

# Procesamiento de señal de video analógica con optimización de ancho de banda para ser transmitido por internet

Carlos R. Soria-Cano  
Salvador Álvarez-Ballesteros

Instituto Politécnico Nacional  
Escuela Superior de Ingeniería Mecánica y Eléctrica  
Sección de Estudios de Posgrado e Investigación (Zacatenco)  
Ciudad de México, CP 07738.  
MÉXICO

correos electrónicos (emails):  
rcanos02@hotmail.com  
salvarez@ipn.mx

Recibido 16-09-2016, aceptado 13-12-2017.

## Resumen

En este artículo se detalla el procesamiento de la señal de video y audio analógico para ser transmitidos como stream de video sobre IP, respetando cualquier información de metadata en cada procesamiento de la señal, tomando en cuenta desde que se realiza una grabación en banda base, señal analógica y su conversión a señal digital con audio embebido (SDI), en calidad estándar SD-SDI y en alta definición HD-SDI, así como la selección del proceso de compresión para reducir y optimizar el ancho de banda de cada señal; se analiza códec MPEG y H.264, se realizan pruebas de ancho de banda con diferentes tasa de bits de compresión para tener calidad de señal, sin *lipsync* y la menor latencia posible. Para poder transmitir un mayor número de señales en normas de calidad de video se realiza una multiplexación; obteniendo una señal ASI (*Asynchronous Serial Interface*). Se muestra la forma en que el stream puede ser transmitido mediante la nube (internet) para que el usuario final lo pueda observar en diferentes dispositivos móviles (smartphone, tablet, pantalla, PC, laptop).

**Palabras clave:** señal de video analógico, señal de video SD.SDI, HD-SDI, MPEG, H.264, señal de video ASI.

## Abstract (Video Analog Signal Processing with Bandwidth Optimization to Broadcast by Internet)

This article describes the processing of the video analog signal and audio to broadcast as stream video over IP is detailed, taking into account since recording is done in baseband analog signal and converting it to a digital signal with embedded audio (SDI), standard quality SD-SDI and high definition HD-SDI, it shows the selection of the compression process to reduce and optimize the bandwidth of each signal; MPEG and H.264 were analyzed, bandwidth tests are performed with different compression bit rate to have signal quality without lipsync and the lowest possible latency. In order to transmit a greater number of signals in video quality standards multiplexing is performed; obtaining a signal ASI (asynchronous serial interface). The way the stream can be transmitted through the cloud (Internet) so that the end user can observe on various mobile devices (smartphone, tablet, screen, pc, laptop) is shown.

**Index terms:** analog video signal, video signal SD.SDI, HD-SDI, MPEG, H.264, video signal ASI.

## 1. Introducción

El mundo está experimentando un espectacular cambio en la recepción de televisión. Las cifras fueron reveladas por la Unión Internacional de Telecomunicaciones, UIT, que hace un balance de los servicios de comunicaciones en 157 países. El informe indica que en los países desarrollados, hasta 91 por ciento de todos los hogares con TV reciben las señales a través de tecnologías digitales, gracias a los avances efectuados en el paso a la televisión por IP[1].

Para entender cómo funciona la transmisión de video y audio se necesita comprender cada uno de los procesos que se llevan a cabo en la señal original. Un primer paso clave en la prestación de servicios de video y audio es la conversión de las señales de voz y video analógicos en un formato digital (digitalización).

### 1.1. Señal en banda base analógico

Se denomina banda base al conjunto de señales que no sufren ningún proceso de modulación a la salida de la fuente que las origina, es decir son señales que son transmitidas en su frecuencia original, dichas señales se pueden codificar. Por ello, podemos decir que la banda base describe el estado de la señal antes de la modulación y de la multiplexación y después de la demultiplexación y demodulación. Las ondas de video y sonido son analógicos en su origen, son capturadas por una cámara y un micrófono y convertidas en una pequeña variación analógica de tensión denominada señal de video y señal de audio. Esta tensión varía de manera continua a medida que cambia el volumen, frecuencia del sonido, la luminancia y el cromina [2].

### 1.2. Conversión de video y audio analógico a digital

Para convertir la señal de componente analógica en un valor digital, enviamos la señal analógica en un convertidor analógico a digital (ADC), el cual convierte la forma de onda analógica en una palabra binaria, en este caso una palabra binaria de 10 bits. La figura 1 se muestra de manera resumida la conversión analógica a digital.

## 2. Señal de video Digital SD/SDI y HD/SDI

*Serial Digital Interface* (SDI) es una interfaz de video digital estandarizada. Utilizada principalmente para la transmisión de señal de video sin compresión (video RGB) y sin encriptación (incluyendo opcionalmente audio). También se utiliza para la transmisión de paquetes de datos. Las especificaciones de este formato están recogidas en la recomendación ITU-R-BT 656 para la transmisión de señales de vídeo en componentes digitales, utilizando un flujo de 270 Mbps (véase Figura 2). Esta cifra se deriva de [3]:

- Cuadros (o frames) por segundo
- Líneas de cada cuadro
- Muestras de cada línea
- Bits de cada muestra

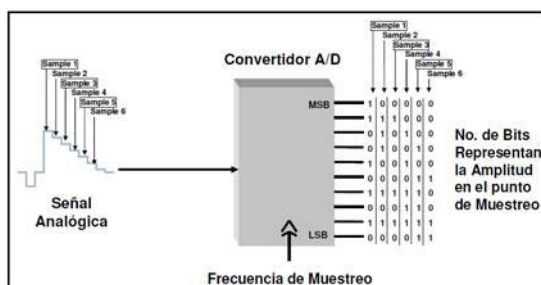


Fig. 1. Convertidor A/D.

Si la señal original de video es PAL (*Phase Alternating Line*, línea de fase alternada, nombre con el que se designa al sistema de codificación utilizado en la transmisión de señales principalmente en Europa) tendrá un número de líneas: 625 de las cuales las líneas activas (resolución vertical efectiva): 576, frecuencia de cuadro: 25 Hz (40 ms), frecuencia de campo: 50 Hz (20 ms, de los cuales 18,4 ms activos). Si la señal original de video es NTSC (llamado así por las siglas de *National Television System Committee*, Comisión Nacional de Sistema de Televisión, utilizado principalmente en América) cada trama o cuadro se compone de dos campos, cada uno de los cuales constan de 262,5 líneas de exploración, para un total de 525 líneas [4].

La tasa de transferencia de una señal PAL es 177 Mbps, señal NTSC es 143Mbps, señal digital estándar 270Mbps y una señal digital de alta definición es 1.485 Gbps. De ahí la necesidad de reducir el ancho de banda para poder ser enviado por internet, con una tasa de transferencia promedio de 3 Mbps. Para ello se necesita comprimir la señal, teniendo especial cuidado en la calidad de la señal.

## 3. Compresión MPEG 2 vs H264

Se necesita compresión porque sin compresión una imagen bitmap = 2,300,000 bytes, mientras que un archivo comprimido como imagen JPEG = 159,000 bytes; es decir, la reducción de información es de 14.5 veces solo para una imagen.

Con compresión MPEG2:

- ~39MB por minuto de video comprimido
- Película completa (~120 minutos) en 1 DVD.

Sin compresión MPEG2

- ~11GB por minuto de video sin compresión
- Película completa (~120 minutos) en ~280 DVDs.

MPEG-4 AVC, también llamado H.264 y MPEG-4Parte 10, está siendo rápidamente adoptado en todos los segmentos de la industria de video en red, ya que ahorra aproximadamente el

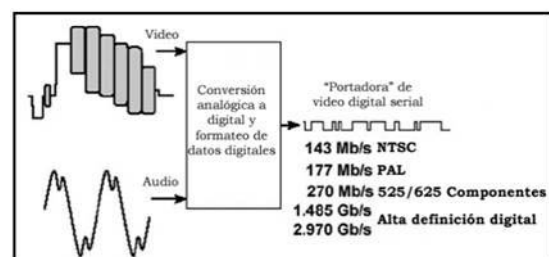


Fig. 2. Tasa de bits de la señal de video.

60% del ancho de banda [6]. Es ahora el estándar aceptado para las comunicaciones, difusión, y las aplicaciones de *streaming*. Se utiliza dentro de Flash, Silverlight, QuickTime, iPhones, iPod, PlayStation, Nero, HD DVD, y las tecnologías de discos Blu-ray.

Un grupo de imágenes (GOP) es el número de fotogramas de un I-Frame a otro. Así, en un GOP, hay un marco completo, y un número de fotogramas que predicen el movimiento asociado con ese fotograma completo. Los niveles más altos de compresión (menos ancho de banda para un clip determinado) se logran a través de mayor tamaño de GOP y predicción de cuadro más profundo (es decir, utilizando una estructura de compresión IBP en lugar de una estructura de IP), ver figura 3 Sin embargo, las compresiones más profundas requieren más tiempo para codificar, resultando en latencias más altas [5].

#### 4. Codificación IP

Para poder enviar por IP la señal de video SDI (audio y video digital embebido), la señal debe de ser comprimida para utilizar el menor ancho de banda posible sin perder la calidad de video y audio. El codificador de señal SDI al estándar de compresión H.264 tiene que tener FEC (factor de corrección de errores) para no tener pérdidas de bits considerables que afecten la calidad de la señal.

Dentro de la configuración del codificador se pueden reproducir secuencias H.264 en resolución completa y plena velocidad de fotogramas de hasta 15 Mbps, al reducir la escala y la velocidad de fotogramas se puede emitir secuencias tan pequeñas de hasta 150 kbps. Dentro del estándar se tiene que verificar que los servicios de METADATA no se pierdan en la codificación. Dentro del Stream de video se pueden tener todos los servicios de datos auxiliares, como *cue tone* (audio), GPI, V-chip, *closed caption*, *raiting*, estos se encuentran dentro de la señal en líneas reservadas para datos auxiliares. Estas líneas no son visibles para el usuario [7].

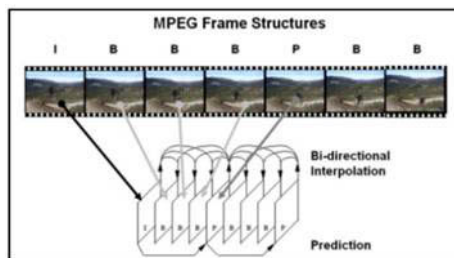


Fig. 3. Estructura MPEG.

La señal H.264 puede ser transmitida en multicast, donde cualquier usuario que este dentro de la red podrá ver la señal. Si se requiere poner diferentes tipos de usuario y contraseña, para que usuarios no permitidos no puedan ver la señal, se necesita un servidor. Con él se podrán administrar las señales: tener una lista de programación de la señal en vivo o del storage disponible en el disco. Vista previa de la señal, cambio de señal (canal). Se tendrá almacenado videos o películas, para que el usuario las pueda ver cuando lo requiera (video bajo demanda). Se tiene una encriptación AES para seguridad de la red.

Los administradores pueden ajustar la característica operativa del cualquier usuario pueden enviar mensajes a todos los usuarios que visualizan la señal. Informes de audiencia, permite a los administradores tener el log exacto de cualquier usuario con detalle. Por ejemplo, pueden saber si alguien tiene *mute* o minimiza la ventana de visualización [8].

Un aspecto importante, principalmente para señales de video en vivo es la latencia, desde el lugar de grabación hasta el dispositivo del usuario final, con este tipo de codificadores se tiene una latencia de 70 milisegundos de extremo a extremo (véase Figura 4).

Se tiene un equipo codificador HAIVISION para poder comprimir la señal de video HD/SDI. Los parámetros que se deben de configurar con respecto al video son:

Input: BNC (SDI)

Input Format: HD o SD dependiendo de la señal.

Color space: Auto

Aspect Ratio: 16:9

Display Logo: Ok

Eneable: High

Profile: Default

Resolución: 1920\*1080

Framing: IP

GOP y Bitrate son lo que determina el ancho de banda del video.

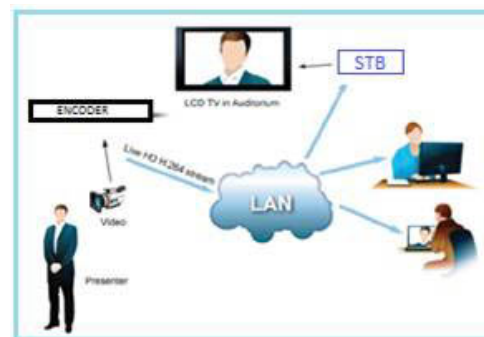


Fig. 4. Diagrama de red.

## 5. Optimización de ancho de banda

Como se ha estudiado los *settings* que afectan principalmente al ancho de banda son: bit rate de video, de audio y GOP. Se realizarán pruebas modificando cada uno de estos parámetros para obtener la optimización, y que el usuario final pueda observar el streaming en normas de calidad y video. La metodología consiste en dejar fijo el bitrate del video y audio y comenzar a mover el GOP para obtener una latencia por debajo de los 100 ms. Después dejar fijo el bitrate de video y el GOP y modificar el bitrate de video para obtener el menor valor sin que la señal presente problemas de *Lipsync*. Por ultimo se deja fijo el GOP y el bitrate de audio y se modifica el bitrate del video para obtener una señal con normas de calidad.

Parámetros configurados:

GOP: 10  
Resolución: 1920\*1080  
Tasa de Muestreo: 48 kHz  
Bitrate de video: 500 kbps  
Bitrate de audio: 128 kbps  
Ancho de banda utilizado 750 kbps.

La figura 5 muestra la salida de video con la configuración antes mencionada del codificador.

En la figura 5 se nota que el video no está dentro de normas de calidad, se observan macroblocking y digitalizaciones constantes. Al revisar la señal en un monitor forma de onda Tektronix WFW 8300, se dice que está en normas de video cuando cumple con las especificaciones marcadas en ITU-R BT. 1700 [9].

Se observa una señal fuera de normas en Luminancia (mayor de 700mV), en crominancia se encuentra en niveles +/-350 mV, figura 6.

En la figura 7 se observa la señal en niveles de Luminancia pero baja en crominancia, por lo cual la señal se ve opaca y sin color.



Fig. 5. Salida de codificador HAIVISION.

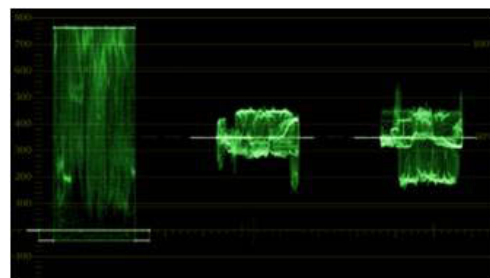


Fig. 6. Señal de video con luminancia saturada.

En la figura 8 se observa la señal configurada con el convertidor de normas, en la cual tiene los niveles en el rango permitido.

Parámetros configurados para siguiente prueba:

GOP: 100  
Resolución: 1920\*1080  
Tasa de Muestreo: 48 KHz  
Bitrate de video: 1000 Kbps  
Bitrate de audio: 64 Kbps  
Ancho de banda utilizado 1.3 Mbps.

La figura 9 es salida del codificador con los parámetros antes configurados.

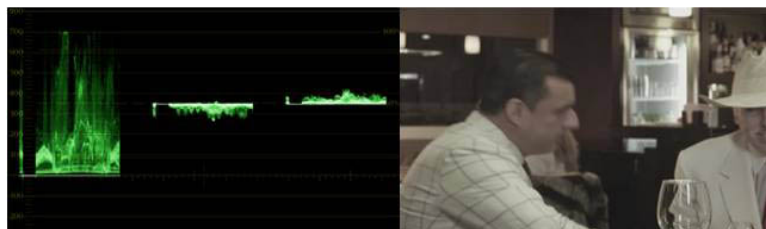


Fig. 7. Señal de video con baja ganancia en crominancia.



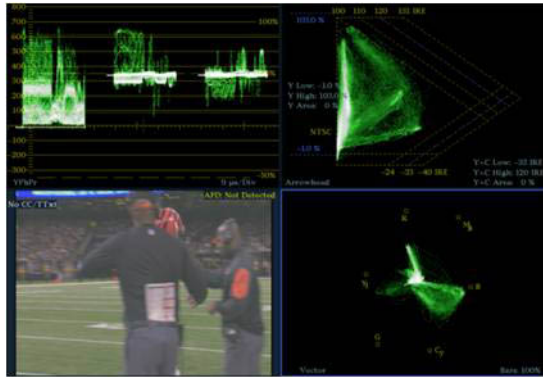


Fig. 8. Salida de codificador HAIVISION.

En la figura 9 se nota que el video no está dentro de normas de calidad, se observan cromas pastosos y digitalizaciones esporádicas, al aumentar el GOP aumenta la latencia (2 segundos con respecto al video original), esto significa que el video tiene un delay de 2 segundos con respecto a la entrada del codificador.

Parámetros configurados:

GOP: 15

Resolución: 1920\*1080

Tasa de Muestreo: 48 kHz

Bitrate de video: 1700 kbps

Bitrate de audio: 128 kbps

Ancho de banda utilizado 2.3 Mbps

En la figura 10 se observa señal de video salida de HAIVISION. Se nota que el video está dentro de normas de calidad. Se tiene latencia baja (70 segundos) pero se tiene un ancho de banda de 3.1 Mbps, este AB se puede optimizar sin perder calidad en imagen.

Se realizan pruebas cambiando cada uno de los parámetros para obtener una señal comprimida en H.264 dentro de nor-



Fig. 9. Señal de video comprimida, salida del codificador HAIVISION.



Fig. 10. Señal de video salida HAIVISION.

mas de calidad, con un bajo ancho de banda, baja latencia (delay de video de salida con respecto a video de entrada).

## 5.1 Resultados

Después de realizar varias pruebas con diferentes parámetros, la optimización con respecto ancho de banda y calidad de imagen tienen que ver principalmente con el bitrate del video, se obtiene video sin digitalizaciones y en normas de video [8], al optimizar el GOP obtenemos una latencia menor a 70 ms sin tener perdida de información en el video (digitalizaciones), al optimizar el bitrate de audio se tiene el video y el audio emfasado sin presentar problemas de *Lipsync* (audio adelantado o atrasado con respecto al video). La configuración optimizada queda de la siguiente manera.

Parámetros óptimos:

GOP: 15

Resolución: 1920\*1080

Tasa de Muestreo: 48 KHz

Bitrate de video: 1500 Kbps

Bitrate de audio: 128 Kbps

Ancho de banda utilizado 2 Mbps

La figura 11 muestra video configurado con los parámetros anteriores.



Fig. 11. Señal de video salida HAIVISION.

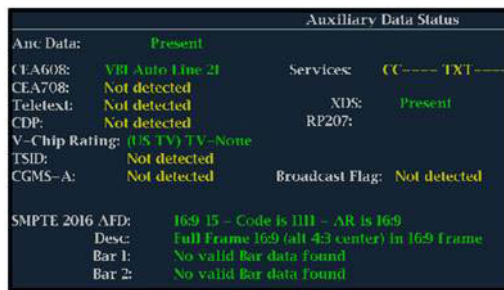


Fig. 12. Metadata.

La señal de video SD/SDI al ser codificada con codificación H.264: se comprime la señal a 2Mbps, con todas las características de video, audio y metadata. Al decodificar el streaming para volver a tener el video original, se conservan todos los servicios de metadata: Cue tone (Audio), GPI, V-chip, Closed Caption, Raiting, AFD como se muestra en el monitor forma de onda de la figura 12.

## 6. Conclusiones

En este trabajo se ha propuesto el uso de la red para la transmisión de cualquier tipo de video y audio en diferentes formatos, códec, wrapper, realizándole la conversión y compresión para ser transmitidos por IP, y posteriormente aplicando el proceso inverso para recuperar la señal original y poder ser reproducida en el dispositivo electrónico que se desee.

Para broadcast es una alternativa confiable, se realizaron diferentes tipos de pruebas, configurando cada uno de sus parámetros y modificándolos para obtener su mejor rendimiento ancho de banda, lipsync, latencia. Se obtienen los parámetros

optimizados: GOP de 15, resolución: 1920\*1080, tasa de muestreo 48 KHz, bitrate de video = 2500 Kbps. Con los parámetros utilizados se obtiene un ancho de banda reducido de 2Mbps, una latencia baja (menor a 70 milisegundos con respecto al video original), luminancia y crominancia en normas de calidad de video, no se tienen problemas de *Lipsync*, no existen digitalizaciones ni macroblocking. Se realiza la codificación sin perder los servicios embebidos en metadata: *cue tone* (audio), GPI, V-chip, *closed caption*, *raiting*, AFD.

## Referencias

- [1] International Telecommunication Union, *ITU-T Recommendations and selected Handbooks*, 2014, [<http://www.itu.int>.]
- [2] J. Osorio, *Fundamentos y mediciones de video y audio digital en SD y HD*. Tektronix Latinoamérica, 2009.
- [3] C. Becker C. *Mediciones críticas para video SD y HD sin compresión*. Tektronix Latinoamérica, 2009.
- [4] Tektronix, *MPEG Fundamentals and Protocol Analysis*, 2007. Disponible en <http://www.tek.com/video-test>
- [5] Tektronix, *A Guide to Compression Systems*, 2007. Disponible en <http://www.tek.com/video>.
- [6] Tektronix, *A Guide to Standard and High-Definition Digital Video Measurements*, 2007. Disponible en <http://www.tek.com/video>
- [7] F. Ibrahim, *Guide to Television and Video Technology*, Newness, 2009.
- [8] Tektronix, *A Guide to IPTV The Technologies, the Challenges and How to Test IPTV*, 2007. <http://www.tek.com/video>
- [9] International Telecommunication Union, *Recommendations ITU-R BT. 1700*, 2017.