

Códecs con tecnología de video afluente: un análisis de desempeño en el marco IPTV

Codecs in the Video Streaming Technology: An Analysis of Performance in the IPTV Framework

García-Vázquez M.S.

Instituto Politécnico Nacional-CITEDI, Tijuana
E-mail: mgarciav@citedi.mx

Ramírez-Acosta A.A.

Instituto Politécnico Nacional-CITEDI, Tijuana
E-mail: alvaro@citedi.mx

García-Garduño V.

*Departamento de Telecomunicaciones,
Facultad de Ingeniería
Universidad Nacional Autónoma de México*
E-mail: francia@marconi.fi-b.unam.mx

Información del artículo: recibido: enero de 2008, reevaluado: diciembre de 2008 y agosto de 2009, aceptado: abril de 2010

Resumen

Gracias a la evolución permanente de los algoritmos de codificación de video, la televisión digital ha encontrado un prometedor medio de transmisión a través de Internet. En este artículo, se analiza el desempeño de los codificadores comerciales: RealNetworks, Apple y VC-1 (Microsoft) con tecnología de video afluente (*video streaming*), todos estos basados en los estándares MPEG-4 y AVC/H.264. Estos codificadores representan hoy día la solución para la difusión de la televisión digital a través de la red IP. Después de presentar el esquema general de la tecnología de video afluente, se analiza el desempeño de algunos códecs en función del PSNR de la componente de luminancia (PSNR-YYUV) y compresión del video con relación a la velocidad binaria de transmisión.

Abstract

Thanks to the permanent evolution of codecs for video and the coding algorithms, every time with better performance, the television has found a new way of transmission across Internet. Taking into account these evolutions, the performance of the most commercial codecs based on MPEG-4 AVC/H.264 (RealNetworks, Apple) and VC-1 (Microsoft) standards is analyzed in this article. Nowadays, these codecs are the solutions for the diffusion of the television through IP network using video streaming technology. After presenting the general outline of the streaming video technology analyzes the performance of some codecs in terms of their most sensitive parameters.

Descriptores

- códecs
- video afluente
- video streaming
- MPEG-4 AVC/H.264
- educación

Keywords

- codecs
- video streaming
- MPEG-4 AVC, H.264
- education

Introducción

Debido al crecimiento explosivo de Internet y al incremento en la demanda de la información multimedia, el video afuente o *video streaming* ha generado un gran interés en la industria y en la academia durante la última década (García *et al.*, 2006). Efectivamente, esta tecnología ha revolucionado el mundo de las comunicaciones; ya que permite la descarga y visualización simultánea de un contenido de video sin la necesidad de guardar la información visual (Michel, 2002), (David, 2004). Sin embargo, y a pesar de las ventajas señaladas, el mayor inconveniente que presenta el video afuente al desplegar el video en tiempo real, es que el ancho de banda utilizado es directamente proporcional a la calidad del video en términos de marcos/segundo y resolución. Esto es, entre más alta es la calidad del video se requiere un ancho de banda más importante. En el caso de una conexión ISDN (Hermann *et al.*, 1991) o ADSL (Starr *et al.*, 1999), el ancho de banda sería fácilmente saturado al querer transmitir video sobre Internet. Por ello, es imprescindible usar algoritmos de compresión de video que permitan maximizar el ancho de banda. La transmisión eficiente de datos es asegurada por un sistema específico llamado códec (Lajos *et al.*, 2007), (Michael, 2002) (palabra construida por las palabras CODificador/DECodificador). El códec tiene como tareas principales:

- 1) Codificar los flujos de video nativo¹ a través de una serie de etapas, a un tamaño de datos más compacto.
- 2) Decodificar los datos para reconstruir el video original.

Algunos de éstos, proveen un flujo de datos en paquete para ser enviados a través de Internet y en la fase de decodificación reconstruye el flujo de datos provenientes de Internet. Estos componentes son llamados códecs de video afuente, los cuales son usados en diversas aplicaciones tales como; la videoconferencia, la radiodifusión en vivo (broadcast) y el video sobre demanda (VoD) (David, 2004). Estos códecs ofrecen una buena relación entre la calidad del video y el ancho de banda requerido para su transmisión.

El objetivo de este artículo es dar a conocer el desempeño de tres códecs comerciales con tecnología de video afuente, evaluándolos en función de los principales parámetros que esta tecnología requiere (cf. 3). Este artículo está organizado de la manera siguiente: se inicia con una breve descripción de la tecnología de video afuente;

posteriormente se describen las principales características de los códecs de video afuente utilizados en este análisis de desempeño; en la sección siguiente se describe la metodología empleada para la evaluación de los códecs; en la penúltima sección se muestran los resultados experimentales del desempeño de los códecs y finalmente las conclusiones de este trabajo.

Tecnología de video afuente

El término afuente (*streaming*) (David, 2004), agrupa un conjunto de tecnologías que permiten a través de una computadora o de un módulo de conexión que interactúa con el protocolo IP (*Internet Protocol*) denominado STB-IPTV (*Set Top Box-Internet Protocol TeleVision*), escuchar un contenido de audio o de observar un contenido de video de forma continua y sin la necesidad de esperar que la totalidad de la información sea descargada. El contenido es leído de una memoria provisional de almacenamiento. El afuente está regido bajo el modelo cliente/servidor.

El principio general de este conjunto de tecnologías se basa en la codificación de contenido audio/video de acuerdo a un formato y a una razón de transmisión, el cual es enviado vía Internet. El flujo de audio/video es fragmentado en paquetes para permitir su transporte hacia el usuario vía la red IP (Siyan *et al.*, 1997), (CalSoft, 2001).

Un cliente accede al contenido a través de un reproductor de audio/video. Este dispositivo de lectura memoriza un segmento de audio o de video, por lo general un segmento de 6 a 10 segundos de contenido (David, 2004), lo despliega mientras está memorizando otro segmento de contenido y así sucesivamente. El tiempo de duración de los segmentos está definido por la codificación para permitir absorber los retardos eventuales de espera generados por la red IP. Es importante recordar que todos los paquetes creados por el codificador no siguen el mismo camino y deben ser ordenados para reconstruir el contenido inicial (Siyan *et al.*, 1997), (CalSoft, 2001).

El esquema convencional para la instalación de un servicio de video afuente tiene dos actividades fundamentales (figura 1):

- a) La elaboración de contenidos en un formato digital utilizando procedimientos de compresión.
- b) La distribución de los contenidos por la red a los clientes o usuarios finales.

Para la elaboración de contenidos existe generalmente una primera fase de captura de audio/video, ya sea para eventos en directo o de contenidos grabados (cintas

¹ Se le llama video nativo a la señal de video sin comprimir.

de video, archivos multimedia) y una fase de compresión en la que se procesa el audio y el video.

La distribución de contenidos incluye la difusión de referencias, generalmente mediante URLs (*Uniform Resource Locator*) de los mismos incluidos en páginas web. Un servidor afluente (*streaming server*) puede almacenar y/o distribuir los contenidos a los clientes.

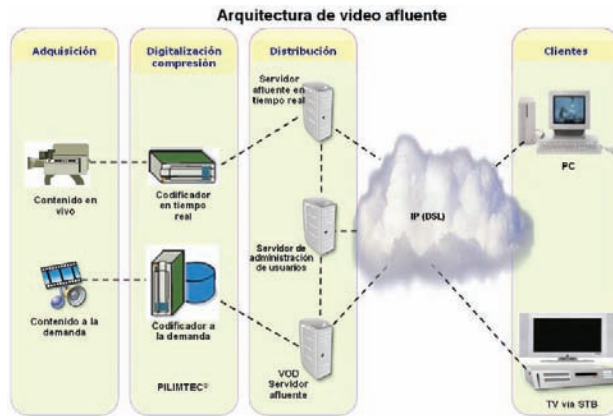


Figura 1. Arquitectura de video afluente (video streaming)

Los servidores afluentes pueden proporcionar dos tipos de contenido (David, 2004):

Afluente a la demanda (*streaming on demand*)

Los archivos multimedia provienen de un fichero digitalizado y almacenado en el servidor. Los usuarios pueden solicitar su visualización a través de Internet en cualquier momento.

Afluente en vivo (*streaming live*)

El contenido proveniente de alguna fuente multimedia, se digitaliza y se retransmite en tiempo real a Internet. Los usuarios pueden seguir el desarrollo de un evento en el mismo momento que éste se está produciendo.

Códecs de video para la difusión afluente

Las restricciones usuales de ancho de banda en las redes IP limitan la cantidad de información enviada entre servidor y cliente. La selección de un codificador que represente la mejor calidad de video para una velocidad binaria o *bitrate* determinada es crucial, si se desea ofrecer un mejor servicio a los clientes de video afluente.

Actualmente, los códecs empleados para la difusión de video por Internet están basados en el estándar MPEG-4 AVC/H.264 (ITU-T, 2010); (Iain, 2003); (Gary, 2004). Estos códecs mantienen una complejidad relativamente baja, tanto en hardware como en software,

para permitir su uso en aplicaciones interactivas de gráficos y multimedia.

El estándar MPEG-4 AVC/H.264 especifica perfiles y niveles, que tienen por objeto facilitar la interoperabilidad entre aplicaciones del estándar (Gary, 2004). Un perfil, define un conjunto de herramientas de codificación a utilizar para la generación de un flujo de bits. Por su parte, el nivel establece las limitaciones del flujo de datos como tamaño de imagen y velocidad binaria. El estándar MPEG-4 AVC/H.264 en su primera edición aprobada en Mayo del 2003 contaba con 3 perfiles: *Baseline*, *Main* y *Extended* (la edición 1.0 corresponde a la versión 1). En su segunda edición aprobada en Marzo del 2005 se le agregaron 4 nuevos perfiles referidos como *High*, *High 10*, *High 4:2:2* y *High 4:4:4* para mejorar las capacidades en calidad del video y extender el rango de aplicaciones de este estándar (la edición 2.0 corresponde a la versión 3). En su tercera edición aprobada en Noviembre del 2007 se le agregaron otros 5 nuevos perfiles el *High 10 Intra*, *High 4:2:2 Intra*, *High 4:4:4 Intra*, *CAVLC 4:4:4 Intra* y el *High 4:4:4 Predictive* para aplicaciones profesionales y otros 3 nuevos perfiles el *Scalable Baseline*, *Scalable High* y el *Scalable High Intra* para codificación de video escalable (la edición 3.0 corresponde a la versión 8). En esta edición 3.0 se quitó el perfil *High 4:4:4*. En su cuarta edición aprobada en Marzo del 2009 se le agregaron otros 2 nuevos perfiles el *Multiview High* para codificación de video multi-vistas y el *Constrained Baseline* (la edición 4.0 corresponde a la versión 11). En su quinta edición aprobada en Marzo del 2010 se le agregó un nuevo perfil el *Stereo High* para codificación de video de dos vistas (la edición 5.0 corresponde a la versión 13) (ITU-T, 2010). En resumen, en la actualidad el estándar MPEG-4 AVC/H.264 tiene 17 perfiles.

Con la aparición del estándar MPEG-4 AVC/H.264, las grandes compañías desarrolladoras de software fabricaron códecs, que si bien cumplen con las especificaciones de dicho estándar, su desempeño varía según la eficiencia de los algoritmos implementados por cada desarrollador. Muchas soluciones de video afluente están disponibles actualmente en el mercado, las cuales usan por ejemplo los códecs de *RealNetworks*, *Microsoft* y *Apple*. Estas soluciones de video afluente constan de servidor afluente, codificador y reproductor de audio/video. Para todos estos actores del mercado de video afluente, el compromiso de ser líder y de ser el primero en el mercado ha originado una verdadera y sana competencia. El precursor de esta tecnología *RealNetworks* ha perdido terreno frente al ascenso fulgurante de la solución de *Microsoft* y a la presencia de la solución de *Apple*.

Dada la importancia de estas soluciones en el mercado, los códecs evaluados son:

RealNetworks (RealNetworks, Inc. 2007):

Real Producer Plus version 11.0 rv40, programa de codificación multimedia.

RealVideo 10, códec, las características técnicas de este formato no son públicas, éstas están basadas en un prototipo del estándar H.264 (ITU-T, edition 2.0:2005), (King *et al.*, 1999), (Iain, 2003).

Microsoft (Microsoft, 2008):

Microsoft Windows Media Encoder 9 Series versión 9.00.00.3352, programa de codificación multimedia-estándar VC-1, SMPTE 421M (SMPTE, 2006).

Apple (Apple, Inc., 2008):

QuickTime 7.3.1 Pro, programa de codificación multimedia-estándar H.264.

Metodología de evaluación

La evaluación comparativa de los códecs se desarrolla con el objetivo de determinar el códec más apropiado para aplicaciones de difusión y almacenamiento de video afluente. El desempeño es evaluado en función de la calidad del video reconstruido de las secuencias con diferente tipo de movimiento y diferentes condiciones de velocidad binaria.

Parámetros de codificación

Los códecs evaluados se configuraron mediante una serie de parámetros que a continuación se enlistan:

- Velocidad binaria: número de bits procesados por unidad de tiempo en el códec. Para los tres códecs evaluados se utilizó la codificación con velocidad binaria constante (CBR).
- Pasos de codificación: se utilizó un solo paso. El contenido pasa a través del codificador una sola vez.
- Resolución: se emplearon las resoluciones QCIF (176×144) y CIF (352×288).
- GOP (Group Of Picture): IBBP.
- Formato: YUV 4:2:0.
- Marcos por segundo: 30.

Además de los parámetros mencionados, algunos otros parámetros se configuraron por omisión (definidos previamente por el constructor) para una óptima codificación.

Proceso de codificación

Las secuencias de video nativo se codificaron a diferentes velocidades binarias (tabla 1).

Los principales formatos de video afluente son *Real Media* (.rm), *Windows Media* (.wmv, .asf) y *QuickTime* (.mov).

Métrica de evaluación

La evaluación de los tres códecs (*RealNetworks*, *Microsoft* y *Apple*) se basó en el análisis de su desempeño y compromiso con respecto a la relación de *calidad de reconstrucción/velocidad binaria de transmisión* y a la relación de *tamaño de archivo comprimido generado/ velocidad binaria de transmisión*.

Para comparar la calidad de las secuencias de video reconstruidas, se utiliza el criterio de medida de calidad más utilizado en procesamiento digital de imágenes: el *Peak Signal to Noise Ratio* (PSNR) (Winkler, 2005), (Winkler, 1999), el cual está definido por la ecuación (2). Esta razón señal/ruido está medida bajo una escala logarítmica y dependiente de la media de la Suma del Cuadrado de las Diferencias (SSD) (Tuukka, 2002), vea la ecuación (1). Entre más grande sea esta razón mayor será la calidad de la imagen. Hay que hacer notar que una comparación basada en la medida PSNR entre dos imágenes diferentes debe ser tratada con precaución, ya que una imagen que posee un PSNR bajo puede, sin embargo, poseer una mejor calidad subjetiva a simple vista, sobre una imagen que posee un mayor PSNR.

$$SSD = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (X_{i,j} - Y_{i,j})^2 \quad (1)$$

$$PSNR_{db} = 10 \log_{10} \frac{(2^n - 1)N \times M}{SSD} \quad (2)$$

Tabla 1. Rango de velocidades binarias para la codificación

Resolución	Velocidad binaria (kbps)										
	20	56	64	128	150	256	350	500	750	1000	2000
QCIF	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
CIF		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

La ecuación del PSNR se calcula para una imagen con resolución $N \times M$ píxeles, n representa el número de bits por componente, 8 bits en este caso. En la ecuación (1), X e Y representan la imagen original y la imagen procesada, respectivamente. El PSNR calculado en este artículo es el PSNR de la componente de luminancia (PSNR-YYUV), ya que tiene un mayor efecto en la percepción del sistema de visión humano, a este valor de PSNR se le conoce como PSNR relativo. En las últimas décadas se ha utilizado como criterio principal para evaluar el desempeño de los códecs respecto a la calidad visual de las secuencias de video.

Biblioteca de las secuencias de prueba

La biblioteca de prueba está constituida por 8 secuencias de video en formato nativo YUV 4:2:0 (tablas 2 y 3). Estas secuencias son utilizadas ampliamente en el campo de la compresión de video debido a la riqueza en la información contenida: diferentes tipos de movimiento y variedad de texturas (ITU-R., 1994).

Resultados experimentales

En la sección anterior se explicó bajo qué condiciones se realizaron las pruebas, es decir, los parámetros de codificación para los tres códecs, el proceso de codificación y la métrica de evaluación empleada. En resumen, se utilizaron 8 secuencias de prueba para analizar el desempeño de los tres códecs en cuanto a la razón de compresión y la calidad de reconstrucción cuantitativa PSNR-YYUV. Dado el proceso de codificación/decodificación, los resultados por secuencia son los siguientes:

Secuencia *Carphone*: resolución tipo videoconferencia, no presenta movimiento de la cámara. Sin embargo, el movimiento facial es importante por lo cercano de la cámara. Otro elemento a considerar, es la parte derecha del fondo de la imagen en donde se presenta un movimiento importante debido a que el auto en donde se lleva a cabo la adquisición de la secuencia, se encuentra en marcha. Se puede observar en la figura 2, que para velocidades binarias entre 20-350 kbps, el códec Windows presenta un mayor PSNR-YYUV. Para velocidades binarias mayores que 350 kbps, el códec

Tabla 2. Biblioteca de secuencias de prueba

Resolución QCIF (176×144 píxeles), aplicación: video conferencia, video móvil		
Secuencia	Frecuencia marcos/seg.	Cantidad de marcos
Carphone	30	382
Highway	30	2000
Hall	30	300
Container	30	300

Tabla 3. Biblioteca de secuencias de prueba

Resolución CIF (352×288 píxeles), aplicación: monitoreo de video, video web		
Secuencia	Frecuencia marcos/seg.	Cantidad de marcos
Foreman	30	300
Mobile	30	300
News	30	300
Stefan	30	090

QuickTime presenta un mejor desempeño con relación al PSNR-YYUV. Con respecto a la compresión, el códec que mejor comprime es el códec RealMedia para todas las velocidades binarias. Note que la diferencia en compresión entre los códecs se marca más a partir de 750 kbps. Se puede observar que para valores entre 300 y 1000 kbps el desempeño del códec RealMedia no varía

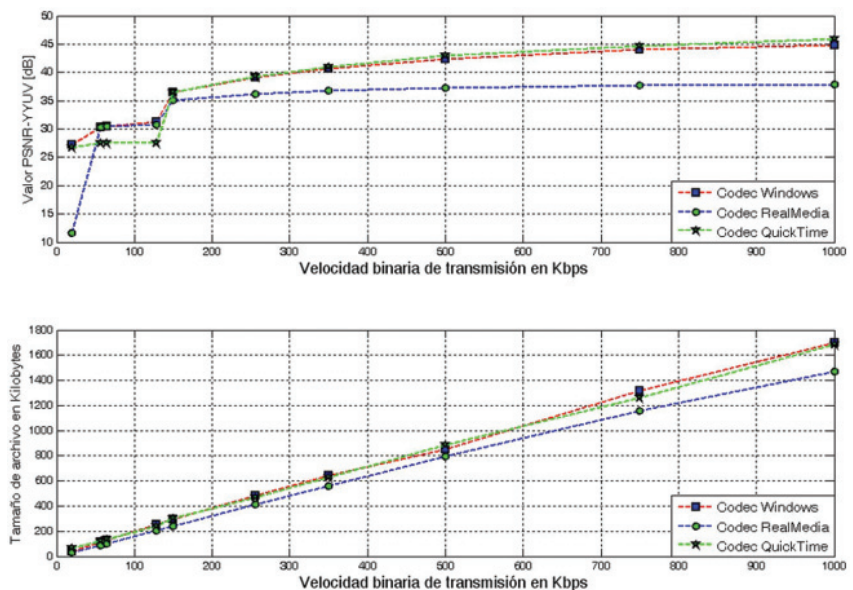


Figura 2. Secuencia Carphone, QCIF. Desempeño de los códecs en PSNR-YYUV y compresión del video con relación a la velocidad binaria de transmisión

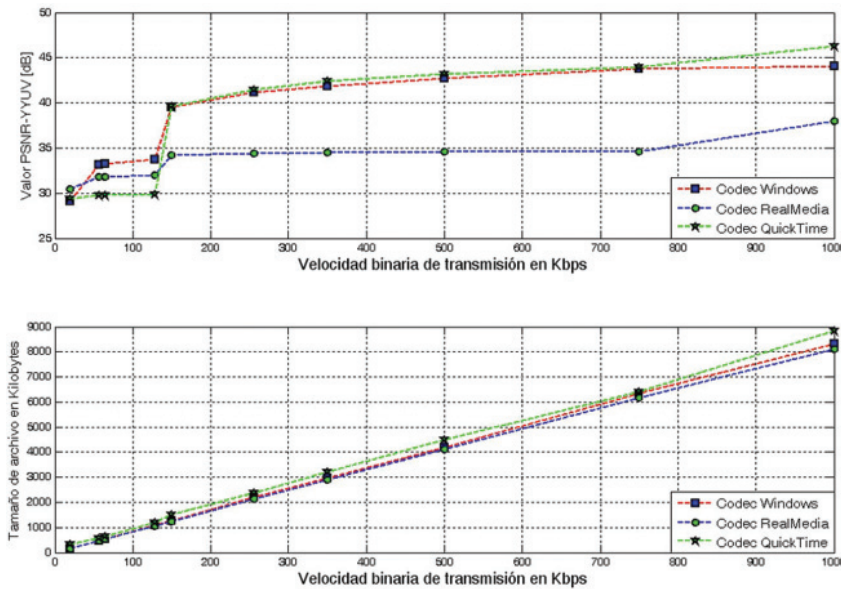


Figura 3. Secuencia Highway, QCIF. Desempeño de los códecs en PSNR-YYUV y compresión del video con relación a la velocidad binaria de transmisión

mucho. Esto es debido a su bajo desempeño de sus algoritmos en el modelo espacial del códec.

Secuencia *Highway*: tiene un acercamiento de cámara y en el fondo presenta cambios de textura por la maleza en la carretera. En este caso, a 20 kbps el códec RealMedia presenta un PSNR-YYUV ligeramente mayor a los otros dos códecs (figura 3). Para velocidades binarias entre 56-256 kbps, el códec Windows sigue presentando un mayor PSNR-YYUV. El códec QuickTime sigue presentando un mejor desempeño de PSNR-YYUV para velocidades binarias mayores que 256 kbps. Con respecto a la compresión, el códec RealMedia es ligeramente superior a los otros dos códecs. Para esta secuencia, se observa que el códec RealMedia tiene el más bajo PSNR-YYUV en general.

Secuencia *Hall*: una persona se desplaza hacia el fondo del pasillo dando la espalda a la cámara, la cual no presenta movimiento. Aquí, la dificultad se presenta al estimar el movimiento correcto de la persona. En esta secuencia, el códec RealMedia tiene mejor desempeño a 20 kbps. El códec Windows da un mejor desempeño de PSNR-YYUV en el rango de 56-350 kbps. Para valores mayores que 350 kbps, el que mejor se

desempeña es el códec QuickTime (figura 4). Con relación a la compresión, los tres códecs presentan valores similares en el rango de 20-500 kbps. Para valores mayores que 500 kbps, el que presenta un mejor desempeño en compresión de archivo es el códec RealMedia.

Secuencia *Container*: presenta un movimiento de cámara de traslación, además del movimiento del barco que se desplaza en la misma dirección de la cámara. Para una velocidad binaria de 20 kbps, el códec RealMedia

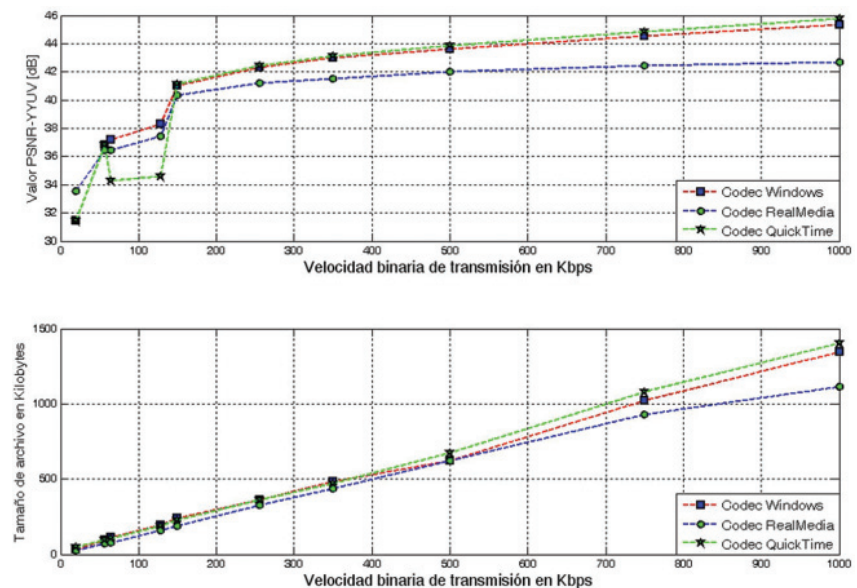


Figura 4. Secuencia Hall, QCIF. Desempeño de los códecs en PSNR-YYUV y compresión del video con relación a la velocidad binaria de transmisión

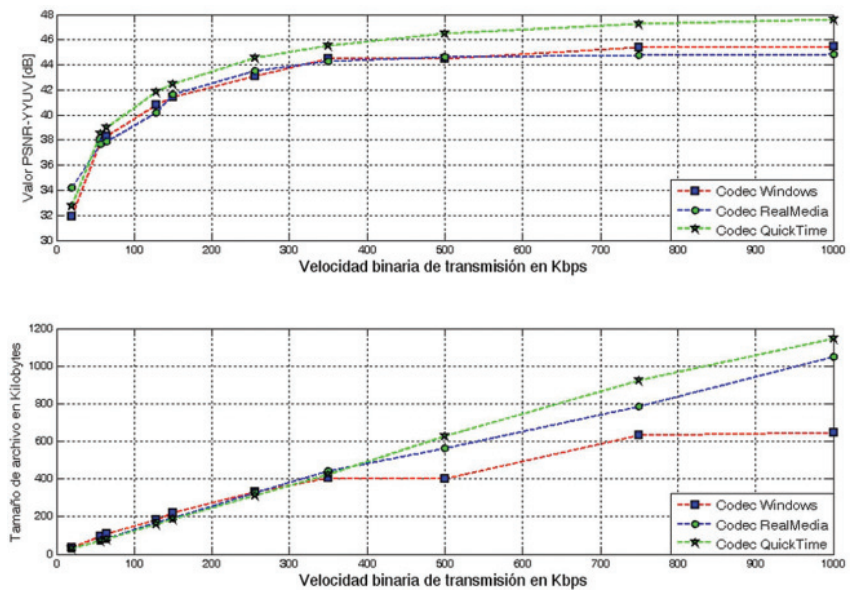


Figura 5. Secuencia Container, QCIF. Desempeño de los códecs en PSNR-YYUV y compresión del video con relación a la velocidad binaria de transmisión

presenta un mayor desempeño de PSNR-YYUV. El códec QuickTime tiene el mayor desempeño de PSNR-YYUV a partir de los 56 kbps (figura 5). Con respecto a compresión, en el rango de 20-350 kbps, los tres códecs presentan un desempeño similar. Para valores mayores a 350 kbps el códec Windows presenta la mayor compresión.

Sridhar y colaboradores en (Sridhar *et al.*, 2004) realizaron un análisis del desempeño del códec Windows Media serie 9, donde utilizan la secuencia Container y Stefan en sus pruebas. El PSNR es también calculado en la parte de luminancia de las secuencia. Utilizan 15 imágenes por segundo y sólo imágenes I y P. En este artículo, se utilizaron imágenes I, B y P, a 30 imágenes por segundo. Dadas estas diferencias, no se realizó una comparación de fondo de resultados del artículo de Sridhar y colaboradores, con los presentados en este artículo. Sin embargo, podemos mencionar que los resultados para las secuencias Container y Stefan en este artículo, tienen un mejor desempeño con respecto a los dados por Sridhar. Esto se debe principalmente, a que en este artículo se utilizó la última actualización del códec de Windows (2008) y ellos emplearon una versión del 2003. Otro factor que hace la diferencia en los resultados, es que los parámetros de codificación son diferentes. De esta manera, vemos que el desempeño del códec de Windows ha evolucionado.

En resumen, para resoluciones QCIF y los diferentes grados de dificultad para el procesamiento de estas cuatro secuencias, se puede observar que para una velocidad binaria baja de 20 kbps, el códec RealMedia es el que mejor se comporta en términos de PSNR-YYUV

y compresión. Para velocidades binarias en el rango de 56-350 kbps, el códec Windows es el que mejor desempeño tiene en términos de PSNR-YYUV y compresión, con excepción de la secuencia Container, donde su PSNR-YYUV es menor a la del códec QuickTime. Para velocidades binarias mayores que 350 kbps, el códec QuickTime da el mejor comportamiento en PSNR-YYUV. Sin embargo, el desempeño de éste, en compresión, es inferior con respecto a los otros dos códecs. El códec RealMedia es el que mejor se desempeña en términos de compresión, con excepción de la secuencia Container, donde el tamaño de archivo comprimido es superior al del códec Windows. Se puede resaltar que los códecs Windows y Quicktime no tienen un buen desempeño en cuanto a compresión para secuencias que presentan movimientos importantes (Carphone y Hall), esto es debido al bajo desempeño de sus algoritmos que participan en la etapa de estimación y compensación de movimiento y en la etapa de entropía. Sin embargo, cabe mencionar que el códec Windows tiene un muy buen desempeño en cuanto a compresión para secuencias con poco movimiento (Container).

Secuencia *Foreman*: existe un movimiento importante de la cara de la persona; debido al acercamiento de la cámara. En este caso, los 3 códecs presentan desempeños similares en PSNR-YYUV. A una velocidad binaria de 128 kbps, el códec QuickTime presenta un desempeño en PSNR-YYUV muy bajo con respecto a los otros dos códecs (figura 6). Los tres códecs presentan una compresión similar en el rango de 56-350 kbps. Para velocidades binarias mayores que 350 kbps, el códec Real-

Media tiene una mejor compresión. En esta secuencia para velocidades binarias mayores que 750 kbps, el códec Windows presenta una compresión inferior a los otros dos códec.

Secuencia *Mobile*: presenta la característica de diferentes tipos de movimiento de cámara. Además, los obje-

tos presentes en ella, tienen movimiento en diferentes direcciones. En este caso, el códec RealMedia presenta un buen desempeño con relación al PSNR-YYUV en todas las velocidades binarias, seguido muy de cerca por los otros dos códec. El códec QuickTime presenta un desempeño en PSNR-YYUV muy bajo a una velocidad binaria de 128 kbps (figura 7). Para velocidades binarias inferiores a 750 kbps, los códec RealMedia y QuickTime presentan una mejor compresión. Para valores mayores que 750 kbps el códec RealMedia tiene un mejor desempeño en compresión y el códec Windows presenta el menor desempeño.

Secuencia *News*: no presenta movimiento de cámara. Sin embargo, el fondo contiene movimiento debido a escenas de ballet. Los objetos de enfrente (2 personas) poseen movimiento ligero de cara. En esta secuencia, en el rango de 56-64 kbps, el códec RealMedia tiene un mayor PSNR-YYUV. Para valores entre 128-150 el códec RealMedia y el códec Windows presentan valores de PSNR-YYUV muy similares. Para velocidades binarias mayores que 256 kbps, los tres códec presentan valores de PSNR-YYUV muy similares (figura 8). Referente a la compresión, para velocidades binarias inferiores a 1000 kbps, el códec RealMedia tiene una mejor compresión. Para valores mayores que 1000 kbps, el códec Windows presenta una alta compresión con respecto a los otros dos códec.

Secuencia *Stefan*: presenta diferentes movimientos de cámara (translación y acercamiento), además del movimiento del objeto principal. La textura del fondo es importante. Para velocidades binarias entre 56-64 kbps, los tres códec presentan valores de PSNR-YYUV similares. Para 128 kbps, el códec Windows y el códec RealMedia tienen valores de PSNR-YYUV similares. Para 150 kbps, el mayor PSNR-YYUV lo logra el códec Windows. Para velocidades binarias mayores que 150

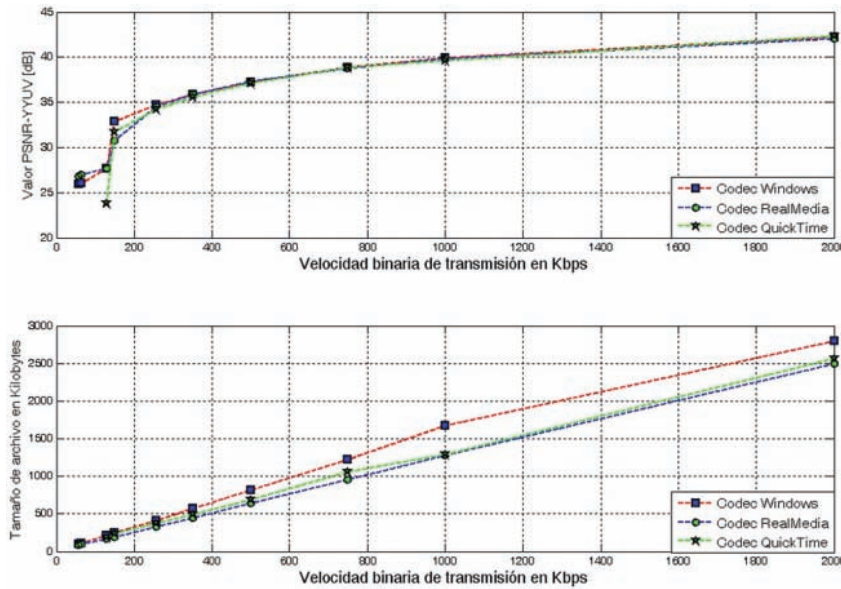


Figura 6. Secuencia Foreman, CIF. Desempeño de los códec en PSNR-YYUV y compresión del video con relación a la velocidad binaria de transmisión

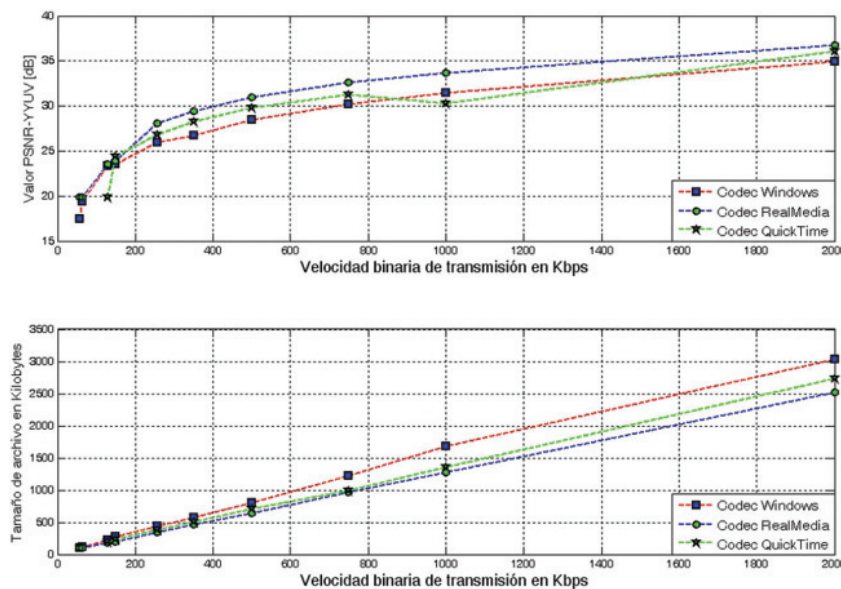


Figura 7. Secuencia Mobile, CIF. Desempeño de los códec en PSNR-YYUV y compresión del video con relación a la velocidad binaria de transmisión

kbps, el que mejor respuesta tiene en PSNR-YYUV es el códec Windows, seguido muy de cerca por los otros dos códec (figura 9). Los códec RealMedia y QuickTime tienen una mayor compresión que el códec Windows para todas las velocidades binarias.

En resumen, para resoluciones CIF y los diferentes grados de dificultad para el procesamiento de estas otras cuatro secuencias, se observa que para velocidades binarias en el rango de 56-350 kbps el códec RealMedia es el que mejor desempeño tiene en términos de PSNR-YYUV y compresión. Sin embargo, una excepción se presenta en el valor de 150 kbps para las secuencias *Foreman* y *Stefan*, donde el códec Windows tiene un mejor desempeño de PSNR-YYUV. El códec QuickTime presenta en todas las secuencias un mal desempeño de PSNR-YYUV para una velocidad binaria de 128 kbps. Para velocidades binarias mayores que 350 kbps, el códec Windows da el mejor comportamiento en PSNR-YYUV. Con relación a la compresión, el códec Windows es inferior respecto a los otros dos códec. Una excepción se marca en la secuencia *News* donde presenta una compresión mayor a los otros dos códec a partir de 1000 kbps. Con los resultados de estas otras cuatro secuencias, se reafirma que los códec Windows y QuickTime siguen presentando un no muy buen desempeño en cuanto a compresión para secuencias que presentan movimientos importantes (*Mobile* y *Stefan*). El códec Windows sigue presentando un muy buen desempeño en cuanto a compresión para secuencias con poco movimiento, en este caso, la secuencia *News* (figura 8) tiene muy poco movimiento; por lo que el códec Windows se desempeña mejor que los otros dos, presentando una mejor compresión para velocidades de transmisión altas. Las pruebas realizadas a la secuencia *News* mostraron que para velocidades de transmisión mayores a 2Mbps el tamaño del archivo ya no varía de ma-

nera importante con respecto al tamaño del archivo obtenido en la figura 8 y se estabiliza a un cierto tamaño después de cierta velocidad de transmisión; así que no por aumentar la velocidad de transmisión el tamaño del archivo seguirá disminuyendo. Este comportamiento del códec Windows es debido a la eficiencia de sus algoritmos en el modelo espacial del códec.

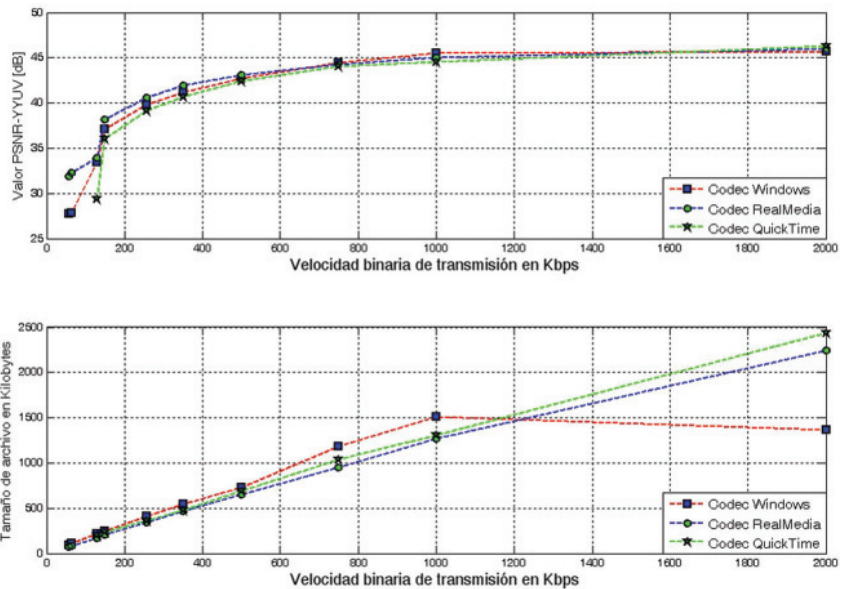


Figura 8. Secuencia News, CIF. Desempeño de los códec en PSNR-YYUV y compresión del video con relación a la velocidad binaria de transmisión

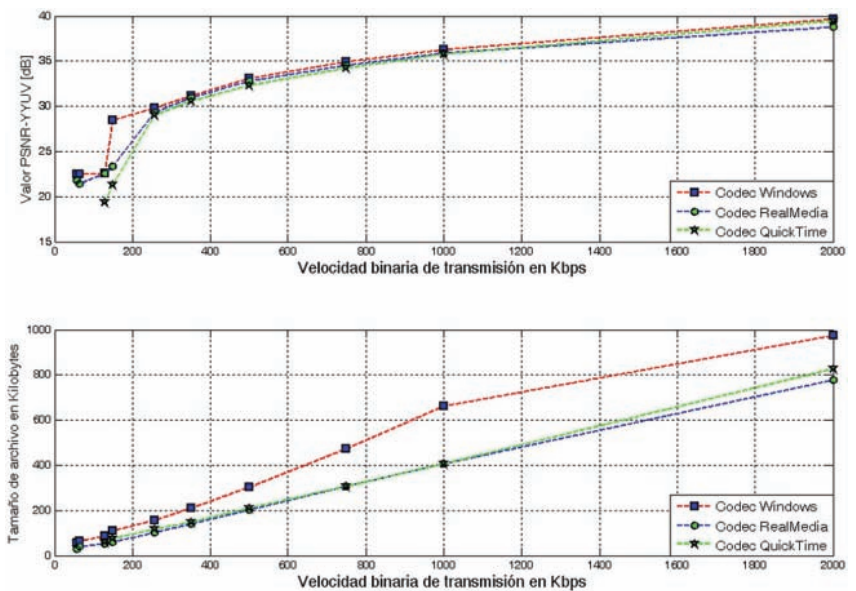


Figura 9. Secuencia Stefan, CIF. Desempeño de los códec en PSNR-YYUV y compresión del video con relación a la velocidad binaria de transmisión

Conclusiones

En este artículo se realizó el análisis del desempeño de los tres principales códecs comerciales, utilizados en la tecnología de video afluente, en función de la compresión y de la calidad de reconstrucción cuantitativa. En resumen, se puede concluir que para una velocidad binaria de 20 kbps y resoluciones QCIF, el códec RealMedia es el más adecuado a utilizar, dados los resultados de *calidad cuantitativa/compresión*; sin embargo, para velocidades binarias en el rango de 56-350 kbps con resoluciones QCIF y CIF el códec Windows y el códec RealMedia mostraron un mejor desempeño. Para velocidades binarias mayores que 350 kbps con resolución QCIF, el códec QuickTime es el más adecuado a utilizar si se desea una buena *calidad cuantitativa* y el códec RealMedia es el indicado si se desea una *compresión* alta. Para resoluciones CIF y velocidades binarias mayores que 350 kbps, el códec Windows es el mejor si se desea una buena calidad y el códec RealMedia si se desea una compresión alta.

Podemos concluir que no existe un códec universal que sea el óptimo para todos los valores de velocidad binaria y cualquier tipo de secuencia de video a procesar; por lo que se recomienda utilizar el códec específico a la velocidad binaria, al contenido a procesar y el compromiso *calidad cuantitativa/compresión* que se desee obtener.

Agradecimientos

Este trabajo fue realizado dentro del marco del proyecto SIP2009 bajo el número SIP20090041.

Referencias

- Apple, Inc. 2008 [en línea]. Disponible en: <http://www.apple.com>
- Apple, Inc. 2008 [en línea]. Disponible en: <http://www.apple.com/quicktime/pro/>
- CalSoft Labs. 2001. Basic Streaming Technology and RTSP Protocol. California Software Corporation [en línea]. Disponible en: www.cswl.com
- David-Austerberry. *Technology of Video and Audio Streaming*. United States of America. Elsevier Science & Technology Books. 2004.
- García M.S., Ramírez A.A. Video sobre IP: introducción. 2o Encuentro Regional Académico. Tijuana, Baja California, México. Noviembre 2006.
- Gary-Sullivan J. The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions. Conference on Applications of Digital Image Processing XXVII, 2004.
- Hermann J.H. *Integrated Services Digital Networks-Architectures/Protocols/Standards*. Addison Wesley. 1991.
- Iain-E.G.R. *H.264 and MPEG-4 Video Compression, Video Coding for Next Generation Multimedia*. John Wiley & Sons. 2003.
- ITU-T Recommendation H.264 & ISO/IEC 14496-10 (MPEG-4) AVC. Series H: Audiovisual and Multimedia Systems. Infrastructure of Audiovisual Services-Coding of Moving Video. Advanced Video Coding for Generic Audiovisual Services. (Edition 1.0:2003, Edition 2.0:2005, Edition 3.0:2007, Edition 4.0:2009). Current Edition 5.0:2010. 2010.
- ITU-R. Encoding Parameters of Digital Television for Studios. Recommendation BT. 601-4. 1994.
- King N.N., Thomas M., Douglas Ch. *Advanced Video Coding: Principles and Techniques*. New York, NY, USA. Elsevier Science Inc. 1999. ISBN 044482667X.
- Lajos H., Peter Ch., Jurgen S. *Video Compression and Communications: from Basics to H.261, H.263, H.264, MPEG4 for DVB and HSDPA-Style Adaptive Turbo-Transceivers*. Wiley-IEEE Press. 2007. ISBN 978-0-470-51849-6.
- Michael T. *Streaming Media Demystified*. United States of America. McGraw-Hill. 2002.
- Microsoft 2008 [en línea]. Disponible en: <http://www.microsoft.com/en/us/default.aspx>
- Microsoft 2008 [en línea]. Disponible en: <http://www.microsoft.com/windows/windowsmedia/forpros/encoder/default.msp>
- MSU Video Quality Measurement Tool 2010 [en línea]. Disponible en: <http://compression.ru>.
- Koenen R. Overview of the MPEG-4 Standard. International Organization for Standardization 2002 [en línea]. Disponible en: www.chiariglione.org/mpeg/standards
- RealNetworks, Inc. 2007 [en línea]. Disponible en: <http://www.realnetworks.com>
- RealNetworks, Inc. 2007 [en línea]. Disponible en: <http://www.realnetworks.com/products/producer/index.html>
- Siyan K. *Inside TCP/IP*. New Riders Publishing. 1997. ISBN 1-56205-714-6.
- SMPTE 421M. VC-1 Compressed Video Bitstream Format and Decoding Process. 2006.
- Starr T., John M.C., Peter J.S. *Understanding Digital Subscriber Line Technology*. Upper saddle River. NJ. Prentice-Hall, Inc. 1999.
- Sridhar-Srinivasan P., Hsu J., Holcomb T., Kunal-Mukerjee., Shankar L.R., Bruce-Lin. J.L., Ming-Chieh L., Ribas-Corbera J. 2004. Windows Media Video 9: Overview and applications. Signal Processing Image Communications 19, 2004, 851-875. Elsevier.

- Tuukka-Toivonen Number Theoretic Transform-Based Block Motion Estimation. PhD dissertation, Department of Electrical Engineering, University of Oulu, Oulu, Finland. 2002.
- Wikipedia (Jan. 2008). The free encyclopedia. Streaming media [en línea]. Disponible en: http://en.wikipedia.org/wiki/Streaming_media
- Wikipedia (Jan. 2008). The free encyclopedia. Video on demand [en línea]. Disponible en: http://en.wikipedia.org/wiki/Video_on_demand
- Winkler-Stefan. *Digital Video Quality*. Wiley. March. 2005.
- Winkler S. A Perceptual Distortion Metric for Digital Color Video. *Proc. SPIE*, 3644:175-184. 1999.

Semblanza de los autores

Mireya Sarai García-Vázquez. Recibió el grado de doctor en señales y telecomunicaciones en 2004 por el Institute of Computer Science and Random System Research (IRISA) Rennes, Francia. Realizó en 2005 sus estudios de posdoctorado en el Laboratorio de Investigación y Desarrollo de PILIMTEC en la incubadora de empresas de l'École Nationale Supérieure des Télécommunications de Bretagne (ENST-Bretagne), Brest, Francia. Actualmente es profesor en el CITEDI-IPN, Tijuana, México. Sus áreas de investigación son procesamiento de imagen, compresión de video, codificación de audio, comunicaciones, IPTV y TV móvil, en particular, estimación de movimiento, técnicas de cancelamiento de error y estándares de compresión de video. Es miembro del Sistema Nacional de Investigadores (SNI) de CONACYT.

Alejandro Álvaro Ramírez-Acosta. Recibió el grado de doctor en electrónica y telecomunicaciones en 2000 de l'École Nationale Supérieure des Télécommunications de Bretagne (ENST-Bretagne) y de la Université de Bretagne Occidentale (UBO), Brest, Francia. Realizó en 2001 sus estudios de posdoctorado en el Laboratorio de Investigación y Desarrollo de FRANCE TÉLÉCOM, Lannion, Francia. Actualmente es profesor-investigador de CITEDI-IPN. Sus áreas de investigación son codificación de video/audio/voz, reconocimiento de patrones, estándares de compresión de audio/video, codificación conjunta fuente-canal, desarrollo de Frameworks Multimedia OpenMAX, desarrollo de sistemas de comunicaciones con la familia TI TMS320C6000, técnicas de cancelamiento de error, cuantificación vectorial, IPTV, VoIP y TV móvil, teoría y aplicaciones.

Victor García-Garduño. Recibió el grado de doctor en señales y telecomunicaciones en 1995 del Institute of Computer Science and Random System Research (IRISA) Rennes, Francia. Actualmente es profesor en la Facultad de Ingeniería de la Universidad Nacional Autónoma de México (UNAM). Sus áreas de investigación son compresión y transmisión de video digital, análisis de secuencias de imágenes: estudio de alternativas a las normas de comunicación MPEG-2 y H.261.