

PERFIL MULTIVISIÓN MPEG-2 PARA TELEVISIÓN ESTEREOSCÓPICA

(1998)

1 Introducción al perfil multivisión (MVP, *multi-view profile*) MPEG-2

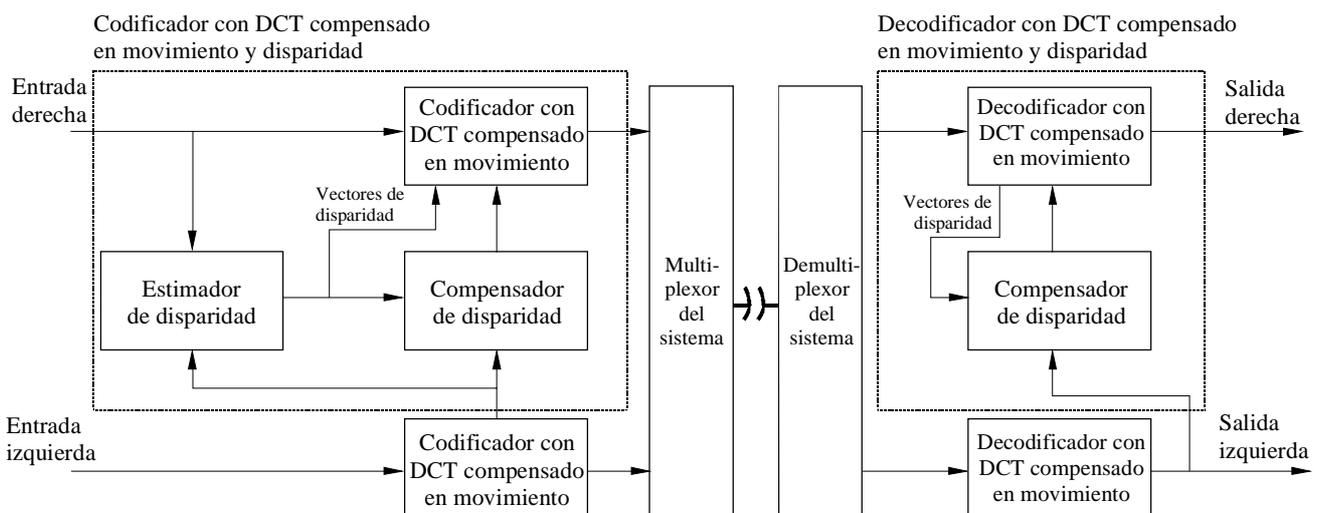
La ampliación de la norma de vídeo MPEG-2 (Recomendación UIT-T H.262 | ISO/CEI 13818-2: Tecnología de la información – Codificación genérica de imágenes en movimiento e información de audio asociada: Vídeo) en previsión de las aplicaciones multivisión (por ejemplo las utilizadas en el vídeo estereoscópico) ha sido elevada al rango de Norma Internacional final en la reunión ISO/CEI JTC 1/SC 29/GT 11 de septiembre de 1996 (Enmienda 3, GT 11 N1366). Se considera que el perfil multivisión (MVP) es idóneo para las aplicaciones que necesitan muchos puntos de visión en el contexto de la norma de vídeo MPEG-2. El MVP admite imágenes estereoscópicas como imágenes fuente para una amplia gama de resoluciones y calidades de imagen, que dependen de las necesidades de las aplicaciones de que se trate.

1.1 Esquema de codificación para el MVP

La Fig. 1 muestra un diagrama de bloques del modelo de códec de referencia para el MVP. Sus principales características son codificación monoscópica en su capa base a efectos de compatibilidad y predicción híbrida de movimientos y disparidad a fin de aumentar la eficacia de la compresión. Para codificar una capa de mejora se utilizan herramientas de escalonabilidad temporal.

FIGURA 1

Modelo de códec de referencia para el MVP



DCT: transformación discreta en coseno (*discrete cosine transform*)

Rap 2017-01

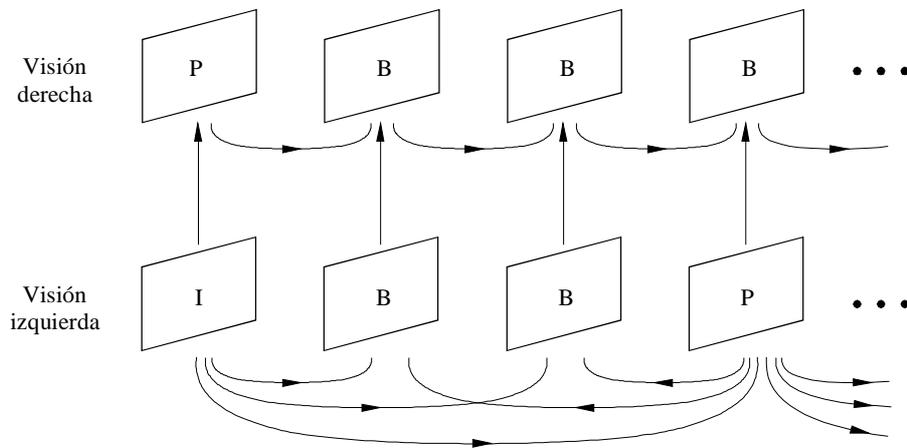
La Fig. 2 muestra una configuración de modos de predicción. A la capa base se le aplica una codificación monoscópica con las mismas herramientas que al perfil principal (MP, *main profile*), incluida la Norma ISO/CEI 11172-2. Se asigna una capa base de MVP a la visión izquierda y una capa de mejora a la visión derecha. La capa de mejora se codifica utilizando herramientas de escalonabilidad temporal y en la capa mejorada puede aplicarse la predicción híbrida de movimientos y disparidad. Se prevé una mayor compresión de la visión derecha del vídeo estereoscópico a causa del parecido entre la visión izquierda y la visión derecha.

El MVP, uno de los perfiles escalonables en términos de capas de múltiples puntos de visión, tiene las mismas características de compatibilidad que otros perfiles escalonables, entre ellas, la compatibilidad con el MP. Por ejemplo:

- decodificadores que se ajustan al MVP a un cierto nivel pueden decodificar trenes de bits que se ajustan al MP al nivel correspondiente (es decir, compatibilidad hacia adelante),
- decodificadores que se ajustan al MP a un cierto nivel pueden decodificar los trenes de bits de la capa base del MVP (es decir, compatibilidad hacia atrás).

FIGURA 2

Ejemplo de configuración de predicción con codificación $M = 3$ de la visión izquierda, imagen de trama de la visión derecha codificada mediante predicción de disparidad respecto a la visión izquierda y predicción de movimientos con respecto a sí misma



Rap 2017-02

1.2 Valores de los parámetros del MVP

Los niveles del MVP son alto, alto-1440, principal y bajo. La escalonabilidad temporal comporta dos capas, una capa base y una capa de mejora. Ambas capas tienen la misma resolución espacial a la misma frecuencia de trama. Los Cuadros 1 a 4 indican los límites a los que se han de atener las velocidades de muestreo, las velocidades de los pels de luminancia, las velocidades binarias y los tamaños de memoria tampón del MVP.

CUADRO 1

Límites superiores de la densidad de muestreo

Nivel	Capa de resolución espacial		Perfil
			Multivisión
Alto	Mejorada (visión derecha)	Muestras/línea Líneas/trama Tramas/s	1 920 1152 60
	Inferior (visión izquierda)	Muestras/línea Líneas/trama Tramas/s	1 920 1152 60
Alto-1440	Mejorada (visión derecha)	Muestras/línea Líneas/trama Tramas/s	1 440 1152 60
	Inferior (visión izquierda)	Muestras/línea Líneas/trama Tramas/s	1 440 1152 60
Principal	Mejorada (visión derecha)	Muestras/línea Líneas/trama Tramas/s	720 576 30
	Inferior (visión izquierda)	Muestras/línea Líneas/trama Tramas/s	720 576 30
Bajo	Mejorada (visión derecha)	Muestras/línea Líneas/trama Tramas/s	352 288 30
	Inferior (visión izquierda)	Muestras/línea Líneas/trama Tramas/s	352 288 30

CUADRO 2

Límites superiores de la velocidad de muestreo de luminancia (muestras/s)

Nivel	Capa de resolución espacial	Perfil
		Multivisión
Alto	Mejorada (visión derecha)	62 668 800
	Inferior (visión izquierda)	62 668 800
Alto-1440	Mejorada (visión derecha)	47 001 600
	Inferior (visión izquierda)	47 001 600
Principal	Mejorada (visión derecha)	10 368 000
	Inferior (visión izquierda)	10 368 000
Bajo	Mejorada (visión derecha)	3 041 280
	Inferior (visión izquierda)	3 041 280

CUADRO 3

Límites superiores de las velocidades binarias (Mbit/s)

Nivel	Perfil
	Multivisión
Alto	130 ambas capas
	80 capa base
Alto-1440	100 ambas capas
	60 capa base
Principal	25 ambas capas
	15 capa base
Bajo	8 ambas capas
	4 capa base

CUADRO 4

Requisitos en cuanto a tamaño de memoria (bits)

Nivel	Capa	Perfil
		Multivisión
Alto	Mejorada	15 898 480
	Base	9 787 248
Alto-1440	Mejorada	12 222 464
	Base	7 340 032
Principal	Mejorada	3 047 424
	Base	1 835 008
Bajo	Mejorada	950 272
	Base	475 136

1.3 Ampliación de los parámetros de cámara

Se ha introducido en el MVP una ampliación para dar cabida a la información relativa a la cámara. La ampliación permite especificar la altura del dispositivo de imagen, la longitud focal, el número F, el ángulo vertical del campo de visión, la posición y dirección de la cámara y la dirección superior de la misma.

2 Pruebas de evaluación del MVP

Las pruebas de verificación del MVP se llevaron a cabo en tres emplazamientos de prueba diferentes situados en Japón, Alemania y Canadá. Los resultados de dichas pruebas se presentaron en la reunión del Grupo de Trabajo 11 celebrada en Chicago (GT 11 N1373) septiembre de 1996. Test and video subgroup. Results of MPEG-2 multi-view profile verification test. Los resultados de los distintos emplazamientos de pruebas son coherentes entre sí y ponen de manifiesto que en líneas generales, a las velocidades binarias utilizadas, los observadores opinaron que el esquema de codificación del perfil multivisión MPEG-2 no introducía ninguna perturbación.

2.1 Método de prueba

Se utilizó el método de escala de degradación con doble estímulo (variante II) de la Recomendación UIT-R BT.500. Para obtener evaluaciones más precisas se utilizó una escala continua en vez de la escala discreta recomendada por el UIT-R.

2.2 Condiciones de las pruebas

Se utilizaron las secuencias de prueba generadas durante el intercambio de trenes binarios. El Cuadro 5 resume las condiciones de las pruebas. En cada emplazamiento de prueba se utilizó un sistema de visualización diferente.

CUADRO 5

Resumen de las condiciones de las pruebas subjetivas

Secuencias	«Street organ (Organillo)», «Flower pot (Maceta)», «Trapeze (Trapecio)» (525/60) «Fun fair (Feria)» (625/50)
Algoritmos y velocidades binarias (visión izquierda/derecha)	MVP@ML: 6/3 Mbit/s, 9/4 Mbit/s Simulcast de perfil principal en el nivel principal (MP@ML): 4,5/4,5 Mbit/s, 6,5/6,5 Mbit/s Simulcast de MP@ML como anclaje inferior: 2,5/2,5 Mbit/s (para «Street organ», «Fun fair»), 1,5/1,5 Mbit/s (para «Flower pot», «Trapeze») Original/original como anclaje superior
Método de prueba	Método de escala de degradación con doble estímulo (variante II) descrito en la Recomendación UIT-R BT.500, con escala continua
Sistema de visualización estereoscópica (tamaño de la imagen, distancia de observación)	HHI: Sistema de visualización de doble espejo (19 cm × 14 cm, 5 H) CRC: Visualización secuencial en el tiempo y gafas con obturador LCD (40,6 cm × 30,5 cm, 4 H) NHK: Proyectores de televisión de alta definición (TVAD) con LCD y gafas polarizantes (82 cm × 57 cm, 5 H)
Observadores	HHI: 24 espectadores sin experiencia CRC: 18 espectadores sin experiencia NHK: 19 espectadores sin experiencia (se rechazó un observador en la selección realizada en base a la Recomendación UIT-R BT.500)

HHI: Heinrich-Hertz-Institut für Nachrichtentechnik (Alemania)

CRC: Communications Research Center (Canadá)

NHK: Nippon Hoso (Kyokai (Japón))

2.3 Resultados de las pruebas de evaluación subjetivas

Para cada condición de prueba se calcularon las notas medias y los intervalos de confianza del 95%. Los resultados de las pruebas de HHI, CRC y NHK se presentan en el Cuadro 6 y en la Fig. 3. HHI1 y HHI2 son los resultados obtenidos en HHI en dos partes distintas de la misma secuencia. HHI no pudo probar las secuencias en su totalidad porque no tenía suficiente memoria de pantalla.

CUADRO 6

Notas medias e intervalos de confianza del 95%

a) Secuencia: Street organ (Organillo)

	Fuente	MVP (9/4 Mbit/s)	MP × 2 (6,5/6,5 Mbit/s)	MVP (6/3 Mbit/s)	MP × 2 (4,5/4,5 Mbit/s)	Anclaje inferior
NHK	4,71 ±0,17	4,18 ±0,27	4,40 ±0,26	4,06 ±0,39	3,51 ±0,32	1,74 ±0,33
CRC	4,24 ±0,37	4,19 ±0,33	4,33 ±0,29	4,27 ±0,34	4,07 ±0,35	2,19 ±0,35
HHI1	4,89 ±0,12	4,55 ±0,21	4,58 ±0,22	4,23 ±0,26	3,63 ±0,35	1,30 ±0,19
HHI2	4,86 ±0,13	4,68 ±0,19	4,85 ±0,13	4,44 ±0,24	4,24 ±0,32	1,80 ±0,23

b) Secuencia: Flower pot (Maceta)

	Fuente	MVP (9/4 Mbit/s)	MP × 2 (6,5/6,5 Mbit/s)	MVP (6/3 Mbit/s)	MP × 2 (4,5/4,5 Mbit/s)	Anclaje inferior
NHK	4,79 ±0,16	4,03 ±0,44	4,28 ±0,25	4,07 ±0,33	4,13 ±0,37	2,28 ±0,32
CRC	4,53 ±0,14	4,57 ±0,20	4,45 ±0,22	4,40 ±0,20	4,40 ±0,21	2,70 ±0,34
HHI1	4,81 ±0,19	4,49 ±0,25	4,52 ±0,26	4,33 ±0,24	4,46 ±0,23	1,96 ±0,25
HHI2	4,83 ±0,14	4,48 ±0,21	4,33 ±0,22	4,08 ±0,26	4,16 ±0,25	1,69 ±0,24

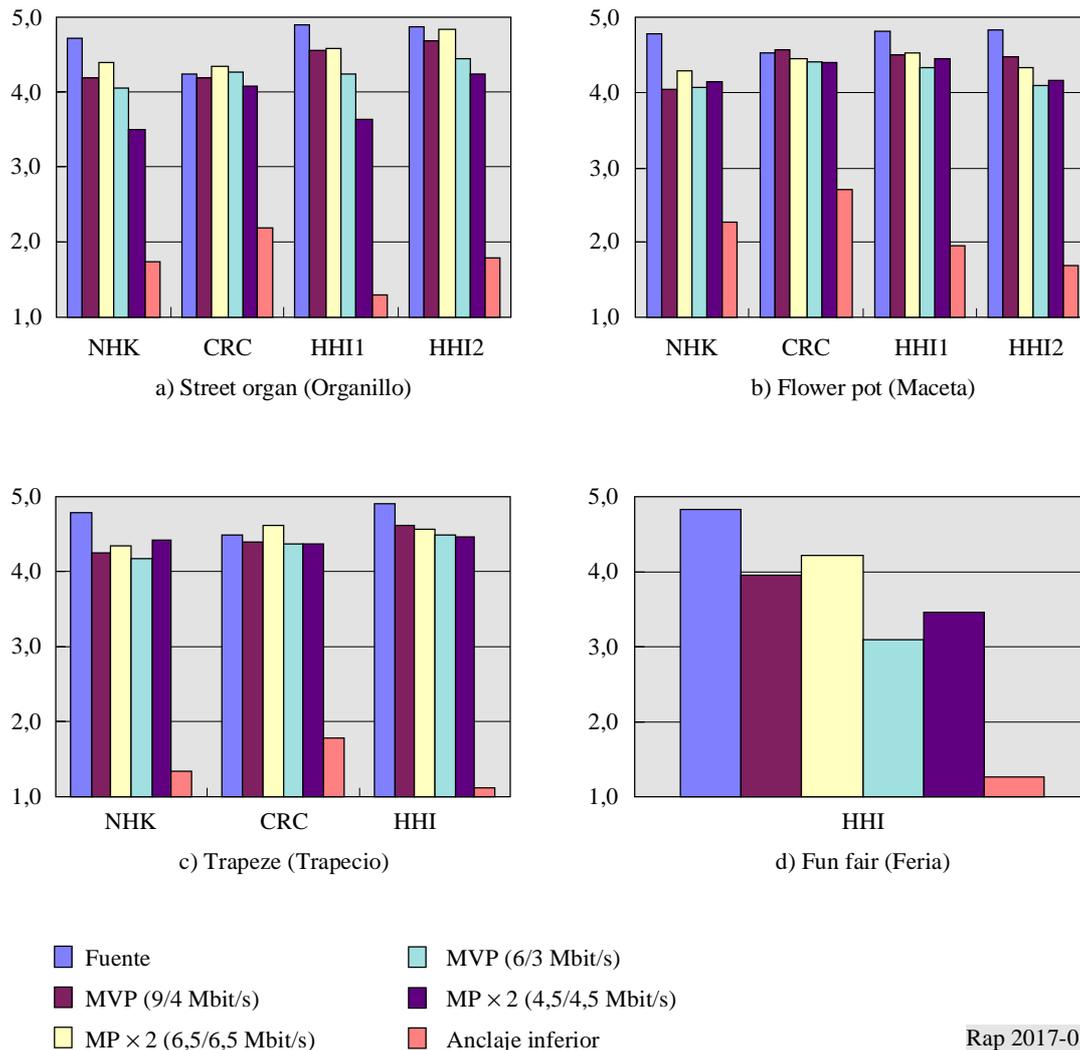
c) Secuencia: Trapeze (Trapezio)

	Fuente	MVP (9/4 Mbit/s)	MP × 2 (6,5/6,5 Mbit/s)	MVP (6/3 Mbit/s)	MP × 2 (4,5/4,5 Mbit/s)	Anclaje inferior
NHK	4,77 ±0,13	4,24 ±0,25	4,34 ±0,38	4,16 ±0,24	4,41 ±0,23	1,33 ±0,18
CRC	4,48 ±0,22	4,38 ±0,24	4,62 ±0,14	4,37 ±0,23	4,36 ±0,24	1,78 ±0,31
HHI1	4,90 ±0,11	4,60 ±0,19	4,55 ±0,25	4,48 ±0,27	4,46 ±0,28	1,13 ±0,14

d) Secuencia: Fun fair (Feria)

	Fuente	MVP (9/4 Mbit/s)	MP × 2 (6,5/6,5 Mbit/s)	MVP (6/3 Mbit/s)	MP × 2 (4,5/4,5 Mbit/s)	Anclaje inferior
HHI1	4,83 ±0,14	3,96 ±0,32	4,23 ±0,29	3,10 ±0,35	3,46 ±0,27	1,27 ±0,20

FIGURA 3
Notas medias de la evaluación subjetiva



Vale la pena mencionar algunos aspectos de los resultados de estas pruebas:

- En cada una de las cuatro secuencias, la nota media de la secuencia MVP a la velocidad binaria de 9/4 Mbit/s no presenta una diferencia importante con respecto a la nota media del simulcast de los MP a la velocidad binaria de 6,5/6,5 Mbit/s. Asimismo, la nota media de la secuencia MVP a la velocidad binaria de 6/3 Mbit/s no se diferencia sensiblemente de la nota media del simulcast de los MP a la velocidad binaria de 4,5/4,5 Mbit/s, excepto el par de la secuencia «Street organ». Para «Street organ», la calidad del MVP es superior a la del simulcast de los MP. Estos resultados ponen de manifiesto que las diferencias en la evaluación subjetiva entre el MVP y simulcast de los MP son muy pequeñas a velocidades binarias superiores para imágenes de poco movimiento («Flower pot» y «Trapeze») y/o con diferencia de luminancia importante entre las visiones izquierda y derecha («Fun fair»).
- «Fun fair» es la escena con notas medias más dispares. En esta escena puede observarse un número mayor de movimientos (cambios en el contenido de imagen con respecto a la trama siguiente) que en las demás escenas. En «Fun fair» especialmente, en los objetos en movimiento cubren la mayor parte de la imagen.

3 Trabajos futuros sobre televisión estereoscópica

Los progresos realizados hasta la fecha han puesto de manifiesto que la televisión estereoscópica es técnicamente viable. El perfil multivisión MPEG recientemente aprobado ofrece una base para la codificación y compresión de las secuencias de vídeo estereoscópico. Las pruebas de evaluación de calidad llevadas a cabo también han evidenciado que, dentro de

los límites de los parámetros de prueba seleccionados, puede lograrse una calidad de imagen cuya percepción subjetiva sea satisfactoria. No obstante, quedan muchas cuestiones por resolver. Algunos de los temas en los que hay que profundizar son los siguientes:

3.1 Requisitos

- Sería conveniente que los futuros sistemas de televisión estereoscópica fuesen compatibles con los sistemas de televisión digital monoscópica que están apareciendo en la actualidad, y que la velocidad binaria adicional fuera lo más reducida posible.
- La calidad de la imagen principal monoscópica de una pantalla de televisión monoscópica debe ser lo más próxima posible a la de la imagen monoscópica que utilice toda la capacidad del canal.

3.2 Información necesaria tanto para la televisión digital con definición convencional (TVDC) como para la TVAD

- El grado posible de asimetría de la velocidad binaria asignada a las imágenes izquierda y derecha correspondientes a una secuencia de vídeo estereoscópico para reducir al mínimo la degradación de la calidad de imagen del nivel de base.
- La repercusión de la asimetría de la velocidad binaria asignada a las imágenes de visión izquierda y derecha sobre las perturbaciones debidas a la codificación y la compresión percibidas subjetivamente y la calidad global de la secuencia de vídeo estereoscópico.
- Los factores que pueden provocar fatiga en el espectador; y las medidas que pudieran reducir o suprimir dicha fatiga.
- La gama de velocidades binarias que se precisa para conseguir una calidad satisfactoria percibida subjetivamente tanto de la imagen estereoscópica como de la imagen monoscópica proporcionada por la imagen del nivel de base. Obtenida mediante pruebas de evaluación adicionales con numerosas secuencias de vídeo que representen una gran variedad de material de programación y una amplia gama de velocidades binarias.
- Métodos de pruebas adecuados para evaluar las imágenes estereoscópicas.
- Algoritmos de codificación con los que se consiga una compresión más eficaz de las señales de televisión estereoscópica.

Los estudios al respecto deben realizarse en coordinación con el GT 11B, el GMT 10-11Q y demás Grupos de Trabajo y organismos pertinentes.
