

PRUEBAS OBJETIVAS DE CALIDAD DE VIDEO PARA TELEVISIÓN DE DEFINICIÓN ESTÁNDAR.

Osmany Yaunner Núñez¹, Orlando Landrove Gámez², Rufino R. Cabrera Álvarez³

LACETEL, Instituto de Investigación y Desarrollo de Telecomunicaciones, Ave. Independencia, Km. 14½, No. 34818, Boyeros, La Habana.

¹e-mail: osmany@lacetel.cu

²e-mail: landrove@lacetel.cu

³e-mail: rufino@lacetel.cu

RESUMEN

Las técnicas de estimación de calidad de video son de gran utilidad para los operadores de servicio y generadores de contenido, dado que permiten monitorear y evaluar la imagen recibida. En un sistema de procesamiento y transmisión de video, la señal siempre está expuesta a degradaciones, distorsiones, pérdidas de cuadros y otros fenómenos indeseables que afectan la calidad de la imagen percibida por los usuarios. Lo anterior se pone de manifiesto en el esquema de implementación actual de la Televisión Digital en Cuba. La calidad percibida en diferentes programas con el mismo contenido en el mismo momento varía notablemente. Para cuantificar y poder incidir sobre el efecto anterior, **LACETEL**, Instituto de Investigación y Desarrollo de Telecomunicaciones, como parte de la asimilación de las tecnologías asociadas al despliegue de la TVD, ha ejecutado pruebas objetivas de calidad de video. Se grabaron y analizaron fragmentos de los programas Cubavisión, Cubavisión Internacional e Infantil/Canal Habana en definición estándar mientras transmitían el mismo contenido. Mediante la herramienta de análisis de calidad Elecard se obtuvieron valores procesados matemáticamente con los que se determinan las diferencias objetivas de calidad de los videos analizados. Los resultados objetivos de los métodos y algoritmos de medición permiten cuantificar y avalan las diferencias de percepción visual subjetiva en la comparación de estos programas.

PALABRAS CLAVES: estimación, calidad de video, métodos objetivos, sistema de Televisión Digital Terrestre, Definición Estándar.

ABSTRACT

The video quality estimation techniques are very useful for service operators and content generators, to monitor and assess the received image. On a video processing/transmission system, the signal is always exposed to degradations, distortions, frame losses among other unwanted phenomena that affect the image quality perceived by users. These effects are included on the current implementation of Digital Television in Cuba. The perceived quality on different programs with the same content at the same time has a noticeable variation. In order to quantify and influence on the previous effect, **LACETEL**, Research and Development Telecommunications Institute, as part of the technologies assimilation related to DTV deployment, has carried out objective video quality tests. Fragments of "Cubavisión", "Cubavisión Internacional" and "Infantil/Canal Habana" programs in standard definition were recorded and analyzed

while transmitting the same content. Mathematically processed values were obtained with the Elecard quality analysis tool, which are used to determine the quality objective differences of analyzed videos. The objective results of measurement methods and algorithms allow to quantify and validate the differences of subjective visual perception comparing these programs.

KEYWORDS: estimation, video quality, objective methods, Terrestrial Digital Television system, Standard Definition.

INTRODUCCIÓN

Cuba se encuentra en estos momentos en el proceso de despliegue de la TDT. La mayoría del territorio nacional tiene cobertura de la señal digital, constituyendo el proceso de masificación de los receptores el paso clave para estar listo para el llamado apagón analógico. Sin embargo, poco a poco la penetración de estos va en aumento, y actualmente más de 500 000 usuarios pueden disfrutar de las ventajas de este servicio.

Entre otras bondades, esta tecnología en nuestro país ofrece a los usuarios una calidad uniforme en la recepción para todos los programas digitales. El acceso a nuevos servicios constituye un elemento importante, donde aparecen los servicios de EPG (Electronic Program Guide, por sus siglas en inglés, Guía Electrónica de Programa), Radiodifusión de Datos, posibilidad de grabación del programa mientras se visualiza otro, posibilidad de pausar y luego continuar la recepción mediante la grabación y la presentación de la información como procesos independientes todo en tiempo real. También es posible y de hecho en nuestras transmisiones de TVD se insertan programas de radio multiplexados con los programas de televisión. Además, la calidad del audio y el video debe ser mayor en comparación con el servicio tradicional de Televisión Analógica.

El video digital, distribuido a través de la red de telecomunicaciones, sufre de varios tipos de distorsión durante el proceso de adquisición, compresión, procesamiento, transmisión y reproducción. Las técnicas usadas en el proceso de codificación digital del video introducen pérdidas de la información con el propósito de reducir el ancho de banda necesario para ajustarse dentro del canal de transmisión disponible.

Resulta importante, principalmente para el proveedor de servicio, conocer la calidad de las aplicaciones utilizadas por sus usuarios. En Cuba la aplicación sería la recepción de los programas de televisión y de radio, y el servicio de datos. Pero específicamente para audio y video, es deseable medir la calidad percibida, a los efectos de brindar mayor información útil con una utilización más óptima de la infraestructura disponible. Generalmente los operadores utilizan una realimentación de la información, desde el receptor al emisor, de manera que se tenga total control de la calidad perceptual, y toman acciones determinadas para aumentar o disminuir ancho de banda de un flujo de video cuando las condiciones lo impongan.

Los métodos de medición de la calidad de video se dividen en dos tipos, los métodos subjetivos y los objetivos. Los métodos subjetivos están basados en la percepción visual humana, y serán los sistemas de medidas más confiables, al ser el resultado de la opinión directa de los observadores. La opinión de los

evaluadores es promediada, obteniéndose el parámetro MOS (Mean Opinion Score, por sus siglas en inglés, Puntuación de Opinión Media). Existen estándares que indican como realizar estas evaluaciones. Sin embargo, su realización es costosa y compleja ya que es necesario implementarlas en ambientes controlados, disponer de un número importante de evaluadores y de contenidos multimedia especialmente preparados. Adicionalmente no responde a la necesidad de evaluar en el momento determinado en que se requiera, además que es difícil medir la repetitividad de la prueba.

Los métodos objetivos son sistemas de evaluación basados en métodos y algoritmos matemáticos que permiten estimar la calidad percibida por los usuarios. Un sistema ideal de estimación de calidad percibida debería dar como resultado una calificación casi idéntica a la que se obtendría en pruebas subjetivas promediando los resultados de un gran número de individuos. VQEG (Video Quality Experts Group, por sus siglas en inglés, Grupo de Expertos de Calidad de Video) se ha destacado en el proceso de análisis y comparación de modelos de estimación de calidad de video y las estandarizaciones de los procedimientos de pruebas. Los métodos objetivos pueden ser usados en aplicaciones en tiempo real para la monitorización y evaluación. Pueden ser repetitivas buscando un término medio en sus resultados. El costo de utilizarlo puede variar en dependencia del nivel de precisión buscado, pero en muchos casos son eficaces para determinar conclusiones, con un esfuerzo mucho menor que las evaluaciones subjetivas. [1]

Los métodos objetivos de análisis del contenido de video se pueden clasificar según la disponibilidad total, parcial o nula de la señal original. Los de disponibilidad total se denominarán FR (Full Reference, por sus siglas en inglés, Referencia Completa) y se basan en poseer la señal original, la que puede ser comparada con la señal degradada cuadro a cuadro. Estos métodos pueden ser importantes para categorizar de forma objetiva un sistema de transmisión, un códec, el efecto de un ancho de banda reducido u otros factores que degraden la señal. Generalmente se utilizan en ambientes controlados, como dedicados a experimentos. Por tanto, estos no son adecuados para aplicaciones en tiempo real (como conferencias, por ejemplo) ya que es difícil disponer de las señales originales junto con las señales degradadas. [2]

Los métodos de disponibilidad parcial de la señal original se llaman métodos RR (Reduced Reference, por sus siglas en inglés, Referencia Reducida) y se basan en enviar, en un canal auxiliar, junto con el video codificado, algunos parámetros que caractericen a la señal. De esta manera sirve como referencia en el receptor para poder estimar la calidad percibida. [2]

Los modelos del tipo NR (No Reference, por sus siglas en inglés, Sin Referencia) intentan estimar la calidad percibida basándose únicamente en el análisis de la señal recibida. Son métodos más complejos de implementar, pero no requieren otra información que la propia señal de video. Ya que los seres humanos no necesitan de referencia para juzgar al menos de forma general la calidad de un video, sino que se basan en la experiencia y expectativas, estos sistemas buscan realizar el mismo comportamiento, por tanto son los más complejos a utilizar. [2]

Los algoritmos usados en las técnicas FR se pueden categorizar en parámetros básicos y parámetros perceptuales. Los primeros calculan la diferencia píxel a píxel entre las imágenes original y la degradada. Entre estos métodos aparecen MSE (Mean Square Error, por sus siglas en inglés, Error Cuadrático Medio), PSNR (Peak Signal to Noise Ratio, por sus siglas en inglés, Razón Señal a Ruido Pico), MSAD, DELTA. Las segundas son más sofisticadas y tratan de buscar una percepción más parecida al sistema visual humano. Algunas de ellas son SSIM (Structural Similarity, por sus siglas en inglés, Similitud

Estructural), NQI (New Quality Index, por sus siglas en inglés, Nuevo Índice de Calidad), VQM (Video Quality Metric, por sus siglas en inglés, Métrica de Calidad de Video). [3]

El MSE da la diferencia cuadrática media entre la imagen de prueba y la referencia. MSE sugiere mayor similitud cuando más pequeño sea este valor. Este parámetro, como mismo PSNR, se basa en comparación píxel a píxel entre las imágenes. A veces esta medida no brinda total correspondencia con el sistema visual humano. En ocasiones el sistema visual humano reconoce mejoría en imágenes en que se añade cierto ruido. En ese caso la medición MSE arrojaría mayor valor, por tanto, sugeriría menor correspondencia entre las imágenes, cuando la práctica visual demuestra lo contrario. MSE es usado regularmente para evaluar objetivamente la calidad de video. Su resultado, junto con el de otros métodos similares, da una idea más completa para comparar la correspondencia entre imágenes. [4]

PSNR es una de las más utilizadas en la práctica para evaluar calidad de codificación y decodificación. Es una expresión de relación entre la máxima energía posible de una señal y el ruido que afecta la calidad de su representación. PSNR generalmente está representado en términos de escala logarítmica en decibeles. Sin embargo, la propia naturaleza de este método hace que en ocasiones no se ajusten completamente al resultado arrojado a partir de una observación subjetiva. [4]

MSAD se define como la diferencia máxima absoluta entre los componentes de color. En la figura 1 se muestra un ejemplo visual de su método, a partir del resultado entre las imágenes de referencia y la procesada. [4]



Figura 1: Representación visual de MSAD

DELTA define la diferencia máxima entre los componentes de color. En la figura 2 se muestra el resultado entre la comparación de una imagen de referencia y su procesado. [4]



Figura 2: Representación visual de DELTA

Entre las técnicas perceptuales, SSIM está basada en la medición de tres componentes: luminancia, contraste y similitudes estructurales. SSIM requiere que dos imágenes a comparar sean debidamente alineadas y escaladas para que pueda ser comprobado punto a punto. El cómputo se realiza a partir de una ventana gaussiana deslizando. SSIM es altamente sensible a degradaciones en estructuras espaciales de luminancia. [4]



Figura 3: Representación visual de SSIM

En la figura 3 se muestra un ejemplo del comportamiento de SSIM para una imagen de referencia, el procesado y el resultado.

NQI define la distorsión de la imagen como una combinación de tres factores: diferencia entre características comunes, luminancia y contraste. [4]

VQM es una medida basada en el modelo humano de contraste espacial-temporal. Este método calcula la distorsión de un video comprimido mediante cuatro pasos. Utiliza la Transformada del Coseno Discreto con bloques de la imagen original y bloques de la imagen distorsionada. Después convierte los coeficientes transformados en valores locales de contraste usando los coeficientes de DC de cada bloque. A través de funciones de sensibilidad de contraste espacial se convierten los valores locales en valores de diferencia, con menos energía, los cuales se restan de las secuencias originales y comprimidas. Esta diferencia es procesada matemáticamente, apareciendo VQM. [4]

En la práctica los softwares de medición dan como valores cercanos a cero como mejor similitud para las imágenes. Asimismo, declaran resultado igual a cero para imágenes completamente idénticas.

DESCRIPCIÓN DE LAS PRUEBAS OBJETIVAS

Obtención de videos de prueba

Con el objetivo de estimar la degradación de calidad correspondiente a los videos transmitidos en los tres programas de interés: Cubavisión, Cubavisión Internacional e Infantil/Canal Habana¹, se obtuvieron las secuencias de video como se describe en la figura 4.

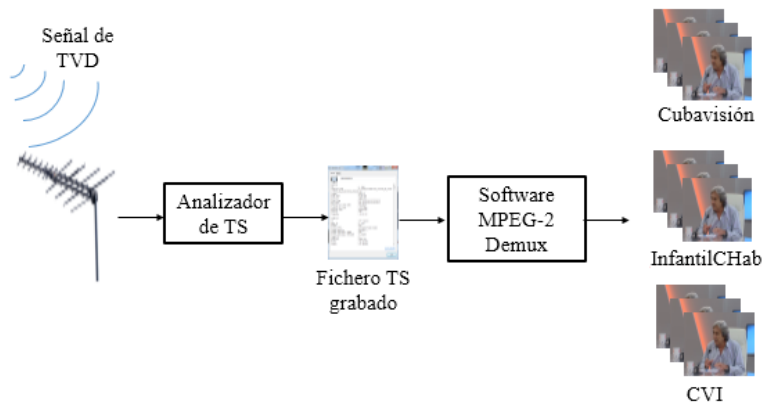


Figura 4: Videos de prueba obtenidos a partir de la señal de televisión digital

Se grabó un fichero .ts (por sus siglas en inglés, Transport Stream) mediante un Analizador de Tramas TS a partir de la señal de televisión digital proveniente del servicio de radiodifusión terrestre. Posteriormente fueron extraídos los tres videos a través del software de demultiplexación “MPEG-2 Demux” desarrollado en **LACETEL** [5] cuyas características principales se pueden observar en la tabla 1.

Tabla 1: Características de videos de prueba

Duración	29 mn 59 s
Codificación	AVC Perfil Principal Nivel 3.0
Resolución	720x480 píxeles
Relación de Aspecto	4:3
Frecuencia de Imagen	29.97 fps
Estándar	NTSC
Espacio de Color	YUV

¹ En el resto del trabajo se emplearán los términos video Cubavisión, video CVI y video InfantilCHab como referencia a los videos correspondientes a los programas Cubavisión, Cubavisión Internacional e Infantil/Canal Habana respectivamente.

Representación de las pruebas

Dado que desde un inicio el video Cubavisión mostró subjetivamente mayor calidad con respecto a los otros dos videos, se decidió seleccionarlo como video de referencia, para así calcular la diferencia de calidad de los videos CVI e Infantil CHab relativa a Cubavisión. Para ello se empleó la herramienta software de Estimación de Calidad de Video del Elecard (EVQE, por sus siglas en inglés, Elecard Video Quality Estimator), el cual permite calcular parámetros de calidad de video, tales como APSNR, NQI, VQM, SSIM, DELTA, MSE y MSAD, entre otras. La figura 5 ilustra la forma en que se le aplicaron los videos de prueba al EVQE, para obtener los parámetros correspondientes.

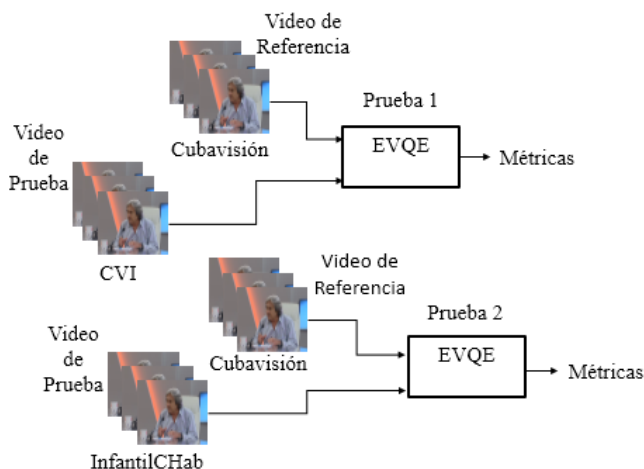


Figura 5: Esquema de las pruebas realizadas

RESULTADOS

La figura 6 muestra los valores de los parámetros calculados por el EVQE. El RMSE se refiere a la raíz cuadrada del MSE. El APSNR se basa en el PSNR convencional y tiene mayor desempeño y exactitud ya que tiene la capacidad de detección de pérdidas de cuadros. En la figura también se aprecia el área de recorte establecida en la esquina superior izquierda de cada video con el objetivo de evitar que el logo presente en el video CVI afectara la estimación de calidad.

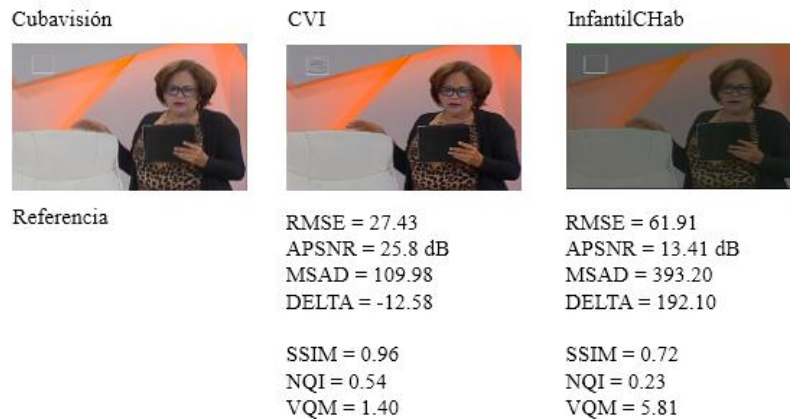


Figura 6: Valores de parámetros obtenidos por el EVQE

Los gráficos de las figuras 7a y 7b muestran los resultados de forma comparativa. Debido a que los parámetros RMSE, MSAD y DELTA se basan en diferencias entre las imágenes de referencia y de prueba, mientras más próximos a cero sean sus valores, menor diferencia existirá entre dichas imágenes, lo cual representa mayor calidad. En el caso del APSNR, un mayor valor indica mayor calidad. Los valores de índices SSIM y NQI más cercanos a uno, señalan mayor calidad. Mientras menor sea el valor de VQM, mayor será la calidad la cual es máxima para valores iguales a cero. El video CVI tiene menor RMSE y MSAD que el video InfantilCHab respecto a Cubavisión, lo cual tiene correlación con la percepción visual subjetiva (InfantilCHab tiene mayor diferencia con respecto a Cubavisión).

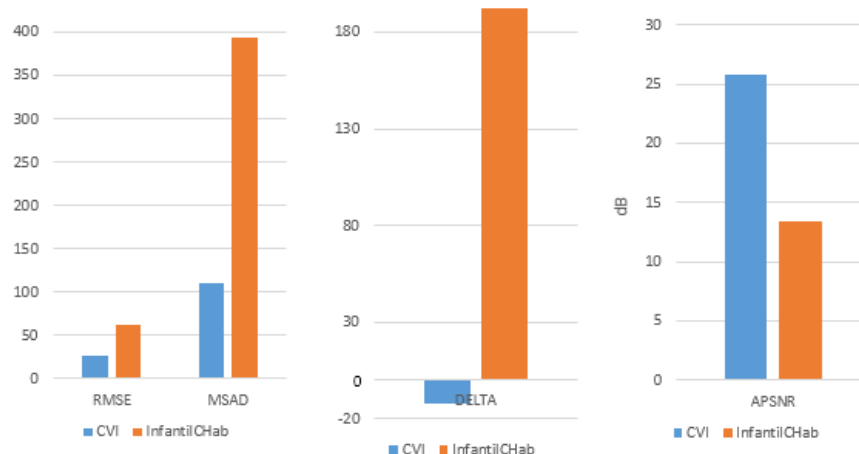


Figura 7a: Resultados comparativos para parámetros básicos

Aunque no es perceptible, el DELTA indica que en determinadas regiones el video CVI presenta componentes de color con mayor nivel que el video Cubavisión. El APSNR, los índices SSIM, NQI y VQM también corroboran que CVI tiene mayor calidad que InfantilCHab.

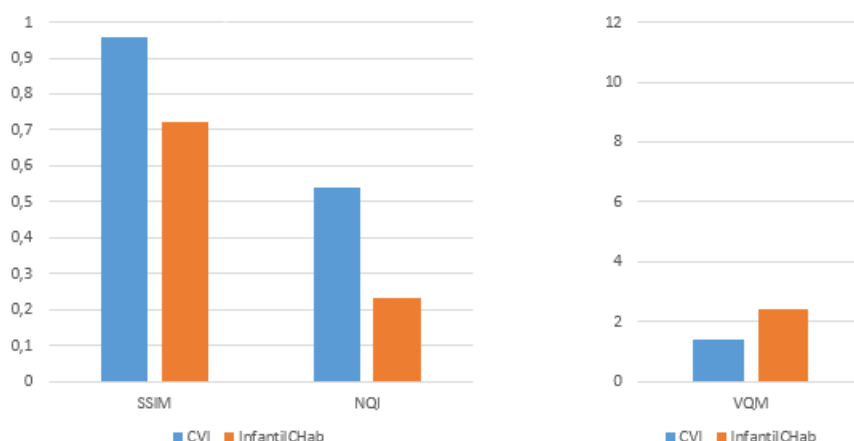


Figura 7b: Resultados comparativos para parámetros perceptuales

CONCLUSIONES.

Los parámetros objetivos de referencia completa calculados demuestran cuantitativamente que una misma fuente de video transmitida en tres diferentes programas del actual sistema de Televisión de Definición Estándar, experimenta una significativa degradación de calidad. El video transmitido en el programa Infantil/Canal Habana tiene menor calidad que el transmitido en Cubavisión Internacional.

REFERENCIAS.

1. JOSKOWICS, José; SOTELO Rafael. "Modelo de Estimación de Calidad de Video: Video Quality Experts Group". Memoria de Trabajos de Difusión Científica y Técnica num.10, 2012.
2. Grupo Técnico del Foro de la Televisión de Alta Definición en España. "Calidad de Video en Alta Definición". Universidad Politécnica de Madrid.
3. KOTEVSKI, Zoran; MITREVSKI Pece. "Comparación Experimental de Métricas PSNR Y SSIM para Estimación de Calidad de Video". Facultad de Ciencias Técnicas de Bitola, Macedonia.
4. KOTEVSKI, Zoran; MITREVSKI Pece. "Evaluación de Desempeño de Métricas para Estimación de Calidad de Video". ICEST 2010, Ohrid, Macedonia.
5. RODRÍGUEZ PORTAS, Iván; PÉREZ ESTÉVEZ Lázaro M. "Desarrollo de una Aplicación para la Demultiplexación de una Trama de Transporte MPEG-2". Tesis de Pregrado, Instituto de Investigación y Desarrollo de Telecomunicaciones e Instituto Superior Politécnico José Antonio Echeverría, La Habana, 2014.