la tarjeta es el convertidor A/D, sin el cual no se podrían realizar las conversiones de la señal de vídeo analógico a la velocidad que se requiere. La calidad de los resultados es buena tomando en cuenta las limitaciones de la sencillez, el tamaño y el costo. Lógicamente, dicho circuito no puede competir directamente con las tarjetas que existen actualmente en el mercado, ya sea de captura de vídeo o de aplicaciones específicas. Sin embargo, gracias a nuestro desarrollo, nos podemos dar una idea que la tarjeta podría llegar a tener cierta aplicación comercial, por ejemplo, como parte de algún sistema que no requiera de mucha velocidad.

Por otra parte hay que decir que existen circuitos integrados que realizan gran parte del trabajo hecho por la tarjeta. Sin embargo, la adquisición en México es complicada y hay que recordar que el objetivo del presente trabajo es el mostrar una posible solución al problema desarrollando la solución de acuerdo con las limitantes que se tienen. Nuestra tarjeta ofrece una solución simple y didáctica del proceso de digitalización de imágenes.

En lo que respecta a los resultados obtenidos, se puede ver que la limitación más importante de la tarjeta es el tiempo en el que se lleva a cabo toda la operación, por lo que se deben buscar nuevos caminos para extender el rendimiento de la tarjeta. Un caso específico podría ser el incorporarla dentro de un sistema basado en un DSP (Digital Signal Processing), el cual puede mejorar la velocidad en procesos donde se realice la compresión o algún procesamiento de manera necesaria.

Otro punto importante que puede tener ajustes es el hecho de que la tarjeta trabaja con el bus ISA, por lo que la adquisición de la imágenes de la tarjeta a la PC queda limitada primeramente por este detalle. La posible solución podría ser el rediseñar la tarjeta para trabajar con el bus PCI, el cual parece ser el estándar que poco a poco substituirá al ISA por su mayor velocidad y capacidad.

En cuanto al tamaño, por el gran espacio que ocupan las componentes convencionales, no es posible hacer el circuito impreso como si se tratara de una tarjeta de expansión común y corriente. La solución podría ser el utilizar circuitos que substituyeran a un gran número de los que se emplean actualmente, por ejemplo, el usar circuitos GAL. Otra posible solución, pero demasiado ambiciosa, podría ser el diseñar el circuito impreso empleando la tecnología de montaje superficial.

Por otro lado, se puede ver que las componentes empleadas son relativamente baratas, pero el diseño y la fabricación de la tarjeta fue muy complicada, lo que hace que la tarjeta tenga un valor comercial elevado. Sin embargo, como en la mayoría de los diseños, una vez de que se produce uno, el costo de las siguientes copias se va reduciendo.

El montaje de la tarjeta en WIRE WRAP se eligió porque es una forma muy práctica para la presentación de la tarjeta, ya que como se ha mencionado anteriormente, este circuito es parte de una investigación más compleja, es decir, se busca incorporar un DSP

en el circuito, lo cual traería evidentemente modificaciones en la tarjeta, por lo que no tendría caso desarrollar una máscara de circuito impreso por el momento.

En lo personal, como futuro ingeniero, el presente trabajo tiene muchos puntos importantes dentro del desarrollo profesional, ya que cuenta con muchas características que por cuestión de tiempo o de temario no son vistas con profundidad a lo largo de la carrera. Podría mencionar a modo de ejemplo la teoría de la señal de video compuesto analógico, la cual considero que es muy interesante que engloba a su vez varios puntos interesantes y didácticos como el multiplexaje de señales o de la sincronía de color. Otro punto podría ser el proceso completo de digitalización de señales.

Juan Rene López Esqueda

El principal objetivo en el desarrollo de este tema, fue el de diseñar una tarjeta capaz de digitalizar imágenes para el procesamiento digital. Dicha tarjeta digitalizaría a 8 bits y en tonos de grises la mayor cantidad posible de cuadros aplicando algoritmos de procesamiento digital a estas imágenes o simplemente desplegarlas secuencialmente sobre el monitor de la computadora formando una señal de vídeo. Este cometido fue cumplido, no sin antes pasar por una serie de problemas y conflictos como los que a continuación mencionaré.

Durante el desarrollo del presente trabajo hemos recorrido una serie de términos y conceptos, que van de terminología empleada cotidianamente a conceptos técnicos y complejos, como puede ser hablar de una imagen hasta llegar a los algoritmos de procesamiento digital de imágenes, pasando por el vídeo analógico y digital, así como hablar de los tipos de cámaras existentes. Todo con el fin de lograr nuestro propósito, el poder obtener un diseño de tarjeta que nos permitiera obtener las imágenes que queríamos, de acuerdo a nuestra necesidad, para que una vez capturada esa imagen, entonces si aplicar sobre ésta algunos algoritmos de procesamiento digital.

Ahora recuerdo los inicios de este proyecto, cuando se nos planteó la idea de desarrollar una tarjeta que fuese capaz de obtener una sola imagen y poderla llevar al monitor de una PC, en estos tiempos nos encontrábamos al final de la carrera y comenzábamos el servicio social, nos sentíamos capaces de desarrollar los proyecto de electrónica que se nos encomendaran, fue así que decidimos caminar por este sendero. No sabíamos los problemas a los que nos íbamos a enfrentar, pues éste ya no era un trabajo escolar en el cual tendríamos todos los guías que saben o han desarrollado proyectos semejantes, sino que éste es un trabajo que poco se ha difundido, en el que todas las ideas y decisiones tendríamos que tomarlas nosotros mismos.

El primer tropiezo que tuvimos fue que no sabíamos mucho relacionado al tema, ¿a que me refiero?, la idea original era capturar una imagen y transmitirla vía el puerto paralelo a la PC; hasta aquí todo perfecto, pero como digitalizar la imagen, pues ésta tenía una forma de onda que no era muy conocida en los proyectos escolares. Fue éste el principio de nuestro estudio, aquí se puede justificar el CAPÍTULO 1 concerniente al vídeo analógico en este capítulo logramos conocer a fondo las características de la señal de vídeo, las partes que la conforman, como pueden ser las señales de control que necesitábamos para manejar la tarjeta. Cuando terminamos el análisis nos dimos cuenta que el puerto paralelo era demasiado lento para nuestro propósito, un nuevo problema se presentó, optamos por usar el bus de la PC y el que teníamos a la mano y del que más se ha desarrollado hardware hasta ahora fue el BUS ISA, el cual trabaja a una frecuencia superior que el puerto paralelo, fue entonces que decidimos no sólo capturar un cuadro, sino una serie de estos para que en una exposición secuencial sobre el monitor nos diera la similitud al vídeo digital. Punto de arranque para el desarrollo de lo que ahora es nuestra tarjeta. Se hizo el diseño en tres módulos: la parte analógica, la parte digital y lo que sería la transmisión a la PC con el despliegue en pantalla. Cada etapa tenía su complejidad pero dadas las bases que nos dejaron los cursos que a lo largo de 5 años nos dio la Facultad de Ingeniería se nos hizo más fácil comprender y desarrollar muchos temas que hasta ese momento la información que teníamos no era abundante, mucho menos transparente y clara.

Durante el desarrollo de este trabajo y bajo una serie de pruebas que se hicieron, como por ejemplo aplicar algoritmos de procesamiento de imágenes, nos dimos cuenta que el despliegue de las imágenes se volvía más lento a medida que los algoritmos se tornaban más complejos, además de las operaciones matemáticas que se tenían que hacer para lograr el procesamiento, el tipo de máquina que se empleaba para el trabajo también redituaba en tiempo de despliegue (como se ha mencionado en el capítulo de resultados). Es por ello que se ha pensado en hacer esta tarjeta más poderosa agregando nuevos dispositivos que realicen el procesamiento independientemente a el procesador de la PC incrementando con esto el número de cuadros por segundo en pantalla.

Muy importante se ha vuelto el hablar del video digital y tal parece que éste es lo que en un futuro se tendrá tanto en zonas de trabajo como centros de enseñanza y el hogar, para no ir tan lejos ya todos hemos oído hablar de la televisión digital; es por eso que el desarrollo de técnicas de adquisición de video están siendo analizadas y estudiadas muy afondo con el fin de lograr una calidad excelente. El desarrollo de nuevo equipo para este fin está en vías de convertirse en una fuente de trabajo muy redituable, dado que, siendo esto apenas el inicio los especialistas en esta rama de la ingeniería que se requerirán serán bastantes y deberán estar lo suficientemente preparados para superar todos los obstáculos que se presenten a su paso. Considerando esto creo que la capacidad de los ingenieros de la Facultad de Ingeniería han demostrado con creces su capacidad en innumerables ocasiones, por lo que profesionistas de esta escuela estarán sin duda trabajando en estos y muchos otros proyectos.

Finalmente, se puede decir que los resultados obtenidos al final del trabajo son muy satisfactorios, pues el lograr diseñar nuestra propia tarjeta de captura de video nos ha permitido abrir nuevas puertas, es decir, el uso de esta tarjeta tiene innumerables aplicaciones y basándose en ella pueden hacerse una expansión de la misma con fines particulares o generales.

APÉNDICE A

SISTEMAS
COMERCIALES

A.1 VFG 512-8

Esta tarjeta, fabricada por "High Res Technologies Inc.", es un grabador de cuadros con 256 tonos de grises en tiempo real, diseñado para plataformas PC.

Digitaliza imágenes provenientes de los formatos NTSC, PAL y SECAM. Es posible guardar las imágenes en formato TIFF, PCX, GIFF y en Binario.

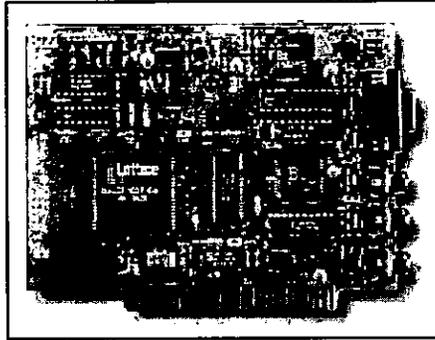


Fig. A.1. Sistema VFG 512-8

Especificaciones:

- Trabaja sobre el bus ISA.
- Cuenta con un búfer de 256K.
- Resolución de 512 x 512.
- Entrada de vídeo NTSC, PAL y SECAM.
- Despliegue de color usando OLUT (output lookup table).
- Ganancia y offset ajustable.
- Fácil instalación, no usa interrupciones ni los canales DMA.
- "Trigger" externo.
- Múltiples zonas de almacenamiento de imagen.
- Múltiples entradas de vídeo.
- Reloj de pixel externo.
- Hardware implementado para operaciones lógicas y aritméticas en tiempo real entre cuadros o grupos de pixeles.

Conectores:

- Entrada de vídeo compuesto.
- Salida de vídeo compuesto.
- Salida de color RGB (conector DB9).

Esta tarjeta tiene un costo en dólares de:

Versión para: Windows 3.1 \$299, Windows NT \$499, Windows 95 \$499.

High Res Technologies Inc.
PO Box 76
Lewiston NY 14092
TEL 416-497-6493 FAX 416-497-1636
EMAIL:HRT@PLANETEER.COM

A.2 CG-7

La tarjeta fabricada por "Scion Corporation" es un grabador de cuadros con tres canales RGB, diseñado expresamente para trabajar con cámaras CCD. Sus tres canales de digitalización independientes producen excelentes imágenes con 24 bits por color, 8 bits para cada canal. Soporta ambos tipos de señal NTSC y PAL en forma RGB con resolución de 640 x 480 y 768 x 576 pixeles respectivamente. La tarjeta puede funcionar igualmente con cámaras monocromáticas, siendo conectadas cuatro al mismo tiempo.

Contiene un sistema de alto desempeño con el DMA, esta tarjeta es capaz de desplegar imágenes con 24 bits en tiempo real sobre el monitor de una computadora. No es necesario equipo externo para enfocar o ajustar la imagen. Sin embargo, este sistema provee salida de vídeo RGB del vídeo digitalizado para ser usado con vídeo grabadores u otros periféricos de vídeo.

Es compatible con equipos Macintosh y PC sobre un slot PCI. Este sistema no puede ser usado con señal compuesta de vídeo en color.

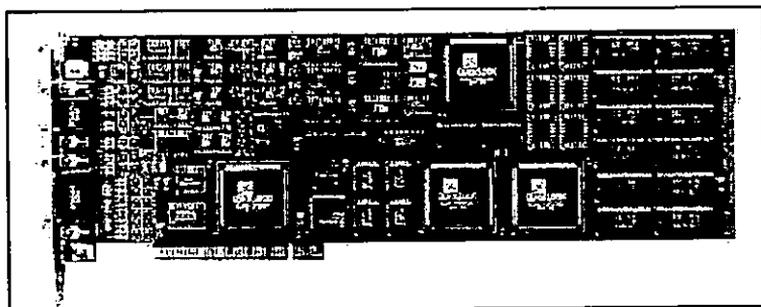


Fig. A.2 Sistema CG-7.

Características:

- Digitalizador de precisión tres canales RGB para el bus PCI.
- Despliegue en tiempo real.
- Entradas de vídeo soportadas, NTSC, PAL y RGB.
- Modo de sobremuestreo para imágenes con alta resolución.
- Ganancia y offset individual para cada canal.
- Igual desempeño en aplicaciones con escala de grises.
- Entrada y salida analógica y digital, además de dos puertos seriales.
- Compatible con Macintosh y PC.

Especificaciones:

- Captura 30 cuadros/seg. del formato NTSC, 25 cuadros/seg. del formato PAL.
- Frecuencia máxima de despliegue en NTSC 30 campos/seg., en PAL 25 campos/seg.
- Profundidad de pixeles en la captura: 24 bits en color, 8 bits en escala de grises.
- Transferencia DMA con 8, 15, 16 o 24 bits/pixeles.
- Resolución en NTSC: 640 x 480, 1280 x 960.
- Resolución en PAL: 768 x 576, 1536 x 1152.
- Relación de aspecto 1:1.
- Relación señal a ruido: 47 dB.
- Factor de ganancia de vídeo 3.7.
- Rango de digitalización 0 a 5 V.
- 3 salidas de vídeo en tonos de grises y 1 en RGB.
- Memoria de 3 Mbytes sobre la tarjeta.
- 2 puertos seriales RS-232C.
- "Trigger" externo.
- Conexión sobre el bus PCI.
- Rango de operación de 0 a 70° C.
- Consumo máximo, 15 W.

Requisitos:

Para la operación de esta tarjeta se requiere un slot PCI Macintosh o PC. Requiere una fuente de vídeo RGB o en Blanco y Negro con características NTSC o PAL. Esta tarjeta no puede trabajar con una señal compuesta. El sistema fue diseñado para trabajar principalmente con el uso de cámaras CCD y no provee ninguna corrección en la señal de vídeo.

Para la aplicación del software se requiere un equipo compatible Macintosh o PC con 16 Mbytes de memoria. Para mejorar el desempeño en el despliegue en DMA una tarjeta de despliegue de vídeo PCI podría ser usada en lugar del circuito interno de vídeo en la tarjeta madre.

Esta tarjeta tiene un costo de \$2195 dólares en:

Scion Corporation
82 Worman's Mill Ct. Suite H.
Frederick, Maryland 21701
Tel: 301 695-7870 Fax: 301 695-0035
www.scioncorp.com

A.3 MOTION PICTURE

Esta es una tarjeta de captura de imágenes de alta calidad para PC. Captura imágenes provenientes del formato NTSC y PAL. Usada para sistemas de videoconferencia, arte, gráficos, seguridad y cualquier aplicación que demande máxima

calidad al menor costo. Fabricada por la compañía "WonderWorks" especialistas en tecnología multimedia y soluciones en Internet.

El sistema es compatible con televisores, videograbadoras, camcorders, reproductores de disco láser, decodificadores de satélite, televisión en cable, cámaras de seguridad, de hecho cualquier fuente que tenga salida NTSC o PAL.

Trabaja en Windows 3.1 o Windows 95, el software permite congelar una imagen y procesarla a colores con 24 bits. Se tiene un control de balance del color, contraste, brillo e igual para los colores rojo, verde y azul.

Cuenta con un sistema de edición que te permite guardar imágenes en formato BMP o TIFF en archivos con 16.8 millones de colores, cambiar la escala a las imágenes, voltearlas horizontal y verticalmente, invertir el color.

Puede almacenar video en formato AVI.

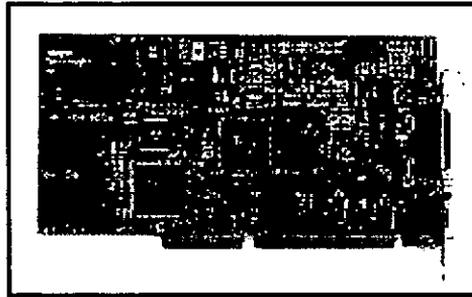


Fig. A.3 Motion Picture.

Especificaciones generales:

Captura de cuadros

- Tarjeta conectada al bus ISA.
- Reloj máximo de 10 MHz.
- Limitado a 16 bits.
- Segunda generación de DSP en la decodificación de color.
- 450 mA de consumo a 5 V.
- Captura en formato BMP y TIFF.
- Colores con 24 bits (hasta 16.8 millones).
- Archivo de mayor tamaño sin compresión 1.2 Mbytes.
- Frecuencia de captura máxima de 20 cuadros/seg.
- El tiempo de procesamiento depende del desempeño de la maquina.

Vídeo en formato AVI

- Captura de vídeo en formato AVI en tamaño de 192 x 129 y 384 x 258.
- Tamaño de los archivos de vídeo 370 kbytes/seg. y 1.4 Mbytes/seg. respectivamente.
- Despliega en 256 tonos de grises.
- Frecuencia de exposición 15 campos/seg. típicamente.

Requisitos mínimos:

- Procesador 386.
- Sistema operativo Windows3.1 o Windows 95.
- Controlador de gráficos de 256 colores en cualquier resolución.

El costo de esta tarjeta es de \$169.00 dólares en:

Wonderworks
11 Market Place
Alnwick
Northumberland
NE66 1HS
UK
phone: +44 1665 602868
fax: +44 1665 605911
email: sales@wonderworks.co.uk

A.4 AV2001

El AV2001 es un sistema de video conferencia a través de Internet fabricado por "ACE Vision", la tarjeta permite introducir un sistema de captura de vídeo en una PC. Es una tarjeta "Plug & Play", basada en el bus PCI y diseñada para cualquier fuente de vídeo compuesto. Sólo se conecta una cámara o una fuente de vídeo cualquiera y al instante se tiene listo el sistema para transmitir y recibir vídeo de cualquier parte del mundo.

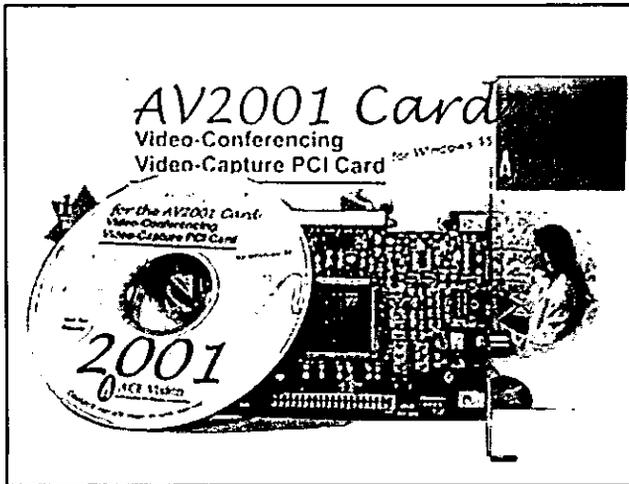


Fig. A.4 SistemaAV2001.

Características:

- Entrada de vídeo: NTSC, PAL o SECAM.
- Entrada de cámara, vídeo cámara, camcorder, etc.
- Captura a 30 cuadros/seg.
- Resolución de 758 x 576 pixeles.
- Soporta 16.8 millones de colores.

Especificaciones:

Hardware	Interface Plug & Play sobre el bus PCI.
Captura de cuadros	30 cuadros/seg. Para NTSC, 25 para PAL o SECAM.
Vídeo capturado	640 x 480 (NTSC) y 768 x 576 (PAL o SECAM), 16.8 millones de colores.
Entrada de vídeo	Vídeo compuesto (conector RCA).
Sistema RISC	+5 Volts para alimentar a la cámara.
Ajuste	Brillo, contraste, saturación del color, matiz.
Compatibilidad con VGA	En resoluciones arriba de 1600 x 1280, todos los colores se codifican con 24 bits.

Tabla A.1 Especificaciones AV2001.

Requisitos mínimos:

- Sistema Operativo Windows '95.
- Pentium 90 MHz tipo PC.
- Slot PCI.
- Fuente de vídeo compuesto.
- 100 Mbytes de espacio libre en disco.
- 16 Mbytes en RAM.
- CD Rom Drive.
- Tarjeta de sonido compatible con Windows.
- Tarjeta de vídeo PCI 1 Mbyte.

Esta tarjeta tiene un costo de \$175.00 dólares en: acevision@acevision.com

A.5 ULTRA II

El sistema de grabación de cuadros "Ultra II" fabricado por Coreco. Es el único habilitado con dos zonas de almacenamiento (*búfer*), el primero dedicado a operaciones de gráficos y el segundo reservado para la captura de imágenes. Corre sobre Windows con 8, 16 o 24 bits por pixel, independientemente de con cuantos colores o tonos de grises ha sido adquirida la imagen. Cuenta con un búfer especial de vídeo del cual puede desplegarse la imagen en escala de grises a 8 bits o en color a 16 o 24 bits.

La tarjeta trabaja y despliega sobre Windows simultáneamente. Además, cuenta con una salida entrelazada, para grabar imágenes directamente a una vídeo grabadora.



Fig. A.5 Sistema Ultra II.

Características:

- Adquisición y despliegue con 8, 16 y 24 bits.
- Formato de vídeo de entrada NTSC, PAL, RS-170 y CCIR.
- Adquisición RGB.
- Modulo de adquisición digital para cámaras de salida digital.
- Soporta resoluciones de 1024 x 768 sobre tarjetas VGA.
- Salida de vídeo para vídeo grabadora, NTSC y PAL.
- Trabaja en Windows '95, Windows 3.11 y Windows NT.
- Instancia en pantalla sencillo o doble.

Aplicaciones:

- Visión en maquinaria.
- Análisis de imágenes.
- Procesos de inspección y control.
- Visualización médica.
- Robótica.
- Análisis Biomédico.
- Seguridad.
- Imágenes científicas.
- Microscopía.

Interfacc a la computadora

Esta tarjeta trabaja sobre el bus PCI, lo cual incrementa su funcionalidad y facilita la instalación. Sobre la mismo sistema se tiene un tarjeta VGA y se cuenta con búfer de cuadros dedicados para garantizar la compatibilidad con VGA. El búfer de cuadro de vídeo tiene la suficiente capacidad para garantizar el fácil manejo a 30 HZ. El sistema puede manejar imágenes a 30 cuadros por segundo (NTSC) y 25 cuadros por segundo (PAL).

Adquisición

Cuenta con un desempeño superior, al incorporar un decodificador digital de color, ofrece colores puros y una alta relación señal a ruido.

Una variedad de formatos de entrada puede ser elegido, incluyendo monocromo en 8 bits, RGB (15 o 24 bits), YUV 4:1:1 y YUV 4:2:2.

Puede adquirir la imagen de una gran variedad de cámaras con estándar: NTSC, PAL, Y/C, RS-170 y CCIR.

Tipo de adquisición	Profundidad de pixel
RS-170	8 bits/pixel.
CCIR	8 bits/pixel.
NTSC	16 o 24 bits/pixel.
Y/C	16 o 24 bits/pixel.
PAL	16 o 24 bits/pixel.
RGB	24 bits/pixel.

Tabla A.1 Bits/pixel en los distintos formatos.

Especificaciones:

- Entrada de vídeo 1 Vpp, 75 ohms.
- Procesamiento sobre la tarjeta, el software carga las tablas de Clut permitiendo ajustes del vídeo.
- Rango de offset, programable por software, 0.7 volts.
- Exactitud de digitalización, 45 dB en la relación señal a ruido.
- Frecuencia de captura de 30 cuadros/seg. NTSC o 25 cuadros/seg. PAL.
- Profundidad de pixel 8, 16 y 24 bits.
- Decodificador de color para mejora la resolución en color y decrementa el arrastre de croma.
- Formato de entrada: NTSC, PAL, Y/C, RS-170 o CCIR (entrada opcional RGB).
- Resolución: 640 x 480 NTSC o 768 x 576 PAL.
- 1 o 2 Mbytes de memoria VGA o 1, 2 o 4 Mbytes de memoria de vídeo.
- Resolución sobre VGA 640 x 480, 800 x 600 o 1024 x 768.
- Formatos de salida: SVGA, RS-170, CCIR, NTSC, PAL o RGB.
- Transferencia sobre el bus PCI.

http://www.yahoo.com/Business_and_Economy/Companies/Computers/Hardware/Components/Video_Cards/Video_Capture/

A.6 RT300

La tarjeta de captura de vídeo "Video Blaster RT300" es una tarjeta de alto desempeño fabricada por "Creative Labs".

Emplea un procesador de vídeo, el i750 de intel. Puede elegir de 3 fuentes de vídeo posibles. Admite el formato NTSC, PAL y S-Vídeo. Así, ésta puede capturar vídeo a 30 campos por segundo con una resolución de 320 x 240 pixeles. Realiza compresión en tiempo real. Cuenta además con un software de edición.

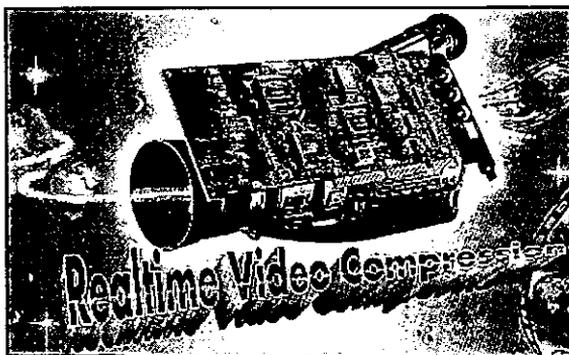


Fig. A.6 Vídeo Blaster RT300.

Características:

- Puede seleccionar de tres posibles fuentes de entrada de vídeo compuesto o de una entrada de S-Vídeo.
- Soporta los estándares NTSC, PAL y S-Vídeo.
- Procesador i750 sobre la tarjeta para alta calidad en la captura de vídeo.
- Captura de vídeo a 30 cuadros por segundo con resolución de 320 x 240 o 160 x 120.
- Reproducción de vídeo a 30 cuadros por segundo a resolución de 320 x 240.
- Incluye Adobe Premier para edición.

La velocidad de procesamiento para la captura de vídeo en formato AVI a disco duro depende de las características del disco, de la arquitectura del bus y del sistema del usuario.

Especificaciones:

- Trabaja en Windows 3.x y Windows 95.
- 3 entradas RCA hembra para la señal compuesta de vídeo.
- 1 conector S-Vídeo para la entrada de vídeo S-Vídeo.
- Salida hacia monitor VGA.

Requisitos mínimos:

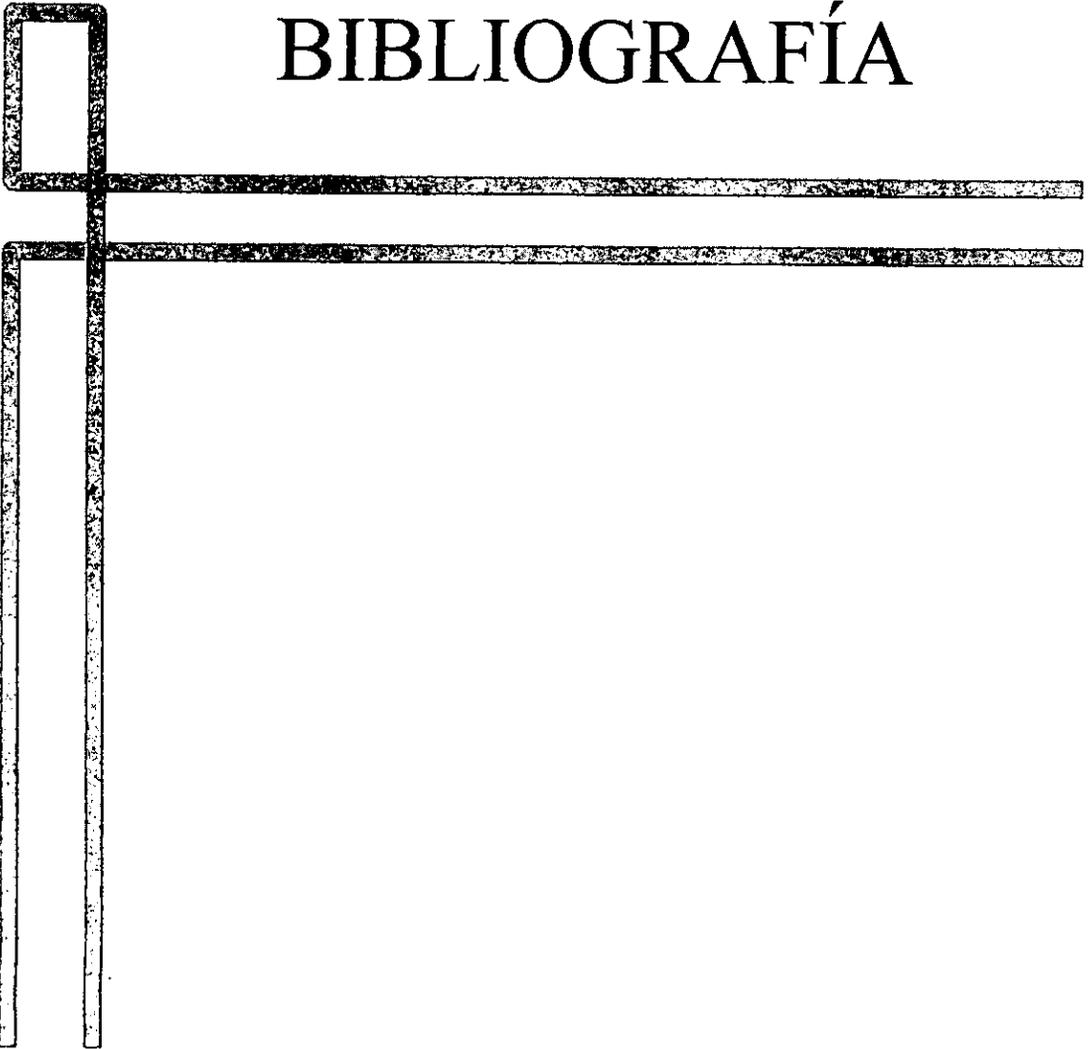
- Procesador 386DX a 33MHz .
- 4 Mb en RAM.
- 1 slot ISA de 16 bits.

- MS-DOS 3.1 o superior.
- Windows 3.1 o superior.

Esta tarjeta tiene un costo en el mercado de \$199 dolares:
http://www.cle.creaf.com/germany/products/video/rt_tech.html

Creative Labs.
Genetec S.A. de C.V.
telefono 225 07 20
fax 255 09 56.

BIBLIOGRAFÍA



- Benson, K. B., and Whitaker, J. C.
"Television Engineering Handbook"
New York, McGraw-Hill, 1992.
- Benson, K. B., and Whitaker, J. C.
"Television and Audio Handbook"
U.S.A., McGraw-Hill, 1988
- Saintjit K. Mitra - James F. Kaiser
"Handbook for Digital Signal Processing"
U.S.A., John Wiley & sons, 1993.
- Luther, Arch C.
"Digital Video in the PC Environment"
New York, McGraw-Hill, 1991
- Luther, Arch C.
"Principles of Digital Audio and Video"
U.S.A., Artch House Inc., 1997
- Inglis, Andrew F.
"Video Engineering"
Second Edition, U.S.A., McGraw-Hill, 1996.
- Boris Bertelson & Mathias Rasch
"PC al Limite"
España, Computec Marcombo: Alfaomega, 1995.
- Barry B. Brey
"Los Microprocesadores Intel"
Tercera Edición, México, Prentice Hall Hispanoamericana, 1995
- Herbert Schildt
"Turbo C/C++ 3.1 Manual de Referencia"
España, Osborne/McGraw-Hill, 1994
- Motorola
"Analog/Interface ICs Device Data"
DL128/D REV 6 Vol 1
U.S.A., Motorola, 1996

- Majid Rabbani, and Paul W. Jones
"Digital Image Compression Techniques"
U.S.A., Spie Optical Engineering Press, 1991
- Y. Mahdavih, and R.C. Gonzalez
"Advances in Image Analysis"
U.S.A., Spie Optical Engineering Press, 1992.
- Michael Tischer
"PC INTERNO (2.0)"
España, Computec MARCOMBO,1995.
- Robert Boylestad - Louis Nashelsky
"Electrónica, Teoría de Circuitos"
cuarta edición, México, Prentice Hall Hispanoamericana s.a., 1989.
- National Semiconductor
"National Application Specific Analog Products Databook"
National Semiconductor 1995.
- Denton J. Dailey
"Operational Amplifiers and Linear Integrated Circuits, Theory and Applications"
Singapore, McGraw-Hill, 1989
- Professeur Joseph Ronsin
"Image Compression 1"
Institut National des Sciences Appliquées
35043 Rennes cedex- France
- Publicaciones e Internet