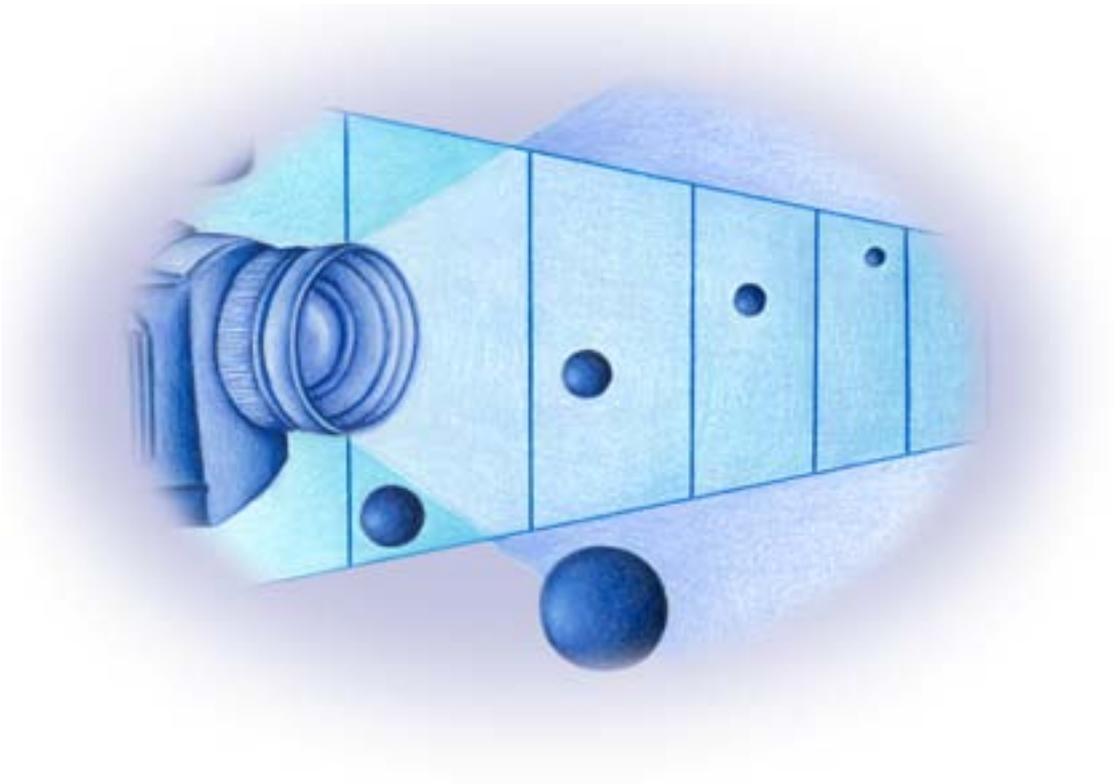


Conversores y normas



Índice

Etapa 1: Digitalización de la señal de vídeo	5
Introducción	5
Digitalización de señales de vídeo analógicas	5
Módulo de acondicionamiento de señales	6
Número de líneas de una señal digital procedente del formato PAL.....	7
Muestreo de la señal de iluminación PAL	8
Número de líneas de señales procedentes del estándar NTSC	8
Las imágenes como matrices.....	8
La señal de vídeo como secuencia de matrices	9
Formatos digitales: el estándar ITU-601	9
Formatos digitales de píxel cuadrado	11
Submuestreo de la información de color	12
Otras relaciones de submuestreo	14
Formatos digitales reducidos	15
Formatos digitales de alta definición	17
Cuantificación de la señal de vídeo	17
Cuantificadores uniformes.....	18
Codificación PCM	19
Conversión de las palabras código a muestras analógicas	21
Selección del número de bits	22
Codificación de imágenes en color	22
Ancho de banda de la señal de vídeo digital	23
Etapa 2: Tarjetas y sistemas de digitalización	25
Introducción	25
Tarjetas digitalizadoras y sintonizadoras de señal de televisión	26
Principios de funcionamiento del sistema de digitalización de vídeo en el entorno Windows	28
Etapa 3: Interfaces de vídeo digital	31
Introducción	31
Serial digital interface (SMPTE 259M)	31
Estándar IEEE 1394	32

Etapa 1: Digitalización de la señal de vídeo

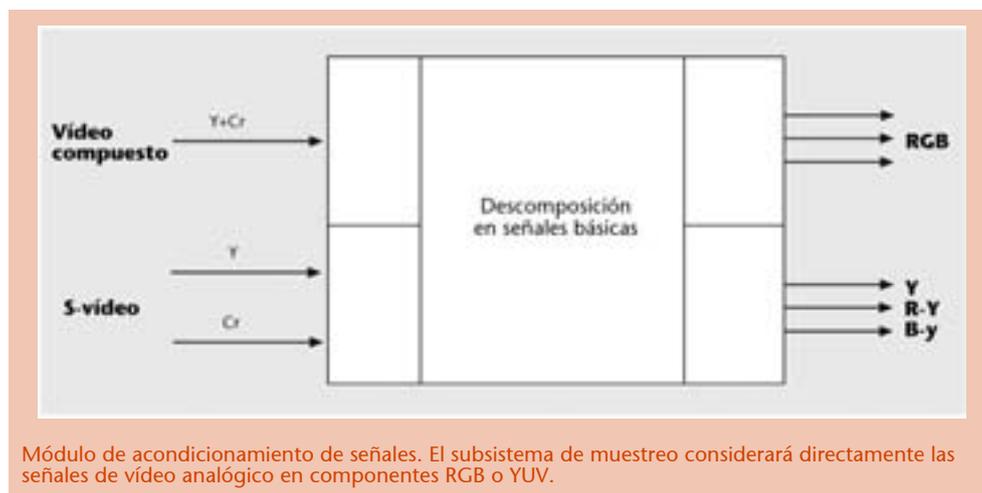
Introducción

Cada vez es más habitual la presencia de sistemas puramente digitales en cualquier tipo de instalación de vídeo profesional o doméstico. La señal digital aporta considerables ventajas con respecto a su contrapartida analógica, tanto desde el punto de vista de inmunidad al ruido, transmisión y almacenamiento más fiables, posibilidad de corregir errores producidos durante la transmisión o el almacenamiento, como desde el de la facilidad para manipular y acceder a los contenidos almacenados en forma digital.

No obstante, el punto de partida de la mayor parte de los grandes sistemas de vídeo todavía sigue siendo una señal analógica proporcionada por la cámara. En los últimos tiempos se están implantando algunas cámaras que proporcionan señales directamente digitales, pero su uso se limita por el momento a sistemas de inspección industrial, *webcams* de baja calidad o aplicaciones de vídeo profesional muy concretas. La mayoría de los sistemas digitales existentes en los estudios y productoras de vídeo todavía parten de un cuerpo de cámara que proporciona las señales en formato analógico. Asimismo, tened en cuenta que aunque las *camcorder* digitales (DV, miniDV, DVCPRO) almacenan las señales en formato digital, realizan una conversión interna del formato analógico que proporciona el cuerpo de la cámara (óptica, sensor CCD y unidad de control) al formato digital de almacenamiento.

En consecuencia, es importante considerar con cierto detalle el proceso de conversión de las señales analógicas a los formatos de vídeo digital, teniendo en cuenta todas sus limitaciones y peculiaridades.

Digitalización de señales de vídeo analógicas



Módulo de acondicionamiento de señales. El subsistema de muestreo considerará directamente las señales de vídeo analógico en componentes RGB o YUV.

La digitalización de cualquier tipo de señal exige tener en cuenta dos parámetros básicos:

- El número de muestras necesario para que la digitalización represente toda la información útil existente en la señal (teorema del muestreo).
- El número de bits con el que hay que codificar cada muestra.

Considerando el problema de la señal de vídeo en color, la información que deseamos convertir en digital puede provenir de distintos formatos. En el formato de vídeo compuesto, la información de luminancia y la de color son proporcionadas de forma conjunta mediante una única señal que hay que decodificar para separar los tres componentes de color. En el formato S-vídeo las señales de croma y de luminancia se proporcionan mediante cables separados, pero también es necesario procesar la información de color para separar cada uno de sus componentes. En las cámaras profesionales, las señales son proporcionadas por tres canales independientes que facilitan las informaciones de los tres componentes de color. En algunos sistemas se pueden proporcionar las señales de luminancia y diferencia de color (YUV) por separado.

En cualquier caso, resulta evidente que en el proceso de conversión de la señal de vídeo a formatos digitales deberemos tener en cuenta los tres componentes de color, ya sea en un formato directo (como el RGB), ya en uno indirecto (como el YUV).

Módulo de acondicionamiento de señales

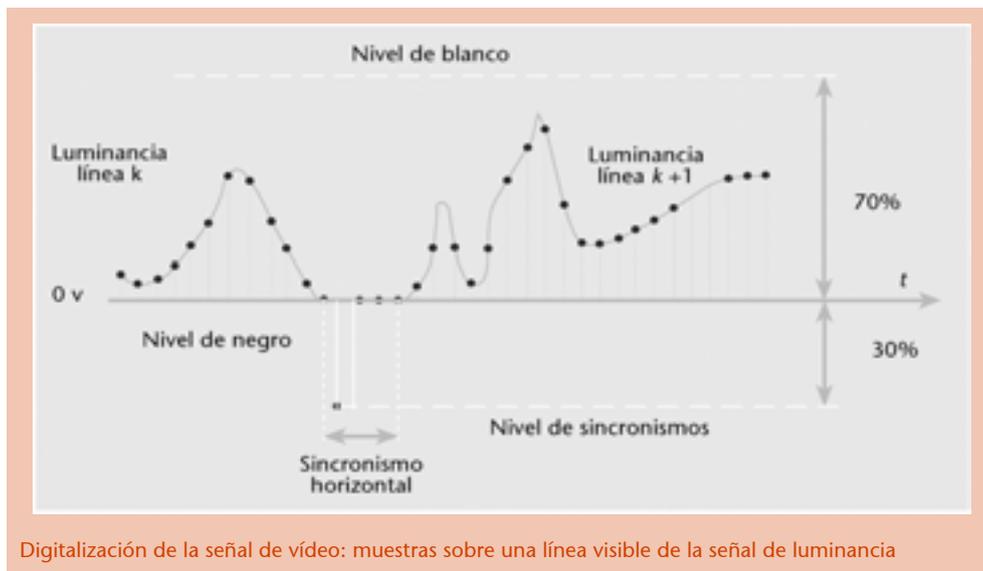
Será necesario que dispongamos de un módulo de acondicionamiento de señales previo que lleve a cabo la decodificación de las señales de vídeo y nos proporcione tres señales independientes.

Las señales que nos proporcione el módulo de acondicionamiento serán digitalizadas posteriormente. En la figura hallamos una representación conceptual del módulo de acondicionamiento; para el caso particular que las señales de vídeo, se proporcionan en S-vídeo o vídeo compuesto. Los casos en que las señales analógicas se representan directamente en componentes RGB o YUV son, con toda evidencia, directos.

En el caso de señales en color, lo más habitual es tratar la señal analógica original mediante sistemas analógicos hasta descomponerla en tres señales independientes que, en conjunto, aportan la información referente al color. Las dos descomposiciones más utilizadas son la R, G y B y la Y, U y V. Una vez realizada la descomposición en tres señales independientes, procederemos al muestreo de cada una de ellas de forma independiente (como si se tratase de señales en blanco y negro). La información asociada a un determinado elemento de imagen deberá tener en cuenta estos tres componentes.

El uso de una u otra descomposición depende, en gran medida, de la aplicación final a la que se destina el sistema digital. En sistemas de visión artificial se suele utilizar una descomposición RGB para obtener un mejor rendimiento del espacio de color de la imagen original, mientras que en sistemas de vídeo, es habitual el uso de una codificación YUV que exige un menor ancho de banda.

Número de líneas de una señal digital procedente del formato PAL



Una vez separada la señal de vídeo en sus tres componentes básicos podemos tratar el proceso del muestreo de cada una de ellas por separado. Si consideramos la señal proporcionada por una cámara PAL, obtendremos una frecuencia de imagen de 25 Hz (25 imágenes por segundo), con dos campos entrelazados y un total de 625 líneas por imagen.

De estas 625 líneas sólo son visibles 575 (287,5 por cada campo); el resto de líneas (25 líneas por cada campo) se utilizan para enviar las señales de sincronismo de campo y algunas señales auxiliares como el teletexto, señales para el control de calidad de la transmisión, etc. Por lo tanto, un sistema digital que tenga que trabajar con señales procedentes del formato PAL, debe ser capaz de digitalizar todas las líneas visibles, de manera que no se produzca una pérdida de información relevante.

En la práctica, como cada campo está formado por 287 líneas más media línea, se toman muestras en las 288 líneas (tomando como valores negros la mitad de la línea final o inicial del campo).

La señal digital estará formada por un total de $2 \times 288 = 576$ líneas que representan la información contenida en las líneas visibles de la señal PAL original.

Muestreo de la señal de iluminación PAL

En la figura se presenta el proceso de muestreo de uno de los componentes de la señal PAL, la luminancia. Las muestras que se toman sobre cada una de las líneas de la señal determinarán la resolución horizontal del sistema digital. Como es evidente, no tiene demasiado sentido que la resolución horizontal sea mucho mayor o menor que la resolución vertical, puesto que el sistema visual humano posee las mismas características de agudeza visual en ambos sentidos. En consecuencia, el número de muestras que tomemos para una única línea deberá ser proporcional, de manera aproximada, al número de líneas de la imagen. La relación de proporcionalidad vendrá determinada por la relación de aspecto de la señal de televisión que en sistemas convencionales es de 4:3 (anchura : altura).

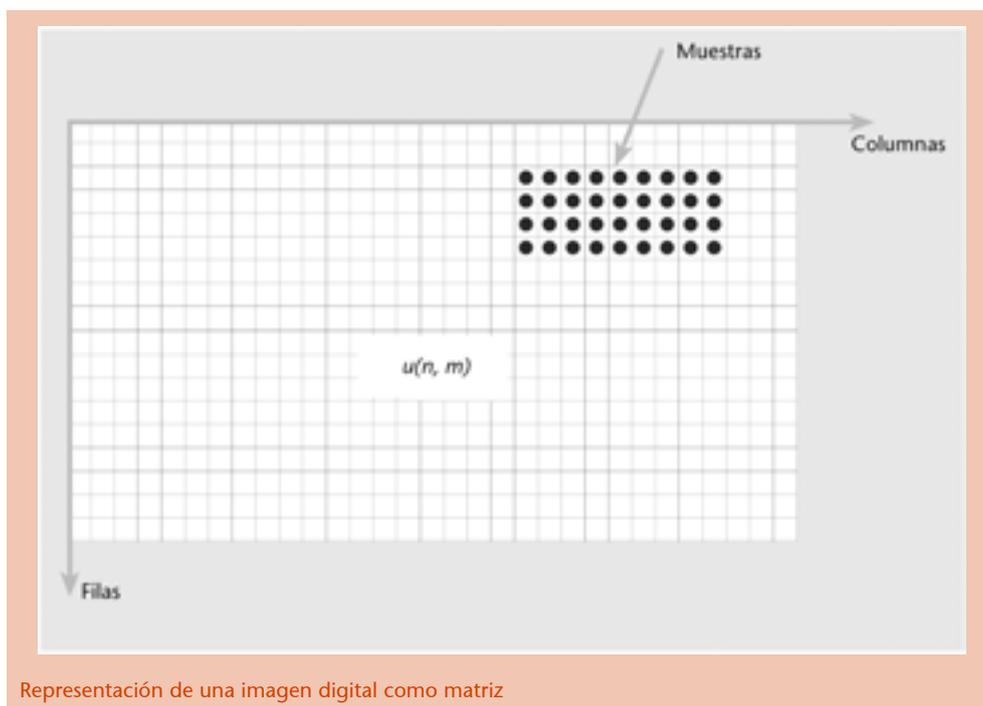
Número de líneas de señales procedentes del estándar NTSC

El estándar NTSC se utiliza principalmente en EE.UU. y Japón, y cada imagen está formada por un total de 525 líneas de las cuales sólo 479 resultan visibles. Las imágenes se transmiten a una frecuencia de 30 Hz y también se encuentran divididas en 2 campos.

El número de líneas que resultan al muestrear completamente la información contenida en una señal NTSC es de 480.

En este caso la mitad de línea de cada campo se muestrea completa. La diferencia observada con respecto al formato PAL es importante y justifica, como veremos, la existencia de algunos formatos utilizados en aplicaciones informáticas.

Las imágenes como matrices



Representación de una imagen digital como matriz

La digitalización de una imagen da lugar a un conjunto de muestras que se pueden ordenar de forma matricial. En cada fila y columna de la matriz hemos obtenido el valor de luminancia (o cualquier otro componente) de la imagen.

Las filas de esta matriz se corresponden con las líneas visibles de la señal de televisión, mientras que las columnas se asocian a las muestras que hemos tomado en cada línea. Es importante mantener un buen sincronismo entre los impulsos de inicio de línea y campo con la señal de muestreo, a fin de que las muestras se correspondan con las de una retícula rectangular como la que se ilustra en la figura. Esta matriz correspondería a los valores numéricos de la luminancia para un fotograma determinado; para los componentes U y V dispondremos de matrices con características parecidas.

La señal de vídeo como secuencia de matrices

La digitalización de toda la secuencia de vídeo genera, a su vez, una secuencia de matrices que deberemos almacenar, transmitir o procesar. Los equipos dedicados a la codificación, transmisión o almacenamiento de vídeo digital suelen disponer de un *buffer* de memoria en el que podemos almacenar las últimas imágenes digitalizadas. Estas imágenes se procesan y con posterioridad se transmiten o almacenan en un soporte masivo de forma continua, con lo que se libera la memoria para la incorporación de las nuevas imágenes. El flujo medio de salida de datos del *buffer* de memoria tendrá que ser superior al flujo de entrada, para, de esta manera, no rebasar la capacidad del sistema durante el funcionamiento continuo. Al disponer de un conjunto de imágenes correspondientes a distintos instantes de tiempo, es posible realizar operaciones de compresión de datos o de tratamiento de imágenes que tengan en cuenta no sólo las características espaciales de cada imagen, sino también la evolución temporal de la escena.

Formatos digitales: el estándar ITU-601

Para finalizar con el proceso de digitalización de la señal de vídeo nos disponemos a definir el número de muestras que se toman para cada una de las líneas. En el caso de señales de televisión digitales con definición convencional existe un estándar aceptado internacionalmente que se conoce como el ITU-601.

Según el estándar ITU-601, el número de muestras visibles en cada línea de la imagen de televisión es de 720.

Este número de muestras es común tanto para el formato de televisión europeo (PAL), como para el americano o japonés (NTSC) y se obtiene mediante el uso de una frecuencia de muestreo de 13,5 MHz.

En 1982, el ITU-R (anteriormente CCIR), en su recomendación 601, desarrolló un conjunto de especificaciones para señales de televisión digital en estudios de televisión o de producción de vídeo. El objetivo de esta recomendación era facilitar el intercambio de programas a escala internacional. Las recomendaciones definen una gran cantidad de parámetros comunes entre el formato americano y el europeo con el objetivo de que los fabricantes puedan incluir varios módulos comunes en equipos para 525 y 625 líneas. El uso de este formato permite la interconexión entre distintos equipos digitales.

Existen, básicamente, dos variantes principales conocidas como 4:2:2 y 4:4:4, en las que los dígitos indican la proporción entre las muestras dedicadas a la luminancia y las dedicadas a los componentes de color. El sistema más utilizado es el 4:2:2, donde se encuentran codificadas la luminancia y las señales de diferencia de color; el formato 4:4:4 también se puede utilizar con componentes RGB. En la siguiente tabla se detallan las características más significativas del formato 4:4:4, que proporciona la base de las especificaciones del ITU-601. Aunque en estas tablas se proporcionan los datos del número de niveles de cuantificación con el que se codifica cada elemento de imagen, consideraremos con más detalle el proceso de cuantificación en los siguientes apartados.

Tabla resumen del estándar digital ITU-601 4:4:4		
ITU R-601	Parámetros de codificación en componentes 4:4:4	
Parámetros	Sistemas 525/60	Sistemas 625/50
Señales codificadas: Y, CR, CB o R, G, B.	Las señales se denotan como EY, ER-EY, EB-EB o como ER, EG, EB. Su relación con los componentes RGB del sistema NTSC aparece comentada en el texto.	
Número de muestras por línea completa de cada señal.	858	864
Estructura de muestreo.	Retícula de muestreo ortogonal con repetición sobre la base de la línea, el campo y el cuadro. La estructura de muestreo hace que los tres componentes coincidan y que también coincidan con las posiciones de las muestras de la luminancia en el sistema 4:2:2.	
Frecuencia de muestreo para cada señal.	13,5 MHz	
Tipo de codificación.	Cuantificador uniforme de un mínimo de 8 bits tipo PCM (8 o 12).	
Duración de la línea activa expresada en número de muestras	Mínimo de 720	
Correspondencia entre los niveles de señal de vídeo con los 8 bits más significativos:	<p>0-255 220 niveles de cuantificación, en los que el nivel de negro se corresponde con el nivel 16 y el nivel de blanco, con el 235. El nivel de la señal puede superar, en algunas ocasiones, el nivel 235.</p> <p>225 niveles de cuantificación situados en la parte central de la escala del cuantificador con el nivel de señal cero asociado al valor 128.</p>	
<ul style="list-style-type: none"> ● escala; ● R, G, B o señal de luminancia; ● señales de diferencia de color. 		

Tabla resumen del estándar digital ITU-601 4:4:4

De acuerdo con este estándar, el número de filas y columnas en las que se descompone una imagen de televisión digital en el sistema PAL es de 576×720 (filas \times columnas). En el sistema NTSC, el número de píxeles de la imagen de televisión digital es de 480×720 (filas \times columnas).

Formatos digitales de píxel cuadrado

El formato ITU-601 está ampliamente extendido en sistemas de vídeo y televisión digital (MPEG-2, DVD, etc.) No obstante, en algunas aplicaciones concretas (por ejemplo, almacenamiento de vídeo en disco duro, grabación de vídeo digital para edición, captura de imágenes para sistemas de visión artificial, etc.) resulta habitual el uso de ligeras modificaciones del formato debido a que la relación de aspecto del elemento de imagen en el estándar ITU-601 no es cuadrada. Y, con esto, queremos decir que la densidad de píxeles en el eje vertical no es idéntica a la del eje horizontal.

En efecto, como la relación de aspecto de la televisión convencional es 4:3 (la relación entre la anchura y la altura de la pantalla), la densidad de píxeles en el eje vertical $576/3 = 192$ y la del eje horizontal $720/4 = 180$ no coinciden (en el sistema PAL existe una densidad mayor de filas que de columnas, mientras que en el NTSC –podemos comprobarlo– sucede lo contrario).

El problema será significativo en aquellos sistemas que deban llevar a cabo operaciones matemáticas sobre los elementos de imagen (por ejemplo, sistemas de reconocimiento de imagen), en los que, para facilitar la simetría de los tratamientos numéricos en el eje vertical y horizontal, es conveniente que las densidades de ambos sean idénticas. Por este motivo, del ITU-601 se derivan los denominados formatos digitales de píxel cuadrado que se obtienen modificando el número de columnas para forzar esta relación cuadrada.

En el caso del PAL, el formato de píxel cuadrado será de 576×768 (filas por columnas) ya que 768 es el resultado de multiplicar 576 por la relación de aspecto 4/3. Tened en cuenta que para forzar el hecho de que el elemento de imagen sea cuadrado debemos modificar el número de columnas y no el de filas, ya que éste viene predeterminado por el número de líneas que se han definido para el estándar analógico. En el NTSC, utilizando las mismas relaciones algebraicas, obtenemos 480×640 (filas por columnas). Notad que el formato de píxel cuadrado en el NTSC coincide con el formato VGA definido en aplicaciones informáticas. Así pues, éste será un claro ejemplo de la conveniencia de utilizar formatos de píxel cuadrado en sistemas en los que sea susceptible realizar un tratamiento de la imagen.

Resulta evidente que si modificamos el número de columnas de la imagen digital, la frecuencia de muestreo no se mantiene. Podemos calcular fácilmente cuál es la nueva frecuencia de muestreo para el sistema derivado del PAL, ya que sólo tendremos que considerar que el tiempo de línea activa es de 52 ms, en los que se toman un total

de 768 muestras. El periodo de tiempo entre dos muestras consecutivas será de $52/768 = 0,068$ ms, lo que da lugar a una frecuencia de muestreo de 14,769 MHz.

Submuestreo de la información de color



En el formato que hemos definido hasta ahora, los tres componentes R, G y B o las Y, U y V se muestrean con las mismas frecuencias, obteniendo una secuencia de matrices del mismo tamaño para cada uno de los componentes. Este estándar (4:4:4 del ITU-601) se utiliza, sobre todo, en aplicaciones en las que deban intervenir los componentes R, G y B.

No obstante, en aplicaciones de televisión también podemos sacar algún provecho de las características del sistema visual, el cual tiene una menor resolución espacial a los componentes de color que a los de luminancia. Podemos sacar partido a esta circunstancia y reducir el flujo de datos asociado a los componentes U y V, sin que por ello se produzca una pérdida de calidad apreciable. La reducción del flujo de datos se suele obtener descartando algunas de las muestras de los componentes U y V siguiendo un patrón de submuestreo predeterminado.

Los formatos más utilizados son el 4:2:2 y el 4:2:0. El 4:2:2 es el formato que se usa en los magnetoscopios profesionales tipo Betacam Digital, DV, DV-Cam, DVC-Pro, etc. El patrón de submuestreo para los componentes U y V consiste, tal y como se ilustra en la figura que vemos al principio de este subapartado, en no almacenar la información de los componentes diferencia de color correspondientes a las columnas pares. De esta manera, el número de muestras que tendremos que almacenar para cada uno de estos componentes se reduce a la mitad.

A pesar de que los datos de croma que se almacenan o transmiten en el formato 4:2:2 se encuentran submuestreados, los dispositivos finales para la representación gráfica deben recomponer las muestras que faltan para que cada elemento posea la informa-

ción de color completa que le corresponde. El cálculo de las muestras de croma que no se han codificado suele realizarse mediante una interpolación lineal (promedio entre las dos muestras adyacentes), lo que significa un ligero incremento en la carga computacional del subsistema de representación gráfica. Este coste adicional se ve compensado por la reducción en el volumen de información. En la siguiente tabla hallaremos un resumen de las características del formato 4:2:2.

Tabla resumen de las especificaciones del formato 4:2:2		
ITU-601	Parámetros de codificación en componentes 4:2:2	
Parámetros	Sistemas 525/60 Sistemas 625/50	
Señales codificadas: Y, CR,CB.	Las coordenadas utilizadas se denotan EY, ER-Y y EB-Y. Su relación con las coordenadas del sistema NTSC es discutida en el texto.	
Número de muestras por línea completa: ● luminancia, ● señales diferencia de color.	858 429	864 432
Estructura de muestreo.	La retícula de muestreo es ortogonal y se repite según un patrón de línea, campo y cuadro. Las muestras de los componentes diferencia de color se corresponden con las muestras impares de la luminancia (1, 3, 5, etc.).	
Frecuencias de muestreo: ● luminancia, ● señales diferencia de color.	13,5 MHz 6,75 MHz La tolerancia para la frecuencia de muestreo debe coincidir con la tolerancia para la frecuencia de línea del estándar de televisión en color analógico de partida. Las frecuencias de muestreo de 13,5 MHz y 6,75 MHz son múltiplos enteros de 2,25 MHz, que es el mínimo común múltiplo de las frecuencias de línea de los sistemas de 525 y 625 líneas.	
Codificación de las muestras.	Se utiliza un cuantificador uniforme de 8 bits por muestra del tipo PCM, tanto para la señal de luminancia como para las señales diferencia de color.	
Número de muestras digitales por línea activa: ● luminancia, ● señales diferencia de color.	720 360	
Relación de tiempo horizontal de las señales analógica y digital: ● desde el fin de la línea activa a OH.	16 periodos de reloj de luminancia. 12 periodos de reloj de luminancia.	
Correspondencia entre los niveles de la señal de vídeo y los niveles de cuantificación: ● escala, ● señal de luminancia, ● señales diferencia de color.	0 a 255 220 niveles de cuantificación, donde el nivel de negro se corresponde con el nivel 16 y el nivel de blanco, con el 235. El nivel de la señal puede superar, en algunas ocasiones, el nivel 235. 225 niveles de cuantificación situados en la parte central de la escala del cuantificador con el nivel de señal cero asociado al valor 128.	
Uso de palabras código.	Las palabras código corresponden a los niveles 0 y 255 y se utilizan únicamente con propósitos de sincronización. Los niveles 1 a 254 se encuentran disponibles para la señal de vídeo.	

Tabla resumen de las especificaciones del formato 4:2:2

El formato 4:2:0 introduce una reducción adicional que en principio parece totalmente lógica. En efecto, si una disminución del número de columnas de la imagen no afecta a su calidad, también será posible reducir el número de filas sin que la imagen se degrade. Este resultado intuitivo es totalmente correcto y se aprecian pocas diferencias entre una imagen codificada con los formatos 4:2:2 y el 4:2:0, cuya distribución espacial de las muestras de luminancia y croma aparece representado en la figura adjunta.



En esta gráfica las muestras de croma se representan ligeramente desplazadas (en el sentido vertical) con las de luminancia, hecho que se debe a que las muestras de croma se obtienen a partir de las muestras del formato 4:2:2, promediando dos filas consecutivas. En este formato, la reducción del número de muestras de cada componente de croma es de un factor 4 con respecto al formato 4:4:4. El ámbito de aplicación del 4:2:0 es muy amplio e incluye la transmisión de señales de TV digital con MPEG-2, formatos multimedia MPEG1, MPEG4, codificación de vídeo para Internet, etc. En algunos de estos casos, el tamaño de las imágenes transmitidas es menor que el especificado por el ITU-601, aunque el formato de submuestreo entre luminancia y croma sigue manteniéndose.

Otras relaciones de submuestreo

Además del formato 4:2:2 y el 4:2:0, también están definidos otros formatos de submuestreo de los componentes de croma como el 4:1:1 y el 4:1:0, que a menudo se confunden con los primeros. En la tabla adjunta proporcionamos los tamaños de las imágenes utilizados en cada uno de ellos y los factores de diezmado de la señal de croma utilizados en cada caso. El formato 4:1:0 fue utilizado durante bastante tiempo por Intel en el DVI (*digital video interactive*).

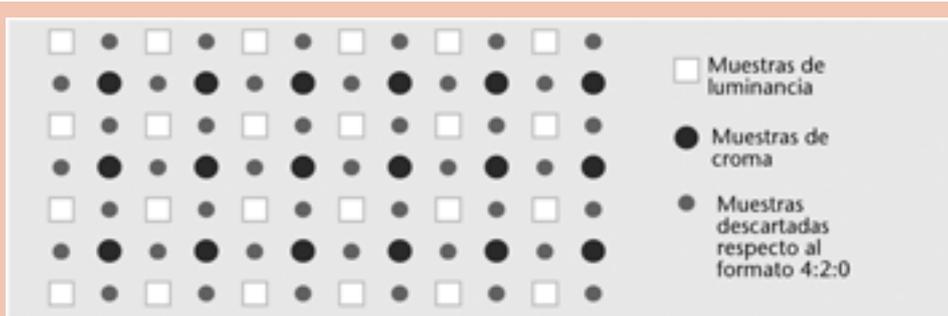
También existe una normativa definida para los formatos 3:2:2, 3:1:1 y 3:1:0, pero prácticamente no se han utilizado. Además de todas estas relaciones de submuestreo, contamos con la presencia de otras relaciones propietarias que surgen, sobre

todo, en el entorno de aplicaciones informáticas para la digitalización de secuencias de vídeo.

Formato	Muestras Y línea	Muestras Y columna	Muestras C línea	Muestras C columna	Submuestreo horizontal	Submuestreo vertical
4:4:4	720	576/480	720	576/480	---	---
4:2:2	720	576/480	360	576/480	2:1	---
4:2:0	720	576/480	360	288/240	2:1	2:1
4:1:1	720	576/480	180	576/480	4:1	---
4:1:0	720	576/480	180	144/120	4:1	4:1

Relaciones de submuestreo de los componentes croma más importantes

Formatos digitales reducidos



Disposición de las muestras en el formato SIF con respecto al 4:2:0

En algunas aplicaciones en las que no resulta necesaria una excesiva calidad de la imagen de vídeo digital, es habitual el uso de reducciones sobre el tamaño de la imagen, con lo que se consigue una importante reducción con respecto a la tasa de bits original, que puede hacer factible la codificación de la señal en soportes de baja densidad o canales de reducido ancho de banda. Un ejemplo típico de aplicaciones que utilizan formatos reducidos son los ficheros de vídeo para Windows AVI. Otro ejemplo es el MPEG-1, diseñado para codificar señal de vídeo sobre un soporte CD-ROM, aunque en principio fue diseñado sólo para almacenar información de audio. Todas estas aplicaciones suelen partir de una reducción del tamaño de las imágenes y, además, acostumbran a aplicar codificadores que comprimen la información de vídeo.

Uno de los formatos de vídeo reducido más populares es el SIF (*source intermediate format*) que, en esencia, consiste en un submuestreo de un factor 2 del formato 4:2:0 que se aplica tanto a los componentes de luminancia como a los de croma. Los tamaños de las matrices de luminancia para el formato SIF son de 360×288 , en el estándar de 625 líneas y 360×240 , en el de 525 líneas. Las matrices de croma también

se submuestran en un factor 2 en cada dirección respecto de las matrices de croma del 4:2:0 (180×144 para 625 líneas y 180×120 para 525). Asimismo, se realiza una reducción de la frecuencia de imagen a 25 Hz para el sistema europeo y a 30 Hz, para el americano. Tras haber realizado esta serie de operaciones, las imágenes resultantes no son entrelazadas; en estas condiciones, obtenemos una calidad equivalente al formato de vídeo analógico VHS.

Podemos obtener las muestras espaciales del formato SIF aplicando filtros espaciales y temporales sobre el formato 4:2:0, para eliminar los componentes de frecuencia elevada y, con posterioridad, submuestreando los componentes de luminancia y de croma resultantes.

La posición relativa entre los componentes de luminancia y croma, junto con los componentes descartados aparecen representados en la figura anterior. La posición de las muestras de croma con respecto a las de luminancia se debe al uso de los puntos centrales de las señales filtradas. De modo que, obtenemos una muestra de croma centrada en cada cuatro muestras de luminancia, que se utilizará para recomponer todas los componentes de color (RGB) de la imagen.

El formato CIF (*common intermediate format*) es un compromiso entre el formato SIF para 625 y 525 líneas, que Utiliza 360×288 muestras de resolución de luminancia (europeo) y una frecuencia de refresco de 30 Hz (americano).

Los formatos QSIF y QCIF (*quater*) se obtienen reduciendo, de nuevo, la resolución espacial en un factor 4 (factor 2 en cada dirección) y la resolución temporal en un factor de 2 o 4. Estos formatos se suelen utilizar para la transmisión de señales de vídeo telefonía con el estándar de compresión H261 o para la transmisión de vídeo en directo por Internet. Los submuestreos espaciales se realizan filtrando las señales SIF (o directamente la 4:2:0). Las posiciones espaciales de los componentes de luminancia y croma resultantes son parecidas a las del formato SIF.

En todos los formatos considerados, el número de líneas de la imagen es un submúltiplo del número de líneas de los formatos de televisión analógicos (NTSC o PAL), lo cual se debe a que con esta restricción el subsistema de muestreo podemos actuar directamente muestreando las líneas de la señal analógica asociada. Así, una imagen con 288 filas en el formato PAL se puede obtener muestreando las líneas de un único campo de la imagen analógica y descartando las muestras obtenidas para el otro campo. De manera análoga, obtendremos una imagen de 144 líneas mediante el muestreo las líneas de uno de los campos de forma alternada (una línea sí y otra no).

Como es evidente, existen tarjetas digitalizadoras que pueden proporcionar otros tamaños de imagen que no corresponden directamente a ninguno de los anteriores. En estos casos, la propia tarjeta digitalizadora o sus *drivers* de *software* asocia-

dos, calculan de forma sistemática las líneas de la imagen de las que se toman muestras o redimensionan las matrices resultantes a los tamaños especificados por el usuario.

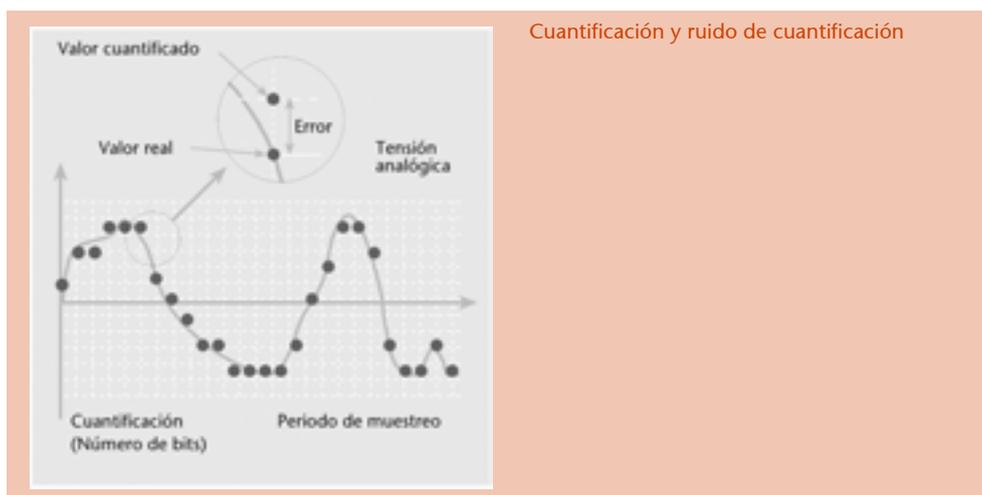
	Europeo TV	Americano TV	Europeo cuadrado	Americano cuadrado
4:2:0	720 x 576	720 x 480	768 x 576	640 x 480
SIF	360 x 288	360 x 240	384 x 288	320 x 240
CIF	360 x 288	360 x 288	384 x 288	384 x 288
QSIF	180 x 144	180 x 120	192 x 144	160 x 120

Tabla comparativa de los formatos digitales más utilizados

Formatos digitales de alta definición

También están definidos los formatos de alta definición digitales, que consisten, básicamente, en doblar el número de píxeles en cada dirección. En el estándar europeo, el tamaño de imagen de alta definición está definido en 1440×1152 , para relaciones de aspecto de 4:3, y en 1920×1152 , para relaciones de aspecto panorámicas.

Cuantificación de la señal de vídeo

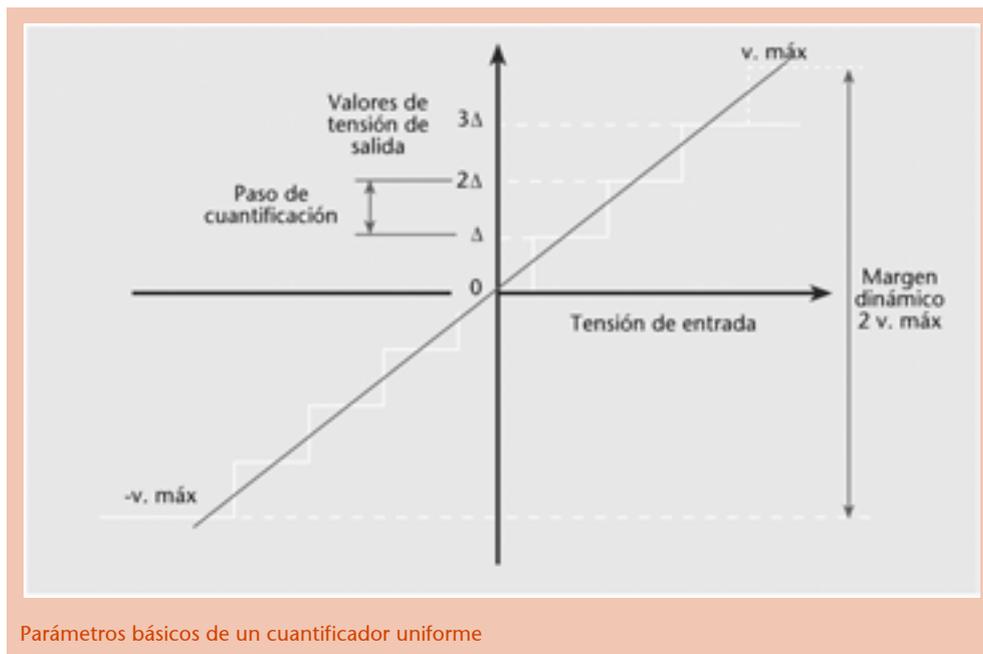


El paso de cuantificación define la precisión con la que se codifican las muestras de la señal y está directamente relacionado con el número de bits que se le ha asignado a cada muestra. La figura anterior muestra el proceso de discretización de la amplitud de la señal e indica cómo se introduce un error entre el valor real de la señal analógica y el valor con que se codificará la muestra una vez digitalizada.

El proceso de digitalización introduce, por tanto, un error aleatorio en la amplitud de la señal que es equivalente a la adición de un componente de ruido. En efecto, podemos suponer que el valor cuantificado corresponde al de la señal original más un ruido “virtual”, que se ha superpuesto con la señal, dando lugar al valor que en realidad adquiriremos. El ruido puede ser tanto positivo como negativo y su valor máximo es igual a la mitad del paso de cuantificación.

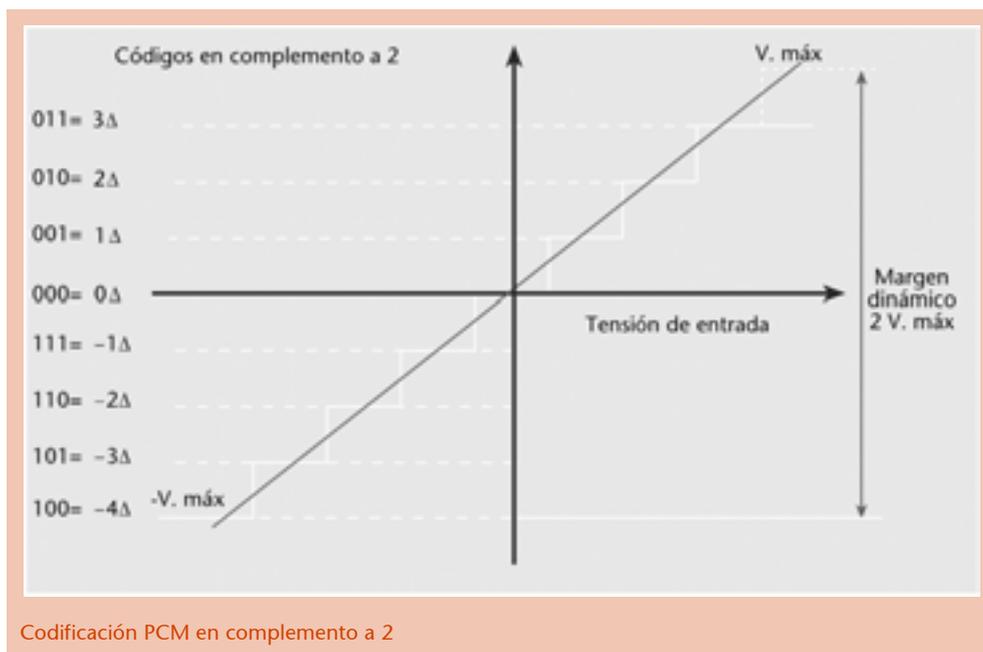
En consecuencia, si disminuimos el paso de cuantificación, estaremos reduciendo, al mismo tiempo, el nivel de ruido existente en la señal digitalizada. Así pues, si deseamos que la versión digital de la señal reproduzca con fidelidad el contenido de la información analógica, tendremos que ajustar el paso de cuantificación a las necesidades del sistema al que finalmente se dirige la señal. En el caso de señales de audio y de vídeo, el paso de cuantificación deberá ser reducido en la medida suficiente como para que el sistema auditivo o visual no puedan distinguir entre la señal original y la digitalizada.

Cuantificadores uniformes



El proceso de discretizar los valores de amplitud de la señal de entrada se conoce como la **cuantificación**. El cuantificador aproxima cada valor de tensión (variable real) de la señal de entrada a un valor de salida dentro de un subconjunto finito de posibles valores. En la figura adjunta mostramos un diagrama típico de entrada y salida para un cuantificador uniforme. En este ejemplo, sólo se admite un total de 8 posibles valores de tensión de salida. Se dice que el cuantificador es uniforme cuando el paso de cuantificación es constante en todo el margen dinámico de la señal. Los cuantificadores uniformes se utilizan en prácticamente todos los sistemas de audio digital para alta fidelidad y en la digitalización de señales de vídeo.

Codificación PCM



La codificación PCM (*pulse code modulation*) consiste en asignar un código binario a cada nivel de cuantificación de la señal, con lo que cada una de las muestras queda codificada con una palabra de un número de bits fijo. Este tipo de codificación se utiliza en el sistema *Compact Disc Digital Audio* y en varios sistemas de vídeo digital no comprimidos. En formatos de audio y vídeo comprimidos (ADPCM, minidisc, MP3, MPEG), la codificación de la señal en PCM suele constituir la primera fase del proceso de compresión. Tras haber convertido la señal a PCM, ésta se analiza y procesa con el objeto de reducir el número total de bits con el que se representa, codificándola en formatos alternativos que pueden resultar más o menos complejos y/o eficientes. En cualquier caso, para reproducir una señal comprimida también es habitual pasarla previamente al formato PCM y después convertirla a una señal analógica que podamos aplicar a los altavoces o al sistema de representación gráfica.

En la figura representamos los códigos binarios que le asignaríamos a un cuantificador uniforme. La codificación utilizada se conoce como codificación en complemento a 2 y es la que más se usa en la codificación de audio y vídeo. En el caso de los sistemas de vídeo, la codificación en complemento a 2 se utiliza para representar los componentes U y V (que pueden tomar valores positivos o negativos). Por otra parte, el componente de luminancia, que únicamente toma valores positivos, la representamos, en general, mediante códigos en los que la palabra binaria es directamente proporcional al valor de tensión. Así, una palabra 0000 representa el valor nulo, mientras que 1111 representa el valor máximo.

Los códigos binarios de los niveles de cuantificación positivos empiezan siempre por cero y a continuación indican el número de pasos de cuantificación asignados. De este modo, para codificar el nivel de cuantificación 3Δ utilizaremos el código 011,

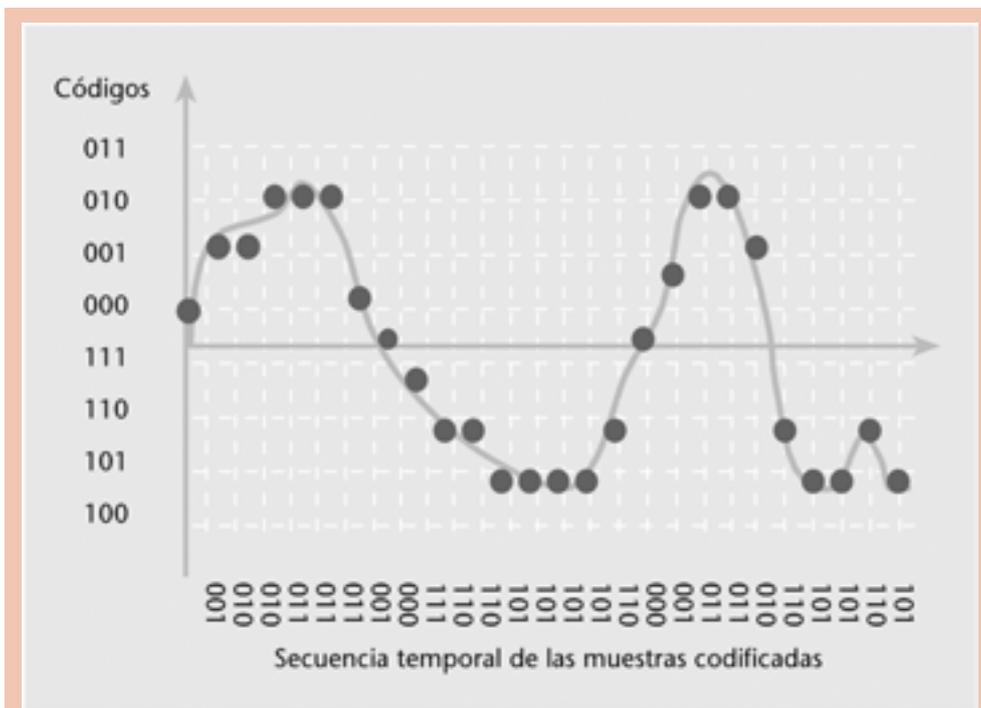
donde el primer 0 indica que el nivel es positivo y los otros dos dígitos codifican el número 3 ($1 \times 2^1 + 1 \times 2^0$).

La codificación de los niveles negativos es un tanto más compleja. Para codificar el nivel $-k\Delta$ expresaremos el de K en binario natural y, con posterioridad, cambiaremos todos los 1 por 0 y viceversa (complementar a 1). Tras esta operación, sumaremos 1 al número resultante (complemento a 2). De este modo, todos los códigos asociados a números negativos empiezan por 1. A modo de ejemplo, vamos a considerar cómo quedaría codificado el nivel -2Δ : el código binario correspondiente al nivel 2 es 010, que complementado a 1 resulta 101. Finalmente, sumando la unidad, obtenemos 110. La tabla adjunta muestra los niveles asignados correspondientes a un código en complemento a 2 de 3 bits.

Código binario	011	010	001	000	111	110	101	100
Nivel	3	2	1	0	-1	-2	-3	-4

Niveles y códigos binarios de 3 bits en complemento a 2

La ventaja de la codificación de datos en complemento a dos es que se pueden realizar las operaciones de suma entre niveles positivos y negativos sin tener en cuenta el signo. En efecto, si sumamos los códigos correspondientes a 2 (010) y a -3 (101), directamente obtenemos el código (111 = 010 + 101) que asociamos con el nivel -1 (2 - 3).



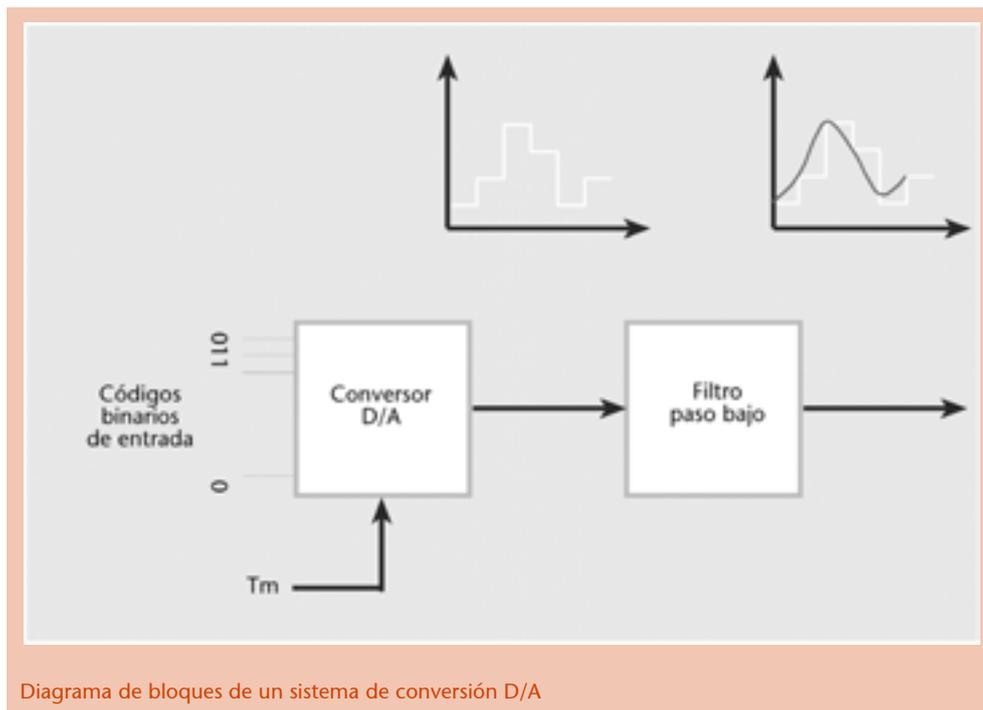
Ejemplo de la codificación de una señal



Ejemplo de codificación de una señal

En la gráfica hemos representado de manera esquemática el proceso de codificación de una señal. A cada muestra se le asigna el código binario correspondiente que representará el nivel de amplitud de la señal. Podremos almacenar los datos resultantes en una memoria para su posterior tratamiento, transmitirlos por algún medio adecuado o registrarlos en un soporte de almacenamiento masivo.

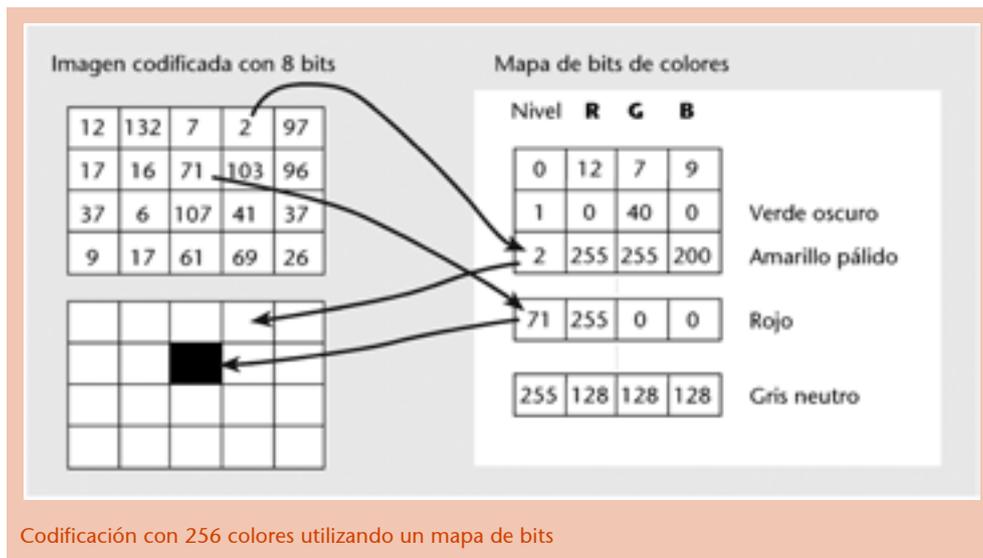
Conversión de las palabras código a muestras analógicas



El proceso de conversión de las palabras digitales a señales analógicas se realiza mediante conversores D/A (digital/analógico) y filtros paso bajo. Desde un punto de vista puramente funcional (no electrónico), los conversores digital analógico se encargan de convertir a valores de tensión los códigos binarios asociados a cada una de las muestras y mantienen este valor en la salida durante todo el periodo de muestreo. Los códigos binarios de las muestras se deben presentar al convertor con la misma cadencia con la que ha sido muestreada la señal. En el supuesto de que las frecuencias de muestreo de grabación y reproducción no coincidan, nos encontraremos ante un efecto parecido al cambio de velocidad de reproducción en un disco de vinilo o en una cinta magnética.

El filtro paso bajo situado en la salida del convertor se encarga de suavizar esta señal. En algunos circuitos integrados, el convertor y el filtro reconstructor de salida están integrados en el mismo sistema. En la figura que vemos a continuación representamos un diagrama de bloques del sistema de conversión D/A y las formas de onda representativas.

Selección del número de bits



El criterio para determinar el número de bits con el que hay que digitalizar la señal de vídeo también se basa en pruebas de calidad subjetiva y aceptación por parte del usuario. Consideraremos, en primer lugar, la digitalización de imágenes en blanco y negro y, más tarde, extenderemos los conceptos a las imágenes en color.

El procedimiento para determinar el número de niveles necesarios para codificar las muestras consiste en presentar varios tipos de imágenes a una serie de espectadores que tienen que valorar cuál es el número mínimo de niveles a partir del que no se aprecia ninguna mejora en la calidad de las imágenes. A partir de esta experiencia solemos obtener que el número de niveles de gris se sitúa entre los 45 y 60, lo cual indica que son necesarios cerca de unos 6 bits para obtener una calidad muy aceptable.

En la práctica, debido a que toda la estructura de las memorias está basada en palabras de 8 bits, se decidió cuantificar los niveles de luminancia con 8 bits, mediante lo que se obtiene un total de 256 niveles posibles que cubren a la perfección las necesidades del sistema visual.

Una imagen en blanco y negro digitalizada se puede interpretar como una matriz (filas y columnas) que en su interior contiene números enteros cuyos valores oscilan entre 0 (negro) y 255 (blanco), suponiendo que está codificada con 8 bits. Desde un punto de vista matemático, la imagen quedará descrita mediante $u[n, m]$, donde n representa la fila, m la columna y $u[n, m]$ el nivel de gris del elemento de imagen situado en la fila n , columna m .

Codificación de imágenes en color

La codificación de imágenes en color mediante componentes RGB es parecida a la de las imágenes en blanco y negro. En este caso, la imagen se encuentra representada por 3 ma-

trices, cada una asociada a uno de los componentes. Los valores de los componentes R, G y B son siempre positivos, por lo que podemos hacer uso de una representación en binario natural como la utilizada para la luminancia. En este caso, el nivel 0 representa el negro y el nivel máximo, la máxima luminancia asociada con cada componente.

El número de bits o niveles de cuantificación utilizado en cada uno de los componentes no está estandarizado. En sistemas de televisión, cada componente se suele codificar con 8 bits, aunque en la actualidad están apareciendo algunos sistemas profesionales que trabajan con resoluciones de 10 hasta 12 bits por componente.



En sistemas multimedia e informáticos se utilizan varias representaciones alternativas que dependen de las características de las tarjetas de digitalización de vídeo. La representación más habitual es la representación de color de 24 bits correspondiente a la cuantificación de cada uno de los componentes con 8 bits ($8 \times 3 = 24$). En este caso, el número total de colores que podemos representar se puede calcular contando todas las posibles combinaciones de colores que podemos obtener al combinar las 3 palabras de 8 bits ($256^3 = 16.777.216$).

La configuración de color en 32 bits, que admiten una gran cantidad de tarjetas gráficas actuales, es, en esencia, la misma que la de 24 bits. En este tipo, cada uno de los componentes RGB se codifican con 8 bits, proporcionando un componente adicional de 8 bits ($24 + 8 = 32$) para codificar la transparencia de la imagen. Este componente toma, por norma general, el valor nulo que indica que la imagen es totalmente opaca. Valores diferentes de cero indican cómo hay que combinar la imagen con el fondo para obtener el nivel de transparencia deseado. Así, un código igual a 128 indicará que cada elemento de imagen se representa como el promedio entre el color de la imagen y el color del fondo; un nivel 255 indicaría que en la representación gráfica sólo aparece el fondo.

La configuración de color de 16 bits reduce el número de bits con el que se codifica cada componente, dedicando 6 bits para el componente de verde (al que el sistema visual presenta una mayor sensibilidad) y 5 para los de rojo y azul. El número de posibles colores representados se reduce a ($64 \times 32 \times 32 = 65.536$), un número de colores que acos-

tombra a proporcionar representaciones de las imágenes en color bastante aceptables.

La representación de imágenes con 256 colores es un tanto más compleja y se utiliza en algunos ficheros de mapas de bits y sistemas informáticos antiguos o actuales simples (ordenadores de bolsillo). En este caso, el fichero o el dispositivo gráfico tienen asociado un mapa de color en el que se definen los tres componentes de color que vamos a utilizar para una determinada palabra código. Los elementos de imagen se codifican con palabras código de 0 a 255. Para asignar el color a un píxel es necesario consultar qué componentes RGB corresponden a esta palabra código; el resultado es que sólo podemos representar 256 colores pero, aun así, podemos ajustar el mapa de colores representados a las características de la imagen. Así, para una imagen con tonos verdes (bosque primaveral), el sistema informático deberá construir un mapa de colores compuesto, de manera predominante, por tonalidades verdes. El problema básico de este tipo de representación es que se requiere bastante cálculo para optimizar el mapa de colores asociado a una imagen. Cuando en la pantalla de un sistema operativo de estas características tenemos que representar de manera simultánea varias imágenes en color, la plantilla de colores utilizada es la de la ventana activa, por lo que es habitual que el resto de imágenes queden con colores falsos. Además, si la imagen se representa utilizando el mapa de colores por defecto del sistema operativo, los colores asignados no se corresponden con los reales. En la siguiente figura podemos ver un diagrama simplificado de la filosofía empleada para codificar imágenes con 256 colores (8 bits por píxel).“

Ancho de banda de la señal de vídeo digital

Vamos a considerar la transmisión o intercambio directo de señales digitales con una calidad parecida al PAL en el estándar ITU-601 y el formato 4:2:2.

La cantidad de bits asociados a la información de luminancia de un fotograma se puede determinar como:

$$576 \text{ (filas)} \times 720 \text{ (columnas)} \times 8 \text{ bits/píxel} = 3,317 \cdot 10^6 \text{ bits (Y)/fotograma}$$

Además de la información de luminancia, será preciso que incorporemos las señales diferencia de color U y V. En el formato 4:2:2 cada uno de estos componentes se codifica con la mitad de muestras que la señal de luminancia. Por lo tanto, el número total de bits asociados a un fotograma será:

$$3,317 \cdot 10^6 \text{ bits (Y)} + 2 \cdot \frac{1}{2} \cdot 3,317 \cdot 10^6 \text{ bits (U/V)} = 6,635 \cdot 10^6 \text{ bits/fotograma}$$

Para finalizar, teniendo en cuenta que tenemos que transmitir un total de 25 fotogramas por segundo, obtenemos:

$$25 \times 6,635 \cdot 10^6 = 165,88 \cdot 10^6 \text{ bits/s}$$

Está claro que a pesar de que se usen códigos y modulaciones digitales de gran eficiencia espectral, el ancho de banda necesario para transmitir esta tasa de bits es muy superior a los 5 MHz que requiere el sistema analógico. Por esta razón, los formatos digitales derivados del ITU 601 se han mantenido como estándares digitales para el intercambio de información entre equipos y terminales en estudios de grabación y producción de vídeo y no se han extendido a sistemas de radiodifusión de la señal de vídeo.

El incremento del ancho de banda asociado al proceso de digitalización de las señales de audio y vídeo representa, sin lugar a dudas, el principal inconveniente de los sistemas digitales. La radiodifusión directa de señales de voz, audio o vídeo no resulta posible en formato PCM debido a que un canal digital ocuparía, como mínimo, un espacio en el espectro de radiofrecuencia de unos 10 canales analógicos, saturando, con mucha rapidez, el número de posibles canales que podríamos distribuir.

Los sistemas actuales de radiodifusión de señales de vídeo y audio por vía terrenal, satélite o cable (DAB, DVB, GSM) han surgido gracias a la posibilidad tecnológica de implementar potentes algoritmos de compresión de información en tiempo real y a bajo coste. Algunos de estos algoritmos proporcionan una excelente calidad de imagen y sonido (prácticamente indistinguible de las señales sin comprimir) con unos factores de compresión superiores a los factores de extensión del ancho de banda con motivo del proceso de digitalización. Por este motivo, el ancho de banda efectivo de las señales comprimidas consigue ser inferior al ancho de banda de sus equivalentes analógicos, pudiendo acomodar varios canales digitales en el espacio espectral que antes ocupaba un canal analógico.

Los métodos matemáticos, así como los algoritmos utilizados para la compresión de audio y vídeo son complejos y su comprensión exige una cierta experiencia en varias técnicas de procesado digital de señales en el dominio del tiempo y de la frecuencia. Veremos los fundamentos de estos métodos de compresión un poco más adelante.

Etapa 2: Tarjetas y sistemas de digitalización

Introducción

Los sistemas de digitalización de señales de vídeo para su almacenamiento en disco duro han tenido un fuerte impacto en estudios de televisión y productoras de vídeo en aplicaciones profesionales. Estos sistemas permiten almacenar varias horas de vídeo, facilitando su edición posterior mediante sistemas informáticos. También se utilizan como servidores de vídeo (por ejemplo en la gestión de anuncios publicitarios) donde la secuencia horaria de los anuncios puede controlarse mediante red local o intranet. Algunos sistemas avanzados se utilizan en hoteles y hospitales para proporcionar servicios de vídeo a la carta (*video on demand* –el usuario pide la película que desea ver en el momento que desea). En general, estos equipos profesionales suelen comprimir la información en formato MPEG-2 o en formatos M-JPEG propietarios.

En el mercado doméstico existen algunas tarjetas que permiten la digitalización y el registro de señales de vídeo en el disco duro del ordenador. Estas tarjetas comprimen la secuencia de vídeo en formato M-JPEG (compresión individual de cada fotograma, sin tener en cuenta la redundancia temporal existente entre fotogramas distintos) y se utilizan, principalmente, para la edición de vídeo y la introducción de efectos especiales. La compresión de la señal de vídeo se realiza mediante un circuito integrado específico que descarga del proceso de compresión en el ordenador principal. Para que el sistema funcione de la manera esperada, tenemos que disponer de un PC de un disco duro de gran capacidad y velocidad, junto con un procesador avanzado.

En general, hay que adquirir las secuencias base que formarán parte de la producción del vídeo y, más tarde, se utilizan programas diseñados específicamente para la edición de vídeo, como puede ser el Adobe Premier, que permite la introducción de titulaciones, efectos de mezcla y transición entre secuencias de vídeo, inclusión de audio, etc. Una vez obtenida la secuencia final podemos recomprimir al formato deseado o enviarlo a la salida de vídeo compuesto para su registro en un magnetoscopio. Esta gama de productos suelen tener unos precios que rondan las 100.000 pesetas y su uso se está popularizando. Las tarjetas comerciales más conocidas pertenecientes a esta categoría son DC-30 (MIRO), AV-MASTER y BROADWAY.

También existen tarjetas de coste muy inferior que no incluyen ningún circuito específico para la compresión de vídeo y que se utilizan, fundamentalmente, para la visualización de programas de televisión en el ordenador. Estas tarjetas incorporan un circuito de sintonización de señal de televisión (la mayoría son circuitos de RF para la captura de señales de radiodifusión terrestre, pero también existen versiones para radiodifusión por satélite) y un circuito de digitalización de la señal de vídeo que suministra directamente los datos digitales al bus PCI del ordenador. Por regla general, dichas tarjetas se basan en alguna variante del circuito integrado de Brooktree Bt848 y no disponen de salida de vídeo analógica.

Estas tarjetas también se pueden utilizar para registrar vídeo en el disco duro, aunque, en este caso, es necesario que sea el ordenador principal el encargado de todas las tareas de compresión de información. Por este motivo, lo más usual es que no podamos disponer de la suficiente potencia de cálculo como para poder trabajar con resoluciones altas (576×768) a 25 imágenes por segundo, con lo que hay que reducir o bien la frecuencia de imagen, o la resolución de las imágenes (y en la mayoría de los casos, ambas). A pesar de todo, para configuraciones típicas de Pentium II o Pentium III, solemos contar con la posibilidad de trabajar con resoluciones CIF (calidad equivalente a VHS) a 15 fotogramas por segundo de forma fiable. Los principales fabricantes de este tipo de tarjetas son MIRO (Easy TV), AVERMEDIA (WinTV) y HAUPAGE.

El reducido coste de estos sistemas hace que su uso se haya popularizado en aplicaciones domésticas y se haya extendido a sistemas de videovigilancia para la captura de secuencias de vídeo en disco duro. En estas aplicaciones, el sistema de captura de vídeo suele estar sincronizado mediante señales de alarma externas (sensores de presencia, apertura de puertas, detectores de movimiento, etc.) de manera que, en el momento en que se activa una alarma, adquirimos una secuencia de vídeo durante unos cuantos segundos (4 o 5 segundos de prealarma y 4 o 5 de postalarma) para capturar la entrada del intruso (tenemos ejemplo típico de esto en las aplicaciones bancarias). Podremos recuperar las imágenes registradas en el disco duro desde un centro de control remoto mediante conexiones tipo TCP/IP. En algunas instalaciones, tan sólo se realiza la compresión de las imágenes y se transmiten directamente al centro de control mediante conexiones de alta velocidad (intranets).

En los apartados que se siguen vamos a tratar con detalle las características de este tipo de tarjetas para la sintonización de señales de televisión, su digitalización y captura. Dedicaremos una especial atención a describir el principal circuito integrado que se encarga de la digitalización de la señal de vídeo en este tipo de tarjetas.

Tarjetas digitalizadoras y sintonizadoras de señal de televisión

La arquitectura básica de este tipo de tarjetas se muestra en la siguiente figura:

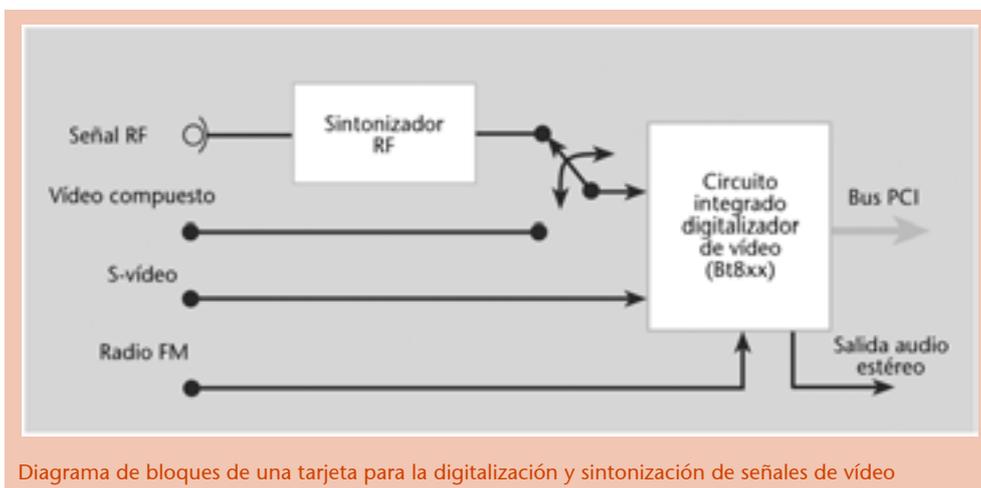


Diagrama de bloques de una tarjeta para la digitalización y sintonización de señales de vídeo

En general, disponemos de varios formatos de entrada para la señal de vídeo:

- toma de antena para señal RF (radiodifusión terrestre);
- señal de vídeo compuesto. Compatible PAL, NTSC SECAM;
- señal S-vídeo.

La señal procedente de la toma de antena se inserta en un circuito sintonizador de señales de televisión que convierte la señal modulada en una de vídeo compuesto, como la que hemos analizado en capítulos anteriores. El circuito de digitalización controla directamente al sintetizador mediante un bus I2C (un protocolo muy utilizado para la comunicación entre circuitos integrados y para programar sus funciones). El usuario puede seleccionar el canal de radiofrecuencia que quiere visualizar mediante un *software* de alto nivel que se comunica directamente con el subsistema de digitalización mediante bus PCI.

Las tomas de entrada en vídeo compuesto o S-vídeo se utilizan, por norma general, para conectar la señal de una cámara de vídeo externa. El formato S-vídeo es parecido al de vídeo compuesto, pero las señales de croma y de luminancia se proporcionan por cables separados. El subsistema de digitalización sólo puede operar con una señal de entrada, por lo que el usuario tiene que seleccionar el tipo de entrada con el que desea trabajar.



“Los fabricantes de circuitos integrados para la digitalización de señales de vídeo más extendidos en esta gama de productos son Philips y Brooktree (Rockwell). Este último ha alcanzado con su gama de productos Bt8xx una situación de dominio considerable en el mercado, un dominio que se debe, fundamentalmente, a la difusión gratuita de *drivers* totalmente compatibles con el *software* de desarrollo proporcionado por Microsoft para plataformas Windows (*Video for Windows*,

DirectShow). Además, se proporciona una excelente documentación acerca de cómo programar el dispositivo, que incluye no sólo el *software*, sino también ejemplos de circuitos impresos para tarjetas y conexiones con otros circuitos integrados. Estas condiciones han propiciado que aparezcan varias aplicaciones de otros desarrolladores (principalmente decodificadores de Canal +) que han favorecido el dominio casi total de este circuito integrado en algunos países.

La presencia de la señal de entrada para antena de radio en FM en la tarjeta depende del modelo de circuito de digitalización utilizado. Algunos modelos del Bt8xx integran un subsistema para la digitalización de audio externo.

La salida de audio estereofónico que aparece representada en el diagrama de bloques de la figura anterior se usa principalmente para conectarla con la entrada de línea de la tarjeta de sonido del ordenador. Gracias a esto es posible utilizar todas las posibilidades que ofrece el sistema de sonido multimedia del ordenador para escuchar la señal de audio de un programa de televisión, de un vídeo procedente de una cámara o un programa de radio FM.

En la figura que vemos a continuación aparece una fotografía de una tarjeta de digitalización muy simple, que no incluye unidad de sintonización de señal de televisión

y que se utiliza en aplicaciones de videoseguridad (captura directa de la señal de vídeo compuesto proporcionada por una cámara). El ejemplo es adecuado para mostrar la potencia del digitalizador Bt8xx, ya que la placa de circuito impreso sólo sirve, prácticamente, como soporte físico del circuito integrado para conectarlo al bus PCI. En la placa también podemos observar los osciladores externos que proporcionan la frecuencia de referencia para detectar señales de vídeo.



Las comunicaciones de datos y programas entre la tarjeta de digitalización y el ordenador se realizan mediante el bus PCI. Este bus admite velocidades de transferencia de datos de hasta 132 *Mbytes* por segundo (tened en cuenta que son *Mbytes* y no *Mbits*), lo que permite un flujo más que suficiente, al menos si todo funciona de forma correcta, para el intercambio de vídeo digital entre la tarjeta y el ordenador.

Principios de funcionamiento del sistema de digitalización de vídeo en el entorno Windows



El principio de funcionamiento del sistema para la programación de la tarjeta y el intercambio de imágenes digitalizadas es bastante sencillo. Cuando el usuario arranca el programa de captura o visionado de vídeo se envía una orden al Bt8xx para que se configure en **modo de programación**.

En este modo, es posible modificar los registros internos del Bt8xx mediante el bus PCI. Hay que tener en cuenta que el Bt8xx dispone de un conjunto de *pins* que están directamente conectados al bus PCI con lo que, por lo tanto, el propio circuito integrado puede actuar como controlador del bus PCI. Los valores que se descargan desde la aplicación a los registros internos del circuito integrado permiten especificar por completo las características del proceso de digitalización. Así, es posible configurar el tamaño de las imágenes, el número de *frames* por segundo que vamos a digitalizar, la fuente de vídeo que utilizaremos, etc. Cuando el usuario entra en una aplicación de captura de vídeo o visualización de señal de televisión puede seleccionar todos estos parámetros que serán cargados en los registros del Bt8xx para configurar cómo se realiza la digitalización. En la figura podemos ver una pantalla de configuración típica en la que seleccionamos la fuente de entrada, el canal de RF que hay que sintonizar y los ajustes de brillo, contraste y saturación del digitalizador.

Cuando el usuario inicia el proceso de digitalización se descargan a los registros internos del Bt8xx todos los parámetros seleccionados. A continuación, se conmuta del modo de configuración al modo de adquisición y el sistema empieza a descargar las muestras de las señales de audio y vídeo en la memoria del PC. El Bt8xx dispone de un controlador de DMA que gestiona la transferencia de datos entre la tarjeta, el bus PCI y la memoria RAM.



Uno de los principales problemas es que otra tarjeta o dispositivo conectado puede solicitar el bus PCI al ordenador, impidiendo, de forma momentánea, la transferencia de datos de vídeo de la tarjeta a la memoria RAM. Para paliar este problema el digitalizador incluye un *buffer* de memoria en el que se pueden almacenar temporalmente las muestras de vídeo y audio. Mientras que no disponemos del control del bus PCI, el *buffer* de memoria se va llenando a la espera de efectuar la transferencia de los datos más adelante. En principio, la tasa de transferencia de datos al bus PCI es elevada en la medida suficiente como para vaciar rápidamente el *buffer* cuando el sistema recupera el control del bus.

No obstante, el *buffer* de vídeo no es demasiado grande (630 *bytes*), por lo que es posible que el flujo de muestras de entrada rebase su capacidad y se sobrescriban las muestras. En estos casos, el Bt8xx gestiona el *buffer* interno descartando todas las muestras de un mismo fotograma. Esta política condiciona la capacidad de la tarjeta de vídeo para cumplir con la configuración de adquisición que ha seleccionado el usuario. Así, es muy probable que si el usuario elige una opción de digitalización de 576×768 píxeles por imagen con 25 imágenes por segundo, se pierdan algunos fotogramas. En este caso, el flujo de datos de vídeo es tan elevado que el *buffer* del sistema se rebasa

con mucha facilidad cuando otro dispositivo adquiere el control del bus PCI. Si la configuración de digitalización se reduce a 288×384 y 15 imágenes por segundo, también se reducirá el flujo de datos de vídeo y, en consecuencia, aumenta la probabilidad de que la digitalización se pueda realizar según las especificaciones del usuario. El máximo flujo de digitalización que podemos obtener con una determinada plataforma PC varía en función de las características del sistema y el tipo de dispositivos que estén conectados al bus PCI. En general, hay que probar varias configuraciones de digitalización hasta determinar la tasa máxima a la que se puede trabajar. Los formatos de 288×384 con 15 imágenes por segundo suelen proporcionar resultados satisfactorios a partir de plataformas del tipo Pentium II.

Podría llegar a pensar que la pérdida de un fotograma no es excesivamente grave y que, en cualquier caso, es mejor utilizar un formato 576×768 con 25 fotogramas, aunque perdamos algún fotograma de vez en cuando, que un formato más reducido. El problema es que cuando el formato de digitalización exige una tasa de transferencia muy alta se suelen perder más de un fotograma consecutivos de forma por completo aleatoria, lo cual provoca que el vídeo obtenido presente frecuentes interrupciones y paradas de imagen que resultan bastante molestas.

Los datos de vídeo se registran en la memoria RAM del sistema en el formato seleccionado por el usuario sin ninguna compresión. La compresión se realiza por *software*, mediante programas que gestionan los datos originales de la RAM, los procesan y los almacenan en un fichero del disco duro. Lo más normal es que podamos seleccionar el algoritmo de compresión y sus parámetros asociados desde la misma aplicación en la que se realiza la

configuración del formato de adquisición. No obstante, estos parámetros no se cargan en el Bt8xx, sino en un programa residente que se encarga de realizar la compresión de la secuencia de vídeo. Estos programas suelen estar relacionados de manera dinámica (DLL) con el gestor de adquisición de vídeo, y los recursos se cargan en el sistema cuando instalamos un determinado compresor (codec).

Etapa 3: Interfaces de vídeo digital

Introducción

La digitalización de la señal de vídeo exige que los distintos equipos que deban tratar con esta señal puedan intercambiar la información digital de forma directa y compatible. El flujo de datos asociado a la señal de vídeo es, como hemos visto, muy elevado, por lo que los sistemas de transporte de este tipo de señales precisan un elevado ancho de banda.

Los estándares más utilizados para intercomunicar equipos de vídeo digital son:

- estándar SMPTE 259M (SDI, SDV),
- estándar IEEE 1394/FireWire.

En los siguientes apartados veremos con cierto detalle las características principales de estos dos estándares.

Serial digital interface (SMPTE 259M)

La recomendación ITU-R 656 establece el estándar de interconexión entre dispositivos que utilizan el formato 4:2:2 (básicamente magnetoscopios digitales, editores y mesas profesionales). En la versión en paralelo de la interfaz se utiliza un conector del tipo DB25, por el que se van transmitiendo las muestras de 8 bits de los componentes de luminancia y croma, multiplexados en el tiempo. El orden de transmisión, es $Cr_1, Y_1, Cb_1, Y_2, Cr_3, Y_3, Cb_3, Y_4, Cr_5, Y_5, Cb_5, Y_6$, etc. Junto con los bits de cada una de las muestras, se transmite un reloj de 27 MHz (un periodo por muestra). También está definida una versión en serie de la interfaz que utiliza un cable coaxial de 75 ohmios con conectores del tipo BNC; en este caso, se utiliza una tasa de 243 Mbps, debido a que, en este modo, la transmisión es de 9 bits por muestra.

No obstante, la mayor parte de los equipos de digitalización y almacenamiento de vídeo digital profesionales han adoptado el estándar SMPTE 259M, que apareció en 1997 y que, de forma conceptual, se corresponde con los estándares del ITU para televisión digital. Se conoce indistintamente con los nombres de SDI (*serial digital interface*), SDV (*serial digital video*) o SMPTE 259M (el nombre del documento en el que se define el estándar). El estándar define las características de las señales, las tasas de bit, el tipo de conectores y cables empleados (coaxial de 75 ohmios), la codificación de canal y el orden de transmisión de los bits.

Las tasas de transmisión utilizadas son de 270 Mbps para TV digital en formato 4:2:2 en componentes con 10 bits por muestra para televisión convencional (relación de aspecto 4:3). En el caso de relaciones de aspecto panorámicas (16:9), la tasa de transmisión utilizada es de 360 Mbps.

Estándar IEEE 1394

Se trata de un estándar para la comunicación entre ordenadores y dispositivos periféricos que Apple desarrolló en un principio con el nombre *FireWire*. Es un bus serie de altas prestaciones y alta velocidad que soporta velocidades de transmisión de 100, 200 y 400 Mbps.

El bus IEEE1394/*FireWire* permite operaciones de entrada/salida de vídeo en tiempo real y admite la posibilidad de conectar otro tipo de periféricos como discos duros, impresoras, escáneres, cámaras digitales y cámaras de vídeo. La alta velocidad del bus permite que el sistema compita directamente con otros estándares ampliamente implantados, por ejemplo el SCSI en aplicaciones de alta velocidad (como el registro de vídeo en disco duro).

Una de las principales ventajas del bus IEEE 1394 es que permite la sustitución o intercambio de periféricos durante la operación del sistema y que no requiere la aplicación de terminaciones del bus (estas dos características son los principales inconvenientes del bus SCSI).



Sony colaboró de manera activa con el IEEE y Apple en la elaboración del estándar para la introducción en tiempo real de señales de vídeo y audio. El estándar se suele conocer con el nombre alternativo "i.Link" en aplicaciones de vídeo digital (DV) desarrolladas por Sony (casi todas las cámaras digitales domésticas disponen de esta entrada/salida di-

gital). Dicho estándar está siendo utilizado cada vez más por sistemas de vídeo semiprofesionales y profesionales.

En la actualidad, muchos ordenadores están incorporando este bus (Macintosh G3 y G4, Sony Vaio, Microsoft Windows y otros fabricantes)."