

APORTE EN LAS METODOLOGÍAS MATEMÁTICAS PARA MEJORAR LA COMPRESIÓN Y CALIDAD DE VIDEO

Contribution in mathematical methodologies to improve video compression and quality

Jorge Robles Bokun

PRESENTACIÓN: 15 DE AGOSTO DE 2018

APROBACIÓN: 19 DE NOVIEMBRE DE 2018

RESUMEN

Esta es una investigación cuasiexperimental aplicada, ya que se manipulan, de forma intencional, variables independientes, como la optimización de la transformada discreta del coseno (DCT) y la implementación de filtros pasa banda superiores. De esta manera, se estudiarán los efectos en otras dependientes, como la calidad de imagen, el retardo, la capacidad de data y la fiabilidad de la información en una imagen o secuencia de video bajo el estándar H.264 sin utilizar un grupo de control predeterminado. Para lograrlo, se utilizará, como instrumento, un simulador matemático (Matlab R2015a o Scilab 6.0) que modifica los códigos matemáticos aplicados a una tarjeta electrónica DSP Prototipo EVM DM series, la cual trabaja entre 550-750MHz con un oscilador controlador de voltaje (VCO) y convertidores análogo-digitales. Todo ello estará conectado a la salida de una señal de televisión y permitirá mostrar el gran avance conseguido en cuanto a calidad de imagen y compresión de video (20 a 25)%.

Palabras clave: transformada discreta del coseno, filtros pasa banda, procesamiento digital de señales, códec de video, transmisión de datos, compresión x.263, H.264

ABSTRACT

This is an applied quasi-experimental research, since it intentionally manipulates independent variables (such as: the optimization of the discrete cosine transform and the implementation of higher bandpass filters) to study the effects on other dependents (such as: image quality, delay-jitter, data capacity and reliability of the information) in an image or video sequence under the H.264 standard, without using a predetermined control group, using as a tool a mathematical simulator (Matlab R2015a or Scilab 6.0) where the mathematical codes are modified applied to an electronic card DSP (Digital Signal Processing) Prototype EVM DM series that works between 550-750MHz with a VCO (Oscillator Voltage Controller) and analog-digital converters connected to the output of a television signal showing a great advance in terms of image quality and video compression (20 to 25)%.

Keywords: discrete transformation of cosine, pass band filters, digital signal processing, video codec, data transmission, compression x.263, H.264

I. INTRODUCCIÓN

La compresión se realiza para mejorar la eficiencia en la transmisión y almacenamiento de información, que tiene aplicaciones en la televisión digital, UHDTV, códecs H.264, MPEG-4, H.265, videoconferencia, emisión en directo, videovigilancia, imágenes médicas, secuencia de imágenes, planos, mapas, etc.

La DCT (Transformada Discreta del Coseno) se expresa en una secuencia de puntos finitos, que son el resultado de una sumatoria de señales cosenoidales con diferentes amplitudes y frecuencias. En el caso de la DFT (Transformada Discreta de Fourier), se trata de la suma de señales exponenciales complejas.

Se deben considerar, siempre, como base, los conceptos sobre matrices, tipos de estas, análisis vectorial, gráfica de histogramas, procesamiento digital de señales, y de radiodifusión y televisión, transformadas según los estándares MPEG-1, MPEG-2, MPEG-4, H.264 y H.265, como se muestra en la figura 1.

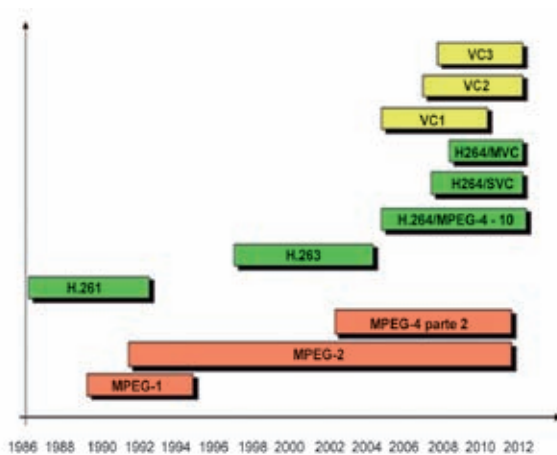


Figura 1. Evolución histórica de los estándares de compresión de video. Adaptado de "Contribución a las metodologías de optimización del tiempo de ejecución de algoritmos de descodificación de video sobre DSPs", por Fernando Pescador, 2011, en http://oa.upm.es/8712/1/FERNANDO_PESCADOR_DEL_OS0.pdf

En los últimos años, el avance tecnológico, en lo que se refiere a la transmisión de video en alta definición con gran resolución, está en su auge, pero muestra una baja calidad en una serie de cuadros por segundo, ya que utiliza más ancho de banda y, a la vez, menos frecuencia.

Esto representa una mayor pérdida de información por dato enviado, y otorga prioridad a los objetos en movimiento e imágenes en 3D, pero olvida la calidad y fidelidad.

Esta necesidad motivó el desarrollo de esta investigación, que busca encontrar una solución equitativa y progresiva, basada en el análisis de los estándares que dieron origen a la compresión, a la baja calidad de video.

En la tesis doctoral presentada por el Mg. Fernando Pescador del Oso [1], el objetivo principal es encontrar una metodología para mejorar la velocidad en tiempo de respuesta en los descodificadores basados en DSPs.

II. METODOLOGÍA DE IMPLEMENTACIÓN

A. TÉCNICAS MATEMÁTICAS DE COMPRESIÓN DE VIDEO PARA MEJORAR LA IMAGEN

A continuación, se presentan algunas técnicas para mejorar la imagen, de las que solo utilizamos las marcadas con asterisco (*) para la presente investigación.

1) Suavización (*): mediante esta técnica matemática sobre las matrices, se reduce el ruido, tanto en la parte central como en los bordes. Es decir, se mejora la claridad del contorno de imagen.

2) Expansión del histograma (*): los valores mayores y menores de grises de la matriz (imagen original) se aumentan a un máximo y reducen a un mínimo, respectivamente, para mejorar la calidad de imagen (su contenido, que puede ser de alta definición).

3) Manipulación de histograma (*): se mejora el contraste y el brillo. Es decir, se suma un valor positivo ($D > 0$) a una matriz para incrementar los valores de los píxeles (aclerar) o uno negativo ($D < 0$) para oscurecer la imagen, píxel por píxel.

4) Zoom tipo1: amplía o agranda la imagen mediante el uso de la repetición en la matriz original en una sola dirección. Su construcción requiere de equipos con tecnología especializada.

5) Zoom tipo2: agranda o amplía la imagen, pero con mayor calidad. Se realiza mediante la técnica del promedio aritmético.

6) La DCT (*): mejora la compresión, y la calidad de imagen y video (secuencia de imágenes) mediante un algoritmo matemático.

1. Suavizado de imágenes y reducción del ruido, que mejora el contorno (*)

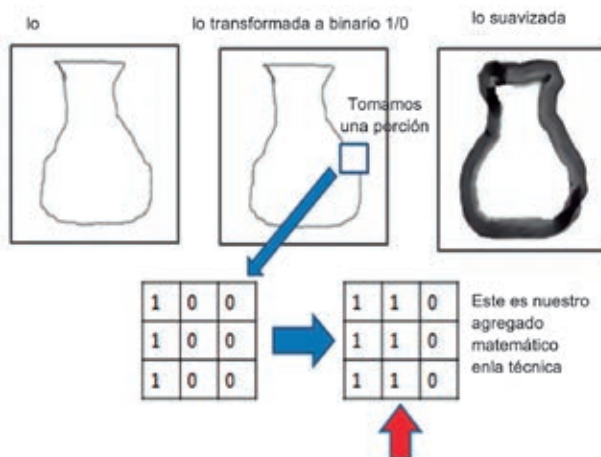


Figura 2. De la matriz, se tomó el mayor valor, que se repite en la segunda columna, para mejorar el contorno o el borde de la imagen. Luego, se ensanchó el borde del jarrón, lo que mejoró, en este caso, entre 13% y 33% la nitidez del contorno. Elaboración propia.

2. Técnica de expansión de histograma y mejora la calidad de imagen (*)

Los valores mayores suben al máximo, los intermedios permanecen y los menores llegan al mínimo, según la siguiente ecuación:

$$I_T(i,j) = \left(\frac{I_o(i,j) - I_o(i,j)_{\min}}{I_o(i,j)_{\max} - I_o(i,j)_{\min}} \right) 2^{N^{\circ}bpp} - 1 - \frac{I_o(i,j)_{\min}}{I_o(i,j)_{\max}} \quad (1)$$



Este factor se agregó para mejorar mucho más la calidad de imagen.

Por ejemplo:

$$\begin{bmatrix} 15 & 15 & 120 \\ 15 & 10 & 120 \\ 4 & 250 & 4 \\ 4 & 4 & 4 \end{bmatrix} \quad 2^{N^{\circ}bpp} = 2^8 = 256 \text{ niveles de gris}$$

4x3 0= negro, 128= gris, 255= blanco

Sin el uso del factor adicional, tomando $I_o=15$, sale $I_t=10.44$. Con $I_o=250$, resulta $I_t=255$.

Con el uso del factor adicional, tomando $I_o=15$, sale $I_t=10.38$ (el mejor resultado, porque es de menor valor que el píxel anterior y se acerca más a 0 —negro—, es decir, a una mejor calidad que la imagen anterior) y con $I_o=250$, resulta $I_t=254.9$ (se mantiene el valor aproximado a 255 —blanco—).

$$\left(\frac{15 - 4}{250 - 4} \right) 256 - 1 - \frac{15}{250} = 10.38$$

$$\left(\frac{250 - 4}{250 - 4} \right) 256 - 1 - \frac{15}{250} = 254.94$$

En conclusión, mejoró la calidad de 1% a 13%.

3. Manipulación de Histograma (*). Frecuencia (píxel) vs. valor (píxel) de una matriz 8bpp = 256 niveles (graficado en PAWNT)

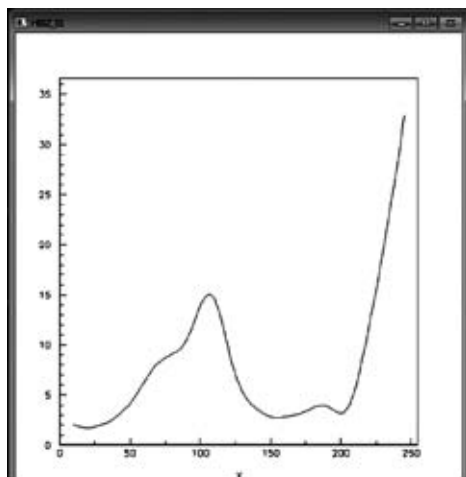


Figura 3. Frecuencia (píxel) vs valor (píxel) de una matriz 8bpp. Elaboración propia en PAWNT.

En la figura 3, considérense los siguientes valores:

- 0 = negro
- 128 = gris
- 255 = blanco

De esa manera, se aprecia el incremento de la frecuencia de los valores del píxel, tanto en la parte intermedia (grises) como en la extrema (blancos) de la matriz o imagen. Eso dota a esta de mayor calidad y nitidez. Asimismo, se desplazó más a la derecha, por lo cual obtuvo mayor claridad.

4. Compresión de video H.264, H.265 mejorado usando la técnica de la DCT (*)

En este estudio, y con el avance de la tecnología, se buscó transmitir contenido de alta definición mejorado, con mayor calidad y menos pérdida. Una de las herramientas matemáticas utilizadas es la DCT (transformada discreta del coseno), que elimina o reduce los datos redundantes e irrelevantes en una matriz antes de ser transmitida y almacenada. Luego, recupera la imagen o secuencia de imágenes original, pero con mayor calidad.

La compresión se realiza para mejorar la eficiencia en la transmisión y almacenamiento de información, que tiene aplicaciones como en la televisión digital, UHDTV, códecs H.264, MPEG-4, H.265, videoconferencia, emisión en directo por internet, videovigilancia, imágenes médicas, secuencia de imágenes, planos, mapas, etc.

B. METODOLOGÍA

a. Tipo y diseño de investigación

a.1. Según el tipo, es una investigación experimental aplicada, ya que se manipula, de forma intencional, diversas variables independientes (que incluyen la optimización de la transformada discreta

del coseno, la perfección de la fórmula matemática DCT, el desarrollo del algoritmo VLC —codificación de longitud variable— y la implementación de filtros pasa banda mejorados) para estudiar los efectos que ejercen sobre otras dependientes (calidad de imagen, retardo-*jitter*, capacidad de data, fiabilidad de la información) en una imagen o secuencia de video bajo el estándar H.264.

Se trata de una investigación aplicada, porque, al conocer y estudiar la teoría sobre procesamiento digital de señales, radiodifusión y tv, sistemas de codificación digital, códecs de video, transformadas, sumatorias matemáticas, y temas afines, se buscan nuevas formas de aplicación y aportes a la ingeniería electrónica, como la perfección de la compresión de video que utilice, como instrumentos de validez, un simulador y las pruebas de laboratorio.

a.2 Además, el presente estudio posee un diseño de investigación cuasi experimental, porque existe una manipulación de variables que modifica los resultados obtenidos, y avanza en gran parte el estudio y diseño sin utilizar un grupo de control predeterminado. Esto se debe a que, en las comparaciones, siempre se necesitó trabajar con variables independientes, como la transformada discreta del coseno, el algoritmo VLC, etc., que cambiaron los valores de la variable dependiente (calidad de imagen, retardo, etc.). Tal es el caso del grupo experimental.

b. Unidad de análisis

La unidad de análisis está compuesta por las imágenes o matrices de una secuencia de video. Este *array* está formado por filas y columnas. Los cruces de estas se denominan celdas, que a su vez son llamados píxeles. En el caso de la investigación, el píxel sería la mínima unidad de medida de una imagen.

c. Población de estudio

En la investigación, se utiliza una secuencia de imágenes, es decir, un pequeño video de veinte fotos en un lapso de cinco segundos.

d. Tamaño de muestra

De este video, se tomaron uno o dos grupos de tres a cuatro imágenes por muestra en un segundo, como se muestra en la figura 4.

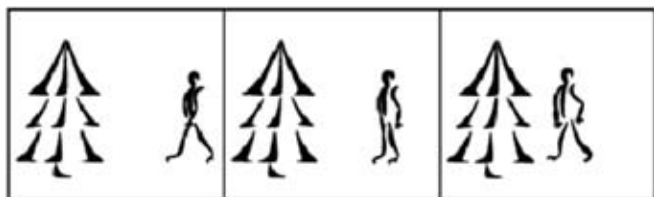


Figura 4. Muestra de una secuencia de tres imágenes bajo el estándar MPEG-4 Parte10. Elaboración propia.

C. SELECCIÓN DE MUESTRA Y TÉCNICAS DE RECOLECCIÓN DE DATOS

D.1 El procedimiento para la selección de muestra está basado en las pruebas de campo y laboratorio aplicadas a la tecnología de la TDT (televisión digital terrestre) en una muestra de tres a cuatro imágenes.

D.2 Como técnicas de recolección de datos, para esta investigación, se utilizó la observación después de hacer una preprueba y una posprueba en que se manipulaba alguna de las variables independientes conocidas (optimización de la DCT, desarrollo del algoritmo VLC, etc.). Para ello, se utilizó, como instrumento, un simulador de redes llamado Matlab R2010, en el que se modificaron los códigos de las fórmulas matemáticas.

Esta teoría de códigos matemáticos modificados se aplica a una tarjeta electrónica DSP (procesamiento digital de señales), del prototipo EVM DM [3]. Son series que trabajan entre 550-750MHz con un VCO (oscilador controlador de voltaje) y convertidores análogo-digitales en la salida de una señal de televisión. De esa manera, muestran una mejora en cuanto a calidad de imagen y compresión de video.

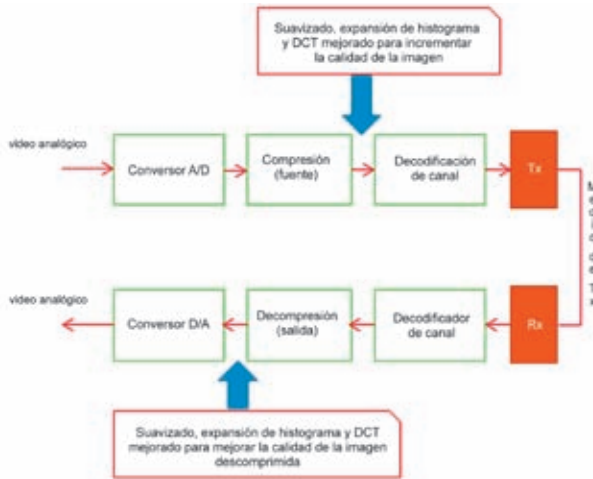


Figura 5. Diagrama de bloques de compresión de video en un sistema digital de televisión, con las metodologías matemáticas para mejorar la compresión y calidad de imagen. Elaboración propia en Power Point 2013.

D. COMPARACIÓN DE LOS ESTÁNDARES DE VIDEO MPEG-2, MPEG-4 (PARTE-2) Y MPEG-4 (PARTE-10) O H.264 AVC USANDO, COMO INSTRUMENTO, EL SIMULADOR MATEMÁTICO MATLAB R2010

Para hacer la comparación de los diferentes resultados de las variables dependientes, como la capacidad o ancho de banda entre los estándares de video en estudio, se realizó una medición en las etapas o elementos de un sistema general de compresión de video como en la figura 5.

1. Filtro Pasa Bandas:

Donde a, b, c y d son vectores:

```
>> a = 1/2;
>> b = [1/4 1/4 1/4 1/4];
>> load count.dat
>> x = count(:,1);
%y: datos filtrados
>> y = filter(b,a,x);
>> figure
```

```

>> t = 1:length(x);
>> plot(t,x,'--',t,y,'-'),grid on
>> legend('Estándar MPEG4 parte10','Estándar MPEG2',2)
>> c = 1/3;
>> d = [1/9 1/9 1/9 1/9];
>> load count.dat
>> x = count(:,1);
%y: datos filtrados
>> y = filter(d,c,x);
>> figure
>> plot(t,x,'--',t,y,'-'),grid on
>> legend('Estándar MPEG-4 parte10','Estándar MPEG-2',2)
>> title('Comparación MPEG-2 vs MPEG-4 parte10 mediante un filtro pasa bandas (Mbps)')

```

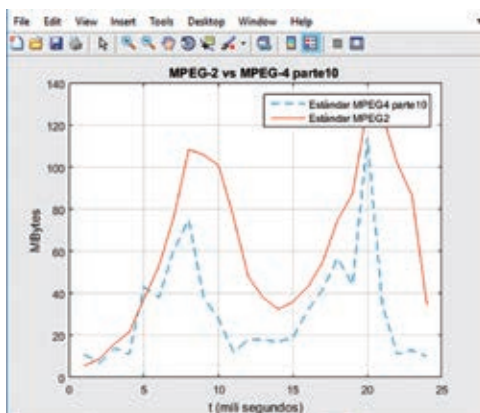


Figura 6. MPEG2 vs MPEG4. Elaboración propia.

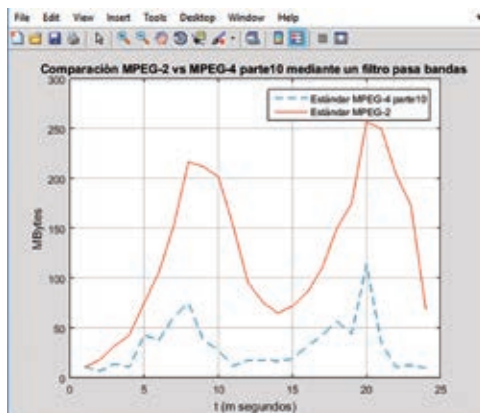


Figura 7. Comparación entre MPEG-2 y MPEG-4 parte 10, mediante filtro pasa bandas. Elaboración propia en Matlab R2010a.

De la figura 7, elaborada en Matlab, concluimos, después de comparar MPEG-2 y MPEG-4 parte10 (estándar de video H.264 AVC), que, al depurar la señal portadora por el filtro pasa banda, la

información en MB del MPEG-4 (H.264 AVC) disminuye; es decir, se comprime más que utilizando el estándar MPEG-2.

2. Retardo de tiempo (jitter)

```
fs = 100; % Frecuencia de muestreo (Hz)
t = 0:1/fs:10-1/fs; % 10 sec muestra
x = (1.3)*sin(2*pi*15*t) ... % 15 Hz componente
+ (1.7)*sin(2*pi*40*(t-2)) ... % 40 Hz componente
+ 2.5*gallery('normaldata',size(t),4); % Ruido Gaussiano
m = length(x); % longitud de la ventana
n = pow2(nextpow2(m)); % longitud de la transformada
y = fft(x,n); % DFT (valores discretos)
f = (0:n-1)*(fs/n); % Rango de la frecuencia
power = y.*conj(y)/n; % Nivel de potencia de la DFT
```

```
plot(f,power)
xlabel('Frequency (Hz)')
ylabel('Power')
title('\bf Periodogram')
```

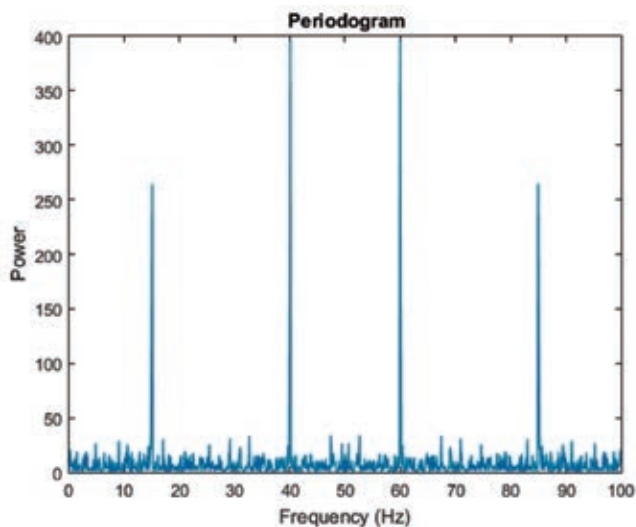


Figura 8. Nivel de Potencia y jitter H.264. Elaboración propia en Matlab R2010a

Luego, agregamos las siguientes sentencias:

```
y0 = fftshift(y); % (Transformada de Fourier) Valores y rangos re-arregables
f0 = (-n/2:n/2-1)*(fs/n); % 0-frecuencia central
power0 = y0.*conj(y0)/n; % 0-Nivel de Potencia central
```

>> plot(f0,power0,'-',f,power,'--'),grid on % Comparación jitter-retardo de tiempo MPEG-2 vs MPEG-4 parte10 (H.264 AVC)

Después de plotear, nos resulta la gráfica de la figura 9.

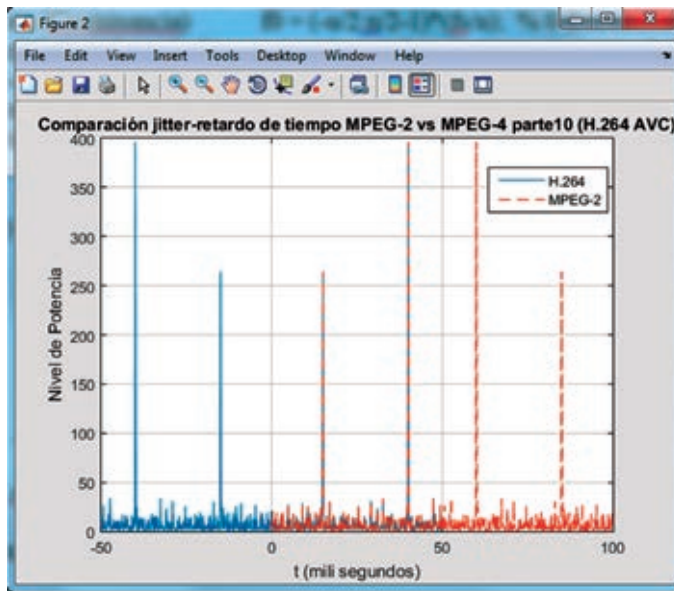


Figura 9. Comparación jitter-retardo de tiempo MPEG-2 vs MPEG-4 parte10 (H.264 AVC).
Elaboración propia en Matlab R2010a.

De los resultados obtenidos de la figura 9, se concluye que, después de muestrear, cuantificar, modular (QAM), filtrar y comprimir la señal de entrada (portadora) con la DFT, los diferentes niveles de potencia del estándar MPEG-2 se retrasan más que los de MPEG-4 parte 10; es decir, su *jitter* es mayor.

III. ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

Después de utilizar Matlab, elaboramos la siguiente tabla:

Variable dependiente de salida	MPEG-2	MPEG-4 parte 10 (aporte con las metodologías en la tesis) H.264 AVC	H.265 HEVC
(Calidad de imagen) Compresión de la portadora después del filtro pasa banda	(120 - 100)MBytes	(80 - 60)MBytes	<= 40MBytes (sigue en estudio, debe mejorar)
(Retardo – Jitter) después de muestrear, cuantizar, modular (modulación QAM), filtrar y comprimir con la DFT la señal de entrada (portadora) en los diferentes niveles de potencia	(0 - 35KHz)	(0 - 20KHz)	<= 20KHz (sigue en estudio, debe mejorar)
(Capacidad - Velocidad de transmisión en Mbps) después de filtrar la señal (portadora) por un filtro digital FIR 3D para las diferentes frecuencias de niveles	(5 - 10) Mbps	(10 - 100) Mbps	>= 100Mbps (sigue en estudio, debe mejorar)
(Fiabilidad de información) después de filtrar la señal (portadora) por un filtro digital FIR 3D para los diferentes niveles de píxeles en (R, G, B)	(15 - 10) % Pérdida de Información	(10 - 5) % Pérdida de Información	<= 5% Pérdida de Información (sigue en estudio, debe mejorar)

Tabla 1. Tabla comparativa entre los estándares de video, según los resultados obtenidos en Matlab r2010a dentro del aporte, en las metodologías matemáticas, para mejorar la compresión de video. *Elaboración propia.*

Con esos resultados, obtenidos en el simulador Matlab, concluimos que, al comparar MPEG-2 y MPEG-4 parte10 (estándar de video H.264 AVC) y después de procesar la señal portadora con el filtro pasa banda, la información del MPEG-4 (H.264 AVC) disminuye en un rango de 80 y 60 MB. Por ello, se comprime más que utilizando el estándar MPEG-2.

Para la compresión de la señal portadora (luminancia y crominancia) de los estándares MPEG-2, MPEG-4(parte-2), MPEG-4(parte-10) o H.264 AVC y H.265, se utilizó como algoritmo o función matemática la DFT (transformada discreta de Fourier) con un filtro digital tipo FIR.

De la tabla 1 se concluye, después de muestrear, cuantizar, modular, filtrar y comprimir (con la DFT) la señal portadora, que los niveles de potencia del estándar MPEG-2 se retrasan más que los del MPEG-4 parte 10 (H.264 AVC) y del H.265 HEVC; es decir, su *jitter* es mayor.

IV. CONCLUSIONES

En conclusión, utilizar la DCT y la DFT como metodologías matemáticas para comprimir video H.264 AVC sería lo más ventajoso, porque con ellas podemos recuperar o restaurar la imagen con la mínima pérdida de información posible. Eso la convierte en una de las técnicas matemáticas más importantes para mejorar la calidad de imagen, como nos muestran los resultados y valores obtenidos en el simulador MatlabR2010a según la tabla 1.

V. APORTES Y LOGROS OBTENIDOS

- Se compararon los estándares de video más importantes, como el MPEG-2, el MPEG-4 y el H.264 AVC, mediante el simulador Matlab. Los resultados nos indican la superioridad, en cuanto a grado de compresión de video de la señal (portadora), del códec H.264 AVC y el H.265 (que sigue en estudio, como se muestra en la tabla 1).
- Se encontró que, como filtro digital, el FIR es el más robusto para comprimir la información de luminancia y crominancia.
- Como transformadas matemáticas, se obtuvo que la DCT y la DFT son las que comprimen más video con la mínima pérdida de información hasta regenerar la señal.

VI. RECOMENDACIONES

- Se debe continuar con el estudio y análisis del estándar de video H.265. Esto contribuiría a mejorar la calidad de señal y las imágenes en movimiento 3D.
- Se sugiere seguir utilizando el simulador Matlab como herramienta de recolección de datos, porque los que obtiene son más precisos cuando comparamos diferentes niveles jerárquicos de códec de audio y video.
- Se recomienda utilizar siempre filtros digitales mejorados para comprimir la señal portadora, ya que estos son más robustos y aportan fiabilidad en la información.

VII. REFERENCIAS

- [1] F. Pescador del Oso, «Contribución a las metodologías de optimización del tiempo de ejecución de algoritmos de descodificación de vídeo sobre DSPs», tesis doctoral. Univ. Politécnica de Madrid, 2011 [En línea]. Disponible en: http://oa.upm.es/8712/1/FERNANDO_PESCADOR_DEL_OSO.pdf. [Accedido: 19-nov-2018]
- [2] M. Mastriani, en *Proceedings of International Journal Conference on Lossy Image Compression Signal Processing*, Buenos Aires, Universidad Nacional de General Sarmiento, 2009. Vol. 5, núm. 2, pp.112-119.
- [3] E. Juárez, F. Pescador, P. Lobo y C. Sanz, *Distortion Energy Analysis on a OMAP Based H.264/SVC Decoder*, Lisboa, 2010.
- [4] R. D. Dony y S. Haykin, en *Proceedings of the IEEE on Neuronal Network Approaches to Image Compression*, IEEE, 1995. Vol. 83, núm. 2, pp. 288-303.
- [5] R. W. Lucky, en *Automatic Equalization for Digital Communication*, Bell Syst. Tech. J, 1965. Vol. 44, Num. 4, pp. 547-588.