



Realización multimedia

UNIVERSIDAD CATOLICA DE MANIZALES

FACULTAD DE INGENIERIA Y ARQUITECTURA

“trabajo de investigación sobre las nuevas alternativas y combinaciones para el transporte de video comprimido el cual contribuirá significativamente al tráfico en el Futuro de Internet y las redes inalámbricas. Por medio de los videos codificados (Comprimidos)”

TOMADO DEL LIBRO

VIDEO TRACES FOR NETWORK
PERFORMANCE EVALUATION

(PATRICK SEELING

Arizona State University, AZ, U.S.A.

FRANK H.P. FITZEK

Aalborg University, Denmark)

PRESENTADO POR:

GUSTAVO ALBERTO ROJAS LOPEZ

0T720091003

Estudiante

ASESOR:

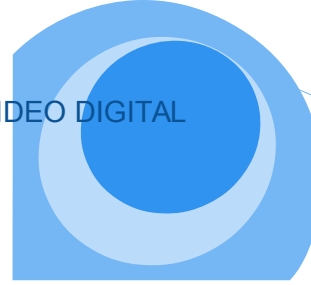
JAIME ALBERTO SEPÚLVEDA

Director Ingeniería Telemática

UCM

MANIZALEZ-CALDAS

11/16/2015



“trabajo de investigación sobre las nuevas alternativas y combinaciones para el transporte de video comprimido el cual contribuirá significativamente al tráfico en el Futuro de Internet y las redes inalámbricas. Por medio de los videos codificados (Comprimidos)”

PLANTEAMIENTO DEL PROBLEMA

En estos momentos en los que nos bombardean con "todo en digital" olvidamos algo muy importante: la compresión. La digitalización proporciona grandes ventajas con respecto a la señal analógica: más robustez frente al ruido, calidad y resolución muy superiores, edición y efectos muy sofisticados, ... pero esta digitalización origina una cantidad de información que supera las capacidades y se ven muy reducidas las posibilidades de proceso de nuestros sistemas.

Debido a la demanda por incorporar datos de video en los servicios de telecomunicaciones, y en especial en la industria del entretenimiento, la tecnología y en especial la de implementar más el video digital ahora es una necesidad. Debido a esto, la transferencia eficiente de este material se convirtió en una problemática que ha inspirado una gran cantidad de investigación.

El gran problema se resume en que el flujo de datos de imagen y video digital alcanza típicamente rangos cercanos a los 150 [Mb/s], lo cual es inaceptable por las tecnologías de comunicación masivas de la actualidad. Por esto, desde fines de los años 80 se ha intentado crear técnicas para reducir el flujo de video, en general, eliminando la redundancia de las imágenes para permitir la transmisión y almacenamiento de manera compacta y eficiente. Es por esto que hoy en día han surgido nuevas investigación y unas series de combinaciones para lograr una mejor calidad en el envío de los videos por medio de la red, donde han dado resultados muy alentadores para seguir en la búsqueda del mejoramiento; una muestra de los frutos de las investigaciones de las últimas décadas son los estándares que están vigentes hoy en día que son los MPEG y los H.26X. Sin embargo no son lo suficientemente buenos y necesitan mejoramiento, lo que hace que la investigación sea constante hasta llegar a la calidad, capacidad, nitidez y velocidad a la que se quiere llegar

ANTECEDENTES DE LA INVESTIGACION

ANTECEDENTES HISTORICOS

El desarrollo de las tecnologías en la última década ha dado un impulso notable a nuevos medios de comunicación, que hasta hace pocos años no pasaban de ser experimentos comunicacionales, con un radio de acción restringido. En este sentido, están llamadas a (PATRICK SEELING Arizona State University, AZ, U.S.A. y a FRANK H.P. FITZEK Aalborg University, Denmark) con su libro VIDEO TRACES FOR NETWORK PERFORMANCE EVALUATION. De donde me eh basado para hacer la documentación. Ellos en el tiempo de arduo trabajo y esfuerzo de investigación, han trabajado sobre vídeo comprimido donde han contribuido significativamente al tráfico en el Futuro de Internet y las redes inalámbricas. Para el transporte a través de redes, el video es típicamente codificado (es decir, comprimida) para reducir los requisitos de ancho de banda. Incluso vídeo comprimido, sin embargo, requiere grandes anchos de banda del orden de 100 kbps o Mbps. Además, secuencias de vídeo comprimidas presentan típicamente Tasas muy variables de bits (VBR), así como la dependencia de largo alcance (LRD) Propiedades. Esto, junto con la calidad estricta de servicio (QoS). Requisitos (pérdidas y retardo) de tráfico de vídeo, hace que el transporte de vídeo.

Donde han llegado a implementar parámetros y estándares actuales, y por medio de estos contribuyen al mejoramiento de la calidad de envío de videos por la red; agregando nuevos resultados de investigación y aportando a los nuevos estándares de codificación elementos que antes no se habían tenido en cuenta y que están dando buenos frutos para llegar al objetivo que hoy en día con el avance de la tecnología nos lo pide a gritos. Y no es de extrañar que prontamente nos sorprendamos con nuevos métodos de envío de video de alta calidad sin tener que omitir parte del contenido.

ANTECEDENTE HISTORICO 2

Un segundo trabajo corresponde a Br. Zapata L., Carlos E. de la Universidad Central de Venezuela donde en su trabajo especial de grado realizo el planteamiento de un modelo de interfaz gráfica de usuario que simule al sistema de gestión y Aprovisionamiento (SGA) del servicio de transporte de señales de TV a través de la red Metro Ethernet de CANTV. Me base en este trabajo de investigación ya que va de la mano

con el avance y aporta de forma positiva a el mejoramiento y la innovación del envío de video por medio de la red, este trabajo se basa en la utilización y el aprovechamiento de la gran telaraña que tenemos hoy en día la cual es el internet, que para el avance y el logro de los objetivos es el complemento más importante para el envío de video comprimido, se deben hacer modificaciones y complementos ya que es el soporte y la plataforma principal para que se cumplan lo deseado por los autores de estos proyectos y de miles de investigadores que quiere cumplir su logro y regalarnos un avance más al medio de la tecnología que hace parte de nuestro diario vivir.

JUSTIFICACION

En un mundo cada vez tan ocupado, lo que se busca es mejorar la calidad de vida, optimizar el tiempo lo mayor posible, e interactuar en tiempo real con las herramientas tecnológicas, es por esta razón que es de vital importancia informar a las sociedad de las diferentes posibilidades que ésta brinda. Es de ahí la importancia de esta investigación ya que hoy en día queremos todo a la mano y de mejor calidad y facilidad. La compresión de videos o mejor dicho video digital hace parte de él gran crecimiento de nuestra tecnología. Con esta investigación se quiere llegar a lo impensable e inimaginable; a quien no le gustaría que se le enviara un video de alta calidad por medio del móvil u otros dispositivos en un tiempo récor y sin omitir datos ni perder información del video; esto ya lo estamos viviendo, y uno de los ejemplos más claro y real es la televisión digital es uno de los temas que ha tomado mayor relevancia hoy en día en cuestión de desarrollo tecnológico, y son muchas las investigaciones y proyectos que se han llevado a cabo relacionadas con éste.

OBJETIVOS

General:

Investigar sobre los medios de compresión que en este documento verán reflejados estudiando la manera de digitalización de vídeo u compresión para que los archivos tengan tamaños razonables. Ya que si el vídeo no está comprimido estas imágenes ocuparán 1 Mb, aproximadamente. Si nuestro vídeo genera una frecuencia de 30 Mbps necesitaremos una capacidad de 1,5 Gb por minuto, es por esto de la necesidad de investigar y aportar a dicho proyecto ya que tendremos la necesidad de enviar videos de alta calidad por medios tecnológicos sin ocupar tanto espacio ni perder datos u información de nuestro archivo; y así poder disfrutar de un producto de alta calidad y eficiente y lo mejor utilizando la nube esto no ahorrara espacio tiempo y nos dará calidad.

Específicos

- Estudiar cada una de la codificación o codecs que reducen considerablemente el tamaño de los archivos de vídeo. Esto los hace más manejables y fácilmente almacenables.
- Ver la manera de reducir y eliminar datos redundantes del vídeo para que el archivo de vídeo digital pueda enviarse y almacenarse de manera eficiente.
- Investigar sobre los algoritmos existentes para poder aportar nuevos elementos de codificación y hacer de la compresión un éxito
- Utilizar cada uno de los algoritmos necesarios y aplicarlos al vídeo original para crear un archivo comprimido y ya listo para ser transmitido o guardando y así estudiar nuevos métodos.
- Tratar de utilizar algoritmos y estándares actuales avanzados ya que cuanto más avanzado es el algoritmo de compresión, mayor es la latencia a igual potencia de procesamiento y así estaremos más cerca de lo que se quiere lograr.
- Utilizar programas de simulación para ver reflejado parte de los resultado de la investigación.

CONTENIDO

1 Introducción	8
Parte I Digital Video	
2 Introducción al vídeo digital	12
2.1 El comienzo de Moving Pictures.	12
2.2 Imagen Digital y Vídeo Representación.	14
2.3 Jerarquía de vídeo.	22
3 Codificación de vídeo	24
3.1 DCT-base de codificación de vídeo.	24
3.1.1 Bloque de barrido.	26
3.1.2 Transformación Discreta de Coseno.	27
3.1.3 Cuantificación.	28
3.1.4 Zig-Zag de barrido.	32
3.1.5 Codificación de longitud variable.	33
3.2 Inter-trama de codificación: Estimación de movimiento y compensación.	34
3.3 Codificación de vídeo escalable.	40
3.3.1 La partición de datos.	41
3.3.2 Escalabilidad Temporal.	42
3.3.3 Escalabilidad espacial.	43
3.3.4 Escalabilidad SNR.	44
3.3.5 Escalabilidad de objetos.	45
3.3.6 Fine Granular Escalabilidad (FGS).	46
3.3.7 Descripción Codificación múltiple (MDC).	48
3.4 Wavelet basado en codificación de vídeo.	49
3.5 Vídeo Estándares de Codificación.	52
4 Métricas y Estadísticas para Trazas de vídeo	55
4.1 Tamaño de fotograma de vídeo.	57
4.1.1 Autocorrelación.	59
4.1.2 Varianza-Tiempo de prueba.	60
4.1.3 R / S Estadística.	60
4.1.4 Periodograma.	62
4.1.5 Diagrama Logscale.	64
4.1.6 Diagrama Multiescala.	64
4.2 Calidad del marco del vídeo.	65
4.3 Correlación entre tamaños de fotogramas de vídeo y cualidades.	70
4.4 Las métricas adicionales para FGS codificaciones.	72
4.5 Métricas adicional para MDC codificaciones.	75
Referencias	77

Introducción

Multimedia aplicaciones de red y, en particular, el transporte de comprimido.

Se espera vídeo a contribuir significativamente al tráfico en el Futuro de Internet y las redes inalámbricas. Para el transporte a través de redes, el video es típicamente codificado (es decir, comprimida) para reducir los requisitos de ancho de banda.

Incluso vídeo comprimido, sin embargo, requiere grandes anchos de banda del orden de 100 kbps o Mbps. Además, secuencias de vídeo comprimidas presentan típicamente Tasas muy variables de bits (VBR), así como la dependencia de largo alcance (LRD) Propiedades. Esto, junto con la calidad estricta de servicio (QoS).

Requisitos (pérdidas y retardo) de tráfico de vídeo, hace que el transporte de vídeo.

Tráfico en las redes de comunicación esto es un problema difícil. Como consecuencia de ello, en la última década, la comunidad de investigación de redes ha sido testigo de una explosión en la investigación sobre todos los aspectos de transporte de vídeo. Las características de tráfico de vídeo, el modelado de tráfico de vídeo, así como protocolos y mecanismos para el transporte eficiente de los flujos de vídeo, han recibido una gran cantidad de interés entre los investigadores de redes y operadores de redes y una gran cantidad de vídeo esquemas de transporte se han desarrollado.

Para el desarrollo y la evaluación de los mecanismos de transporte de vídeo y para la investigación en el vídeo de redes en general, es necesario tener algunas características disponibles del video. En general, hay tres formas diferentes para caracterizar vídeo codificado con el propósito de la creación de redes de investigación: (i) el tráfico de vídeo Modelo, (ii) la corriente de bits de vídeo, y (iii) traza el tráfico de vídeo. Modelos de tráfico de vídeo, tal como [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11] se esfuerzan para capturar las propiedades esenciales del tráfico real, preciso y eficiente computacionalmente con modelos matemáticos. Un modelo de tráfico es típicamente desarrollado en base a la estadística, propiedades de muestras del tráfico real, o, en muchos casos, los rastros de vídeo del tráfico real. Rastros de vídeo suele ser un requisito previo para el modelo del

desarrollo. El modelo de tráfico desarrollado se verifica al comparar el tráfico general con las huellas de vídeo. Si el modelo de tráfico se considera suficientemente Exactitud, se puede utilizar para el análisis matemático de las redes, para el modelo de las simulaciones impulsadas, y también para la generación de los llamados de vídeo (sintético) virtuales de rastros.

La corriente de bits de vídeo se genera utilizando el codificador y contiene la completa información de vídeo. La caracterización del tráfico (por ejemplo, el tamaño del marco) puede ser obtenido mediante la medición del tráfico o mediante el análisis del flujo de bits. El vídeo y su calidad puede determinarse mediante subjetiva (de visualización) Evaluación [12] u objetivo, métodos [13, 14, 15]. La ventaja de la corriente de bits es que permite la creación de redes y experimentos en los que la calidad del video-después de sufrir pérdidas en la red se evalúa, por ejemplo, [16, 17, 18, 19, 20, 21]. Las técnicas presentadas en el capítulo 9 de llevar esta capacidad para evaluar la calidad de vídeo con pérdidas y después transporte de red a las trazas de vídeo. Una limitación de la corriente de bits es que es muy grandes en tamaño; varios GBytes para una hora de vídeo o varios comprimido decenas de GBytes para una hora de vídeo sin comprimir. Otra limitación de bit común y corrientes es que por lo general son de propiedad y / o protegidos por derechos de autor.

Esto limita el acceso de los investigadores a flujos de bits de red, y también los límites, el intercambio de flujos de bits entre los grupos de investigación hay una limitación clave adicional de los flujos de bits los cuales requieren experiencia en la codificación de vídeo. Como En consecuencia, sólo las personas que tengan conocimientos especializados y el equipamiento necesario tanto para la codificación de vídeo y la creación de redes de investigación pueden llevar a cabo la investigación de redes de vídeo con flujos de bits.

Trazas de video son una alternativa atractiva a los modelos de tráfico y flujos de bits ya que representan el tráfico y la calidad de los videos. Mientras que dan los bits reales que llevan la información de vídeo, las huellas sólo dan el número de bits utilizados para la codificación de los cuadros de vídeo individuales y el nivel de calidad (por ejemplo, en

PSNR) de la codificación. Por lo tanto, no hay copyright. Es importante destacar que, trazas de vídeo se extienden significativamente el conjunto de individuos que puede llevar a cabo la investigación de alta calidad de vídeo en red, proporcionando las correspondientes caracterizaciones de vídeo en archivos de seguimiento. Los archivos de seguimiento pueden ser procesados con PCs estándar y utilizarse en simulaciones de red estándar, por lo tanto permitiendo a los investigadores de redes sin experiencia o el equipo de codificación de vídeo para llevar a cabo la investigación de redes de vídeo.

Como huellas de vídeo son una muy conveniente caracterización de vídeo para la creación de redes; la investigación, las huellas a estimular la investigación de redes de vídeo. De hecho, la creación de redes y la investigación experimentó una explosión inicial en la investigación en el transporte de vídeo después de los MPEG-1 rastros [22, 23, 24, 25] se convirtió públicamente, disponible alrededor de 1995. Estos primeros vestigios fueron muy elementales en la que (i) la huellas solo era un pequeño conjunto de vídeos, (ii) los rastros representaban sólo vídeo codificado con el menos eficiente códec de MPEG-1 a un nivel de calidad único, y (iii) no hay rastros de vídeo codificado escalable que estuvieran disponibles. Sin embargo, estos MPEG-1 proporcionaron una base útil para una gran cantidad de redes de vídeo en la investigación.

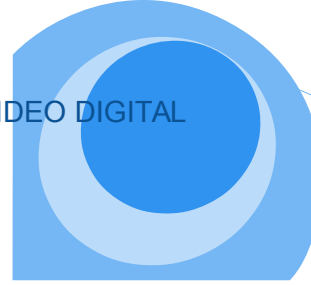
Nuestra participación en la creación y difusión de rastro de vídeo comenzó en 1999 en el instituto de Redes de Telecomunicaciones (NTK) dirigido por el Prof. Adán Wolisz en la Universidad Técnica de Berlín. Creamos rastros de una forma moderada gran número de vídeos con el formato MPEG-4 y H.263 codecs para una gama de diferentes escalas de cuantificación (niveles de calidad). Nos centramos en una sola capa (no escalable) codificada y vídeos e incorporado ambas codificaciones sin y con.

El control de la tasa y su objetivo era una tasa de bits de destino prescrito. Estas huellas fueron primero utilizadas para simular el tráfico de video en el estudio de streaming de vídeo inalámbrico [26] y descrito en el artículo [27]. Las codificaciones con una gama de niveles de calidad permitido y la simulación del transporte de red de diferentes versiones de

el mismo vídeo, ampliando así la gama de escenarios de red que pueden ser examinados utilizando trazas de vídeo. Al mismo tiempo, el análisis estadístico de las huellas revelan una forma característica "joroba" de la variabilidad de la frecuencia de bits representada gráficamente como una función de la calidad de vídeo [28].

Una dirección es importante para trazas de vídeo codificada escalable. La escalabilidad de la codificación de vídeo es prometedora ya que permite la transmisión a través de redes heterogéneas proporcionando velocidades de bits variables y a receptores heterogéneos con diferentes formatos de visualización (tamaño de pantalla) y capacidades de procesamiento con una sola codificación. Hemos generado trazas para la codificación MPEG escalable en capas, así como para la multa granular codificación MPEG escalable que permite la escalabilidad en un poco de granularidad. También generó rastros de codecs basados en una onda básica, que proporcionan escalabilidad muy flexible. Un importante avance conceptual para la trazas de vídeo han sido los rastros de distorsión de compensación que permiten evaluar el video, y su calidad después de transporte de red con pérdidas, una capacidad que antes sólo podía ser logrado a través de experimentos con corrientes reales de bits de vídeo.

Este informe promete una introducción completa a las trazas de video y su utilizar en la creación de redes de investigación. Después de la primera prestación de los conceptos básicos de vídeo digital y la codificación de vídeo, introducimos las huellas de vídeo, que cubre las métricas capturadas en las huellas, la generación de trazas, así como las características y estadísticas del vídeo que caracteriza en los rastros. Luego nos dirigimos nuestra atención a la utilización de las huellas de vídeo en la creación de redes de investigación, el examen de los aspectos prácticos de el transporte de vídeo a través del protocolo de Internet (IP) y redes de la simulación de transporte de vídeo usando los rastros, incluyendo las simulaciones utilizando el desplazamiento trazas distorsión. Las herramientas de software y utilidades que facilitan el uso del video y las huellas en simulaciones de red y otros programas relacionados con la creación de redes de vídeo También se presentan herramientas.



CAPITULO 2

Introducción al vídeo digital

En este capítulo, le damos una introducción al vídeo digital y las diferencias entre los diferentes estándares de vídeo. De los diferentes modos de visualización digital, vídeo, nos fijamos en los diferentes espacios de color y las conversiones entre ellos. Para el Espacio de color YUV, miramos más detallado en las formas de representación para cada píxel individual o grupos de píxeles. A partir de esta unidad más pequeña, que semánticamente integra diferentes niveles de una jerarquía de vídeo a una película completa.

2.1 El comienzo de Moving Pictures

A principios de la década de 1900 las imágenes en movimiento se hizo popular. Una de las primeras realizaciones de imágenes en movimiento fue el libro animado. El libro de tapa es la forma más simple de animación usando una secuencia de imágenes fijas con pequeños cambios de una imagen a la siguiente. Para utilizar el libro animado, el espectador comienza con la primera imagen y voltea a través de las siguientes imágenes. Si la tasa de flipping es suficientemente alta, se crea la ilusión de movimiento, un fenómeno conocido como persistencia de la visión. El fenómeno se basa en el hecho de que la retina del ojo humano es retener cualquier imagen durante un tiempo dado. Por lo tanto, si la tasa de retención es mayor de 16 imágenes por segundo, el cerebro humano superpone las imágenes creando la ilusión de movimiento real. Para describir que el ojo humano es sensible para el número de imágenes por segundo, la llamada "tasa de fusión parpadeo" era

introducida. Se trata de una media estadística para describir cuando el parpadeo de la mudanza de imágenes que se observa. La tasa de parpadeo de fusión varía dependiendo de los diferentes observadores,

Los primeros mecanismos para crear la ilusión de imágenes en movimiento, como por ejemplo el libro de tapa, se basa en la iluminación pasiva. Más tarde, con la introducción de película química, la luz, generada por una fuente activa, se pasó a través de las imágenes. El mayor número de imágenes por segundo de la película química.



(a) Camera



(b) Projector

Fig. 2.1: Traditional film camera and projector.

Hoy en día sería duro y, por tanto, prácticamente imposible. La mayor parte de la película y el estilo de cámaras, como el que se ilustra en la Figura 2.1, disparar la película con 16, 18 o 24 cuadros por segundo. Como esta velocidad de cuadro es muy por debajo de la fusión de parpadeo (Ver Figura 2.1).

El disparador tiene dos objetivos principales. El primer objetivo es suspender la la imagen y tiene que ser eliminado. Además, el obturador muestra las imágenes reales puede ser compensada por lámparas más potentes dentro del proyector de comparación a un proyector de diapositivas. Disparar la película con 16, 18 cuadros por segundo, lo que se denomina obturación (ver Figura 2.2) se introdujo al proyector de cine la luz del proyector cuando la siguiente imagen tiene que ser demostrado y la vieja varias veces, aumentando así la velocidad de fotogramas, por esto. El obturador es nada mas una tasa virtual de los 48 fotogramas por segundo. La razón de 18 cuadros por segundo es suprimir el parpadeo, el inconveniente es que el brillo se reduce. Este más que un disco giratorio con áreas en que se detiene para que la luz pase a través de la tasa de parpadeo, pero por razones de grabación

de audio. Por tres persianas aumentan la velocidad de fotogramas de 24 y 16 fotogramas por segundo hasta suspenderlo. Las persianas típicas son dos y tres postigos de las alas. Los dos y tres postigos de las alas aumentan la velocidad de fotogramas de 24 y 16 fotogramas por segundo para una tasa virtual de los 48 fotogramas por segundo. La razón de 18 cuadros por segundo es no motivado por la tasa de parpadeo, pero por razones de grabación de audio. Por suprimir el parpadeo, el inconveniente es que el brillo se reduce. Esto puede ser compensado por las lámparas más potentes dentro del proyector comparado con un proyector de diapositivas.

2.2 Imagen Digital y Vídeo Representación

El video consiste en una secuencia de tramas o imágenes de vídeo individuales, que se muestra a una cierta velocidad de fotogramas. La pantalla consecutiva de imágenes a continuación, crea el efecto de movimiento capturado. En la grabación de vídeo tradicional en rollos de película, por ejemplo, la cámara captura a una velocidad de 24 cuadros.

2.2 Digital Picture and Video Representation



Fig. 2.2: A shutter in a film projector.



Fig. 2.3: Concept of capturing motion in film.

por segundo (fps). Esta velocidad de cuadro da la impresión de movimiento continuo para el ojo humano cuando se reproducen. La pantalla basada en película tradicional se ilustra en la Figura 2.3.

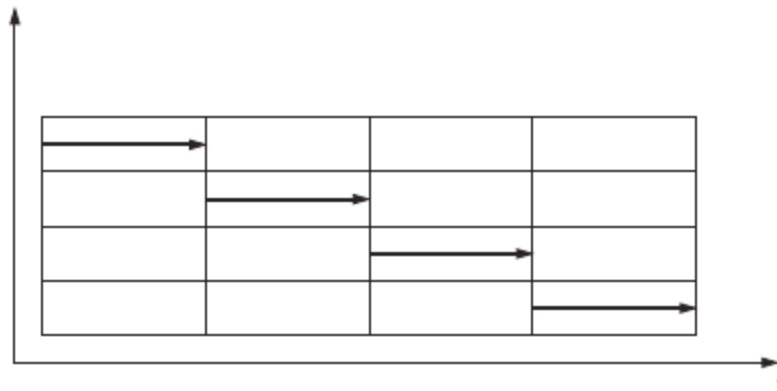


Fig. 2.4: Line display in progressive video.

Existen diferentes estándares para velocidades de fotogramas. Vídeo progresivo dibuja todas las líneas individuales de una imagen en secuencia y los muestra como en la Figura

2.4. El (NTSC) Formato de velocidad de fotogramas National Television Standards Committee es fijado en 29,97, o aproximadamente 30 fps. El Alternating Line Fase (PAL) estándar utiliza 25fps. La velocidad de fotogramas de Super8 es 18fps. En la televisión normal, conjuntos, sin embargo, la visualización de los fotogramas de vídeo se realiza con el doble de esa frecuencia, por lo que los cambios en las imágenes se capturaron con sólo el envío de la mitad de las líneas que comprenden la resolución de la pantalla de televisión completa. Este concepto se llama entrelazado, y mediante la combinación de las líneas entrelazadas, la velocidad de fotogramas se da de vídeo no entrelazado o progresivo. La Figura 2.5 ilustra el concepto entrelazado de línea sabio. El concepto de entrelazado es que el cerebro humano y los ojos trabajan juntos para eliminar las discrepancias que son causados por el mecanismo de entrelazado.

Con el advenimiento de la representación digital de cine y vídeo y la codificación del material de origen, el enfoque progresivo determina la frecuencia en términos de fotogramas por segundo que se utiliza a lo largo de este informe. Con Codificadores y decodificadores de vídeo que trabajan juntos en las computadoras y decodificadores antes de enviar la señal al aparato de televisión conectado (y la introducción artificial entrelazada), podemos suponer que el video se procesa y se muestra Sobre una base de fotograma completo, por lo que la visualización de las imágenes individuales se produce en los puntos en el tiempo dado por la velocidad de cuadro elegido. Nos ilustran el concepto que asumimos lo largo de este informe en la Figura 2.6.

Cada fotograma de vídeo individual consiste de elementos de imagen (por lo general a que se refiere como píxeles o elementos de imagen). El formato de trama especifica el tamaño del individuo marcos en términos de píxeles. El formato de la UIT-R / CCIR-601 (la TV común tiene formato) tiene 720×480 píxeles (es decir, 720 píxeles en la dirección horizontal y 480 píxeles en la dirección vertical), mientras que el formato Common Intermediate.

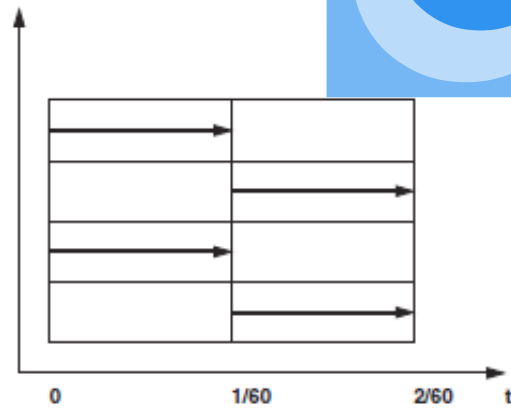


Fig. 2.5: Line display in interlaced video.

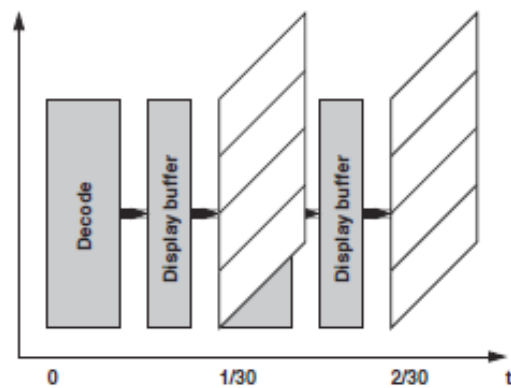


Fig. 2.6: Decoding, storing in frame buffer of the graphics card and display of digitally encoded video.

Formato (CIF) tiene 352×288 píxeles y el formato Barrio CIF (QCIF) tiene 176×144 píxeles; los formatos CIF y QCIF típicamente se consideran en la red estudios relacionados. Diferentes espacios de color se utilizan para representar un solo píxel. El común espacio de color en el dominio de la computadora es la representación sobre la base de los tres colores componentes rojo, verde y azul (RGB). Con estos tres componentes el color de un píxel puede ser definido. En la transmisión de vídeo, los píxeles están representados por tres componentes diferentes, el componente de luminancia (Y), y el dos componentes de crominancia tonalidad (U) e intensidad (V). Esta representación se remonta a los primeros días de la televisión en color blanco y negro y más tarde. La representación de color YUV era necesaria para transmitir señales de televisión en color

12 al tiempo que permite a los viejos aparatos de televisión en blanco y negro para funcionar sin modificaciones.

Como la información de luminancia se encuentra en diferentes bandas de frecuencia, Los viejos sintonizadores son capaces de sintonizar en la señal Y solo. La conversión entre estos dos espacios de color está definido por una matriz de conversión fijo. La matriz de conversión general para la conversión de RGB a YUV se da como en

La ecuación (2.1)

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.1)$$

Estos valores se utilizan para convertir los valores RGB al espacio de color YUV utilizado en los sistemas PAL. Para los sistemas NTSC, la conversión se da como en la Ecuación (2,2).

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.595716 & -0.274453 & -0.321263 \\ 0.211456 & -0.522591 & 0.311135 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.2)$$

Del mismo modo, los valores RGB se pueden convertir de YUV o YIQ. Para todos la conversión fines, los valores tienen que ser asignada a la gama típica de [0 . . 255] que se utiliza en entornos digitales de 8 bits. Para obtener más conversiones y de color espacios, que se refieren a [29]. A continuación, nos centramos en la representación YUV que se utiliza normalmente en sistemas de compresión de vídeo.

Existen varios diferentes formatos YUV. Se pueden clasificar por el sub- de muestreo que se usa entre los diferentes componentes y la forma en que el pixel individual y este formato se conoce como YUV 4: 4: 4 y se ilustra en la Figura 2.7. El ojo humano, sin embargo, es

mucho más sensible a los cambios en luminancia que a los otros componentes. Por tanto, es común para reducir la información que se almacena por imagen por submuestreo de crominancia. Con sub-muestreo de la relación de crominancia de bytes de luminancia se reduce. Más específicamente, sub-muestreo representa un grupo de por lo general de cuatro píxeles por su cuatro componentes de luminancia (bytes) y un conjunto de dos valores de crominancia. Cada uno de estos dos valores de crominancia típicamente se obtiene promediando los correspondientes valores de crominancia en el grupo. En el caso de los cuatro píxeles son agrupados como un bloque de 2 x 2 píxeles, el formato es YUV 4: 2: 0. Si las luces agrupadas píxeles están formando una línea de 4×1 , el formato se conoce como YUV 4: 1: 1.

Estos dos formatos más comunes de muestreo YUV se ilustran en las Figuras 2.8 y 2.9. Utilizando el enfoque de promediación para obtener los valores de crominancia para YUV 4: 1: 1 de YUV 4: 4: 4, los valores de matiz se puede calcular como

$$U_{1,1}(411) = \frac{U_{1,1}(444) + U_{2,1}(444) + U_{3,1}(444) + U_{4,1}(444)}{4}, \quad (2.3)$$

y así sucesivamente para los píxeles restantes. Los valores de saturación se calculan en una manera similar. Utilizando el enfoque de promediación para obtener la crominancia los valores se almacenan. Los valores originales Y, U, y V se pueden almacenar para cada

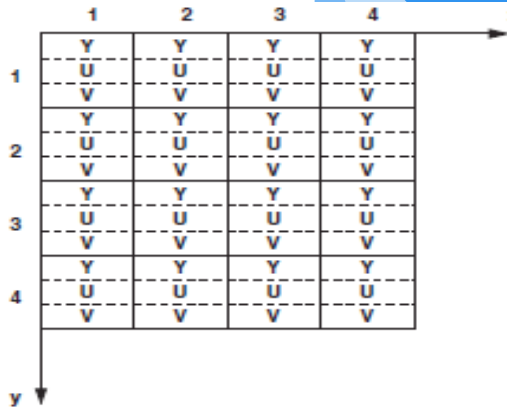


Fig. 2.7: YUV 4:4:4 without any sub-sampling.

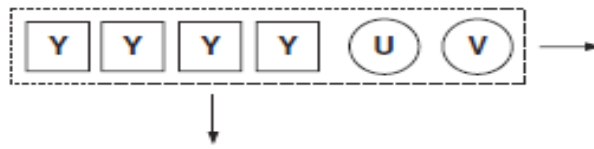


Fig. 2.8: YUV 4:1:1 subsampling.

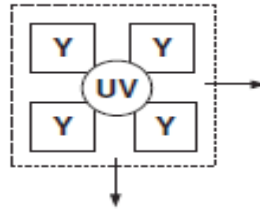


Fig. 2.9: YUV 4:2:0 subsampling.

valores de YUV 4: 2: 0 de YUV 4: 4: 4, los valores de matiz se puede calcular como

$$U_{1,1}(420) = \frac{U_{1,1}(444) + U_{2,1}(444) + U_{2,1}(444) + U_{2,2}(444)}{4}. \quad (2.4)$$

El segundo método de caracterizar los diferentes formatos YUV es por la manera en que las diferentes muestras se almacenan en un archivo para los diferentes componentes. Los valores se pueden almacenar ya sea empaquetado o plana.

Repleto de almacenamiento guarda los valores de píxel-sabio para cada componente del bloque creado, mientras que el método de almacenamiento plana guarda la información del

píxel en arrays siguiente entre sí. El ejemplar YUV 4: 4: 4 (progresiva) en un Formato de empaquetado tanto, sería en un archivo almacenado con los bytes en el orden Y (1,1), U (1,1), V (1,1), Y (2,1), U (2,1), V (2,1),. . . . El YUV más comúnmente usado 4: 2: 0 se almacena en formato plano y se compone de todos los valores de bytes para el componente Y, seguido por los valores de bytes para los componentes U y V. Esto resulta en un archivo (por 176×144 píxeles que comprenden una imagen) similares a Y (1,1), Y (2,1),. . . , Y (172.144), seguido directamente por U (1,1) (420),. . . ,

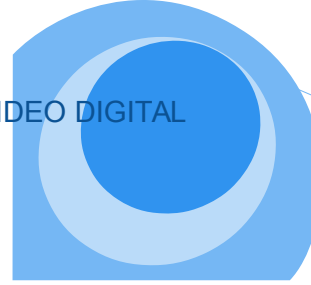
$$U_{(88,72)}(420) = \frac{U_{(176,143)}(444) + U_{(176,143)}(444) + U_{(176,144)}(444) + U_{(176,144)}(444)}{4},$$

que se siguen directamente por V (1,1) (420),. . . , V (88,72) (420). Así, el tamaño de una trama con YUV 4: 2: 0 (o 4: 1: 1) sub crominancia toma de muestras en el formato QCIF (176 columnas de píxeles por 144 filas de píxeles para la componente de luminancia y la mitad de las filas y columnas para cada uno de los dos componentes de crominancia) es

$$176 \cdot 144 \cdot \left(8 \text{ bit} + 2 \cdot \frac{8 \text{ bit}}{4} \right) = 304,128 \text{ bit} = 38,016 \text{ byte.} \quad (2.5)$$

Los tamaños de trama y velocidades de datos para los diferentes formatos de vídeo y el marco las tasas se resumen en la Tabla 3.2. Claramente visible ya desde el pequeño resolución de vídeo y tamaño del marco resultante en la ecuación (2.5), la transmisión de este baja resolución de vídeo a la velocidad de fotogramas NTSC da un enorme ancho de banda

$$304,128 \frac{\text{bit}}{\text{frame}} \cdot \frac{30 \text{ frames}}{\text{sec}} = 9,123,840 \frac{\text{bit}}{\text{sec}} \approx 9.1 \text{ Mbps.} \quad (2.6)$$



2.3 Jerarquía de vídeo

En el vídeo digital en general no se procesa, almacena, comprimido y transmitido sobre una base por píxel, pero en una jerarquía [30], como se ilustra en la Figura 2.10. En la parte superior de esta jerarquía es la secuencia de vídeo, que se divide en individuales escenas. Un ejemplo para una escena podría ser una discusión de varias personas. La escena se puede dividir en varias tomas, por lo que se utilizan vacunas para dramatización Efectos por el Director. Siguiendo el ejemplo, el director podría haber introducido varias tomas de la discusión que muestran las personas con diferentes ángulos en las cámaras. Estos primeros niveles de segmentación de vídeo se deben a la componente artística en las secuencias de vídeo y tienen un significado semántico. Por debajo del nivel de disparo

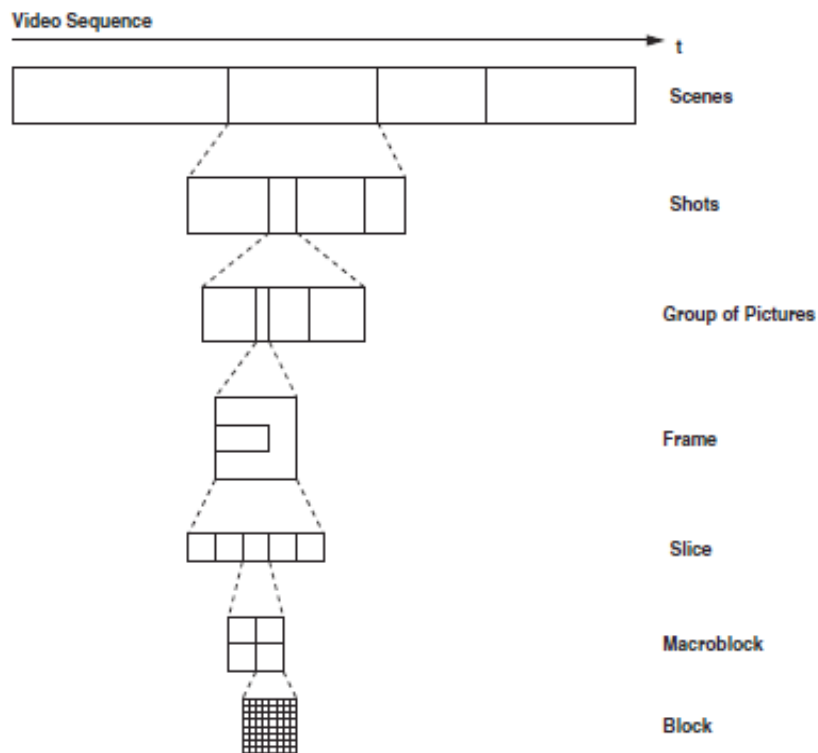


Fig. 2.10: Typical composition of a video sequence.

Vienen los Grupos de imágenes (GOP). Cada GP a su vez se compone de múltiples Fotogramas de vídeo. Una sola trama se divide en rebanadas. Rebanadas se representan independiente en unidades de codificación que pueden ser decodificados sin hacer

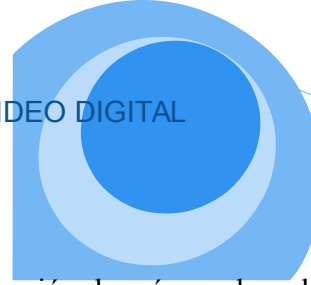
referencia a otras rebanadas del mismo marco. Se componen típicamente de varios macrobloques consecutivos. Rebanar pueden ser utilizados para lograr un mayor error-robustez. Cada segmento consiste de varios macrobloques (MB), cada uno por lo general consta de 4×4 bloques. Cada bloque normalmente consta de 8×8 píxeles. Si bien la segmentación automática de vídeo todavía está experimentando importantes esfuerzos de investigación, los niveles que son más relevantes para la codificación de vídeo y decodificación son del nivel GdP hacia abajo. Diferente vacunas pueden tener diferentes contenidos y por lo tanto el uso de diferentes patrones GdP puede ser beneficiosa, especialmente cuando se considera una transmisión potencialmente con pérdida.

La compresión de vídeo explota generalmente tres tipos de redundancias [30]. Sobre una base por-marco (es decir, una sola imagen), los píxeles vecinos tienden a correlacionarse y así tener redundancia espacial [31]. Codificación intra-frame se emplea para reducir la redundancia espacial en una trama dada. Además, consecutiva en los marcos tienen similitudes y, por tanto, la redundancia temporal. Esto temporalmente los despidos se reducen por medio de técnicas de codificación inter-frame. El resultado de la reducción de estos dos redundancias es una corriente de palabras de código (símbolos) que tiene algo de redundancia en el nivel de símbolo. La redundancia entre estos símbolos se reduce por codificación de longitud variable antes de pasar el código binario a la salida canal.¹ La eliminación de estas redundancias se explica en la siguiente, ya que damos un panorama introductorio de diferente codificación de vídeo 1 esquemas de compresión adicionales, tales como la explotación de reconocimiento de objetos de técnicas, también están en desarrollo, pero no se aplican comúnmente hasta ahora.

principios, nos referimos al lector interesado a [30, 32] para más detalles.

¹ esquemas de compresión adicionales, tales como la explotación de reconocimiento de objetos técnicas, también están en desarrollo, pero no se aplican comúnmente hasta ahora.

Capitulo 3



CODIFICACION DE VIDEO

Se van a mostrar los diferentes métodos de codificación, la más usada es la transformada de coseno discreta (DCT) sin codificación predictiva y los diferentes mecanismos que se utilizan en el proceso de aplicar la DCT en los codificadores de vídeo modernas,

Continuamos con la introducción de los mecanismos de codificación predictiva con sus respectivas complejidades y los diferentes métodos de codificación de vídeo escalable. Una introducción a la codificación de vídeo basado en wavelets y estándares de vídeo actuales concluir este capítulo.

DCT- CODIFICACIÓN DE VÍDEO BASADO

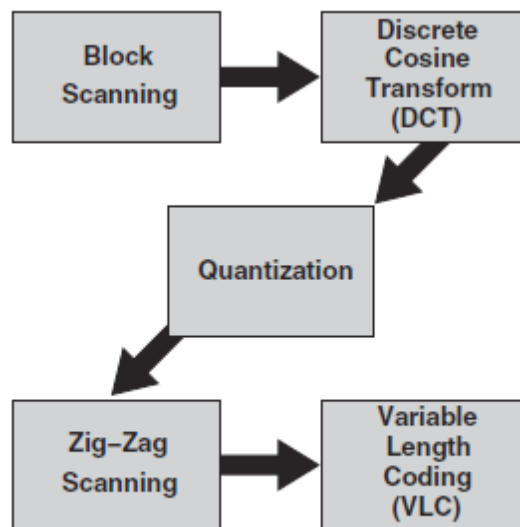
Nos centramos inicialmente en los principios empleados en los estándares MPEG y en una sola capa (no escalable) de codificación de vídeo. Los principios fundamentales de MPEG codificación de vídeo son de codificación utilizando el coseno transformada discreta intra-trama (DCT) e inter-trama de codificación mediante la estimación de movimiento y compensación entre fotogramas de vídeo sucesivas. El enfoque DCT se usa comúnmente en el video codificadores / decodificadores debido a la complejidad asociada con la baja las transformadas.

Como dentro de la codificación solo da una pequeña compresión en la porción, la inter-trama codificación es usada para aumentar la compresión de la porción. Para la codificación de cada fotograma de vídeo intra-trama se divide en bloques de 8×8 muestras de muestras

de Y, U, muestras y muestras V. Se transforma cada bloque utilizando la DCT en un bloque de 8×8 coeficientes de transformación, que representan las componentes de frecuencia espacial en el bloque original. Estos coeficientes de transformación se cuantifican a continuación por una matriz de 8×8 de cuantificación que contiene el tamaño de paso de cuantificación para cada coeficiente. Se obtiene la matriz de cuantificación multiplicando

una matriz de base por una escala de cuantificación. Esta escala de cuantificación se utiliza típicamente para controlar la codificación de vídeo. A escala de cuantificación más grande da una cuantificación más basta que resulta en un tamaño más pequeño (en bits) de la codificada del fotograma de vídeo así como de una calidad inferior. Los coeficientes cuantificados son entonces zigzag escaneada, a nivel de gestión codificado, y de longitud variable codificada para lograr una mayor la compresión.

La intra-codificación (compresión) de un fotograma de vídeo individual. Se basa comúnmente en la transformación discreta del coseno (DCT). (También han surgido esquemas de transformación basado en wavelets. Estudios indican que en el campo de la codificación de vídeo, el enfoque basado en wavelets no mejora la calidad del vídeo transformado significativamente. El intra-trama de codificación produce la partición de la estructura en bloques, también conocida como la exploración de bloque. El tamaño de estos bloques de hoy es típicamente de 8×8 píxeles (anteriormente, también se han utilizado 4×4 y 16×16). La DCT es entonces aplicada a los bloques individuales. Los coeficientes DCT resultantes se cuantifican y zig-zag digitaliza de acuerdo la calidad de la imagen.



3.1.1 Bloque de

Para reducir la necesaria para la subdivide primero

macrobloque es finalmente subdividido en bloques de 4×4 . La utilización de bloques-formas para la codificación es una de las limitaciones para los sistemas de compresión

barrido

potencia del cálculo DCT, la trama original se en macrobloques. Cada

Fig. 3.1: DCT coding concept.

basada en DCT. Las formas de los objetos típicos de las imágenes son irregulares y por lo tanto no se pueden montar en bloques rectangulares, Con el fin de aumentar la eficiencia de la compresión, los diferentes tamaños de los bloque puede utilizado a costa de una mayor complejidad. Con la estandarización de H.264. El estándar H.264 soporta siete modos de macrobloque de subdivisión diferentes, donde cada macrobloque puede ser subdividido en fragmentos más pequeños con el fin de proporcionar una más fina granularidad y de mayor calidad

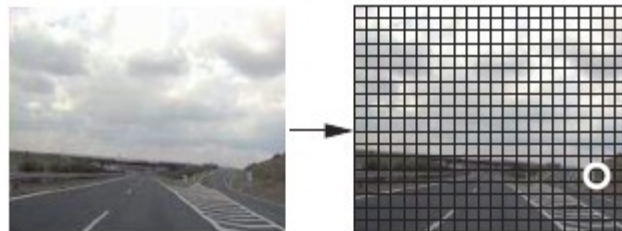


Fig. 3.2: Video frame subdivision into blocks (QCIF format into 22×18 blocks of 8×8 pixels each).

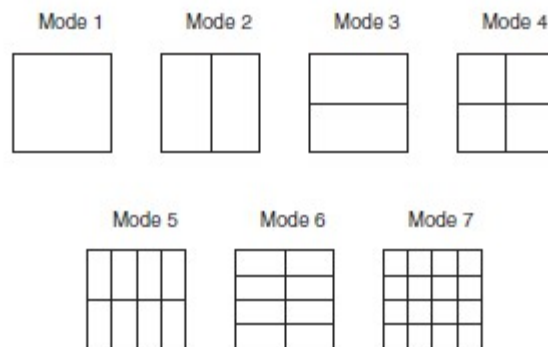


Fig. 3.3: Different macroblock subdivision modes that are supported by the H.264 video coding standard.

3.1.2 Transformación Discreta de Coseno

El DCT se utiliza para convertir un bloque de píxeles en un bloque de coeficientes de transformación. Los coeficientes de transformación representan la frecuencia espacial de

los

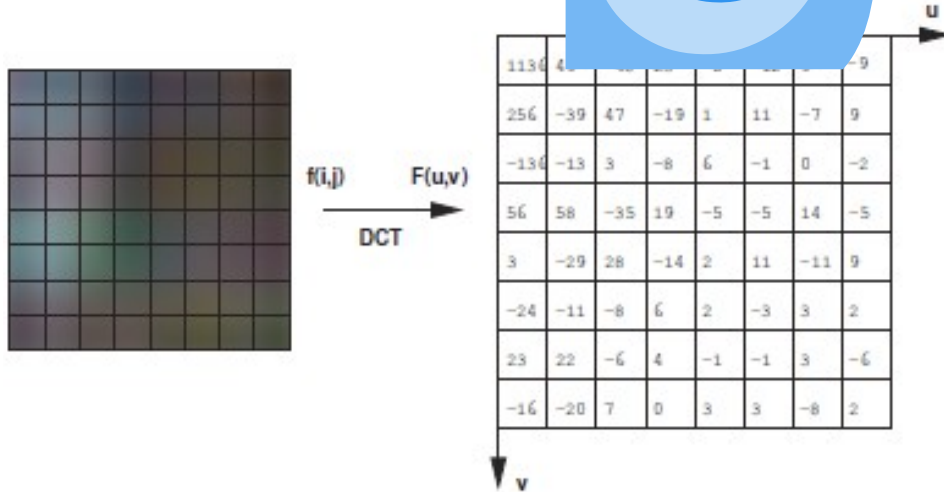


Fig. 3.4: 8 × 8 block of luminance values (visual representation and numerical values) and the resulting DCT transform coefficients (decimal places truncated).

componentes del bloque original. Un ejemplo de esta transformación del bloque marcado en la Figura 3.2 se ilustra en la Figura 3.4.

Esta transformación es sin pérdidas, simplemente cambia la representación del bloque de pixels, o más precisamente el bloque de valores de luminancia (crominancia). El DCT de dos dimensiones para un bloque de N x N de píxeles puede ser descrito como dos DCT unidimensionales consecutivos (es decir, horizontal y vertical). Con f (i, j) denota los valores de píxel y F (u, v) denota los coeficientes de transformación que tener

$$F(u, v) = \frac{2}{N} \cdot C(u) \cdot C(v) \cdot \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i, j) \cos \left(\frac{(2i+1)u\pi}{2N} \right) \cos \left(\frac{(2j+1)v\pi}{2N} \right), \quad (3.1)$$

where

$$C(x) = \begin{cases} \frac{1}{\sqrt{2}}, & x = 0 \\ 1, & \text{otherwise.} \end{cases} \quad (3.2)$$

El coeficiente de orden más bajo se denomina generalmente como el componente de CC, mientras que los otros componentes se denominan AC-componentes.

3.1.3 Cuantificación

En las tramas de video la energía se concentra en los coeficientes de baja frecuencia. Es decir, unos coeficientes con u y v cerca de cero tienen una alta significación para la representación de bloque original. Por otro lado, los más altos coeficientes de frecuencia (es decir, $F(u, v)$'s para ampliar u y v) son pequeños. Dos factores determinan la cantidad de la compresión y la pérdida de información en esta cuantificación:

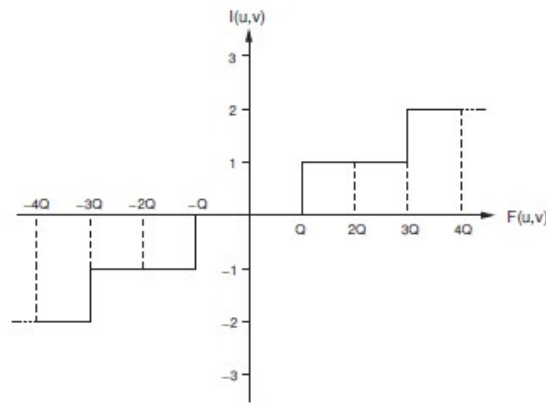
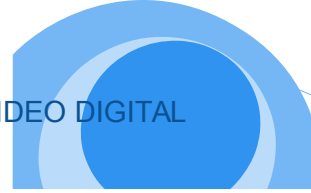


Fig. 3.5: Illustration of quantization, $T = Q$.

1. Coeficientes de $F(u, v)$ con un valor absoluto menor al umbral cuantificador T se ajusta a cero, es decir, que se consideran en la llamada "Zona muerta".
2. Coeficientes de $F(u, v)$ con un valor absoluto mayor que o igual al umbral cuantificador T se divide por dos veces el tamaño de paso cuantificador Q y redondeado al número entero más cercano.

En resumen, los coeficientes DCT cuantificados $I(u, v)$ están dadas por:



$$I(u, v) = \begin{cases} 0 & \text{for } |F(u, v)| < T \\ \left[\frac{F(u, v)}{2Q} \right] & \text{for } |F(u, v)| \geq T, \end{cases}$$

Donde $[\cdot]$ denota redondear al entero más cercano.

Un cuantificador con $T = Q$, se usa típicamente en la práctica,

Table 3.1: Default MPEG-4 quantization matrix for intra-coded video frames (from [39]).

8	17	18	19	21	23	25	27
17	18	19	21	23	25	27	28
20	21	22	23	24	26	28	30
21	22	23	24	26	28	30	32
22	23	24	26	28	30	32	35
23	24	26	28	30	32	35	38
25	26	28	30	32	35	38	41
27	28	30	32	35	38	41	45

En los codificadores de vídeo modernas, por lo general una cuantificación de tamaño optimizado para los coeficientes individuales se utiliza. Las etapas de cuantificación fijos se almacenan

como *matriz de cuantificación* y la aplicada durante el proceso de codificación. La codificación

la calidad es controlada entonces por el factor de *escala de cuantificación* q , que se multiplica con la matriz de cuantificación antes de la cuantificación se lleva a cabo. Una matriz de cuantificación de ejemplo se da en la Tabla 3.1. La compensación en general entre calidad de imagen y compresión (tamaño del fotograma en bytes después de cuantificación) se ilustra para el primer fotograma de la secuencia de prueba a *Foreman* [40], codificado con el software MPEG-4 la referencia [39] en las figuras 3.6, 3.7, y 3.8.

Observe que la calidad de la trama de vídeo disminuye visiblemente como Q aumenta. Además, la limitación de la codificación basada en bloques se hace visible, como la blockiness de la imagen aumenta. Como puede verse a partir de los tamaños de marco, la pérdida de calidad también se refleja en la cantidad de datos necesarios. La relación entre el q escala de cuantificación y el tamaño de la imagen puede ser capturado en una curva de tipo de escala de cuantificación, como se ilustra en la Figura 3.9. Tomamos nota de que la aplicación de unos muy bajos resultados del factor de escala de cuantificación en un tamaño de trama muy grande sin ningún tipo de discapacidad visual, mientras que la aplicación de la escala de cuantificación media tiene factores que pueden producir discapacidad visual, pero el tamaño del fotograma codificado es muy reducido. Esto da lugar a una decisión de compromiso entre tamaño y calidad, que lo haremos discutir más tarde.

Nuestros esquemas anteriores han sido con respecto a la escala de cuantificación controlado. Tasa de control se puede aplicar durante el proceso de codificación ajustando los tamaños de fotograma de vídeo resultantes al ancho de banda disponible. La cubanización se ajusta en un proceso de bucle cerrado (es decir, el resultado de la cuantificación se mide por su tamaño y según se requiera codificada de nuevo con un cuantificador diferente al tamaño de paso) para aplicar una compresión en dependencia de contenido de vídeo y el tamaño del marco resultante. El resultado es una velocidad de bits constante (CBR) flujo de vídeo pero con diferentes cuantificación y por lo tanto la calidad. Lo contrario de la CBR es variable tasa de bits (VBR). Aquí el proceso de cuantificación se mantiene constante, por lo que se conoce como codificación en bucle abierto (es decir, el resultado de la cuantificación y su proceso ya no está sujeta a cambios con el fin de cumplir con los requisitos de ancho de banda). Para lograr una calidad constante, la codificación VBR tiene que ser utilizado [41, 42].



Fig. 3.6: First frame from the *Foreman* test sequence, encoded with MPEG-4 reference encoder, quantization scale $q = 1$ results in 78007 bytes.



Fig. 3.7: First frame from the *Foreman* test sequence, encoded with MPEG-4 reference encoder, quantization scale $q = 15$ results in 9392 bytes.



Fig. 3.8: First frame from the *Foreman* test sequence, encoded with MPEG-4 reference encoder, quantization scale $q = 30$ results in 4461 bytes.

3.1.4 Zig-Zag Escaneo

Los valores de los coeficientes obtenidos en la cuantificación son analizados por partida con el DC-componente y luego continuar a las componentes de alta frecuencia en una forma de zig-zag, como se ilustra en la Figura 3.10. La exploración en zig-zag

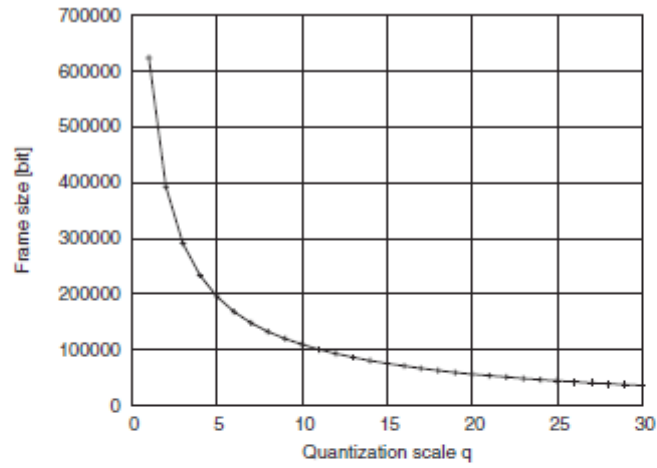
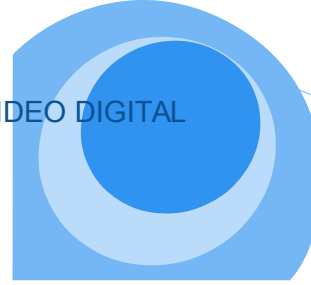


Fig. 3.9: Relationship between the quantization scale factor q and the resulting frame size in bit for the 1st frame of the *Foreman* sequence.

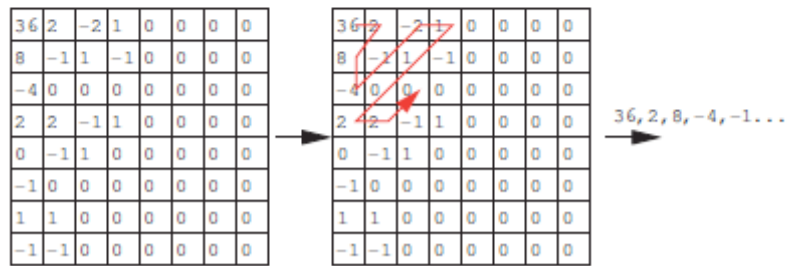


Fig. 3.10: Quantized DCT coefficients ($Q = 16$) and zig-zag scanning pattern.

facilita la codificación de longitud variable posterior al encontrar la más elementos distintos de cero probables primero. Una vez que todos los coeficientes distintos de cero se escanean, la obtenida secuencia de valores se codifica más para reducir la redundancia de palabras de código, véase la Sección 3.1.5. La exploración se puede detener antes de recoger todo coeficientes distintos de cero cuantificados a lograr mayor (con pérdida) de compresión.

3.1.5 Variable Length Coding

El propósito de *la codificación de longitud variable* (VLC) es reducir la redundancia estadística en la secuencia de palabras de código obtenidos a partir de zig-zag escanear un bloque intracodificados (o bloque de diferencias para un bloque previsto). Este es el parte del proceso de codificación de vídeo que proporciona la compresión real. La VLC utiliza un conjunto infinito de palabras de código y se aplica solamente en el mapeo de los símbolos y por lo tanto reduce la necesidad de redefinir las palabras de código [43].

La codificación se basa en una sola tabla, estática de palabras de código que resulta en un proceso de mapeo simple. Palabras de código cortas se asignan a los valores con una alta probabilidad. Palabras de código más largos se asignan a resultados menos probables de la cuantización. El mapeo entre estos valores originales y los símbolos de código que se hace dentro de la codificación de longitud variable (VLC). El mapeo tiene que ser conocido por ambos, el remitente y el receptor. Como se muestra antes, la cuantización y en zig-zag es el resultado de barrido en un gran número de ceros.

Estos valores son codificados utilizando la codificación de nivel de ejecución. Esta codificación sólo transmite el número de ceros en lugar de los ceros individuales. Además, cuando no hay otros valores son detrás de los coeficientes con ceros (y este es el caso más probable), un extremo del Bloque (EOB) de palabras de código se inserta en el flujo de bits resultante. Huffman de codificación [44] y la codificación aritmética [45, 44] y sus respectivos derivados se utilizan para implementar VLC. La codificación de Huffman es bastante simple de implementar, pero logra relaciones de compresión más bajas. Por otro

lado, la codificación aritmética de esquemas son computacionalmente más exigentes, pero lograr una mejor compresión. como la potencia de procesamiento es abundante en muchos de los sistemas actuales más nuevo, codecs su mayoría se aplican codificación aritmética [30, 46].

El *codificador aritmética binaria adaptativa al contexto* (CABAC) [46] es uno de estos técnicas de codificación y utilizada en el estándar de codificación de vídeo H.264. El CABAC enfoque utiliza distribuciones de probabilidad para reducir aún más el espacio necesario para almacenar la trama codificada. Símbolos más cortos se asignan a patrones de bits con una alta probabilidad de ocurrencia y más largos símbolos de patrones de bits con una menor probabilidad de ocurrencia. Este proceso de mapeo logra la compresión sin pérdidas por portar la secuencia de símbolos en un intervalo de números reales entre 0 y 1 con respecto a la probabilidad del símbolo en la fuente. Es Por lo tanto, que la explotación de correlación adicional de símbolos en el lado de codificación para una mayor reducción de datos que se almacena para cada marco.

3.2 Inter-frame Codificación: Estimación de movimiento y Compensación

En las secciones anteriores, hemos introducido los algoritmos empleados típicamente en codificadores de vídeo de hoy en día que se basan en la DCT transformar los marcos individuales. Este proceso, sin embargo, sólo codifica marcos individualmente. Como el vídeo se compone de una serie de fotogramas o imágenes con contenido cambia con el tiempo, los codificadores de vídeo comúnmente emplean codificación entre tramas para reducir la redundancia temporal entre cuadros sucesivos (imágenes).

La idea básica de la codificación entre tramas es que el contenido de un cuadro actual de vídeo dado es normalmente similar a la de una fotograma de vídeo pasado o futuro. El pasado (o futuro) trama se utiliza como una *referencia marco* para predecir el contenido de la trama actual. Esta predicción es típicamente realizada en un macrobloque o bloque de base [47].

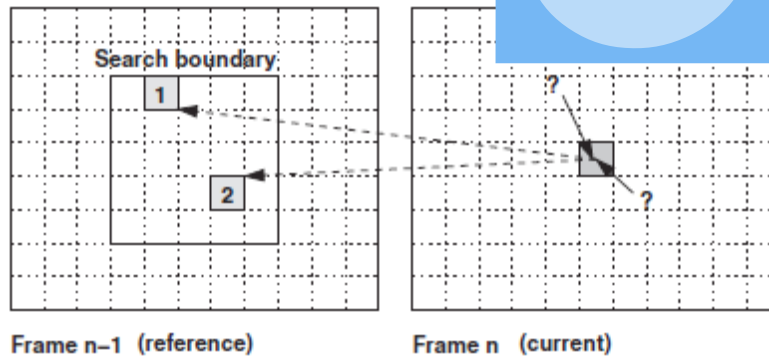


Fig. 3.11: Determination of the motion vector for a macroblock on frame n .

Para la codificación entre tramas MPEG introdujo los tipos de trama codificados-Intra (I), Inter-codificado (P), y Bi-direccional codificada (B). En una trama I todos los bloques son intra-codificados como se indicó anteriormente. Los macrobloques (cuatro bloques de 8×8 muestras por macrobloque) en tramas P son inter-codificados con respecto a la I precedente o P marco, mientras que una trama B se intercoded con respecto a la I anterior o Trama P y el I o P siguiente fotograma. To-código entre un macrobloque dado, el mejor macrobloque correspondiente en el cuadro (s) de referencia es determinado y identificado por un vector de movimiento (referido como estimación de movimiento). Cualquier (típicamente pequeña) diferencia entre el bloque a codificar y el mejor juego el bloque se transformó utilizando la DCT, cuantificado, y se codifica como se describe anteriormente (Denominado compensación de movimiento); si un buen partido no se puede encontrar la macrobloque está codificado intra.

(En el modo opcional 4MV los algoritmos anteriores se aplican a los bloques en lugar de macrobloques.) Para un bloque real dada en el marco actual de un *algoritmo de coincidencia de bloques* (BMA) busca el bloque de predicción más similar en el marco de referencia, como se ilustra en la Figura 3.11. El objetivo de la búsqueda es para determinar el *movimiento vector*, es decir, el vector de desplazamiento de la más similar macro) en el bloque, marco de referencia para la (macro) bloque bajo consideración en el marco actual. Esta búsqueda, que también se conoce como la estimación de movimiento, se lleva a cabo

durante una gama específica alrededor de la ubicación del bloque en la trama actual, como se ilustra en la búsqueda límite. Los algoritmos de búsqueda no se limitan a individuo (Macro) bloques, sino que se expanden en un pixel o base incluso sub-pixel. Varios o diferentes características determinadas (similitud) criterios tales como la función de correlación cruzada, error medio cuadrado, o error medio absoluto se puede aplicar. Como se ilustra en Figura 3.11, varios candidatos a la (macro) bloque bajo consideración pueden existir en el marco de referencia. El candidato con la menor diferencia de la (macro) de bloque en cuestión está seleccionado y el movimiento de acuerdo al vector se utiliza para determinar el desplazamiento. En caso de que la comparación sea:



Fig. 3.12: Illustration of the unrestricted motion estimation mode.

realizado sobre una base de bloque, el proceso produce cuatro vectores de movimiento para cada acrobloque. La mejora de los vectores de movimiento normales es la revocación de la imagen fronteras como límites para la validez de la meta de un vector, también conocidos como sin restricciones o el modo de vector de movimiento extendida. Puesto que no hay contenido y por lo tanto los datos disponibles para el exterior de una imagen, los píxeles en la frontera son simplemente replicada para llenar los valores inexistentes necesarios como referencias. Figura 3.12 ilustra este algoritmo. Para encontrar el mejor partido de búsqueda completo, $(2n + 1)^2$ se requieren comparaciones. Varios esquemas de estimación de movimiento rápido como la *búsqueda de tres pasos* [48] o el *algoritmo de coincidencia bloque jerárquico* [49] han evolucionado para reducir el procesamiento. Una vez que se

determina el vector de movimiento, la diferencia entre el bloque de predicción y el bloque actual se codifica utilizando las técnicas descritas en la sección anterior de codificación intra-cuadro. Estas diferencias pueden deberse a condiciones de iluminación, ángulos, y otros factores que cambiar ligeramente el contenido del bloque (macro). Estas diferencias son típicamente pequeñas y permitir la codificación eficiente con la codificación intra-cuadro (técnicas y técnicas de codificación de longitud variable). El tamaño del paso del cuantificador se puede ajustar de forma independiente para la codificación de estas diferencias. La codificación de estas diferencias cuenta para las diferencias entre el bloque de predicción y el bloque actual, y se conoce como *compensación de movimiento*. La inter-codificados marco está representada por (i) los vectores de movimiento (de estimación de movimiento), y (ii) el error codificado o la diferencia entre la trama actual con el determinado vectores de movimiento y el marco de referencia (compensación de movimiento).

Para el caso de que la estimación de movimiento no dio ninguna coincidencia dentro el límite de búsqueda o si codifica el vector de movimiento y restante diferencia daría lugar a un tamaño más grande que la aplicación de intra-codificación solamente a la (Macro) de bloque, el bloque (macro) se codifica utilizando sólo técnicas intra-codificación. Normas de codificación de vídeo más recientes, como el estándar de codificación de vídeo H.264, permitir múltiples marcos de referencia para un solo fotograma en consideración. Para estimación de movimiento y compensación, el marco en cuestión



Fig. 3.13: Frames 90, 91 and 92 from the *News* test sequence illustrating changing video frame content.

Marcos de referencia tienen que estar disponibles para el codificador y el decodificador. Esto resulta en grandes requisitos de memoria a medida que aumenta la resolución del marco. Como hardware componentes para el codificador y el decodificador, tales como la velocidad de cálculo y memoria, cada vez menos de una restricción (en disponibilidad y precios), la disponibilidad de múltiples marcos de referencia aumenta la probabilidad de encontrar buenos resultados para la (macro) bloques en consideración y por lo tanto aumentar la eficiencia de compresión. Los macrobloques y / o bloques en los fotogramas de vídeo a menudo revelan partes del fondo o escena que no eran visibles antes de la trama real [37]. Movimiento de vectores de estas áreas no pueden, por tanto, se encuentran haciendo referencia del anterior marco, pero sólo considerando también tramas futuras. Nos ilustran esta idea en la Figura 3.13 para los marcos 90, 91 y 92 obtenidos a partir de la secuencia de vídeo *News*.

Los cambios en el contenido de fondo de la estructura 90 para enmarcar 91. Su única referencia marcos del pasado se les permitió, el contenido de fondo cambiada en el marco 91 tendría que ser codificada en el modo intra para este marco. Si la trama siguiente 92 es intra-codificados de forma predeterminada, luego los (macro) bloques intra-codificado adicionales en el marco 91 que reduciría la eficacia de la compresión. Normas de codificación de vídeo han incorporado esta idea e inter-frame de codificación a menudo considera tanto la predicción de los marcos de referencia del pasado, así como el futuro marcos de referencia. Existen tres métodos básicos para la codificación de las imágenes originales en el dominio temporal: **I**ntra codificado marcos, **P**redicted marcos y **B**idireccionalmente predijo marcos, introducido en el estándar MPEG-1 [50].

Estos métodos de codificación se aplican en el marco, macrobloque o nivel de bloque, dependiendo del codec. Un marco intracodificados consiste exclusivamente en intracodificado macrobloques. Por lo tanto, un marco de intra-codificado contiene la imagen comprimida y su información (sin ninguna información de predicción), lo que resulta en un gran tamaño de trama (en comparación con el tamaño de los marcos internacionales o bidireccionales codificados).

Las tramas interconectadas codificado utilizan técnicas de estimación y compensación de movimiento basándose en el marco de inter o intra-codificado anterior. El bi-direccional codificada marcos se basan en una anterior, así como un siguiente intra o inter-coded marco. Esta información de predicción resulta en tamaños de bastidor más pequeños para el P- marcos y tamaños más pequeños de marco para los B-frames. Cuando los marcos B no lo hacen tiene alguna siguientes I- o P-frames que se pueden utilizar como marcos de referencia,

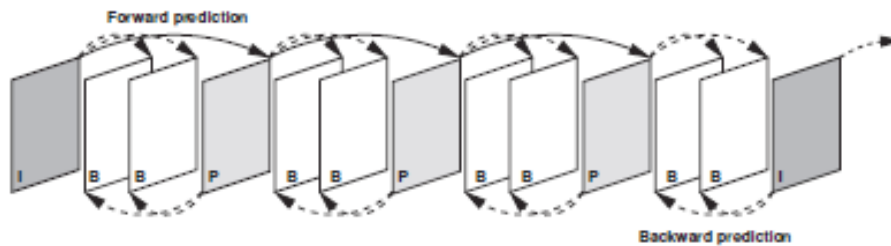


Fig. 3.14: Typical MPEG Group of Pictures (GoP) consisting of I, P and B frames (frames 1-12).

Codificación o decodificación es posible. Marcos o bloques intracodificados no confían en otros cuadros de vídeo y por lo tanto son importantes para detener la propagación de errores. La secuencia de tramas a partir de un marco de intra-codificado, pero no incluyendo la siguiente trama intra-codificados se conoce como un *grupo de imágenes* (GOP).

La relación entre estos diferentes tipos de codificación y cómo marcos confían el uno del otro en una secuencia típica trama MPEG que consta de 12 tramas se ilustra en la Figura 3.14. No es necesario disponer de más de una trama I en el comienzo de la secuencia de vídeo, en cuyo caso la totalidad de la secuencia de vídeo es un solo GP. Otro extremo es no

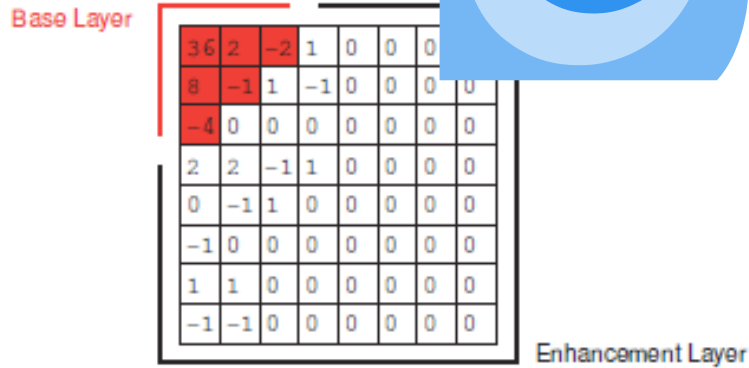
tener cuadros P o B, en cuyo caso la longitud GoP es 1 y cada fotograma de vídeo se codifica de forma independiente, similar a una sola imagen.

3.3 Codificación de vídeo escalable

Con la codificación escalable el codificador produce típicamente múltiples capas. La capa de base proporciona una calidad básica (por ejemplo, la resolución espacial o temporal bajo añadiendo capas de mejora de vídeo) y mejora la calidad de vídeo (por ejemplo, los aumentos y la resolución espacial o velocidad de cuadro). Una variedad de técnicas de codificación escalable se han desarrollado, lo que vamos a presentar en esta sección.

Escalable la codificación es una manera conveniente para adaptarse a la amplia variedad de capacidad de video hardware (por ejemplo, PDAs, teléfonos móviles, ordenadores portátiles, ordenadores de sobremesa) y redes de entrega (Por ejemplo, o inalámbrica vs) [51, 52]. Cada uno de estos dispositivos tiene diferentes limitaciones debidas a la potencia de procesamiento, tamaños de visualización, y así sucesivamente. Scalable codificación puede satisfacer estas diferentes restricciones con una codificación del vídeo. Observamos brevemente que una alternativa a la codificación escalable es para codificar el vídeo en diferentes versiones, cada una con un nivel de calidad diferente, velocidad de bits, espacial / resolución temporal.

La ventaja de tener diferentes versiones es que no requiere los codificadores escalables más sofisticados y no incurre la sobrecarga adicional debido a la codificación escalable. El inconveniente es que las múltiples versiones ocupan más espacio en los servidores y, posiblemente, tienen que ser streaming todos juntos (simulcast) en la red para poder elegir la versión apropiada en cualquier momento dado.



La

Fig. 3.15: Data partitioning by priority break point setting.

transcodificación es otra alternativa a la codificación escalable. La transcodificación puede ser utilizado para adaptarse a las diferentes condiciones de la red, como en [53], o para adaptar a diferentes formatos de vídeo deseado [54]. El enfoque de transcodificación requiere típicamente un alto rendimiento nodo intermedio. Después de haber ofrecido una visión general de las cuestiones de carácter general alrededor de codificación de vídeo escalable, que ahora presentamos los diferentes enfoques de codificación escalable, que remitimos al lector interesado a [55] para más detalles.

3.3.1 particionamiento de datos

Aunque no explícitamente una técnica de codificación escalable, divide particionamiento de datos el flujo de bits de estándares de vídeo no escalables tales como MPEG-2 [56] en dos partes. La capa base contiene datos críticos tales como vectores de movimiento y bajo el ordenamiento de los coeficientes DCT, mientras que la capa de mejora contiene, por ejemplo los más altos coeficientes DCT orden [32]. El *punto de ruptura de prioridad* determina dónde parar en el proceso de cuantificación y escaneo para los coeficientes de la capa base para ser codificado adicional [30] como se muestra en la Figura 3.15. Los coeficientes restantes luego son codificados por reanudar el patrón de exploración en zig-zag en el punto de ruptura y se almacenan en la capa de mejora.

3.3.2 Escalabilidad Temporal

Escalabilidad temporal reduce el número de fotogramas en la capa base. Los marcos extraídos se codifican en la capa de mejora y hacen referencia a la trama de la capa base. Diferentes patrones de combinación de tramas en base de y existe capa de mejora [57]. En la figura 3.16 una capa de refuerzo que consiste de todos los fotogramas B se da como ya se ha utilizado en evaluaciones de trazas de vídeo [58]. No hay otros marcos dependen de la decodificación exitosa de fotogramas B. Si la mejora de la capa no es descifrable, la decodificación de los otros tipos de trama es

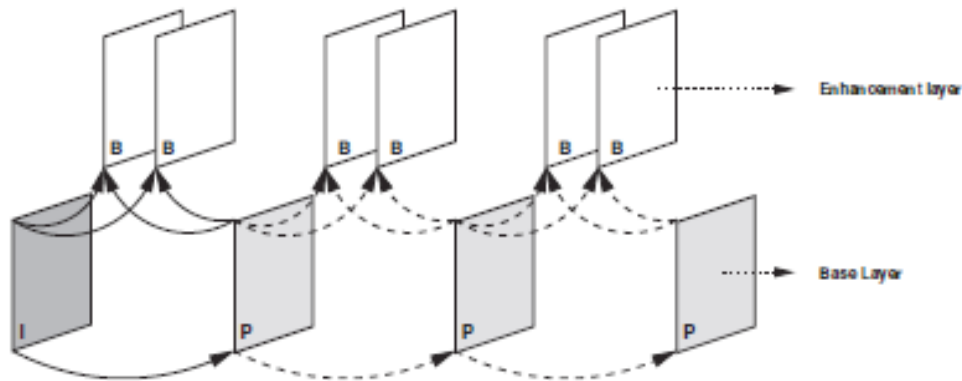


Fig. 3.16: Temporal scalability with all B-frames in enhancement layer.

no se ve afectado. Sin embargo, dado que el número de fotogramas que se reconstruyen, la tasa de imágenes por segundo debe ajustarse en consecuencia para la visión y los métodos de evaluación de la calidad (por ejemplo, el último cuadro decodificado con éxito se muestra durante un período más largo, también llamada *congelación*). Este ajuste está infligiendo la pérdida en la calidad de vídeo visible.

3.3.3 Escalabilidad espacial

Escalabilidad en el dominio espacial es la aplicación de diferentes resoluciones a la base y la capa de mejora. Si, por ejemplo, la secuencia original está en el Formato CIF (352 x 288), la capa de base se muestrea en el formato QCIF (176 x 144) antes de la codificación. Escalabilidad espacial está, por tanto, también conocida como *la codificación pirámide*.

Además de la aplicación de diferentes resoluciones, diferente Estructuras de GOP es usado en las dos capas. El patrón GoP en la capa de mejora se está refiriendo a los fotogramas de la capa base. Un ejemplar disposición de las dependencias resultantes se ilustra en la Figura 3.17. El contenido de la capa de mejora es la diferencia entre las capas, así como la referencia basada en marcos de cuadros anteriores y siguientes de la misma capa. Los

estudio de las características del tráfico y la calidad de escalabilidad temporal y espacial vídeo codificado se da en [58].

3.3.4 SNR Escalabilidad

Escalabilidad SNR proporciona dos (o más) capas de vídeo diferentes de la misma resolución, pero con diferentes calidades. La capa de base se codifica por sí mismo y proporciona una calidad básica en términos de (P) SNR. La capa de mejora es codificada para proporcionar una calidad adicional cuando se vuelve a añadir a la capa base. La la codificación se realiza en dos pasos consecutivos: primero la capa base es codificada con una calidad baja, entonces la diferencia entre la capa base y decodificada como se ilustra en la Figura 3.18. la entrada de vídeo está codificado con ajustes más altos de calidad en una segunda etapa [55],

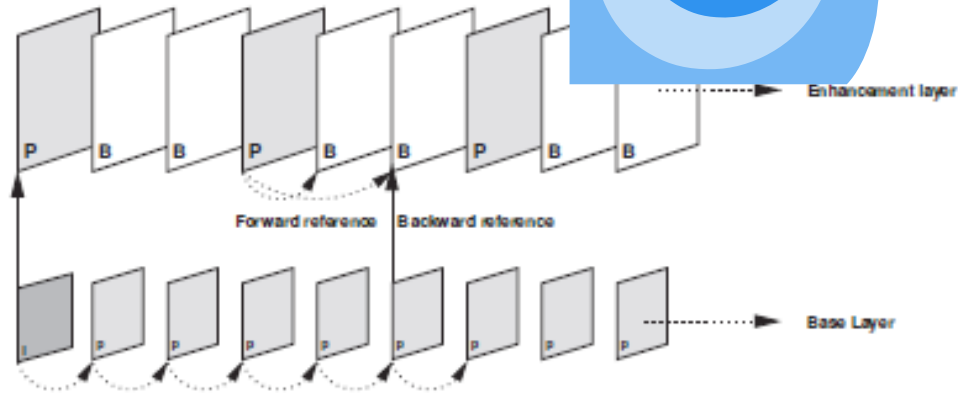


Fig. 3.17: Example for spatial scalability and cross-layer references.

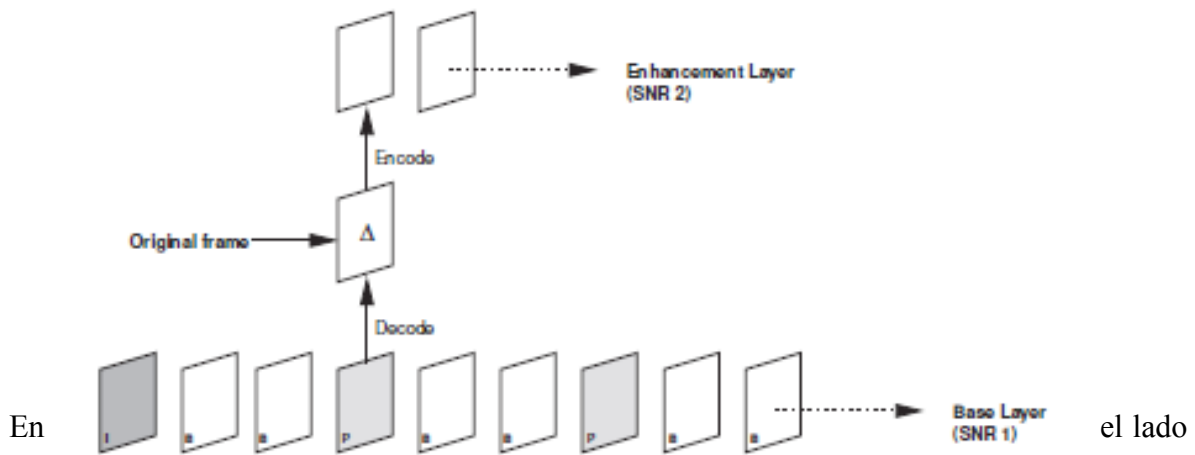


Fig. 3.18: Example for SNR scalability.

receptor de la calidad de base se obtiene simplemente mediante la decodificación de la capa de base. Para obtener una calidad mejorada, la capa mejorada se decodifica y el resultado se añade a la capa de base. No hay necesidad explícita de ambas capas sean codificada por el mismo estándar de compresión de vídeo, aunque para facilidad de uso es aconsejable hacerlo.

Escalabilidad 3.3.5 Objeto

Otra característica de escalabilidad es posible dentro de los estándares de vídeo que apoyan la composición de cuadros de video de varios objetos diferentes, tales como MPEG-4 [57]. La capa de base contiene sólo la información que no pudo ser equipado o identificados

como objetos de vídeo. La capa (s) de mejora se compone de la información respectiva para los objetos de vídeo, tales como la forma y la textura. La



Fig. 3.19: Example for object-based scalability.

example shown in Figure 3.19 presents a case where the background (landscape) and an object (car) were separated. In this case, the background is encoded independently from the object.

3.3.6 Fine Granular Escalabilidad (FGS)

Grano Fino Escalabilidad (FGS) es una forma relativamente nueva de codificación de vídeo escalable [59] que, recientemente, se ha agregado a la norma de codificación de vídeo MPEG-4 [60] con el fin de aumentar la flexibilidad de la transmisión de vídeo. Con FGS, el vídeo se codifica en una capa de base (BL) y una capa de mejora (EL). Similar a la codificación de vídeo escalable convencional, la capa base debe ser recibida hasta que se complete 1 y con el fin de decodificar y mostrar un video básico de calidad. La capa a mejorar tiene la propiedad especial que se puede cortar en cualquier caso y el bit con su

parte recibida de la corriente de la capa de mejora FGS puede ser decodificado con éxito y mejora de la calidad de vídeo básico [59, 61].

FGS por lo tanto elimina la restricción de la convencional de codificación en capas, donde una capa de mejora debe ser completamente recibida para la decodificación con éxito. Similar a la convencional codificación escalable, la capa de mejora FGS es jerárquica en la que "más alto" los bits requieren los bits "inferiores" para decodificación exitosa. Esto significa que cuando hay que cortar el flujo de bits de capa de mejora antes de la transmisión, la parte inferior de la corriente de bits (por debajo del corte) necesita ser transmitida y la mayor parte (por encima del corte) se puede quitar. La capa de mejora FGS puede ser cortar en la granularidad de bits, como se ilustra en la Figura 3.20. La flexibilidad de FGS lo hace atractivo para el streaming de vídeo, como servidores de vídeo pueden adaptar el vídeo transmitido al ancho de banda disponible en tiempo real (sin requerir cualquier computacionalmente exigentes recodificación).

Pero esta flexibilidad tiene expensas de la eficiencia de codificación reducida. Después de la normalización, el refinamiento y la evaluación de la codificación de vídeo FGS ha recibido considerable interés [62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74]. No hay movimiento

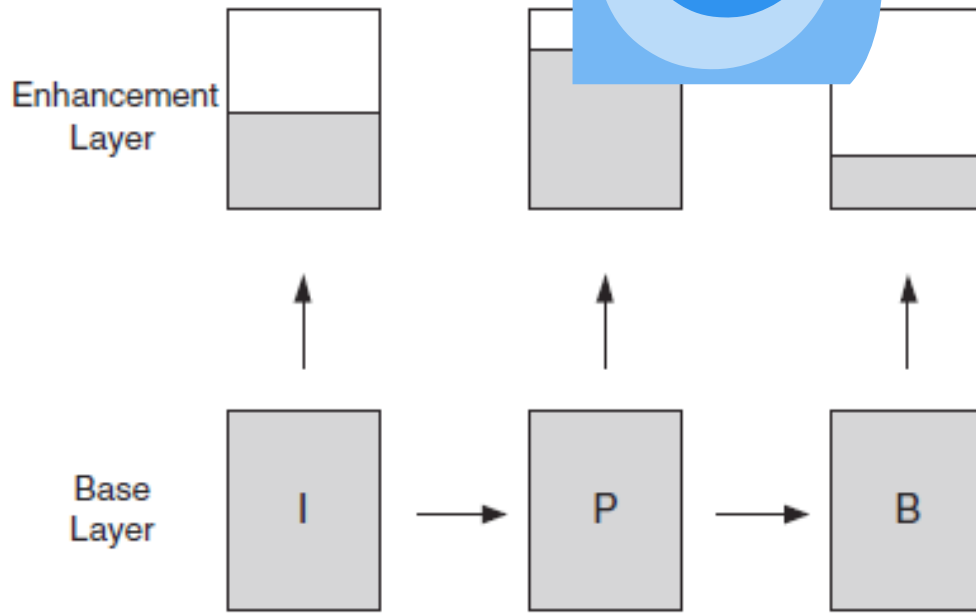


Fig. 3.20: Example of truncating the FGS enhancement layer before transmission.

Compensación dentro de la capa de mejora FGS. Esto hace que la mejora de la capa sea altamente resistente a los errores de transmisión y, posteriormente, bien adaptado a la transmisión a través de redes propensas a errores tales como la Internet de mejor esfuerzo.

Un escenario típico para la transmisión de MPEG-4 FGS codificado videos más Internet ha sido propuesto por el comité de MPEG-4 en [75]. En este escenario de la capa de base se transmite con una alta fiabilidad (consigue a través la asignación de recursos adecuados y / o error de canal de corrección) y el FGS capa de mejora se transmite con baja confiabilidad (es decir, en un mejor esfuerzo forma y sin control de errores). Cerramos esta breve reseña de MPEG-4 FGS codificación señalando que el estándar MPEG-4 incluye varias mejoras en el SNR FGS básica enfoque esbozado anteriormente y también un modo de FGS escalable temporal, que son más allá del alcance de nuestro estudio. (Un mecanismo de transmisión en la adaptación de la de vídeo añadiendo y dejar caer el SNR FGS y capas de

mejora FGS temporales se estudia en [76].) También observamos que un FGS Progresista (PFGS) refinamiento recientemente se ha propuesto [77, 73], pero todavía no estandarizado. En contraste con MPEG-4 FGS, PFGS permite la compensación de movimiento parcial entre el FGS planos de bits, logrando al mismo tiempo la propiedad granularidad fina. Este movimiento compensación típicamente mejora la eficiencia de codificación, pero disminuye el error y capacidad de resistencia de la capa de mejora [78].

3.3.7 Múltiple Descripción Codificación (MDC)

Con múltiples Descripción de codificación (MDC) [79], el vídeo se codifica en varias sub-corrientes que se refiere como *descripciones*. Cada una de las descripciones son transmitidas hacia el receptor. Decodificación más descripciones da un vídeo de mayor calidad al decodificar un subconjunto arbitrario de los resultados en las descripciones en la parte baja de calidad. Las descripciones individuales no tienen jerarquía explícita o dependencia entre ellos, es decir, cualquier combinación de las descripciones se puede combinar y decodificado [80]. Esto es en contraste a los vídeos en capas jerárquicas convencionales donde una capa de mejora recibida es inútil si la capa de base correspondiente no se encuentra como para FGS.

Las ventajas del MDC han sido estudiados para redes multi-hop [81, 82], División de frecuencia ortogonal (OFDM) [83], de entrada múltiple Sistemas (MIMO) de salida múltiple [84], las redes ad-hoc [85], Universal Mobile Telecommunications System (UMTS) [86], Transport Control Protocol (TCP) [87] y Content Delivery Networks (CDN) [88]. MDC es especialmente interesante para el apoyo de los terminales heterogéneos en redes 4G como abogado en [89]. Futuro de cuarta generación (4G) sistemas móviles se prevén para ofrecer servicios inalámbricos a una amplia variedad de terminales móviles que van desde celular teléfonos y asistentes digitales personales (PDA) a los ordenadores portátiles [90]. Estas amplias variedades de terminales móviles se denominan terminales como heterogéneos. Heterogéneo terminales tienen diferente potencia de procesamiento,

memoria, espacio de almacenamiento, capacidades de la vida de la batería y la velocidad de datos. A diferencia de DVB-T y DVB-H, donde el mismo espectro está reservado para el apoyo de cada tecnología en un tiempo de moda multiplex, terminales heterogéneos en 4G debe utilizar el mismo espectro en caso de que los usuarios están interesados en los mismos servicios, para usar el espectro de manera eficiente. En un escenario de multidifusión, terminales de clase alta recibirían un gran número de corrientes, mientras que los terminales de clase baja irían para un número menor.

Terminal de alta clase. La flexibilidad del ancho de banda asignado a cada descriptor y el número de descriptores asignado a los usuarios finales hace MDC un esquema de codificación atractiva para redes 4G. La ventaja de múltiples Descripción de codificación se consigue normalmente en las expensas de la reducción de las ganancias de compresión de vídeo y caracterizaciones de tráfico de vídeo existentes tales como la codificación única y múltiple capa, como se presenta en los capítulos anteriores, no se pueden utilizar tal como se les subestima el ancho de banda requerido.

3.4 Codificación de vídeo Basado en wavelets

Con transformada wavelet de codificación [91] un fotograma de vídeo no se divide en bloques Como con la codificación MPEG basado en DCT. En lugar de ello, toda la trama se codifica en varios sub-bandas que utilizan la transformada wavelet. La Transformada Wavelet tiene muchas ventajas sobre la transformación DCT. El más obvio de todos ellos,

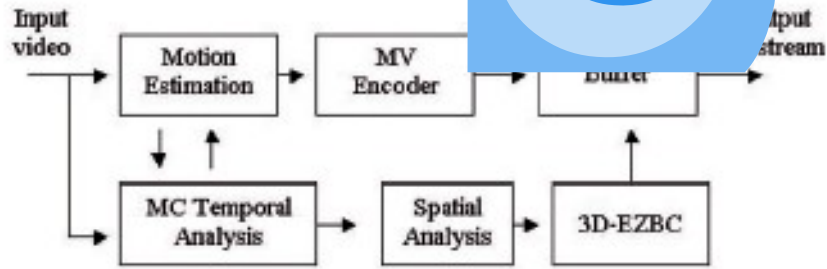


Fig. 3.21: Block Diagram of the MC-3DEZBC wavelet encoder [93].

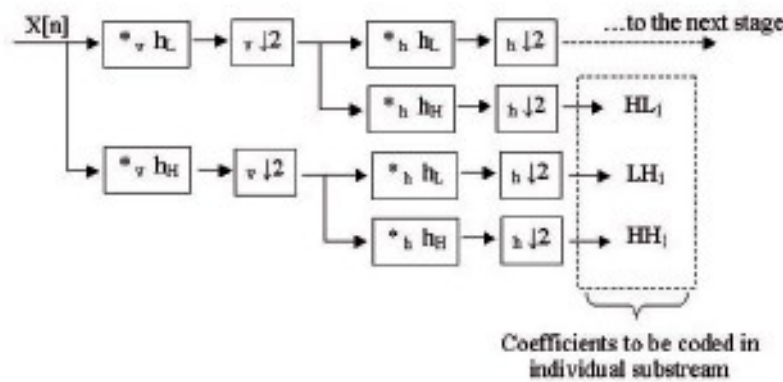


Fig. 3.22: First wavelet decomposition stage.

es la característica compacta de apoyo. Este compacto-soporte permite traducir una función de dominio de tiempo en una representación que no sólo se localiza en frecuencia, pero con el tiempo también. El resultado neto de esto es que la transformada wavelet puede ocurrir en toda la imagen en un plazo razonable bit computacional y presupuesto. Por lo tanto, la ventaja visual obvia es que *los artefactos de bloque*, común se Transforma basado en la DCT, se eliminan en la transformada wavelet. La wavelet transformar codec que presentamos aquí es el MC-3DEZBC [92].

El diagrama de bloques del codec MC-3DEZBC en la Figura 3.21 ilustra el codec completa, incluyendo la descomposición temporal y el movimiento de estimación. Cada

fotograma de vídeo se somete a una descomposición espacial de cuatro etapas, que se lleva a cabo de forma recursiva en la subbanda de frecuencia baja. La primera etapa de de una estructura de banco de filtros utilizado para la descomposición espacial se ilustra en la Figura 3.22. Aquí X_n es la imagen de *entrada*. $*V*$ y h representan convolución en la dirección vertical y de convolución en la dirección horizontal, respectivamente. La respuesta al impulso del filtro de paso bajo y de paso alto Filer están representados

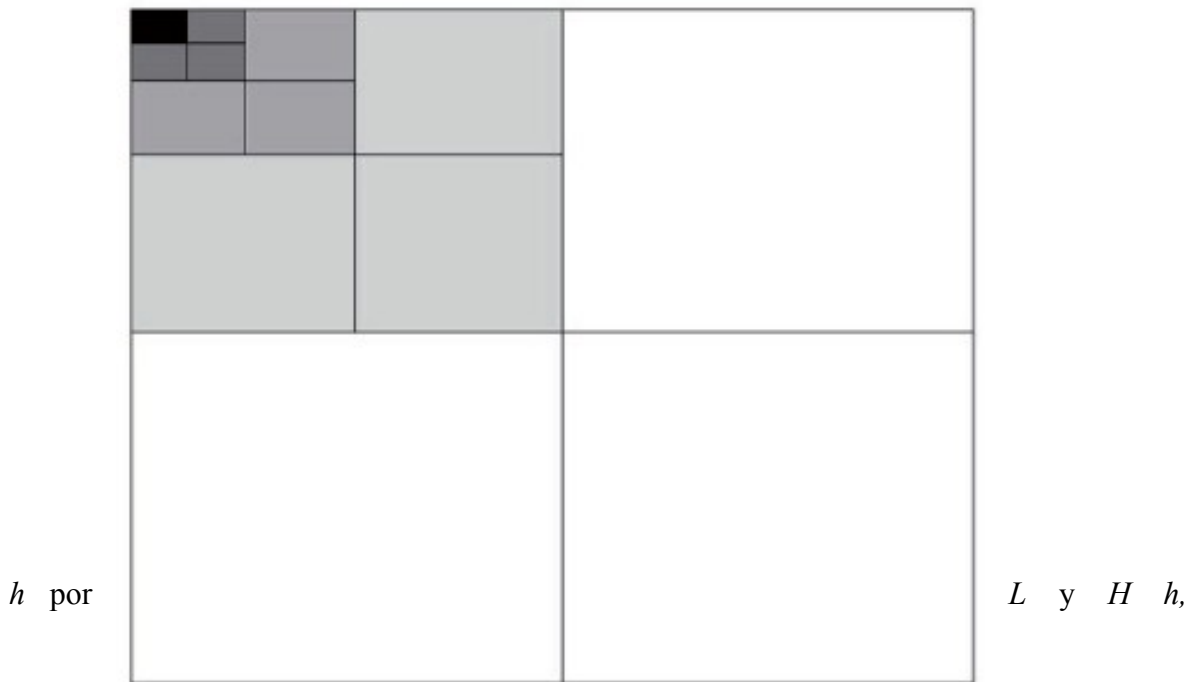


Fig. 3.23: Passband structure for MC-3DEZBC [93].

respectivamente. Una flecha apuntando hacia abajo y seguido el número 2, representa submuestreo por dos en la horizontal o vertical dirección (representado por el subíndice anterior a la flecha). HL 1, LH 1, y HH 1 representan las salidas de los filtros de la primera etapa de descomposición. Cada etapa crea tres sub-bandas, mientras que la cuarta (que es la frecuencia más baja subbanda tanto en la horizontal y las dimensiones verticales) se alimenta en el siguiente etapa de la descomposición espacial. El pro descomposición de cuatro etapas Vides 13 subbandas, como se ilustra en la Figura 3.23. Estos 13 sub-bandas obtenidos desde entonces se codificaron las cuatro etapas de descomposición de forma

individual utilizando la 3- versión D del bloque zerotree incrustado algoritmo de *3D-EZBC* [92]. Esta es una extensión de la codificación de bloque zerotree incorporado (EZBC) algoritmo desarrollado en [94]. Los flujos de bits resultantes de bit avión codificados y combinan para formar un sub-corriente como se ilustra en la Figura 3.24.

Para facilitar la ilustración, cada sub-flujo en la figura 3.24 está codificado por colores de forma que coincida con el color correspondiente en la Figura 3.23. Todas las sub-corrientes de cada cuadro y todas las tramas en el GOP correspondiente se combinan para crear un jerárquico código de la corriente [93]. Cada GP se codifica como un mensaje separado con dependiente del contexto de codificación aritmética. Cada mensaje está integrado, por lo tanto el Flujo de bits se puede truncar en cualquier punto de un presupuesto poco dado. Tasa de control

Compressed substream 1	
Compressed substream 2	
Compressed substream 3	
Compressed substream 4	
Compressed substream 5	

Fig. 3.24: Individually coded sub-bitstreams corresponding to Figure 3.23 [93].

se implementa en cada GP con el presupuesto de bits dado por $R = g N g \cdot r / F$ (bits), donde $N G$ denota el número de tramas en un GOP, R la velocidad de bits dada en bits / seg, y F denota la velocidad de fotogramas de la secuencia de imágenes en cuadros / seg.

3.5 Normas Video Coding

La compresión de vídeo está experimentando cambios constantes nueva codificación / decodificación Sistemas (codec) se están desarrollando y se introducen en el mercado. Sin

embargo, los esquemas de compresión de vídeo normalizados internacionalmente, como el H.26x y MPEG-n las normas, se basan en un conjunto común de fundamental principios de codificación que hemos revisado en los apartados anteriores. Los tamaños de las imágenes en los formatos de vídeo actuales se ilustran en Figura 3.25. Tenga en cuenta que el formato de 601 / CCIR UIT-R (es decir, el televisor común formato de imagen) y el CIF y QCIF tienen la misma proporción entre anchura y altura. En contraste, la Televisión de Alta Definición (HDTV) Formato de imagen tiene un mayor relación de anchura a altura, es decir, se percibe como "amplia". Cada imagen individual se compone de elementos de imagen (por lo general se hace referencia como pixels o pels).

El ancho específico y la altura (en píxeles) en los diferentes formatos se resumen en la tabla 3.2. Hoy en día, los formatos típicos de video inalámbrico son QCIF (176 × 144 pixel) y CIF (352 × 288 píxeles). A pesar de una gran variedad de sistemas de codificación y decodificación de vídeo (por ejemplo, la propietaria real-Media códec etc.), la normalización a nivel internacional se lleva a cabo por dos cuerpos principales: UIT-T e ISO / MPEG. La temprana H.261 codec de la UIT-T se centra en la entrega de vídeo a través de RDSI-redes con una tasa de bits fija de $n \times 64$ kbit / s, donde n indica el número de multiplexado de Líneas RDSI. Desde este punto de partida, codecs fueron desarrollados para diferentes propósitos, tales como el almacenamiento de medios digitales o entrega a través

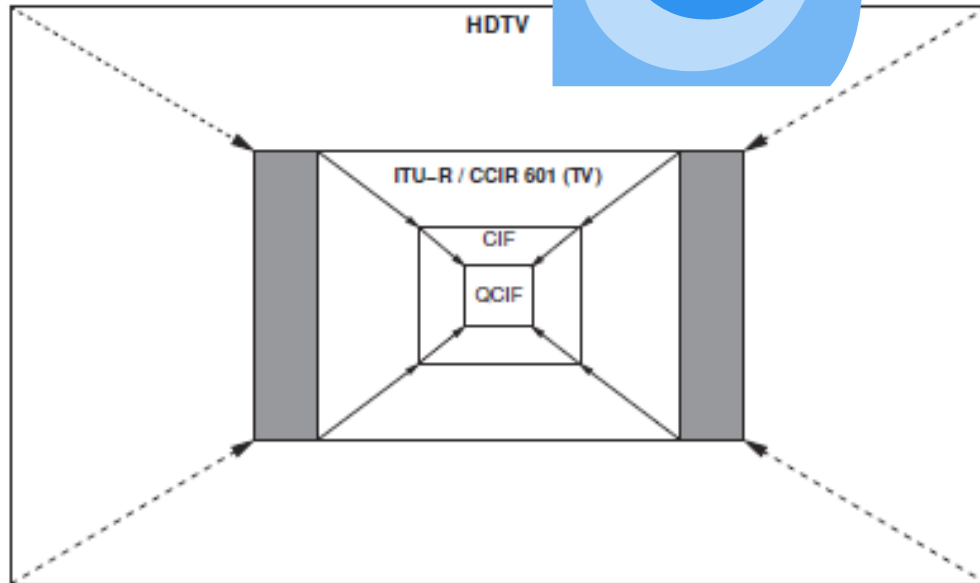


Fig. 3.25: Illustration of image formats.

Table 3.2: Characteristics for different video formats.

Standard	QCIF		CIF		TV		HDTV	
	ITU-T H.261		ITU-T H.261		ITU-R/CCIR-601		ITU-R 709-3	
Format	PAL	NTSC	PAL	NTSC	PAL	NTSC	PAL	NTSC
	[25 Hz]	[30 Hz]	[25 Hz]	[30 Hz]	[25 Hz]	[30 Hz]	[25 Hz]	[30 Hz]
Sub-sampling	4:2:0		4:2:0		4:2:2		4:2:2	
Columns (Y)	176		352		720		1920	
Rows (Y)	144		288		576	480	1080	
Columns (U,V)	88		176		360	360	960	
Rows (U,V)	72		144		576	480	1080	
Frame size [byte]	38016		152064		1244160	1036800	4147200	
Data Rate [Mbit/s]	7.6	9.1	30.4	36.5	248.8	298.6	829.4	995.3

redes orientadas a paquetes. La última H.264 codec de desarrollo (o MPEG-4 Anexo 10, H.264 / AVC) recientemente ha sido ultimado por un Equipo Conjunto de vídeo (JVT) de la UIT-T y los organismos de normalización ISO / MPEG. Las normas en evolución lograrse mejor calidad con velocidades de bits más bajos y por lo tanto un mejor rendimiento tasa-distorsión como tiempo de la Figura 3.26 esboza una visión general del desarrollo de estándares de vídeo hasta la fecha. El estándar de codificación de vídeo H.264 se diferencia de sus predecesores (el UIT-T Familia estándar de vídeo H.26x y estándares MPEG MPEG-2 y MPEG-4) en la prestación de una capa de codificación de vídeo de alta compresión (VCL) para la optimización de almacenamiento así como una capa de adaptación de red (NAL) para la paquetización de la corriente de bits codificada de acuerdo

Con los requisitos de transmisión [95]. Una visión general de estas capas se da en la figura

3.27. La capa de adaptación de red progresada

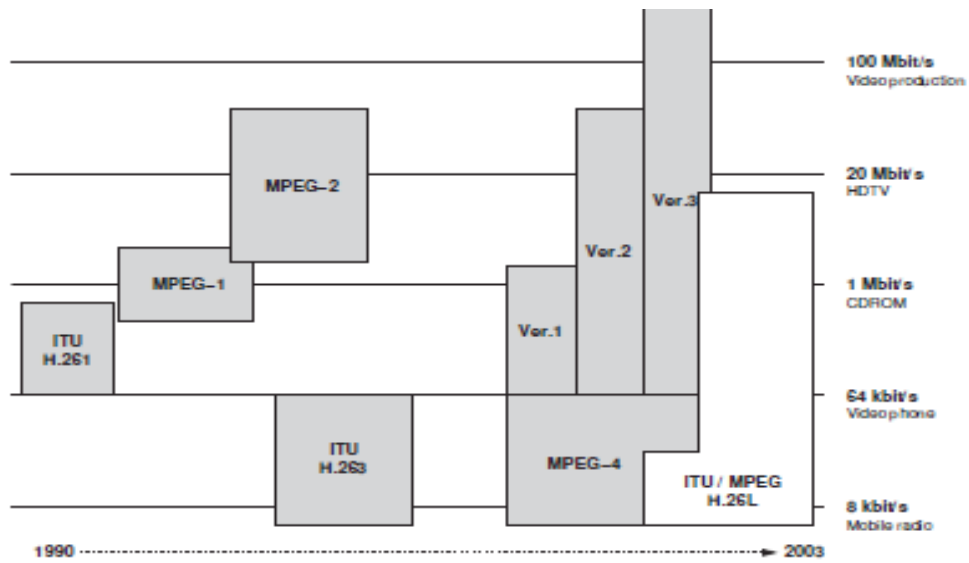


Fig. 3.26: Video coding standards of ITU-T, ISO/MPEG and the Joint Video Team (JVT).

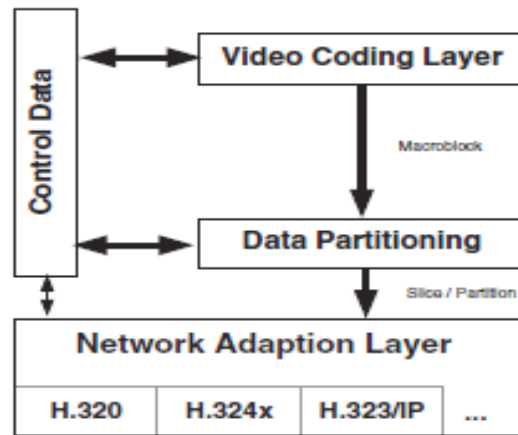
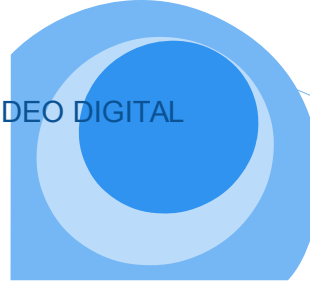


Fig. 3.27: Block diagram of an H.26L coder.

Varía de acuerdo con el tipo de red subyacente (por ejemplo, 802.3, 802.11x, UMTS, y otros). Para manejar los cambios bruscos en el flujo de bits y la pérdida de partes de

imágenes o estructuras, el estándar H.264 ofrece la posibilidad de refrescar los cuadros en un nivel de macrobloque. Además, actualice marcos (intra-imagen actualización) se utiliza para detener el proceso de predicción de tramas que hacen referencia de pérdida o marcos errorneous. Además, la norma establece la posibilidad para cambiar entre varios flujos codificados de manera diferente para evitar la alta computacional esfuerzo (y por lo tanto de alto consumo de energía) para la codificación y decodificación típicamente asociada con la transcodificación. La funcionalidad de conmutación corriente permite la codificación en tiempo no real y en tiempo real, la selección basada en el ancho de banda de las corrientes codificado con diferente de cuantificación y / o la configuración del Partido Republicano. La estimación de movimiento se realiza por múltiples marcos de referencia (ver H.263 estándar ++, anexo U - *predicción memoria a largo plazo*) y funciona más allá de las fronteras de imagen dada por restricciones vectores de movimiento.

El estándar de codificación de vídeo H.264 incluye varias características adicionales y novedades, para los que nos referimos al lector interesado a [36].



Capítulo 4

Métricas y Estadísticas para Trazas de vídeo

En este capítulo, se revisan las definiciones estadísticas y métodos utilizados en los análisis de las huellas de tamaño de vídeo generados. Remitimos al lector interesado a [96, 97] para más detalles sobre las propiedades estadísticas. Y_n el número de fotogramas de vídeo en una traza determinada.

Deje $t_n, n = 0, \dots, N - 1$, denota el periodo de trama (tiempo de visualización) de la trama n . Sea $T_n, n = 1, \dots, N$, denotan la pantalla acumulada tiempo hasta (e incluyendo) el marco $n - 1$, es decir, $T_n = \sum_{k=0}^{n-1} t_k$ (y definir $T_0 = 0$). Sea $X_n, n = 0, \dots, N - 1$, denotar el tamaño del marco (número de bits) del vídeo (comprimido) codificada marco de la trama n . Deje $Q_Y n, n = 0, \dots, N - 1$, denotan la calidad (en términos de el pico de relación señal ruido (PSNR)) del componente de luminancia de la codificada (y posteriormente decodificada) fotograma de vídeo n (en dB). Del mismo modo, dejar $Q_U n$ y $Q_V n, n = 0, \dots, N - 1$, denotan las cualidades de las dos de crominancia componentes de tono (U) y la saturación (V) de la trama de vídeo codificada n (en dB).

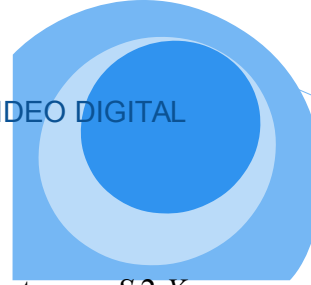
4.1 Vídeo Tamaño de cuadro

La media (aritmética) muestra \bar{X} de una traza tamaño del marco se estima como

$$\bar{X} = \frac{1}{N} \sum_{n=0}^{N-1} X_n. \quad (4.1)$$

La varianza de la muestra S^2_X de una traza tamaño del marco se estima como

$$S^2_X = \frac{1}{N-1} \sum_{n=0}^{N-1} (X_n - \bar{X})^2. \quad (4.2)$$



Una expresión computacionalmente más conveniente para $S^2 X$ es

$$S_X^2 = \frac{1}{N-1} \left[\sum_{n=0}^{N-1} X_n^2 - \frac{1}{N} \left(\sum_{n=0}^{N-1} X_n \right)^2 \right]. \quad (4.3)$$

El coeficiente de variación $CoV X$ de la traza tamaño de trama se define como

$$CoV_X = \frac{S_X}{\bar{X}}. \quad (4.4)$$

El tamaño máximo de la trama X_{\max} se define como

$$X_{\max} = \max_{0 \leq n \leq N-1} X_n. \quad (4.5)$$

Definimos la traza tamaño del marco se agrega con *un* nivel de agregación como

$$X_n^{(a)} = \frac{1}{a} \sum_{j=na}^{(n+1)a-1} X_j, \quad \text{for } n = 0, \dots, N/a - 1, \quad (4.6)$$

es decir, la traza tamaño de trama agregada se obtiene promediando la trama original tamaño de traza X_n , $n = 0, \dots, N-1$, sobre bloques de longitud a que no se solapan.

Definimos el tamaño rastro GoP como

$$Y_m = \sum_{n=mG}^{(m+1)G-1} X_n, \quad \text{for } m = 0, \dots, N/G - 1, \quad (4.7)$$

donde G denota el número de tramas en un GOP (donde típicamente $G = 12$). Tenga en cuenta que $Y m = G \cdot X (G) n$.

4.1.1 Auto correlación

La función de auto correlación [98] se puede utilizar para la detección de la no aleatoriedad en los datos o la identificación de un modelo de series de tiempo apropiado si el dato no es al azar. Un supuesto básico es que las observaciones son equidistantes.

El auto correlación se expresa como un coeficiente de correlación, se refiere como *coeficiente de auto correlación (acc)*. En lugar de calcular la correlación entre dos variables diferentes, como el tamaño y la calidad, la correlación es calculada para los valores de la misma variable en las posiciones n y $n + k$. Cuando la auto correlación se utiliza para detectar no aleatoriedad, es por lo general sólo la primera ($k = 1$ lag) de auto correlación que es de interés. Cuando la auto correlación se utiliza para identificar un modelo de serie de tiempo apropiado, las autocorrelaciones son generalmente trazada para un rango de k rezagos. El coeficiente de autocorrelación $\rho_X(k)$ para lag k , $k = 0, 1, \dots, N - 1$, se estima como

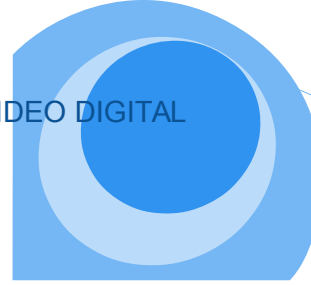
$$\rho_X(k) = \frac{1}{N - k} \sum_{n=0}^{N-k-1} \frac{(X_n - \bar{X})(X_{n+k} - \bar{X})}{S_X^2}. \quad (4.8)$$

```

 $S_X^2 = \frac{1}{N-1} \sum_{n=0}^{N-1} (X_n - \bar{X})^2;$ 
foreach  $a = 12, 24, 48, 96, \dots$  do
     $M = \lfloor N/a \rfloor ;$ 
     $X_n^{(a)} = \frac{1}{a} \sum_{j=na}^{(n+1)a-1} X_j, \quad n = 0, \dots, M - 1 ;$ 
     $S_X^{2(a)} = \frac{1}{M-1} \sum_{n=0}^{M-1} (X_n^{(a)} - \bar{X})^2 ;$ 
    plot point  $(\log_{10} a, \log_{10}(S_X^{2(a)}/S_X^2)) ;$ 
end

```

Algorithm 1: Algorithm for determining the variance-time plot.



4.1.2 Varianza-Tiempo de prueba

El tiempo de varianza parcela [99, 100, 101] se obtiene mediante el trazado de la normalizada varianza de la agregada del rastro $S^2(a)X / S^2X$ como una función de la agregación nivel ("tiempo"), *una* en un gráfico log-log, como se detalla en el algoritmo 1. Rastros sin dependencia a largo plazo eventualmente (por grande a) disminución linealmente con una pendiente de - 1 en el tiempo varianza trama. Las huellas con largo alcance de dependencia, por otra parte, con el tiempo disminuyen linealmente con un aplanador pendiente, es decir, una pendiente mayor que - 1.

Consideramos niveles de agregación que son múltiplos del tamaño de GOP (12 cuadros) para evitar el efecto de la intra-GP correlaciones. Para fines de referencia trazamos una recta con pendiente - 1 a partir del origen. Para la estimación del parámetro de Hurst se estima la pendiente de la parte lineal del tiempo varianza parcela usando un ajuste por mínimos cuadrados. Consideramos los niveles de agregación $a \geq 192$ en esta estimación ya que nuestro tiempo parcelas varianza normalmente son lineales para estos niveles de agregación. El parámetro de Hurst es entonces estimado como $H = \text{pendiente} / 2 + 1$.

4.1.3 R / S Estadística

Utilizamos la estadística R / S [99, 100, 102] para investigar la dependencia de largo alcance características de los restos generados. El estadístico R / S ofrece una heurística enfoque gráfico para estimar el parámetro H Hurst. En términos generales, de largo alcance son procesos estocásticos dependiente, se caracteriza el estadístico R / S por $E [R(n) / S(n)] \sim cn^H$ como $n \rightarrow \infty$ (donde c es una constante finita positiva). El parámetro H Hurst se estima como la pendiente de una gráfica log-log de la R / S estadística. Más formalmente, la *estadística de rango ajustado reescalado* (por sus siglas en R / S estadística) se representa de acuerdo con el algoritmo dado en el Algoritmo 2. El R / S estadísticos $R(t_i, d) / S(t_i, d)$ se calcula para los valores logarítmicamente espaciados del

lag k , comenzando con $d = 12$ (para evitar el efecto de las correlaciones intra-GOP). Para cada

```

foreach  $d = 12, 24, 48, 96, \dots$  do
   $I = K + 1 - \lceil \frac{dK}{N} \rceil$ ;
  foreach  $i = 1, \dots, I$  do
     $t_i = (i - 1)\frac{N}{K} + 1$ ;
     $\bar{X}(t_i, d) = \frac{1}{d} \sum_{j=0}^{d-1} X_{t_i+j}^{(a)}$ ;
     $S^2(t_i, d) = \frac{1}{d} \sum_{j=0}^{d-1} [X_{t_i+j}^{(a)} - \bar{X}(t_i, d)]^2$ ;
     $W(t_i, k) = \left( \sum_{j=0}^{k-1} X_{t_i+j}^{(a)} \right) - k\bar{X}(t_i, d)$ ;
     $R(t_i, d) =$ 
     $\max\{0, \max_{1 \leq k \leq d} W(t_i, k)\} - \min\{0, \min_{1 \leq k \leq d} W(t_i, k)\}$ ;
    plot point  $\left( \log d, \log \frac{R(t_i, d)}{S(t_i, d)} \right)$ ;
  end
end
end

```

Algorithm 2: Algorithm for the R/S statistic plot.

lag valor d tantos como K muestras de R/S se calcula teniendo en cuenta diferentes puntos de partida t_i ; establecimos $K = 10$ en nuestro análisis. Los puntos de partida deben satisfacer $(t_i - 1) + d \leq N$, de ahí el número real de muestras I es menos de K para grandes retrasos d . Trazado de $\log [R(t_i, d) / S(t_i, d)]$ como una función de $\log d$ da la *trama reescalado ajustado rango* (también denominado *diagrama de viruela de R/S*).

Un diagrama típico de la viruela comienza con una zona transitoria representa el corto variar de las características de dependencia de la traza. La trama se establece y fluctúa en torno a una "calle" recta de pendiente H . Si el gráfico presenta este comportamiento asintótico, el *asintótica exponente de Hurst* H se estima a partir de la pendiente de la calle mediante un ajuste por mínimos cuadrados. Para verificar la robustez de la estimación repetimos este procedimiento para cada trazar de diferentes niveles de agregación $a \geq 1$. El parámetro de Hurst, o auto- parámetro de similitud, H , es una medida clave de la auto-

similitud [103, 104]. H es una medida de la persistencia de un fenómeno estadístico y es una medida de la longitud de la dependencia a largo plazo de un proceso estocástico. A

Hurst parámetro de $H = 0.5$ indica ausencia de auto-similitud mientras que $H = 1$ indica el grado de persistencia o un regalo de dependencia a largo plazo.

4.1.4 Periodo grama

Estimamos el parámetro H Hurst usando la heurística de regresión de mínimos cuadrados en el dominio espectral, ver [99, Sec. 4,6] para más detalles. Este enfoque se basa en el periodograma $I(\lambda)$ como aproximación de la densidad espectral, que cerca satisface al origen de sesión $I(\lambda) \approx \text{sesión } c + f \cdot (1 - 2H) \lambda \log k + \log \zeta k$ (4.9) Para estimar el parámetro H Hurst graficamos el periodograma en un diario de registro trama, como se detalla en el Algoritmo 3. (Tenga en cuenta que la expresión dentro de la $|\cdot|$

```

M = floor(N/a);
foreach n = 0, 1, ..., M - 1 do
    | X_n^(a) = 1/a * sum_{j=na}^{(n+1)a-1} X_j;
    | Z_n^(a) = log10 X_n^(a);
end
foreach k = 1, 2, ..., floor((M-1)/2) do
    | lambda_k = 2*pi*k/M;
    | I(lambda_k) = 1/(2*pi*M) * |sum_{n=0}^{M-1} Z_n^(a) * e^{-j*n*lambda_k}|^2;
    | plot point(log10 lambda_k, log10 I(lambda_k));
end

```

Algorithm 3: Algorithm for periodogram.

Corresponde al coeficiente de transformada de Fourier a la frecuencia λk , que puede ser evaluado de manera eficiente usando técnicas de Transformada Rápida de Fourier). Para la Estimación de parámetros Hurst definimos

$$x_k = \log_{10} \lambda_k \quad y_k = \log_{10} I(\lambda_k) \quad (4.10)$$

$$\beta_0 = \log_{10} c_f - 0.577215 \quad \beta_1 = 1 - 2H \quad (4.11)$$

$$e_k = \log_{10} \xi_k + 0.577215 \quad (4.12)$$

With these definitions we can rewrite (4.9) as

$$y_k = \beta_0 + \beta_1 x_k + e_k. \quad (4.13)$$

Estimamos β_0 y β_1 de las muestras (x_k, y_k) , $k = 1, 2, \dots, \lfloor 0.7 \cdot (N/a - 2) \rfloor = K$ utilizando regresión de mínimos cuadrados, es decir,

$$\beta_1 = \frac{K \sum_{k=1}^K x_k y_k - \left(\sum_{k=1}^K x_k \right) \left(\sum_{k=1}^K y_k \right)}{K \left(\sum_{k=1}^K x_k^2 \right) - \left(\sum_{k=1}^K x_k \right)^2} \quad (4.14)$$

Donde: (4.15)

$$\beta_0 = \frac{\sum_{k=1}^K y_k - \beta_1 \sum_{k=1}^K x_k}{K}$$

El parámetro de Hurst se calcula entonces como $H = (1 - \beta_1) / 2$. Se traza el periodograma (junto con la recta $y = \beta_0 + \beta_1 x$) y estimar el Hurst parámetro de esta manera para el nivel de agregación $a = 12, 24, 48, 96, 192, 300, 396, 504, 600, 696, \text{ y } 792$.

4.1.5 Diagrama Logscale

Estimamos conjuntamente el escalamiento parámetros α y c_f utilizando el basado wavelet-enfoque de Veitch y Abry [105], donde α y c_f caracterizan el espectral de densidad

$$(4.16)$$

$$f_X(\lambda) \sim c_f |\lambda|^{-\alpha}, \quad |\lambda| \rightarrow 0.$$

La estimación se basa en el diagrama logscale, que es un gráfico de $\log_2(\mu_j)$ como una función de $\log_2 j$, donde

$$\mu_j = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j, k)|^2 \quad (4.17)$$

es la varianza de la muestra de la wavelet coeficiente $d_X(j, k)$, $k = 1, \dots, N_j$, en octava j . El número de coeficientes wavelet disponibles en octava j es esencialmente $n_j = N/2^j$. Se traza el diagrama logscale para octavas 1 a 14 utilizando el código proporcionado por Veitch y Abry [105]. Utilizamos las daubechies 3 wavelet para eliminar lineales y cuadráticas tendencias [106]. Utilizamos el enfoque choosenewj1 automatizado [105] para determinar el rango de escalas (octavas) para la estimación de los parámetros de escala.

4.1.6 Diagrama Multiescala

Investigamos la multifractal escalada a propiedades [105, 106, 107, 108, 109, 110, 111, 112, 113, 114] con el marco basado en wavelets [109]. En este marco el q°

orden escalado exponente α_q se estima con base en la q° orden logscale diagrama, es decir, una parcela de

$$\log_2(\mu_j^{(q)}) = \log_2 \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j, k)|^q \quad (4.18)$$

como una función de $\log_2 j$. El diagrama multiescala se obtiene entonces por el trazado $\zeta(q) = \alpha_q - q/2$ como una función de q . Una variación del diagrama multiescala, el llamado diagrama de multiescala lineal se obtiene representando $h_q = \alpha_q / q - 1/2$ como una

función de q . Empleamos el código Matlab multiscaling proporcionada por Abry y Veitch [105]. Empleamos las daubechies 3 wavelet. We utilizan la norma L2, sigtype 1, el vector q [0,5, 1, 1,5, 2, 2,5, 3, 3,5, 4]. Utilizamos el enfoque newchoosej1 automatizado formar Abry y de Veitch diagrama logscale código Matlab [105] para determinar la gama de escalas (octavas) para la estimación de los parámetros de escala.

4.2 Calidad de Video Frame

Considere la posibilidad de una secuencia de vídeo con N tramas (imágenes), cada una de las dimensiones $D_x \times D_y$ píxeles y . Sea $I(n, x, y)$, $n = 0, \dots, N - 1$; $x = 1, \dots, D_x$; $y = 1, \dots, D_y$, denotar la luminancia (nivel de gris, o componente Y) valor del píxel en la ubicación (X, y) en el cuadro de video n . El error cuadrático medio (MSE) se define como la media de las diferencias al cuadrado entre los valores de luminancia del vídeo marcos en dos secuencias de vídeo I y \tilde{I} . Específicamente, el MSE para un individuo fotograma de vídeo n se define como

$$M_n = \frac{1}{D_x \cdot D_y} \sum_{x=1}^{D_x} \sum_{y=1}^{D_y} [I(n, x, y) - \tilde{I}(n, x, y)]^2. \quad (4.19)$$

MSE media de una secuencia de N fotograma de vídeo es

$$\bar{M} = \frac{1}{N} \sum_{n=0}^{N-1} M_n. \quad (4.20)$$

El RMSE se define como la raíz cuadrada de la MSE

$$RMSE = \sqrt{MSE}. \quad (4.21)$$

El pico de señal a ruido (PSNR) en decibelios (dB) es generalmente definida como $PSNR = 10 \cdot \log_{10} (p^2 / MSE)$, donde p denota la luminancia máxima el valor de un píxel (255 en 8 bits fotos). Definimos la *calidad* (en dB) de un *fotograma de vídeo* como n

$$Q_n = 10 \cdot \log_{10} \frac{p^2}{M_n}. \quad (4.22)$$

Definimos la *calidad media* (en dB) de una *secuencia de vídeo* que consta de N marcos como

$$\bar{Q} = 10 \cdot \log_{10} \frac{p^2}{M}. \quad (4.23)$$

Nótese que en esta definición de la calidad media, se lleva a cabo el cálculo del promedio con los valores de MSE y la calidad de vídeo se da en términos de la PSNR (en dB). También definimos una *calidad media alternativa* (en dB) de una secuencia de vídeo Como

$$\bar{Q}' = \frac{1}{N} \sum_{n=0}^{N-1} Q_n, \quad (4.24)$$

Donde se realiza el cálculo del promedio en los valores de PSNR directamente. Definimos ahora extensiones naturales de las métricas de calidad superiores. Definimos la varianza de la muestra MSE $S^2 M$ de una secuencia de tramas de vídeo N como

$$S_M^2 = \frac{1}{N-1} \sum_{n=0}^{N-1} (M_n - \bar{M})^2, \quad (4.25)$$

y la desviación estándar MSE S_M como

$$S_M = \sqrt{S_M^2}. \quad (4.26)$$

Definimos la *calidad desviación estándar* S_Q de una secuencia de vídeo como

$$S_Q = 10 \cdot \log_{10} \frac{p^2}{S_M}. \quad (4.27)$$

Se define el *coeficiente de variación de la calidad* $CoQV$ de una secuencia de vídeo como

$$CoQV = \frac{S_Q}{Q}. \quad (4.28)$$

Definimos una *desviación estándar de calidad* como alternativa

$$S'_Q = \sqrt{\frac{1}{N-1} \sum_{n=0}^{N-1} (Q_n - \bar{Q}')^2}, \quad (4.29)$$

y el correspondiente coeficiente de variación alternativa de la calidad como

$$CoQV' = \frac{S'_Q}{Q'}. \quad (4.30)$$

Se define la *gama de la calidad* (en dB) de una secuencia de vídeo como

$$Q_{\min}^{\max} = \max_{0 \leq n \leq N-1} Q_n - \min_{0 \leq n \leq N-1} Q_n. \quad (4.31)$$

Estimamos que el MSE coeficiente de autocorrelación $\rho M(k)$ para el desfase k , $k = 0, \dots, N-1$

$$\rho M(k) = \frac{1}{N-k} \sum_{n=0}^{N-k-1} \frac{(M_n - \bar{M})(M_{n+k} - \bar{M})}{S_M^2}. \quad (4.32)$$

Si bien las definiciones anteriores se centran en las cualidades a nivel de individuo fotogramas de vídeo, que también definen, como extensiones, cualidades de agregados (grupos) de *unos* marcos (con la GoP ser un caso especial de agregación de tramas con $a = G$, donde típicamente $G = 12$). Sea $M(a) m$, $m = 0, \dots, N/a - 1$, denotar el MSE del m -ésimo grupo de tramas, definida como

$$M_m^{(a)} = \frac{1}{a} \sum_{n=ma}^{(m+1)a-1} M_n. \quad (4.33)$$

Deje $Q(a) m$, $m = 0, \dots, N/a - 1$, denotan la calidad PSNR correspondiente (en dB), define como

$$Q_m^{(a)} = 10 \cdot \log_{10} \frac{P^2}{M_m^{(a)}}. \quad (4.34)$$

Se define la varianza de la muestra MSE $S^2(a) M$ de una secuencia de grupos de tramas cada uno como *una*

$$S_M^{2(a)} = \frac{1}{N/a - 1} \sum_{n=0}^{N/a-1} \left(M_n^{(a)} - \bar{M} \right)^2, \quad (4.35)$$

y la correspondiente MSE desviación estándar $S(a) M$ como

$$S_M^{(a)} = \sqrt{S_M^{2(a)}}. \quad (4.36)$$

Definimos la calidad desviación estándar $S(a) Q$ de una secuencia de grupos de *unos* marcos de cada uno como

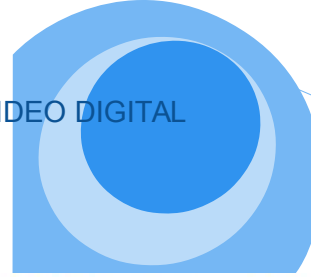
$$S_Q^{(a)} = 10 \cdot \log_{10} \frac{p^2}{S_M^{(a)}}. \quad (4.37)$$

Se define el coeficiente de variación de la calidad $CoQV(a)$ de una secuencia de grupos de tramas cada uno como *una*

$$CoQV^{(a)} = \frac{S_Q^{(a)}}{Q}. \quad (4.38)$$

Se define la *desviación estándar alternativa de calidad* para grupos de *unos* marcos de cada uno como

$$S_Q^{(a)} = \sqrt{\frac{1}{N/a - 1} \sum_{n=0}^{N/a-1} \left(Q_n^{(a)} - Q \right)^2}, \quad (4.39)$$



where $Q_n^{r(a)} = \frac{1}{a} \sum_{n=ma}^{(m+1)a-1} Q_n$. We define the corresponding alternative coefficient of quality variation as

$$C_{oQV}^{r(a)} = \frac{S_Q^{r(a)}}{Q^r}. \quad (4.40)$$

Se define la gama de la calidad (en dB) de una secuencia de grupos de tramas cada uno como *una*

$$Q_{\min}^{\max(a)} = \max_{0 \leq n \leq N/a-1} Q_n^{(a)} - \min_{0 \leq n \leq N/a-1} Q_n^{(a)}. \quad (4.41)$$

Calculamos el coeficiente de autocorrelación MSE para grupos de *unos* marcos de $\rho(a) M$ para lag $k, k = 0, A, 2a, \dots, N/A - 1$ fotogramas como

$$\rho_M^{(a)}(k) = \frac{1}{N/a - k} \sum_{n=0}^{N/a-k-1} \frac{(M_n^{(a)} - \bar{M})(M_{n+k}^{(a)} - \bar{M})}{S_M^{(a)}}. \quad (4.42)$$

4.3 Correlación entre tamaños de fotogramas de vídeo y Calidades

Definimos la covarianza entre el tamaño del marco y de la calidad del marco MSE como

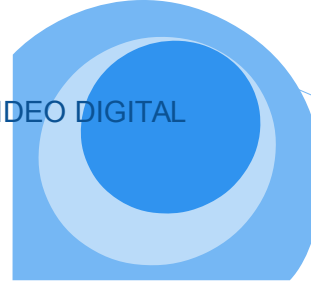
y el

$$S_{XM} = \frac{1}{N-1} \sum_{n=0}^{N-1} (X_n - \bar{X})(M_n - \bar{M}), \quad (4.43)$$

coeficiente de correlación de la calidad tamaño-MSE como

Definimos la covarianza entre el tamaño del marco y (PSNR) calidad de la trama como

$$\rho_{XM} = \frac{S_{XM}}{S_X \cdot S_M}. \quad (4.44)$$



y el

$$S_{XQ} = \frac{1}{N-1} \sum_{n=0}^{N-1} (X_n - \bar{X})(Q_n - \bar{Q}'), \quad (4.45)$$

coeficiente de correlación tamaño calidad como

$$\rho_{XQ} = \frac{S_{XQ}}{S_X \cdot S'_Q}. \quad (4.46)$$

Al igual que en las definiciones de nivel marco anteriores, definimos la covarianza entre los tamaños de bastidor agregados $X^{(a)}_n$, $n = 0, \dots, N/a - 1$, y los agregados MSE cualidades $M^{(a)}_n$, $n = 0, \dots, N/a - 1$

$$S_{XM}^{(a)} = \frac{1}{N/a - 1} \sum_{n=0}^{N/a-1} (X_n^{(a)} - \bar{X})(M_n^{(a)} - \bar{M}), \quad (4.47)$$

y el coeficiente de correlación correspondiente como

$$\rho_{XM}^{(a)} = \frac{S_{XM}^{(a)}}{S_X^{(a)} \cdot S_M^{(a)}}. \quad (4.48)$$

Definimos la covarianza entre agregada tamaño de marco $X^{(a)}_n$, $n = 0, \dots, N/a - 1$ - las cualidades ggregated PSNR) 1, y $Q^{(a)}_n$, $n = 0, \dots, N/a - 1$

$$S_{XQ}^{(a)} = \frac{1}{N/a - 1} \sum_{n=0}^{N/a-1} (X_n^{(a)} - \bar{X})(Q_n^{(a)} - \bar{Q}'), \quad (4.49)$$

y el coeficiente de correlación correspondiente como

$$\rho_{XQ}^{(a)} = \frac{S_{XQ}^{(a)}}{S_X^{(a)} \cdot S_Q^{(a)}} \quad (4.50)$$

4.4 Las métricas adicionales para FGS codificaciones

La capa de base (BL) y la capa de mejora FGS (EL) de la de vídeo son VBR-codificado, con bitrates instantáneos $r_b(t)$ y $r_e(t)$ durante el periodo de trama t , $t = 1, \dots, N$. De acuerdo con la propiedad FGS, la capa de mejora se puede truncar en cualquier lugar antes de la decodificación. Denotamos cualquier parte de la que ha de ser añadido a la BL como *EL subcorriente*. Decimos que una corriente parcial es EL codificado a tasa $C(t) \in [0, r_e(t)]$, cuando el último $T \cdot (r_e(t) - C(t))$ bits de cada trama t , $t = 1, \dots, N$, se han eliminado de la corriente de bits EL originales. El Grupo BL de imágenes (GOP) se compone de 12 imágenes a lo largo de nuestro estudio, y su patrón se fija en IBBPBBPBBPBB.

Suponemos que el vídeo se divide en escenas consecutivas. Sea S el número total de escenas en un vídeo dado de tramas de longitud N . Deje s , $s = 1, \dots, S$, denota el índice de escenas y N_s la longitud (en número de imágenes) de número de escena s (nota que $\sum_{s=1}^S N_s = N$).

Deje $Q_t(C)$, $t = 1, \dots, N$ denotar la calidad de la imagen decodificada t *XIX*, cuando la EL se codifica con tarifa C . Deje $Q_{bt} = Q_t(0)$, denotan la calidad de la misma imagen, cuando se descodifica sólo el BL. Definimos $Q_{et}(C) = Q_t(C) - Q_{bt}$ como la mejora (aumento) en la calidad que se consigue cuando se decodifica la EL, así como la BL, del bastidor t codificado con tasa C .

Las diferentes estadísticas de las cualidades de fotograma de vídeo individuales, se calculan como se da arriba. Nos indican el tamaño total del marco t por $X_t(C) = X_{bt} + X_y(C)$, cuando el EL se codifica con tarifa C . Sean X_{eit} , $i = 1, \dots, 8$,

$t = 1, \dots, N$ denotar el tamaño de plano binario EL i de la trama t e $Y_{e i t}$, $i = 1, \dots, 8$, $t = 1, \dots, Y_n$ el tamaño agregado de los planos de bits $1, \dots, I$ ($I_{e i t} = ij = 1 X_{e j t}$).

Deje $Q_{s, n}(C)$, $s = 1, \dots, S$, $n = 1, \dots, N_s$ denotan la calidad de la n -ésima trama de vídeo decodificada de la escena s , cuando la EL se codifica con la tasa C . En cuanto a $Q_t(C)$, denotamos la calidad de la trama n dentro de la escena s , cuando solamente el BL es decodificado por $Q_{bs, n} = Q_{s, n}(0)$, y la mejora en la calidad alcanzada cuando decodificar el EL es por $Q_{N}(C) = Q_{s, n}(C) - Q_{bs, n}$. La tasa de distorsión (RD) características de cada imagen n dentro de la escena s se obtienen mediante el trazado de las curvas $Q_{s, n}(C)$. La varianza media y la muestra de la calidad de las imágenes dentro de la escena s , $s = 1, \dots, S$, se stimated de la siguiente manera:

$$\bar{Q}_s(C) = \frac{1}{N_s} \sum_{n=1}^{N_s} Q_{s,n}(C), \quad (4.51)$$

$$\sigma_{Q_s}^2(C) = \frac{1}{N_s - 1} \sum_{n=1}^{N_s} [Q_{s,n}(C) - \bar{Q}_s(C)]^2. \quad (4.52)$$

The coefficient of quality variation of scene s , $s = 1, \dots, S$, is given by:

$$Cov_s = \frac{S Q_s(C)}{\bar{Q}_s(C)}. \quad (4.53)$$

Para cada escena s , también denota el tamaño total de la imagen por $n X_{s, n}(C) = X_{bs, N} + X_{es, n}(C)$, cuando la EL se codifica con la tasa C .

Estimamos la media $\bar{X}_s(C)$, la varianza de la muestra $\sigma^2 X_s(C)$ y autocorrelación coeficiente $\rho X_s(C, k)$ de la secuencia del total de los tamaños de imagen $X_{s, n}(C)$, $n = 1, \dots, N_s$, de la misma manera como para calidades de imagen. Se denota la media, la varianza y autocorrelación de los tamaños de marco BL y El As $\bar{X}_{bs}(C)$, $\sigma^2 X_{bs}(C)$, $\rho X_{bs}(C, k)$ y $\bar{X}_{es}(C)$, $\sigma^2 X_{es}(C)$, $\rho X_{es}(C, k)$, respectivamente.

Hacemos un seguimiento de la longitud (en fotogramas de vídeo) de las escenas sucesivas $N S$, $S = 1, \dots, S$. Denotamos la media y la varianza de $N s$ como $\bar{N} = N/n$ y $\sigma^2 N$. La media de la calidad de todas las imágenes individuales de una escena se denota como $\bar{s} Q(C)$ Que $\theta_s(C)$ sea la *calidad total* de número de escena de vídeo s , $s = 1, \dots, N$, cuando

el EL ha sido codificado a tipo C para todas las imágenes de la escena. Al igual que en la medida de la calidad de las imágenes individuales de una escena determinada, definimos $\theta_s(C) = \theta_{bs} + \theta_{es}(C)$, donde $\theta_{bs} = \theta_s(0)$ denota la calidad total de la escena de s cuando sólo la BL se decodifica, y $\theta_{es}(C)$ la mejora en la calidad chieved por el EL codificado a tasa C . Analizamos los media, varianza y coeficientes de autocorrelación de la cualidades de escena, tal como se define por:

$$\bar{\theta}(C) = \frac{1}{S} \sum_{s=1}^S \theta_s(C), \quad (4.54)$$

Para cada escena s , las características de tasa-distorsión se obtienen mediante el trazado de las curvas $\theta_s(C)$. La media y la varianza de las cualidades de las escenas 'dan una idea general de la calidad percibida de todo el video. Sin embargo, la diferencia de calidad de escena no capta las diferencias de calidad entre las escenas de vídeo sucesivas, que degradan la calidad percibida. Para capturar esto, se introduce una nueva métrica, llamada la variabilidad, que se define como:

$$\sigma_{\theta}^2(C) = \frac{1}{S-1} \sum_{s=1}^S [\theta_s(C) - \bar{\theta}(C)]^2, \quad (4.55)$$

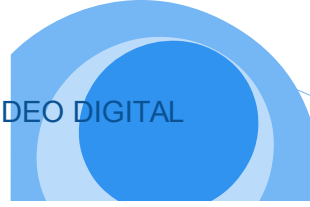
$$\rho_{\theta}(C, k) = \frac{1}{S-k} \sum_{s=1}^{S-k} \frac{[\theta_s(C) - \bar{\theta}(C)][\theta_{s+k}(C) - \bar{\theta}(C)]}{\sigma_{\theta}^2(C)}. \quad (4.56)$$

Nota: Con el fin de explicar las diferencias en la longitud de las sucesivas escenas, también pueden llegar a pesar escenas en función de su respectiva longitud de trama.

Deje $\theta'_s(C)$ indican la medida ponderada de la calidad de la escena, expresado como:

$$\theta'_s(C) = \frac{N_s}{N} \theta_s(C), \quad (4.58)$$

Podemos definir la media y la varianza de la calidad ponderado:



$$\bar{\theta}'(C) = \frac{1}{S} \sum_{s=1}^S \theta'_s(C) = \frac{1}{S} \sum_{s=1}^S \frac{N_s}{N} \theta_s(C) = \frac{1}{N} \sum_{s=1}^S N_s \cdot \theta_s(C), \quad (4.59)$$

$$\sigma_{\theta'}^2(C) = \frac{1}{S-1} \sum_{s=1}^S (\theta'_s(C) - \bar{\theta}'(C))^2 = \frac{1}{S-1} \left[\sum_{s=1}^S (\theta'_s(C))^2 - S(\bar{\theta}'(C))^2 \right], \quad (4.60)$$

Recordando que $\bar{X}_s(C)$ denota el tamaño medio de los marcos dentro de la escena s , el coeficiente de correlación entre el tamaño medio de bastidor $\bar{X}_s(C)$

$$\rho_{X,\theta}(C) = \frac{1}{S-1} \sum_{s=1}^S \frac{(\bar{X}_s(C) - \bar{X}(C))(\theta_s(C) - \bar{\theta}_s(C))}{S_{\bar{X}_s}(C) \cdot S_{\theta}(C)}, \quad (4.61)$$

de una escena y la calidad total $\theta_s(C)$ de una escena se estima como:

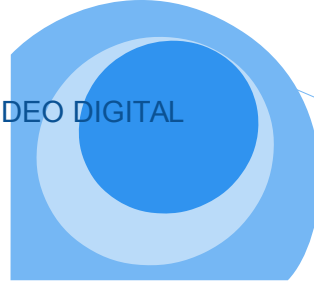
Donde $\bar{X}(C)$ indican la media de los sucesivos tamaños medios de trama de todas las escenas componer el video ($\bar{X}(C) = \sum_{s=1}^S \bar{X}_s(C) / S$).

Por último, se denota el coeficiente de correlación entre la calidad y BL del total (BL + EL) calidad de una escena por $\rho_{b,\theta}(C)$.

4.5 Métricas adicional para MDC codificaciones

Como MDC es la introducción de una sobrecarga de codificación, necesitamos extender la métrica con respecto a esta sobrecarga. La sobrecarga de codificación se define como la cantidad de datos por los cuales las corrientes divididas de una

secuencia de vídeo se incrementan en comparación con el flujo único. La sobrecarga de codificación es diferente a la sobrecarga de la red que viene en la parte superior de cada descriptor. La sobrecarga MDC OH se calcula mediante la suma de todos los tamaños de carcasa $X_{n,j}$ sobre N tramas de J descriptors, dividiéndolo por la suma de todos los marcos de tamaño X_i sobre todos los marcos $N \cdot J$ de la corriente única y resta por 1.



Referencias

1. M. Dai y D. Loguinov, "Wavelet y el dominio del tiempo de modelado de multi-capa de El tráfico de vídeo VBR ", en *Proc. de Packet Video*, Irvine, CA, diciembre de 2004.
2. - "Análisis y modelización de MPEG-4 y H.264 multicapa tráfico de vídeo" en *Proc. de IEEE INFOCOM*, Miami, FL, 03 2005, pp. 2257-2267.
3. I. Dalgic y FA Tobagi, "Caracterización de la calidad y el tráfico para varios esquemas de codificación de vídeo y varios esquemas de control del codificador, "la Universidad de Stanford, Departamentos de Ingeniería Eléctrica y Ciencias de la Computación, Tech. Rep. CSL-TR-96-701, agosto 1996.
4. M. Frey y S. Nguyen-Quang, "Un marco basado gamma para el modelado a tasa variable fuentes de vídeo MPEG: El modelo republicano GBAR, "IEEE / ACM Transacciones de Redes, vol. 8, no. 6, pp. 710-719, diciembre de 2000.
5. M. Garrett y W. Willinger, "Análisis, modelado y generación de auto- el tráfico de vídeo VBR similar ", en *Actas de ACM SIGCOMM*, Londres, Reino Unido, 09 1994, pp. 269-280.
6. DP Heyman y TV Lakshman, "modelos de fuentes para transmisión de video VBR tráfico, "IEEE / ACM Transactions on Networking, vol. 4, pp. 40-48, enero de 1996.
7. M. Krunz y S. Tripathi, "Sobre la caracterización de VBR MPEG arroyos" en *Actas del ACM SIGMETRICS*, Seattle, WA, junio de 1997, pp. 192-202.
8. CH Liew, CK Kodikara y AM Kondo, variables MPEG-codificada " modelado de tráfico de vídeo poco nivel, "IEE Proceedings Communications, vol. 152, no. 5, pp. 749-756, octubre 2005.

9. D. Lucantoni, M. Neuts, y A. Reibman, "Métodos para la evaluación del desempeño de modelos de tráfico de vídeo VBR, " *IEEE / ACM Transactions on Networking*, vol. 2, no. 2, pp. 176-180, abril de 1994.
10. Reino Unido Sarkar, S. Ramakrishnan, y D. Sarkar, "Modelado video completo Utilizando el framework basado en gamma Markov modulada, " *IEEE / ACM Transactions en redes*, vol. 11, no. 4, pp. 638-649, 2003.
11. - "Estudio de larga duración métodos de segmentación MPEG-trace para el desarrollo tamaño del marco modelos de tráfico basados, " *Redes de Computadores*, vol. 44, no. 2, pp. 177-188, 2004.
12. ITU-500-R, "Recomendación BT.500-8 - Metodología para la subjetiva Evaluación de la calidad de las imágenes de televisión, "1998.
13. A. Basso, I. Dalgic, F. Tobagi y C. Lambrecht, "Estudio de codificación MPEG-2 rendimiento basado en una métrica de calidad perceptual ", en *Actas de Imagen Codificación Simposio*, Melbourne, Australia, marzo 1996.
14. A. Webster, C. Jones, M. Pinson, S. Voran, y S. Wolf, "Un video objetivo Sistema de evaluación de la calidad basado en la percepción humana ", en *Actas de SPIE Visión Humana, procesamiento visual y display digital*, Vol. 1913, 1993, pp 15-26.
15. S. Winkler, "Una distorsión perceptual métrica para el vídeo digital en color", en *Actas de SPIE Visión Humana y Electronic Imaging*, Vol. 3644, enero de 1999, pp. 175-184.
16. R. Aravind, M. Civanlar, y A. Reibman, "La pérdida de paquetes resiliencia de MPEG-2 algoritmos, " *IEEE Transactions codificación de Circuitos y Sistemas de video escalable de Tecnología de Video*, vol. 6, no. 5, pp. 426-435, octubre de 1996.

17. H. Liu y ME Zarki, "Rendimiento de la transmisión de vídeo H.263 sobre wireless redes que utilizan ARQ híbrido, " *IEEE Diario en áreas seleccionadas en las Comunicaciones Sociales*, vol. 15, no. 9, pp. 1775-1786, diciembre de 1997. Análisis de los sistemas de ocultación de errores para MPEG-2
18. W. Luo y M. ElZarki " transmisión de vídeo a través de redes basadas en ATM ", en *Actas de SPIE Visual Comunicaciones y Procesado de Imagen 1995*, Taiwán, mayo de 1995, pp. 102-108.
19. - "MPEG2Tool: Un juego de herramientas para el estudio de MPEG-2 de transmisión de vídeo sobre redes basadas en ATM ", Departamento de Ingeniería Eléctrica de la Universidad de Pensilvania, Tech. Rep., 1996.
20. R. Puri, K. Lee, K. Ramachandran, y V. Bharghavan, "una fuente integrada transcodificación y paradigma de control de congestión para la transmisión de vídeo en Internet " *IEEE Transactions on Multimedia*, vol. 3, no. 1, pp. 18-32, marzo 2001.
21. J. Shin, JW Kim, y C.-CJ Kuo, "mecanismo de asignación de calidad de servicio para el vídeo de paquetes en la red de servicios diferenciados, " *IEEE Transactions on Multimedia*, vol. 3, no. 2, pp. 219-231, junio de 2001.
22. W.-C. Feng, *Técnicas de almacenamiento en búfer para la entrega de video comprimido en Video- Sistemas on-Demand*. Kluwer Academic Publisher, 1997.
23. MW Garret, "Aportes hacia los servicios en tiempo real sobre redes de paquetes" Ph.D. tesis, la Universidad de Columbia, mayo de 1993.
24. M. Krunz, R. Sass, y H. Hughes, "Características estadísticas y multiplexación de flujos MPEG ", en *Actas del IEEE Infocom '95*, abril de 1995, pp. 455- 462.
25. O. Rose, "Propiedades estadísticas de tráfico de vídeo MPEG y su impacto en

modelado del tráfico en los sistemas ATM "Universidad de Würzburg, Instituto de Computación Ciencia, Tecnología. Rep. 101, febrero 1995.

26. F. Fitzek y M. Reisslein, "Un protocolo de obtención previa de los medios continuos en streaming en entornos inalámbricos " *IEEE Diario en áreas seleccionadas en las Comunicaciones Sociales*, vol. 19, no. 10, pp. 2015-2028, octubre de 2001.

27. -, "rastros de vídeo MPEG-4 y H.263 para la evaluación del desempeño de la red" *IEEE Network*, vol. 15, no. 6, pp. 40-54, noviembre / diciembre de 2001, el vídeo traza disponible en <http://trace.eas.asu.edu> .

28. P.Seeling y M. Reisslein, "La tasa de variabilidad distorsión (vd) curva de codificada vídeo y su impacto en la multiplexación estadística ", *IEEE Transactions on Radiodifusión*, vol. 51, no. 4, pp. 473-492, diciembre de 2005.

29. K. Jack, *Video Desmitificar: Manual para el Ingeniero digital*, 2ª ed. San Diego, CA: HighText Interactive, 1996.

30. M. Ghanbari, *Video Coding - Una introducción a la Norma Codecs*. The Institution de Ingenieros Eléctricos, 1999.

31. UN Netravali y BG Haskell, *Imágenes Digitales: Representación, compresión, y estándares*. Plenum Press, 1995, cap. 3.

32. T. Sikora, "MPEG vídeo digital Estándares de Codificación" en *Electrónica Digital Guía para el consumidor*. McGraw Hill, 1997.

33. Z. Xiong, K. Ramchandran, MT Orchard, y Y.-P. Zhang, "Un comparativo Estudio de DCT- y basada en ondas de codificación de imágenes, " *IEEE Transactions on Circuitos y Sistemas de Video Technology*, vol. 9, no. 5, agosto de 1999.

34. W. Chen, C. Smith y S. Fralick, "Un algoritmo computacional rápido para la transformada discreta del coseno, " *IEEE Transcripción sobre Comunicaciones*, pp. 1004-1009, septiembre 1977.
35. I. Dinstein, K. Rose, y A. Heimann, "Variable bloque de tamaño transformar la imagen codificador, " *IEEE Transactions on Communications*, pp. 2073-2078, Nov. de 1990.
36. Joint Video Team de la UIT-T e ISO / IEC JTC 1 ", Proyecto de Recomendación UIT-T y Proyecto Final de Norma Internacional de Vídeo Conjunto Especificación (. UIT-T H.264 Rec - ISO / IEC 14496-10 AVC), "2004.
37. MJ Riley y IE Richardson, *digitales de comunicaciones por vídeo*. Artech Casa, 1997.
38. DS Turaga y T. Chen, *vídeo comprimido a través de redes*, ser. La Señal Serie de Procesamiento. Nueva York, NY: Marcel Dekker, Sept. de 2000, ch. Fundamentos Codificación de vídeo: H.263 como ejemplo, pp 3-34..
39. Microsoft, "Visual MPEG-4 codec versión 2.5.0," 08 2004.
40. "El video traza para la evaluación del desempeño de la red," sitio web en línea. [En línea]. Disponible: <http://www.eas.asu.edu/trace>
41. T. Lakshman, A. Ortega, y A. Reibman, "VBR de vídeo: Compensaciones y potencialidades" *Actas de la IEEE*, vol. 86, no. 5, pp. 952-973, mayo de 1998.
42. A. Ortega, *vídeo comprimido a través de redes*. Marcel Dekker, Sept. de 2000, ch. Variable Bit Rate Video Coding, pp. 343-383.
43. K. Dovstam, "Codificación de vídeo en H.26L," Ph.D. disertación, Instituto Real de Tecnología de Estocolmo, Suecia, 2000.
44. D. Lelewer y D. Hirschberg, de "compresión de datos", *ACM Computing Surveys (CSUR)*, vol. 19, no. 3, pp. 261-296, 1987.

45. P. Howard y J. Vitter, "Análisis de codificación para la compresión de datos de la aritmética" *Procesamiento de la Información y Gestión*, vol. 28, no. 6, pp. 749-764, 1992.
46. T. Wiegand, "H.26L Prueba del Modelo Largo Plazo Número 9 (TML-9) draft0" ITUT La Comisión de Estudio 16, diciembre 2001.
47. Y. Shi y H. Sun, *Imagen y compresión de vídeo de Ingeniería Multimedia*. CRC Press, 2000.
48. T. Koga, K. Linuma, A. Hirano, Y. Iijima y T. Ishiguro, "Motion compensado entre tramas de codificación para videoconferencia, "en *Proceedings de Nacional Conferencia de Telecomunicaciones (NTC)*, vol. G5. 3, Nueva Orleans, LA, diciembre . 1981, pp 1-5.
49. M. Bierling, "estimación Desplazamiento por juego bloque jerárquico," *Visual Comunicaciones y Procesamiento de Imágenes*, vol. 1001, pp. 942-951, 1988.
50. MPEG-1, "Codificación de imágenes en movimiento y audio asociado para almacenamiento digital medios de hasta 1.5 mbps, "ISO / IEC 11172, 1993
51. R. Rejaie, D. Estrin, y M. Handley, "la adaptación de Calidad para la congestión controlados reproducción de vídeo a través de Internet ", en *Actas de ACM SIGCOMM*, Cambridge, MA, Sept. 1999, págs. 189-200.
52. D. Saporilla y KW Ross, "la transmisión óptima de vídeo en capas", en *Actas de IEEE INFOCOM*, Tel Aviv, Israel, marzo 2000, pp. 737-746.
53. G. Reyes, AR Reibman, S.-F. Chang, J. y Chuang ", Error elástica *Áreas de transcodificación de vídeo a través de canales inalámbricos, "IEEE Diario sobre la selección: en Comunicación*, vol. 18, no. 6, pp. 1063-1074, junio de 2000.

54. T. Shanableh y M. Ghanbari, "transcodificación de vídeo Heterogénea para bajar resoluciones espaciales y temporales y diferentes formatos de codificación, *"IEEE Transactions en Multimedia*, vol. 2, no. 2, pp. 101-110, junio de 2000.
55. M. Ghanbari, "la codificación por capas", en *vídeo comprimido a través de redes*, M.-T. Sol y AR Reibman, Eds. Marcel Dekker, 2001, pp. 251-308.
56. MPEG-2, "Codificación genérica de imágenes en movimiento e información de audio asociada" ISO / IEC 13818-2, 1994, el proyecto de la norma internacional.
57. F. Pereira y E. Touradj, *MPEG-4 Libro*. Upper Saddle River, Nueva Jersey: Prentice Hall, 2002.
58. M. Reisslein, J. Lassetter, S. Ratnam, O. Lotfallah, F. Fitzek, y S. Panchanathan, "Tráfico y caracterización de la calidad de vídeo codificada escalable: A estudio basado en la huella a gran escala ", la Universidad Estatal de Arizona, Departamento de Ingeniería Eléctrica Ing., Tech. Rep., Diciembre 2002.
59. W. Li, "Visión general de granularidad fina Escalabilidad en MPEG-4 Video Standard" *IEEE Trans. en Circuitos y Sistemas de Video Technology*, vol. 11, no. 3, pp. 301-317, marzo de 2001.
60. ISO IEC JTC 1 / SC29 / WG11 / Tecnología de la información - Codificación genérica de Audio-Visual Objects: Visual ISO / IEC 14496-2 / Amd X, diciembre de 1999.
61. H. Radha, M. van der Schaar, y Y. Chen, "el estándar MPEG-4 de grano fino escalable método para streaming multimedia sobre IP de codificación de vídeo, *"IEEE Transactions en Multimedia*, vol. 3, no. 1, pp. 53-68, marzo 2001.
62. C. Buchner, T. Stockhammer, D. Marpe, G. Blattermann y G. Heising, "Codificación textura Vídeo progresivo", en *Actas del IEEE Conferencia Internacional sobre Acústica, Habla, y Procesamiento de Señales (ICASSP)*, 2001, pp 1813. 1816.

63. H.-C. Huang, C.-N. Wang, y T. Chiang, "Una escalabilidad granularidad fina robusta utilizando fuga predictivo basado en enrejado, " *IEEE Transactions en Circuitos y Sistemas de Video Technology*, vol. 12, no. 6, pp. 372-385, junio de 2002.
64. R. Kalluri y M. Schaar, "escalabilidad de grano fino para vídeo basado en H.26L transmisión ", en *Actas del IEEE Conferencia Internacional sobre Consumidor Electrónica (ICCE)*, 2002, pp. 346-347.
65. E. Lin, C. Podilchuk, A. Jacquin, y E. Delp, "Un códec de vídeo híbrido integrado utilizando información de la capa de base para la codificación de capa de mejora ", en *Actas de IEEE Conferencia Internacional sobre Procesamiento de Imágenes (ICIP)*, 2001, pp. 1005 1008.
66. A. Luthra, R. Gandhi, K. Panusopone, K. McKoen, D. Baylon, y L. Wang, "El rendimiento de MPEG-4 perfiles utilizados para la transmisión de vídeo", en *Actas de Taller y Exposición sobre MPEG-4*, 2001, pp. 103-106.
67. S. Parthasaraty y H. Radha, "métodos de control de velocidad óptimas para granularidad fina vídeo escalable ", en *Actas del IEEE Conferencia Internacional sobre Procesamiento de Imágenes (ICIP)*, vol. 2, Barcelona, España, septiembre de 2003, pp. 805-808.
68. R. Rajendran, M. Schaart, y S.-F. Chang, "FGS +: la optimización de la SNRtemporal conjunta calidad de vídeo en MPEG-4 de grano fino codificación escalable ", en *Actas de IEEE Simposio Internacional de Circuitos y Sistemas (ISCAS)*, 2002, pp. 445-448. La codificación de vídeo escalable

69. AR Reibman, U. Bottou, y A. Basso, basado Dct " con la deriva ", en *Actas del IEEE Conferencia Internacional sobre Procesamiento de Imágenes (ICIP)*, vol. 2, Salónica, Grecia, 10 2001, pp. 989-992.
70. M. Schaar y Y.-T. Lin, "mejora selectiva basado en contenido para streaming video ", en *Actas del IEEE Conferencia Internacional sobre Procesamiento de Imágenes (ICIP)*, 2001, pp. 977-980.
71. M. Schaar y H. Radha, "A-SNR temporal escalabilidad granular fino híbrido de vídeo de Internet, "*IEEE Transactions en Circuitos y Sistemas para Video Tecnología*, vol. 11, no. 3, pp. 318-331, marzo 2001.
72. P. Wang, Z. Xiong, F. Wu, y S. Li, "asignación de velocidad óptima para la progresiva granularidad fina de codificación de vídeo escalable, "*IEEE Signal Processing Letters*, vol. 9, no. 2, pp. 33-39, Feb. de 2002.
73. F. Wu, S. Li, y Y.-P. Zhang, "Un marco para la multa progresiva eficiente granularidad de codificación de vídeo escalable, "*IEEE Transactions on Circuitos y Sistemas de Tecnología de Video*, vol. 11, no. 3, pp. 332-344, marzo 2001.
74. X. Zhao, Y. Él, S. Yang y Y.Zhong ", " asignación de velocidad de la calidad de imagen igual MPEG-4 FGS streaming de vídeo ", en *Actas de Packet Video Workshop*, Pittsburgh, PA, abril 2002.
75. ISO / IEC JTC 1 / SC29 / WG11 N4791, "Informe sobre MPEG-4 visual granularidad fina Herramientas escalabilidad pruebas de verificación, "Mayo de 2002.
76. T. Kim y M. Ammar, "Adaptación óptima calidad MPEG-4 de grano fino vídeo escalable ", en *Actas del IEEE Infocom*, San Francisco, CA, abril de 2003, pp. 641-651.
77. P. Zhang, W. Zhu, y Y.-P. Zhang, "La asignación de recursos para multimedia streaming a través de Internet, "*IEEE Transactions on Multimedia*, vol. 3, no. 3,

pp. 339-355, Sept. de 2001.

78. D. Wu, YT Hou, W. Zhu, Y.-P. Zhang, y JM Peña, "video streaming a través de Internet: Enfoques y direcciones, *"IEEE Transactions on Circuits y Sistemas de Video Technology*, vol. 11, no. 3, pp. 1-20, marzo 2001.

79. V. Goyal, "descripción múltiple de codificación: La compresión se reúne la red"
IEEE Revista Procesamiento de Señales, vol. 18, pp. 74-91, Sept. de 2001.

80. S. Ekmekci y T. Sikora, "desequilibrado cuantificado múltiple descripción del video transmisión utilizando diversidad de trayectoria, "en *IS & T / SPIE de Electronic Imaging 2003*, 2003, santa Clara, CA.

81. N. Gogate, DM Chung, SS Panwar, y Y. Wang, "imagen de apoyo y aplicaciones de vídeo en un entorno de radio de saltos múltiples utilizando diversidad de trayectoria y múltiple Descripción codificación, *"IEEE Transactions on Circuitos y Sistemas para Video Technology*, vol. 12, no. 9, pp. 777-792, Sept. de 2002. Apoyo a las aplicaciones de vídeo / imagen en un móvil

82. N. Gogate y SS Panwar " multihop entorno de radio utilizando diversidad de rutas ", en *Actas del IEEE Conferencia Internacional sobre Comunicaciones*, vol. 3, Vancouver, Canadá, junio 1999, pp. 1701-1706.

83. Z. Ji, Q. Zhang, W. Zhu, J. Lu, y Y.-P. Zhang, "Video de radiodifusión sobre MIMO-OFDM sistemas ", en *Actas del Simposio Internacional IEEE sobre Circuitos y Sistemas*, vol. 2, mayo de 2003, pp. 844-847.

84. Y. Altunbasak, N. Karam, y RM Mersereau, "Multiple Descripción codificación con múltiples antenas de transmisión y recepción para los canales inalámbricos: El caso de modulación digital ", en *Actas del IEEE Conferencia Mundial de Telecomunicaciones (Globecom)*, vol. 6, San Antonio, TX, noviembre 2001, pp. 3.272 hasta 3276.

85. S. Lin, Y.Wang, S.Mao, y S. Panwar, "el transporte de vídeo a través de redes ad-hoc el uso de múltiples caminos ", en *Actas del Simposio Internacional IEEE sobre Circuitos y Sistemas (ISCAS)*, vol. 1, Scottsdale, AZ, mayo de 2002, pp. 57-60.
86. M. Pereira, M. Antonini, y M. Barlaud, "múltiple codificación descripción del video para UMTS ", en *Actas del Simposio de codificación de imagen*, Saint Malo, Francia, 04 2003.
87. R. Puri, KW Lee, K. Ramchandran, y V. Bharghavan, "Un enfoque integrado transcodificación de origen y la congestión del paradigma de control para la transmisión de vídeo en el Internet, "*IEEE Transactions on Multimedia*, vol. 3, no. 1, pp. 18-32, abril 2001.
88. JG Apostolopoulos, W. Tan, y Wee SJ, "Realización de un múltiplo Descripción streaming de medios de comunicación de la red de distribución de contenidos ", en *Actas del IEEE Conferencia Internacional sobre Procesamiento de Imágenes (ICIP)*, vol. 2, Rochester, NY, 09 2002, pp. 189-192.
89. FH Fitzek, H. Yomo, P. Popovski, R. Prasad, y M. Katz, "descriptor Fuente esquemas de selección para la descripción múltiples servicios codificados en comunicación inalámbrica 4G sistemas ", en *Actas de la Primera Internacional Taller IEEE en Sistemas Multimedia y Redes (WMSN05) en conjunto con El 24a IEEE Internacional de Performance Computing and Communications Conference (IPCCC 2005)*, Phoenix, AZ, abril de 2005.
90. R. Prasad y S. Hara, *Técnicas Multicarrier para 4G Mobile Communications*, ser. Universal de comunicaciones personales. Artech House, 2003.
91. DS Taubman y M. Marcelino, *JPEG2000 imagen fundamentos de compresión normas y prácticas*. Kluwer Academic Publishers, julio de 2002.

92. S.-T. Hsiang y J. Woods, "video incrustado de codificación usando un movimiento invertible compensado 3-d banco de filtros de subbanda / wavelet, "*proceso de señales: Imagen de Comunicaciones*, vol. 16, pp. 705-724, mayo de 2001.

93. S.-T. Hsiang, "Altamente escalable subbanda imagen / wavelet y la codificación de vídeo," Ph.D. disertación, el Instituto Politécnico Rensselaer, Nueva York, mayo de 2002.

94. S.-T. Hsiang y JW Woods, "imagen Embedded codificación utilizando zeroblocks de coeficientes subbanda / wavelet y modelado contexto ", en *Actas del IEEE Simposio Internacional de Circuitos y Sistemas (ISCAS)*, vol. 3, Ginebra, Suiza, mayo de 2000, pp. 662-665.

95. G. Sullivan, T. Wiegand, y T. Stockhammer, "Uso del vídeo proyecto H.26L estándar para las aplicaciones móviles de codificación, "en *IEEE Conferencia Internacional sobre Procesamiento de Imágenes (ICIP)*, vol. 3, Salónica, Grecia, 10 2001, pp. 573-576.

Ley

96. AM y WD Kelton, *Simulación, Modelado y análisis*, 3ª ed. McGraw Hill, 2000.

97. C. Chatfield, *el análisis de series de tiempo: Un Intoduction*, 4ª ed. Chapman y Hall, 1989.

98. G. Box y G. Jenkins, *Análisis de Series: Previsión y control*. Holden-Day, 1976.

99. J. Beran, *Estadística para procesos de larga memoria*, J. Beran, Ed. London: Chapman y Hall, 1994.

100. J. Beran, R. Sherman, MS Taqqu, y W. Willinger, "la dependencia de largo alcance en el tráfico de video-velocidad de bits variable, "*IEEE Transactions on Communications*, vol. 43, no. 2/3/4, pp. 1566-1579, febrero / marzo / abril de 1995.

- 101.** M. Krunz, "Sobre las limitaciones de la prueba en tiempo variante para la inferencia de largo dependencia de rango ", en *Actas del IEEE Infocom 2001*, Anchorage, Alaska, Abril 2001, pp. 1254-1211 260.
- 102.** BB Mandelbrot y MS Taqqu, "análisis de R / S robusto de serie a largo plazo correlaciones ", en *Actas del 42° período de sesiones ISI, vol. XLVIII, Libro 2*, 1979, pp. 69-99.
- 103.** H. Hurst, "a largo plazo la capacidad de almacenamiento de los embalses", *Proc. Sociedad Americana de Ingeniería Civil*, vol. 76, no. 11, 1950.
- 104.** Leland, MS Taqq, W. Willinger, y DV Wilson, "Por autosimilaridad naturaleza del tráfico Ethernet "en *SIGCOMM ACM*, DP Sidhu, Ed., San Francisco, California, 1993, pp. 183-193.
- 105.** D. Veitch y P. Abry, "Un estimador conjunto basado wavelet de la parámetros de la dependencia a largo plazo ", *IEEE Transactions on Información Teoría*, vol. 45, no. 3, pp. 878-897, abril de 1999. [En línea]. Disponible: <http://www.emulab.ee.mu.oz.au/~darryl>
- 106.** P. Abry y D. Veitch, "Análisis Wavelet del tráfico de largo alcance-dependiente" *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 2-15, enero de 1998.
- 107.** P. Abry, D. Veitch, y P. Flandrin, "Long-range-dependencia: agregación Revisitando con ondas, " *Diario de Análisis de Series*, vol. 19, no. 3, pp. 253-266, mayo de 1998.
- 108.** M. Roughan, D. Veitch, y P. Abry, "la estimación en tiempo real de los parámetros de dependencia a largo plazo, " *IEEE / ACM Transactions on Networking*, vol. 8, no. 4, pp. 467-478, Agosto de 2000.
- 109.** P. Abry, P. Flandrin, MS Taqqu y D. Veitch, *Auto Similar tráfico de red Análisis y Evaluación del Desempeño*. Nueva York, Nueva York: Wiley, 2000, cap. Wavelets para el análisis, estimación y síntesis de los datos de escala, pp. 39-88.

110. -, *la dependencia de largo alcance: Teoría y Aplicaciones*. Birkhauser, 2002, ch. Auto-similaridad y de largo alcance dependencia a través de la lente de onda.
111. A. Feldmann, AC Gilbert, W. Willinger, y T. Kurtz, "La naturaleza cambiante de tráfico de la red: expansión fenómenos, " *Computer communication Review*, vol. 28, no. 2, abril de 1998.
112. J. Gao y I. Rubin, "modelado multifractal multiplicativo de largo de la GAMA tráfico de la red dependiente ", *Revista Internacional de Sistemas de Comunicación*, vol. 14, pp. 783-801, 2001.
113. AC Gilbert, W. Willinger, y A. Feldmann, "Ampliación de análisis conservador cascadas, con aplicaciones a la red de tráfico, " *IEEE Transactions on Teoría de la Información*, vol. 45, no. 3, pp. 971-991, abril de 1999.
114. RH Riedi, MS Crouse, VJ Ribeiro, y RG Baraniuk, "Un multifractal modelo wavelet con aplicaciones a la red de tráfico, " *IEEE Transactions on Teoría de la Información*, vol. 45, no. 3, pp. 992-1018, abril de 1999.
115. J. Walter, "bttvgrab." [En línea]. Disponible: <http://www.garni.ch/bttvgrab/>
116. F. Bellard ", FFMPEG - sistema multimedia ver. 0.4.9-pre1, "en línea, febrero 2005. [En línea]. Disponible: <http://ffmpeg.sourceforge.net/index.php>
117. I. 14.496, "Video Software de referencia, Microsoft-FDAM1-2.3-001213."
118. Comité TME, "MPEG-2 Video Prueba del Modelo 5, ISO / IEC JTC 1 / SC29WG11 MPEG93 / 457 ", abril 1993.
119. Q. Zhang, W. Zhu, y Y.-P. Zhang, "La asignación de recursos para multimedia streaming a través de Internet, " *IEEE Transactions on Multimedia*, vol. 3, no. 3, pp. 339-355, Sept. de 2001.
120. C. Suehring, "coordinación software H.26L" <http://bs.hhi.de/suehring/tml/>.

121. FH Fitzek y M. Reisslein, "MPEG-4 y vídeo H.263 rastros de rendimiento de la red evaluación ", Universidad Técnica de Berlín, Tech. Rep., 2000, tKN-00-06.