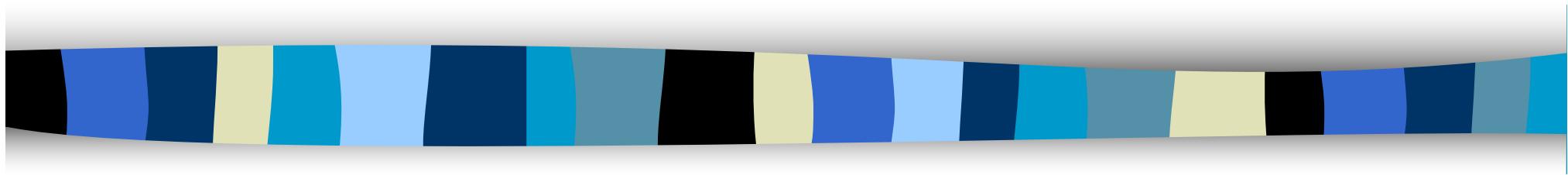


Video Compresión MPEG



Formatos contenedores y QoA

¿De dónde vienen los formatos contenedores?

Evolución del vídeo coding

■ Film

- Creado en el siglo 19th, y todavía se usa en la actualidad



– VHS

- Creado en 1976, ha desaparecido



Evolución de los medios de video

■ DVD

- Creada en 1996, dominante durante una década



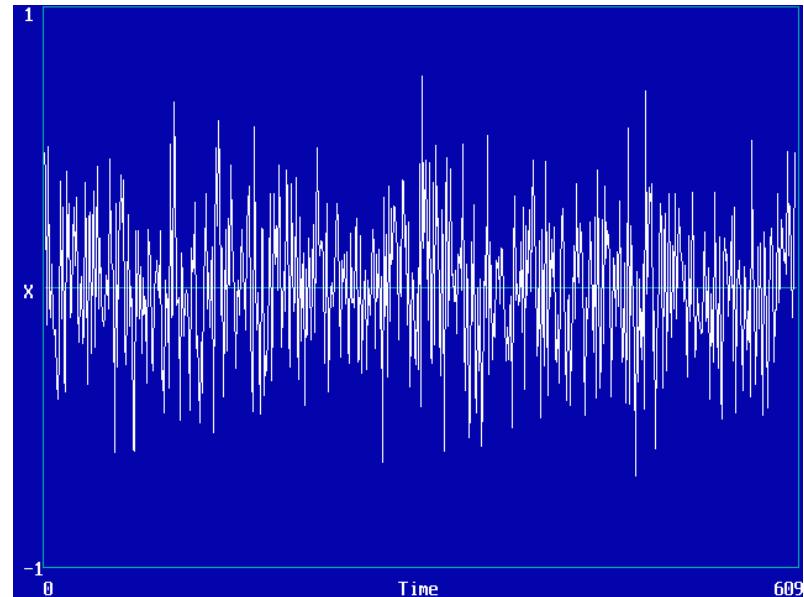
■ Hard Disk

- Diseñado para muchos años, sólo últimamente se ha utilizado para almacenamiento de vídeo (ayudado por el desarrollo de Internet)



Transición de medios analógicos a digitales

- El “N word”
 - Señales analógicas son muy dependientes del tiempo en su degradación
- Económicos
 - Medios ópticos son más baratos que magnéticos
- Se necesita convertir formatos analógicos a digitales



Digitalizar el vídeo

- Nuevas cámaras tienen hardware on-board para capturar directamente formatos digitales
- Viejas películas se pueden leer con máquinas especiales





Video Encoding/Compression

- Once video is in digital format, it makes sense to compress it
- Similarly to image compression, we want to store video data as efficiently as possible
- Again, we want to both maximize quality and minimize storage space and processing resources
- This time, we can exploit correlation in both space and time domains



TMI! (Too Much Information)

- A diferencia de la codificación de imágenes, codificación de video se realiza con poca frecuencia en forma de pérdidas
- No hay medio de almacenamiento con capacidad suficiente para almacenar un archivo de tamaño práctico para vídeo sin pérdida
 - Lossless DVD video - 221 Mbps
 - Compressed DVD video - 4 Mbps
 - 50:1 compression ratio!



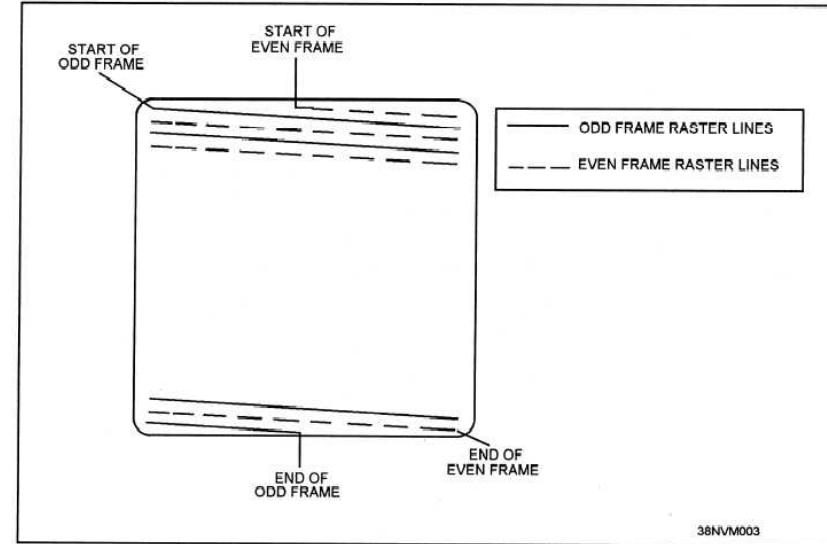
Definiciones

- Bitrate
 - Información almacenada/transmitida por unidad de tiempo
 - Normalmente medida en Mbps (Megabits por segundo)
 - Tasa entre < 1 Mbps y > 40 Mbps
- Resolución
 - Números de pixels por imagen
 - Tasas desde 160x120 a 1920x1080
- FPS (frames per second=imágenes por segundo)
 - Usualmente 24, 25
 - No necesitamos más, a causa de limitaciones del ojo humano

Tipos de escaneo

■ Escaneo entrelazado

- Líneas pares e impares aparecen en marcos alternativos
Inicialmente utilizada para ahorrar ancho de banda en la transmisión de TV
- Cuando se muestra el vídeo entrelazado en una pantalla de exploración progresiva, puede ver el "efecto peine".



Scan types

- Escaneo progresivo
 - Mostrar todas las líneas en cada cuadro
 - Nueva "resolución fija" se muestra (por ejemplo, LCD, Plasma) todos los análisis el uso progresivo
 - Desentrelazado no es una tarea trivial (duplicado de líneas)



Barrido progresivo

Usado en: cámaras Axis como la AXIS 210



[View Full size 640x480](#)

Detalles:



Barrido entrelazado

Usado en: cámaras analógicas de CCTV



[View Full size 704x576](#)

Detalles:



2CIF (con "Duplicado de líneas")

Usado en: DVRs

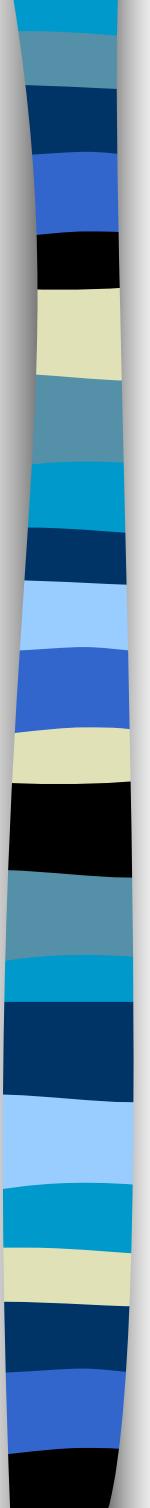


[View Full size 704x576](#)

Detalles:



Nota: En estos ejemplos, las cámaras utilizaron el mismo objetivo. El coche iba a una velocidad de 20 km/h (15 mph).



MPEG (Moving Pictures Expert Group)

- Comité de expertos que desarrolla estándares de codificación de vídeo
- Hasta hace poco, fue el único sistema válido (siendo el más popular, por el momento)
- Adecuado para una amplia gama de videos
 - De baja resolución a alta resolución
 - Movimiento lento de acción rápida
- Se puede implementar en software o hardware



Evolución de MPEG

■ MPEG-1

- Primer estándard de compresión de audio/video
- Usado por VCD's
- MP3 = MPEG-1 audio layer 3
- Objetivo de 1.5 Mb/s bitrate a resolución 352x240
- Sólo soporta imágenes progresivas



Evolución de MPEG

■ MPEG-2

- Actual estándar de facto, ampliamente utilizado en DVD y televisión digital
- Ubicuidad en el hardware implica que va a estar aquí durante mucho tiempo
- Transición a la televisión de alta definición ha tomado más de 10 años y no ha terminado aún
- Diferentes perfiles y niveles de permitir el control de calidad



Evolución de MPEG

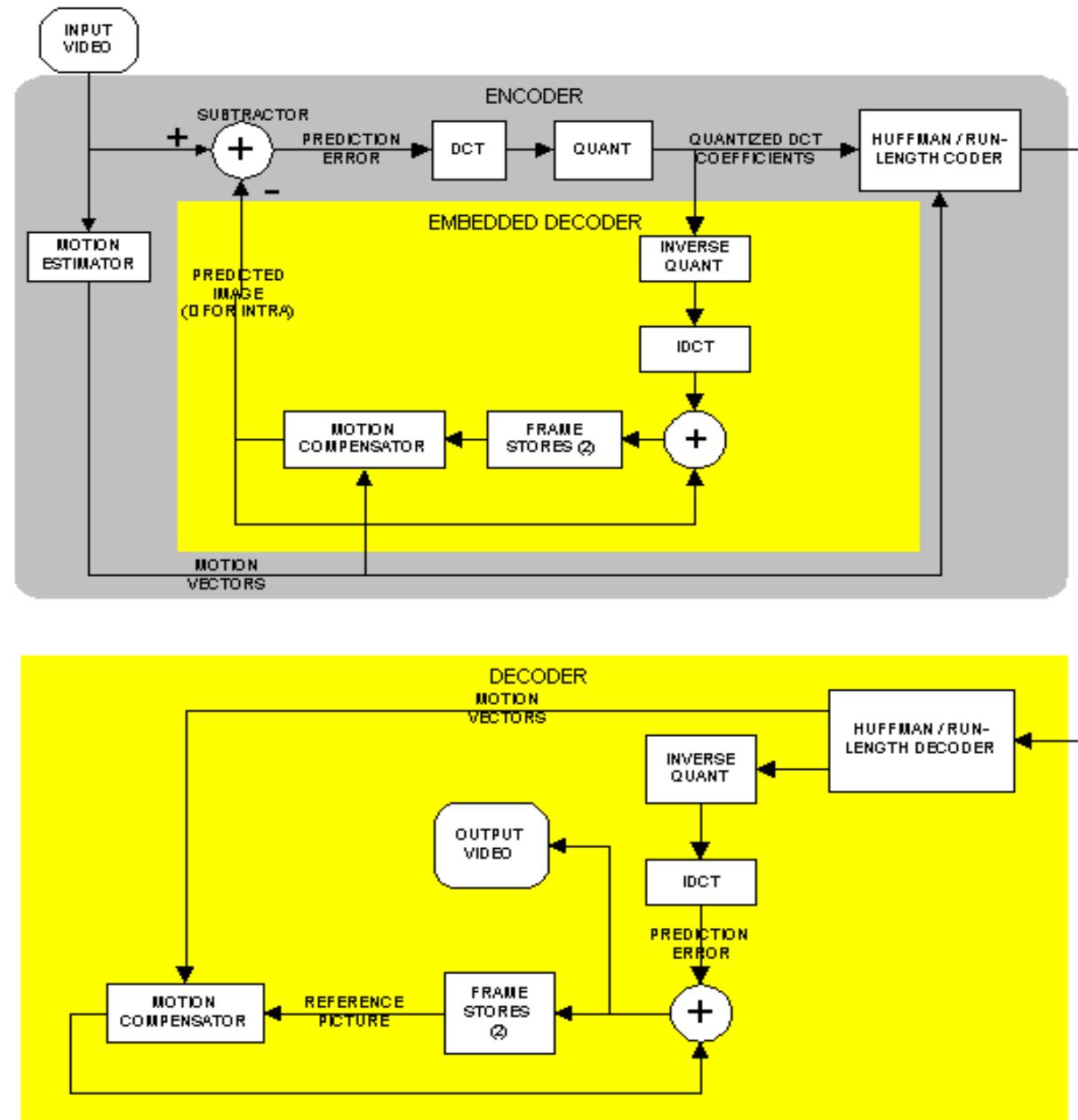
■ MPEG-3

- Originalmente desarrollado para la televisión de alta definición, pero abandonado cuando MPEG-2 se determinó que era suficiente

■ MPEG-4

- Incluye soporte para AV "objetos", contenido 3D, codificación de baja tasa de bits, y DRM
- En la práctica, ofrece la misma calidad que MPEG-2 a una menor tasa de bits, pero a menudo no ofrecen una calidad francamente mejor
- MPEG-4 Parte 10 es H.264, que se utiliza en HD-DVD y Blu-Ray

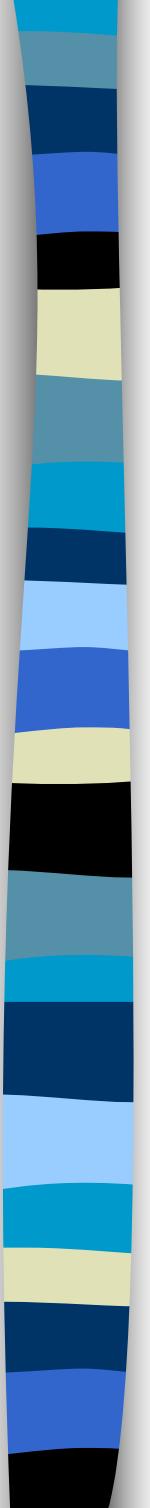
MPEG Block Diagram





MPEG especificación técnica

- Parte 1 - Sistemas - describe la sincronización y la multiplexación de vídeo y audio.
- Parte 2 - Video - codec de compresión de señales de vídeo entrelazado y no entrelazado.
- Parte 3 - Audio - codec de compresión para la codificación perceptual de las señales de audio. Una extensión multicanal con capacidad de audio MPEG-1.
- Parte 4 - Describe los procedimientos para pruebas de conformidad.
- Parte 5 - Describe los sistemas de simulación de software.
- Parte 6 - Describe extensiones para DSM-CC (Digital Storage Media Command and Control).
- Parte 7 - Advanced Audio Coding (AAC)
- Parte 8 – Suprimido
- Parte 9 - Extensión de las interfaces en tiempo real.
- Parte 10 - Extensiones de Conformidad para el DSM-CC.

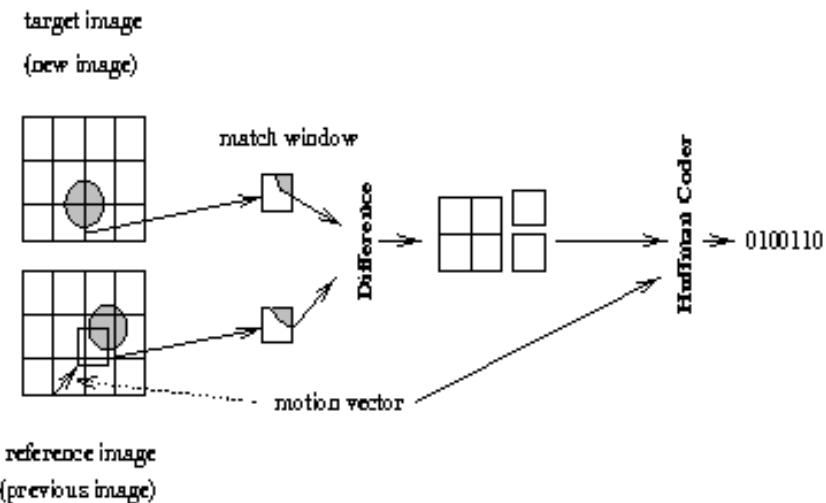
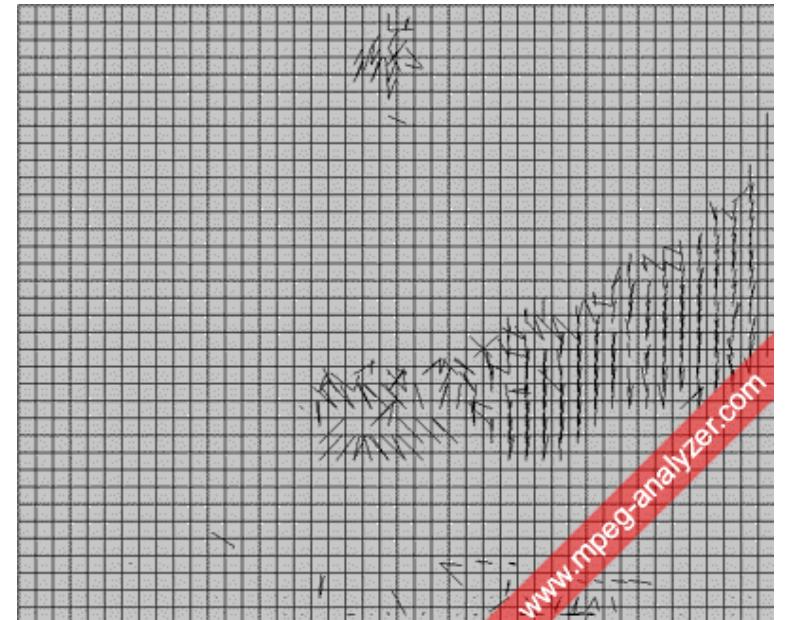


MPEG Procesado de video en el dominio espacial

- Dominio espacial muy parecido a JPEG
 - Convertir RGB a valores de color YUV
 - Dividir marco en bloques de 8x8
 - 2-D DCT de cada bloque
 - Cuantificación de los coeficientes DCT
 - Ejecutar codificación de longitud y entropía

MPEG Procesado en el dominio temporal

- Sistema totalmente nuevo (este concepto no existe en formato JPEG)
- Idea general - Usar vectores de movimiento para especificar cómo se traduce un macrobloque de 16x16 entre la imagen de referencia y la imagen actual, entonces se guarda/transmite la diferencia





Tipos de imágenes

- Imagen I (intra-coded)
 - Codificadas sin referencias a otras imágenes
- Imagen P (predictive-coded)
 - Codificadas con referencias a otras imágenes (bien I o P)
 - Tamaño es habitualmente 1/3 de una I
- Imagen B (bi-directional predictive-coded)
 - Codificadas con referencia a imágenes anteriores y posteriores (bien I o P)
 - Tamaño es habitualmente 1/6 de una I

GOP (Grupo de Imágenes)

- GOP es un conjunto de imágenes consecutivas que pueden ser decodificados sin ningún tipo de imagen de referencia a otras
- Por lo general 12 o 15 imágenes
- Secuencia de transmisión no es la misma que la secuencia de muestra
- Acceso aleatorio a los medios del flujo - Comenzar con I-frame

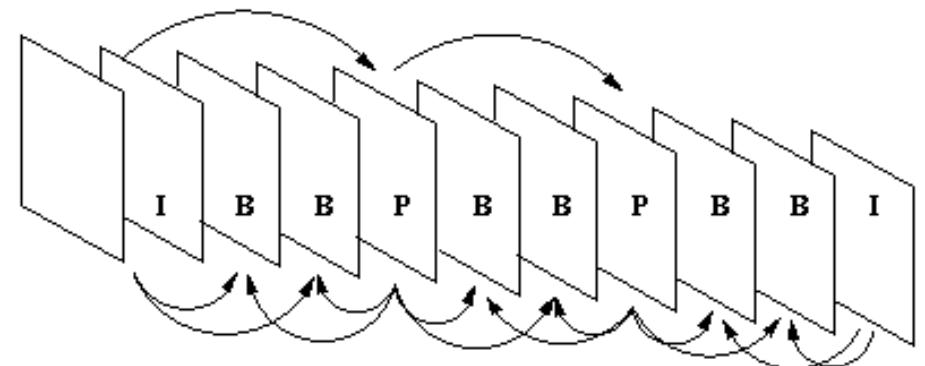


Figure 1: Prediction between MPEG-2 Frames



Datos sobre “predicciones”

- Sólo el uso de vectores de movimiento, si se puede encontrar una marca de "cierre"
 - Evaluar la "proximidad" con métricas MSE u otros
No se puede buscar todos los bloques posibles, por lo tanto necesitan un algoritmo inteligente
 - Si no encuentra referencia posterior adecuada, sólo el código de la macro-bloque con referencia de una I
 - Si un cambio de escena se detecta, empezar un nuevo GoP
- No quiero P después de más imágenes B consecutivas
 - Error de predicción se mantendrá hasta el próximo marco de propagación
 - Retraso en la descodificación



Selección de tipo de Bitrate

■ CBR – Constant BitRate

- Transmisiones Streaming lo usan
- Fácil de implementar

■ VBR – Variable BitRate

- DVD's lo usa
- Normalmente requiere codificación en dos pasos
- Asignar más bits para escenas complejas
 - Esto vale la pena, porque se supone que se codifica una vez, decodificar muchas veces

Selección de tipo de Bitrate





MPEG audio

- MPEG-1: 3 capas de incrementar la calidad, la capa 3 es la más común (MP3)
 - 16 bits
 - Tasa Samping - 32, 44,1 o 48 kHz
 - Bitrate - 32 a 320 kbps
 - De hecho - 44,1 kHz frecuencia de muestreo, tasa de bits 192 kbps
- MPEG-2 - Compatible con > 2 canales, las frecuencias más bajas de muestreo, la mejora de baja tasa de bits
- AAC (Advanced Audio Coding)
 - Más frecuencias de muestreo (8 kHz a 96 kHz)
 - Una mayor eficiencia de la codificación y la más simple Filterbank
 - 96 kbps AAC suena mejor que 128 kbps MP3
- Por lo general, CBR, VBR, pero se puede hacer

Formato contenedor

Container file format
(.ogg, .mkv, .avi, .mpg, .mov, etc.)





Formato contenedor MPEG

- Formato contenedor es un formato de archivo que puede contener datos comprimidos por los codecs estándar
- 2 tipos de MPEG
 - Programa de Stream (PS) - Diseñado para los medios razonablemente fiables, como los discos
 - Transport Stream (TS) - Creado para enlaces con pérdidas, como las redes broadcasting o antenas de transmisión



AV Synchronization

- ¿Quieres flujos de audio y vídeo que se reproducen en sincronía con los demás
- Secuencia de vídeo contiene "marcas de tiempo de presentación"
 - MPEG-2 reloj funciona a 90 kHz
- Bueno para ambos 25 y 30 fps
- PCR (Programa de referencia de reloj) las marcas de tiempo se envían los datos por el remitente
- El receptor utiliza PLL (Phase Lock Loop) para sincronizar los relojes

Codificación de video en tiempo real

- Estimación de movimiento va a ser peor, por lo tanto necesitan mayor tasa de bits para compensar
- Muy difícil de hacer en el software, la necesidad de hardware dedicado o asistencia para el hardware
- TiVo, ReplayTV hacer esto





Transmisión del flujo

- Los tipos más comunes son Flash, Real Video, Quicktime
 - Por lo general, tienen ancho de banda disponible, necesidad de optimizar, como tal,
- Los protocolos de red dedicados para este fin:
 - TCP esperará indefinidamente para la retransmisión, por lo que a menudo no es adecuada

MPEG flujo de datos

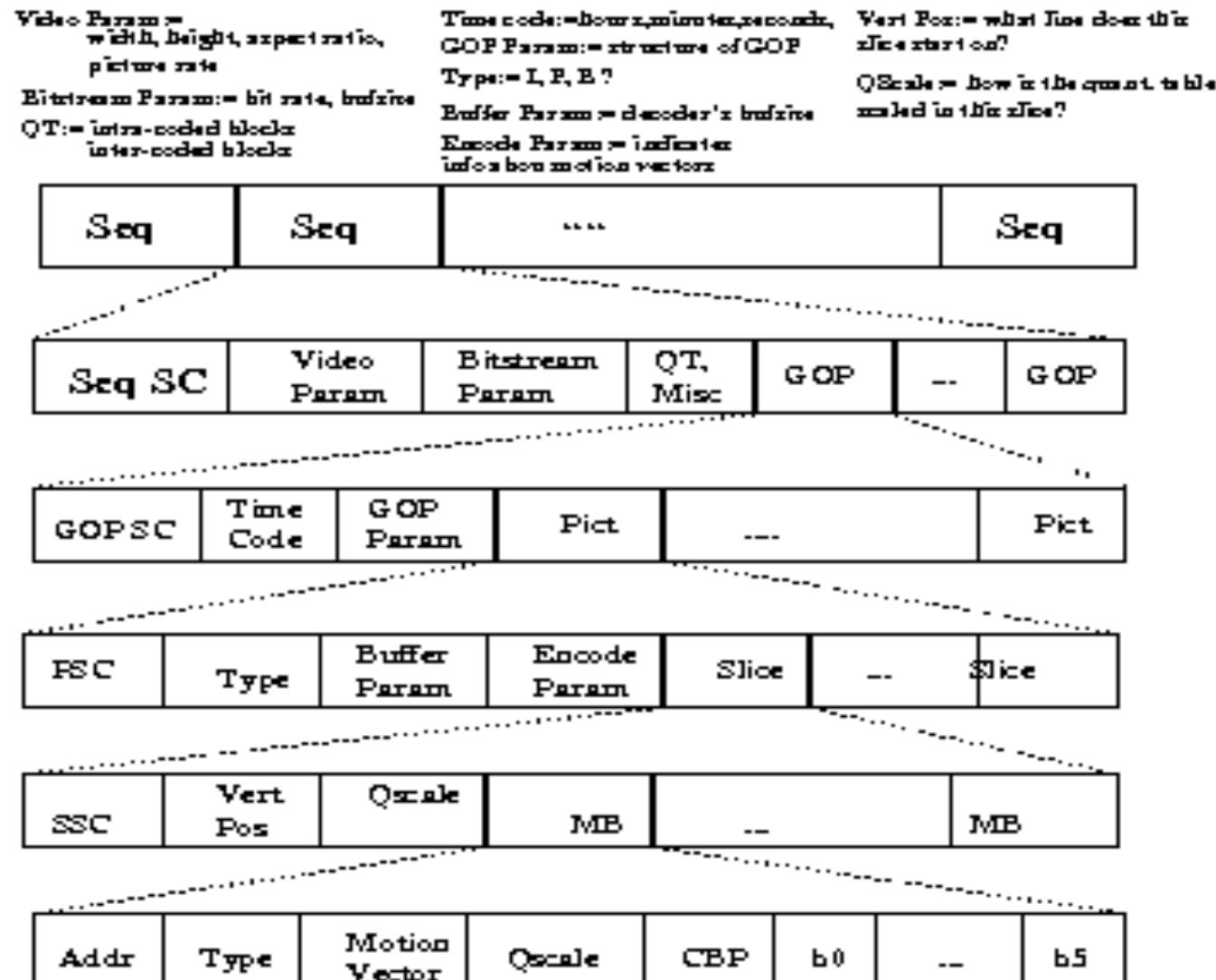
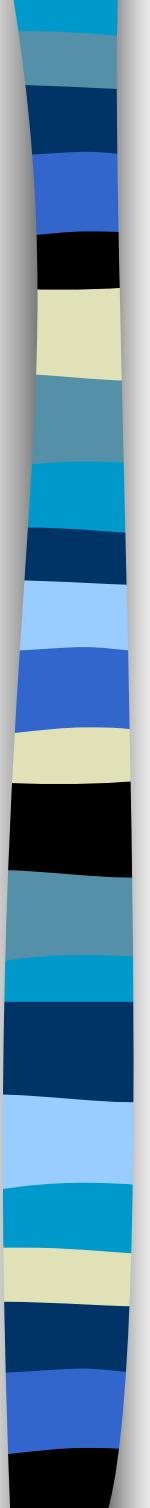


Figure 15: Video Data Stream



Analysis

■ Pros

- Imagen nítida en general
- Audio y vídeo están en sincronía con los demás
- ¿Qué pasaría si la transmisión esta en una red?

■ Cons

- Flashes de imagen, borrosa cuando hay mucho movimiento en la pantalla
 - Mayor tasa de bits a menudo no resuelve este problema

El framework de la ITU para soportar QoS

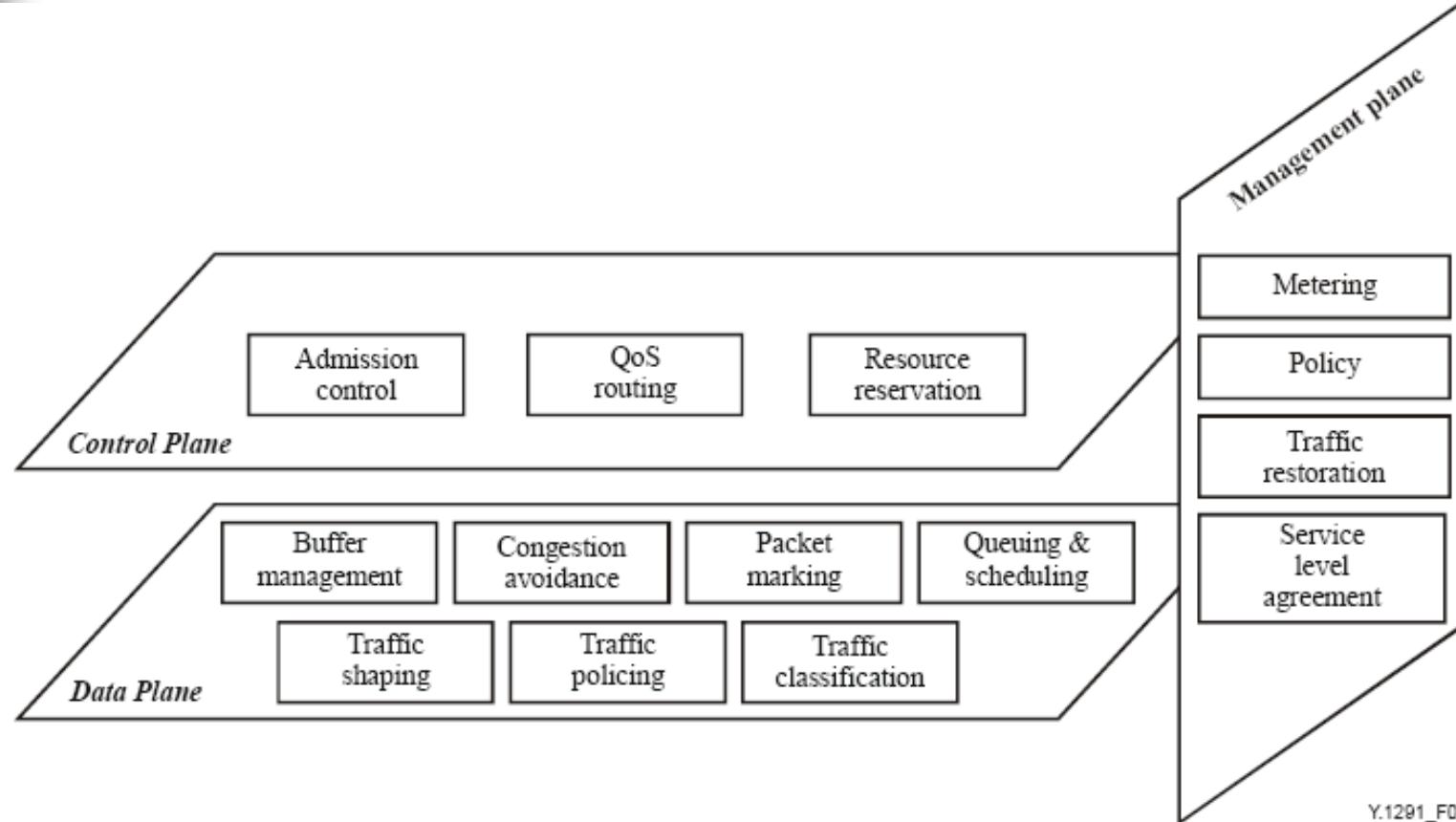
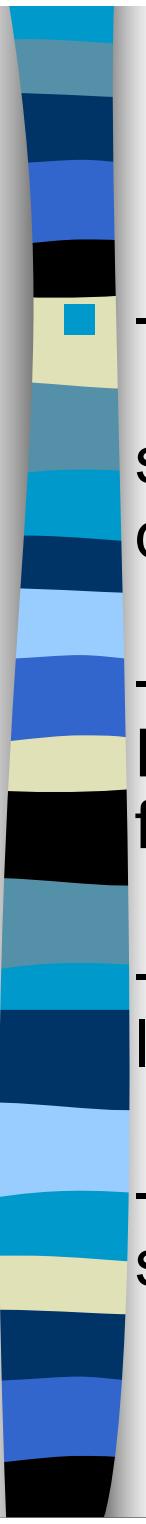


Figure 1/Y.1291 – Architectural framework for QoS support

QoA: Dominios de Evaluación

- Quality of Application(QoA)/Experience (QoE): Caracterización de la accesibilidad a los servicios y la calidad tanto en formas cuantitativas y cualitativas (percepción de los usuarios) considerando distintos parámetros y valores.
- Dominios de evaluación de QoE:
 - **Grado de servicio**
 - **Accesibilidad al servicio**: capacidad de acceder al servicio
 - **Establecimiento de la conexión**: capacidad de conectarse
 - **Transferencia de la información**: Resultado de la calidad de la información
 - **Fiabilidad**: Probabilidad de fallo
 - **Disponibilidad**: Probabilidad de que el servicio esté activo
 - **Supervivencia**: Capacidad de proporcionar servicio en condiciones anormales
 - **Seguridad**: Nivel de protección del sistema e información
 - **Calidad**: Inteligibilidad, audivilidad, visualización de los de los contenidos



Fases de aplicación de la QoA y vistas

- Fases del ciclo de vida útil:
la prestación de servicios, mejora de los servicios, servicios de apoyo, conexión de servicio, la facturación de servicios, gestión de servicios, etc.
- Criterios para la observación de la calidad como:
Disponibilidad, precisión, velocidad, seguridad, fiabilidad, etc.
- Vista del Cliente: los requisitos de calidad de servicio y la percepción
- Vista del Proveedor de servicios: ofrecer calidad de servicio y el logro

Medida de la QoA para servicios de voz

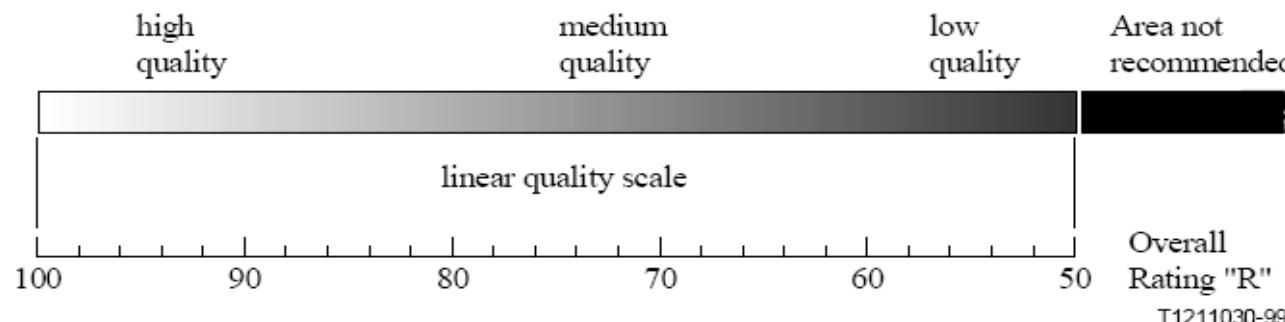
PESQ - (Percentual Evaluation of speech quality). El estándar ITU P.862. Mide la calidad de voz en la percepción del cliente.

PESQ es el estándar de calidad de voz más precisas contra la que muchos otros algoritmos calidad de voz se comparan. Mide los efectos de las distorsiones tales como el ruido y la demora para modelar y predecir la calidad subjetiva. Que produce resultados de medición de calidad de voz basadas en la puntuación de opinión media UIT P.800.1 (MOS) a escala, que sea representativa de las percepciones de los clientes de calidad.

PESQ medidas de extremo a extremo la calidad de voz mediante la comparación de una señal de prueba de entrada con la salida de la señal, y es eficaz en una amplia gama de tipos de redes, incluyendo PSTN, móviles y VoIP.

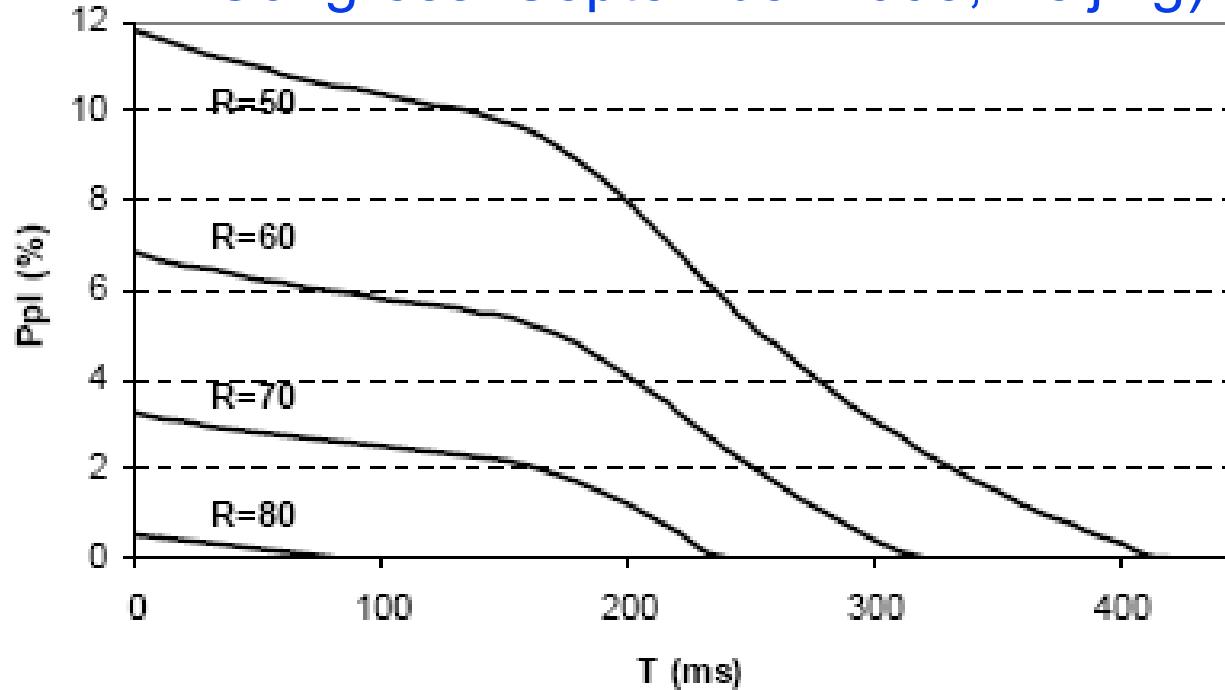
QoA ratios de percepción para servicios de voz

R-value range	Speech transmission quality category	User satisfaction
$90 \leq R < 100$	Best	Very satisfied
$80 \leq R < 90$	High	Satisfied
$70 \leq R < 80$	Medium	Some users dissatisfied
$60 \leq R < 70$	Low	Many users dissatisfied
$50 \leq R < 60$	Poor	Nearly all users dissatisfied



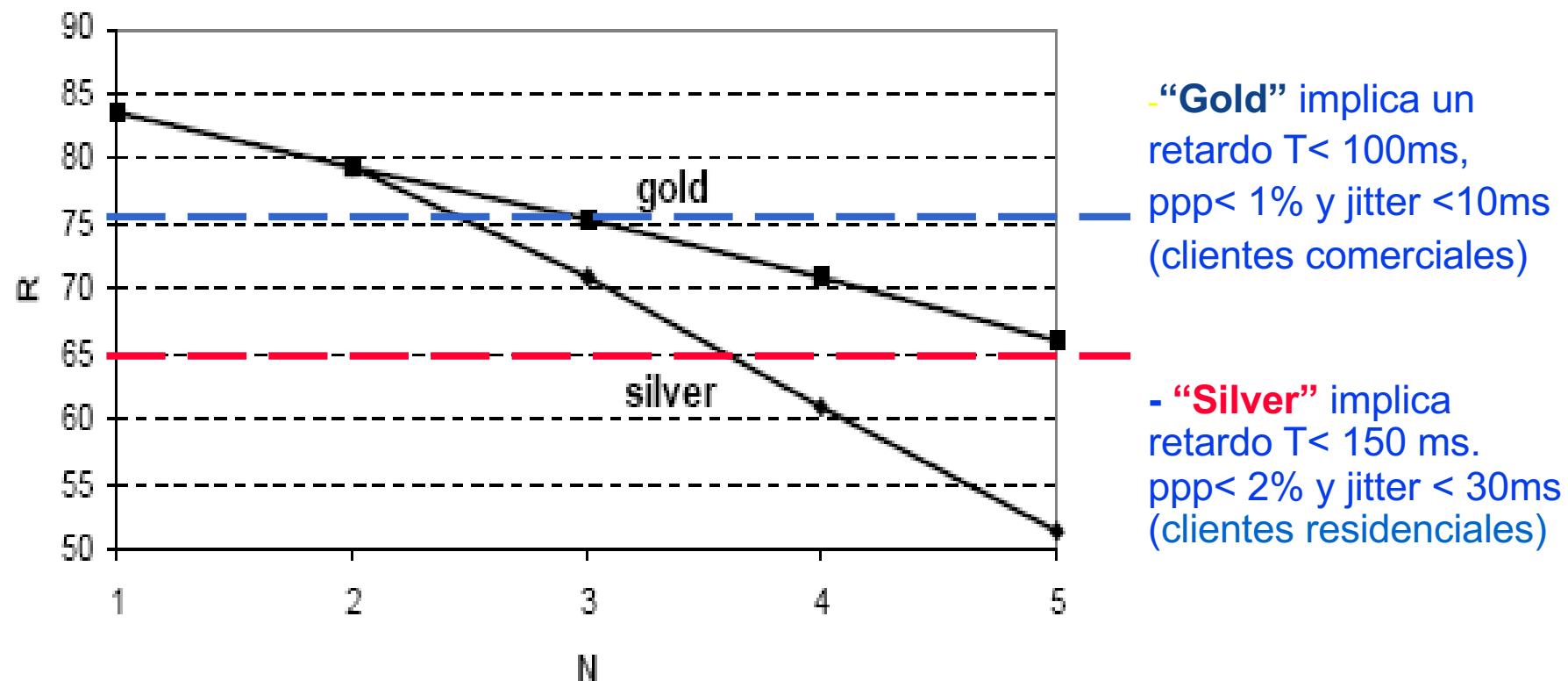
QoA ratios para VoIP como función de pérdidas de paquete y retardos

Las curvas de calidad se pueden presentar como una función de retardo de paquetes y de la probabilidad de pérdidas por ejemplo para G.729 (19th International Teletraffic Congress September 2005, Beijing)



Ratios QoA para VoIP como función de multisaltos cruzados

QoA percibida como función el número de saltos cruzados para el formato G.723 con $\text{ppp} = 0.01$ y gold /silver SLA (19th International Teletraffic Congress September 2005, Beijing)



Ejemplo del patch End-to-End para el cálculo anterior

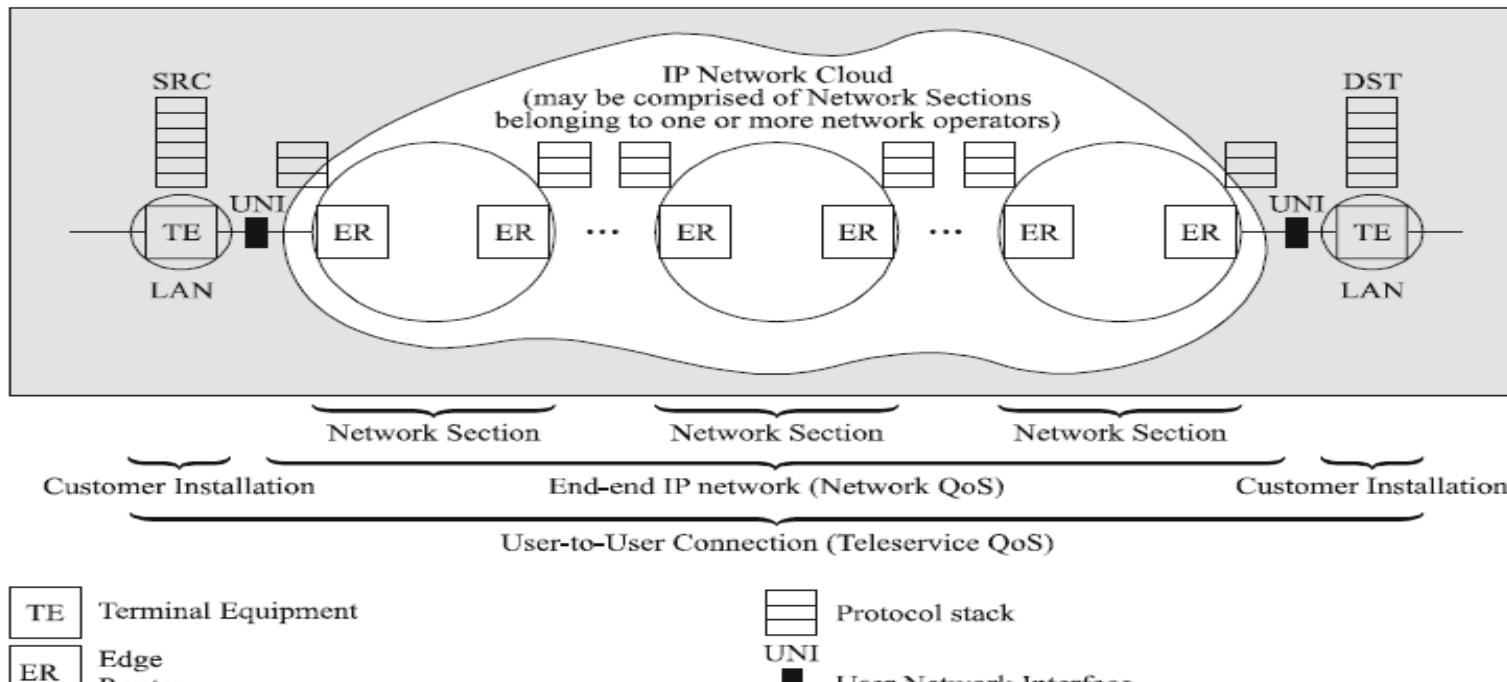


Figure 1/Y.1541 – UNI-to-UNI reference path for network QoS objectives

Puntos de referencia para el análisis del interdominio de la QoA

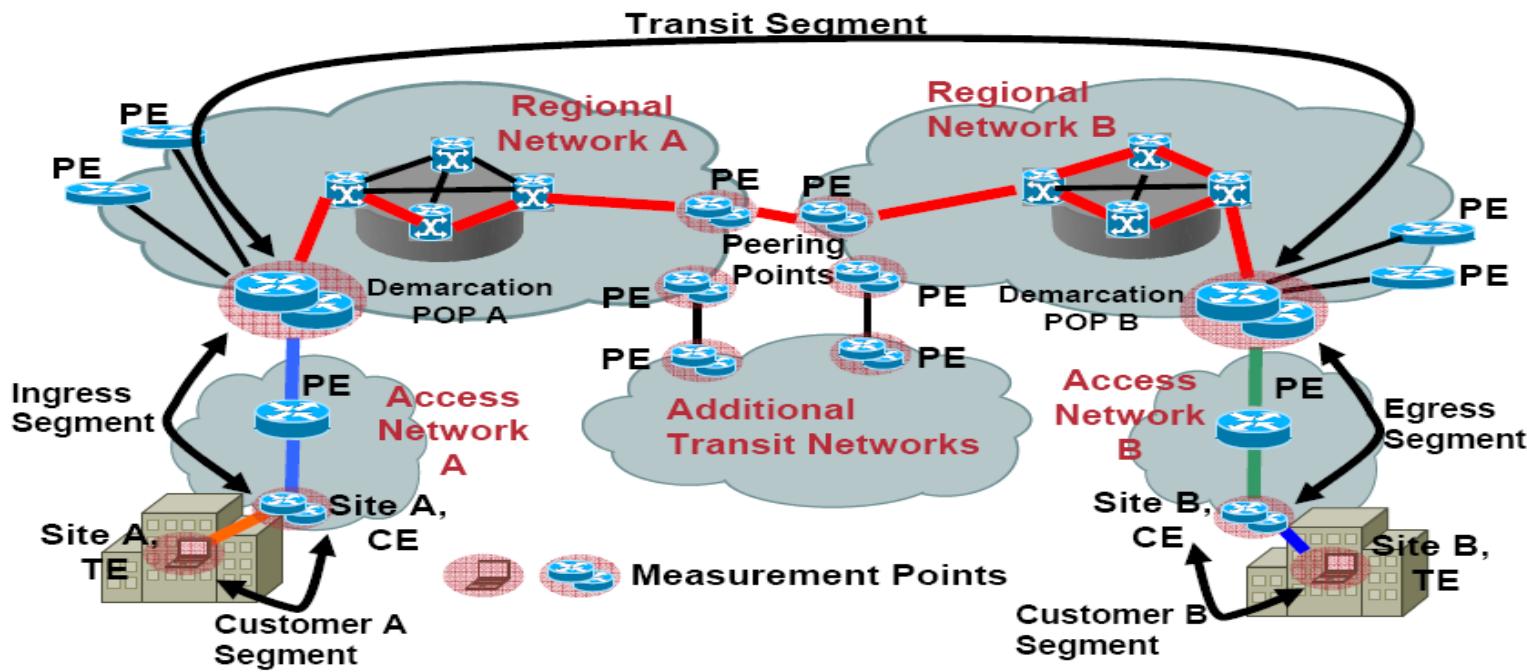


Figure 5/Y.1543 - TE-TE Model

Parámetros QoAy valores para el diseño y planificación de redes

QoS	Prioridad	Probabilidad de pérdida de bit	Probabilidad e pérdida de paquete	Retardo de paquete	Jitter	Disponibilidad
Stream constant	High	<10e-9	<10e-3	<150 ms	<10 ms	>99.999%
Stream Variable	High and medium	<10e-9 <10e-5	<10e-2 <5x10e-2	<150 ms < 400ms	<10 ms <30 ms	>99.999% >99.99%
Elastic	Low	<10e-3	Without guarantee	Without guarantee	Without guarantee	Without guarantee

Guía de referencia de clases de QoS classes para operaciones con redes IP

Table 2/Y.1541 – Guidance for IP QoS classes

QoS class	Applications (examples)	Node mechanisms	Network techniques
0	Real-time, jitter sensitive, high interaction (VoIP, VTC)	Separate queue with preferential servicing, traffic grooming	Constrained routing and distance
1	Real-time, jitter sensitive, interactive (VoIP, VTC).		Less constrained routing and distances
2	Transaction data, highly interactive (Signalling)	Separate queue, drop priority	Constrained routing and distance
3	Transaction data, interactive		Less constrained routing and distances
4	Low loss only (short transactions, bulk data, video streaming)	Long queue, drop priority	Any route/path
5	Traditional applications of default IP networks	Separate queue (lowest priority)	Any route/path

NOTE – Any example application listed in Table 2 could also be used in Class 5 with unspecified performance objectives, as long as the users are willing to accept the level of performance prevalent during their session.

Guía de referencia de clases de QoA para procesado de llamadas

Table IV.1/Y.1530 – Call processing QoS class definitions and performance objectives

Call processing QoS class	Connection set-up parameters			Connection disengagement parameters	
	Connection set-up delay	Connection set- up error probability	Connection set- up failure probability	Connection disconnect delay	CPDP CCFP
QoS Class E (High or Excellent priority level)	Mean < 7500 ms 95%ile < 8450 ms (FFS)	Default (FFS)	Mean < A (FFS)	Defaults Mean = 3500 ms 95%ile = 4250 ms [I.352]	Defaults (FFS)
QoS Class 1 (Ordinary telephone level)	Mean = 7500 ms 95%ile = 8450 ms [I.352]		Mean = A (Value A is FFS) [I.359]		
QoS Class 2 (IP telephone level)	Mean > 7500 ms 95%ile > 8450 ms (FFS)		Mean > A (FFS)		
QoS Class U (Best effort level)	U	U	U	U	U
CPDP Connection Premature Disconnect Probability CCFP Connection Clearing Failure Probability U Unspecified or Unbound FFS For further study					

Guía de referencia de clases de QoA para procesado de llamadas

Table 1/Y.1541 – IP network QoS class definitions and network performance objectives

Network performance parameter	Nature of network performance objective	QoS Classes					
		Class 0	Class 1	Class 2	Class 3	Class 4	Class 5 Unspecified
IPTD	Upper bound on the mean IPTD (Note 1)	100 ms	400 ms	100 ms	400 ms	1 s	U
IPDV	Upper bound on the $1 - 10^{-3}$ quantile of IPTD minus the minimum IPTD (Note 2)	50 ms (Note 3)	50 ms (Note 3)	U	U	U	U
IPLR	Upper bound on the packet loss probability	1×10^{-3} (Note 4)	1×10^{-3} (Note 4)	1×10^{-3}	1×10^{-3}	1×10^{-3}	U
IPER	Upper bound	1×10^{-4} (Note 5)					U



Evaluación de la Calidad de Vídeo

Similitudes con la evaluación de calidad de imagen

- Los ojos más sensibles a la crominancia
- Algunas áreas atraen más atención que otros
- Artefactos de bloqueo / sonido es molesto...

Diferencias

- Dimensión temporal
 - Objetos móviles frente a estáticos
 - Tasa de imagen juega un rol importante
 - Cualquier cosa que se mueva más de 12 veces por segundo es perceptible

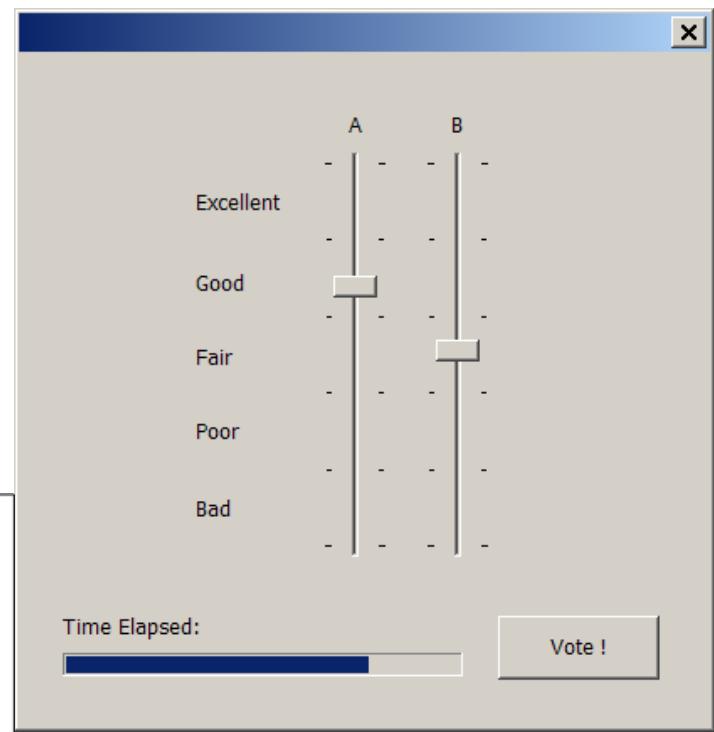
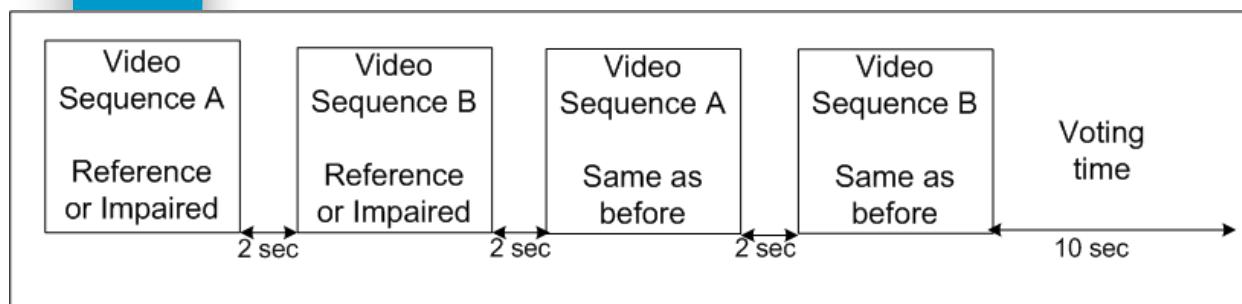
Video QoA

Cálculo Subjetivo

– Un sujeto humano tasa el vídeo en una escala

- Método Escala de calidad continua de doble estímulo

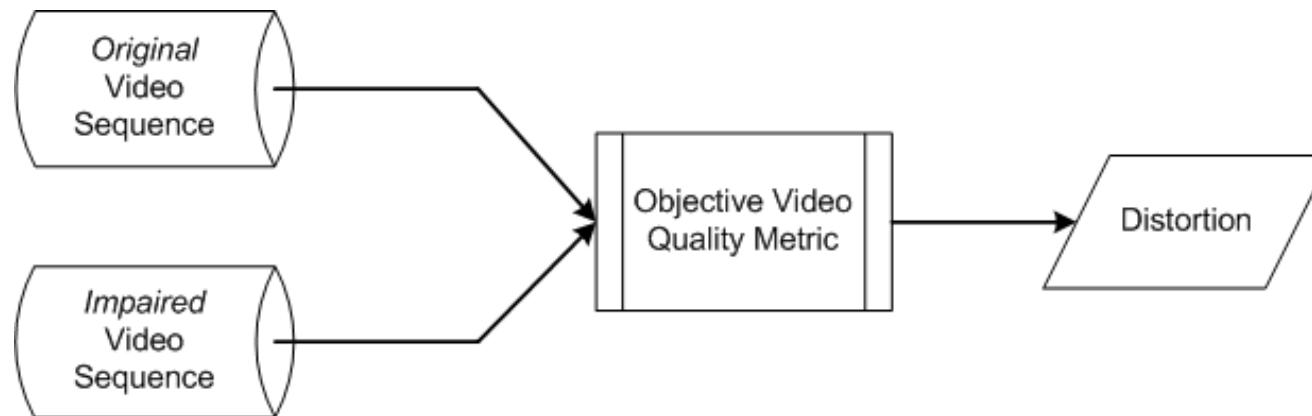
- Escala oculta de 0-100
 - Diferencia es calculada con la actual escala



Video QoA

Objetiva

- Hay algoritmos que juzgan la distorsión entre los videos
- Intenta modelar un observador humano
- Actualmente no hay ningún método estándar



Medida objetiva Video QoA

Peak Signal-To-Noise Ratio (PSNR)

- Ampliamente utilizado para evaluar la codificación
- Diferencia puramente matemática

$$E_{RMS} = \sqrt{\frac{1}{M \times N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [f'(m, n) - f(m, n)]^2}$$

$$PSNR = 20 \times \log_{10} \left(\frac{255}{RMSE} \right)$$

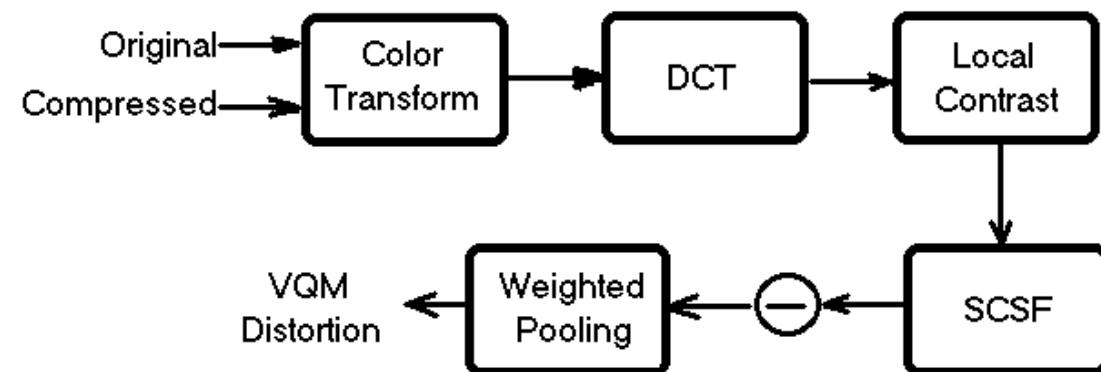
Puede ser engañada fácilmente

Medida objetiva Video QoA

¿Cómo engañar el PSNR?

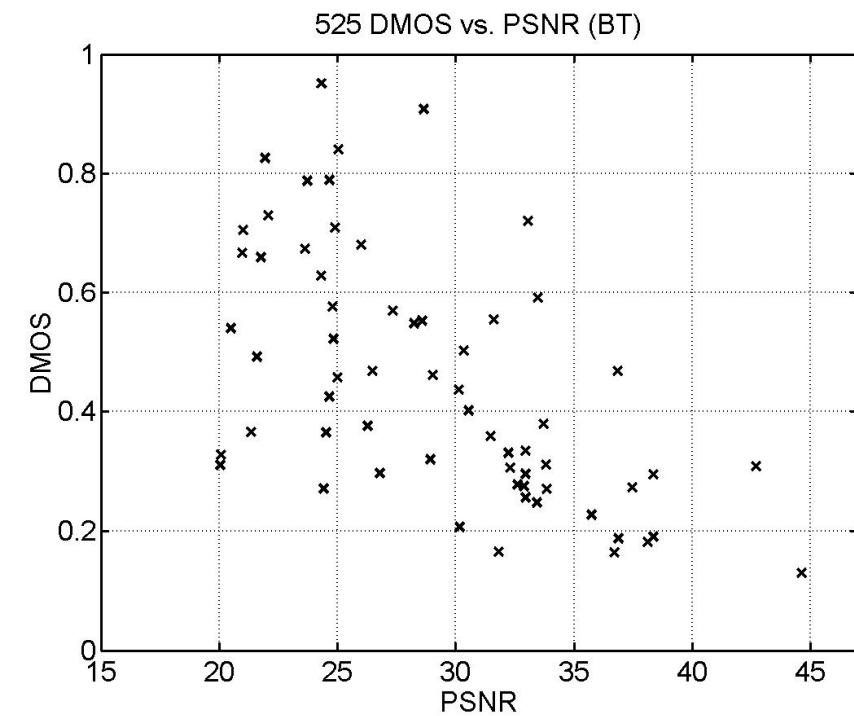
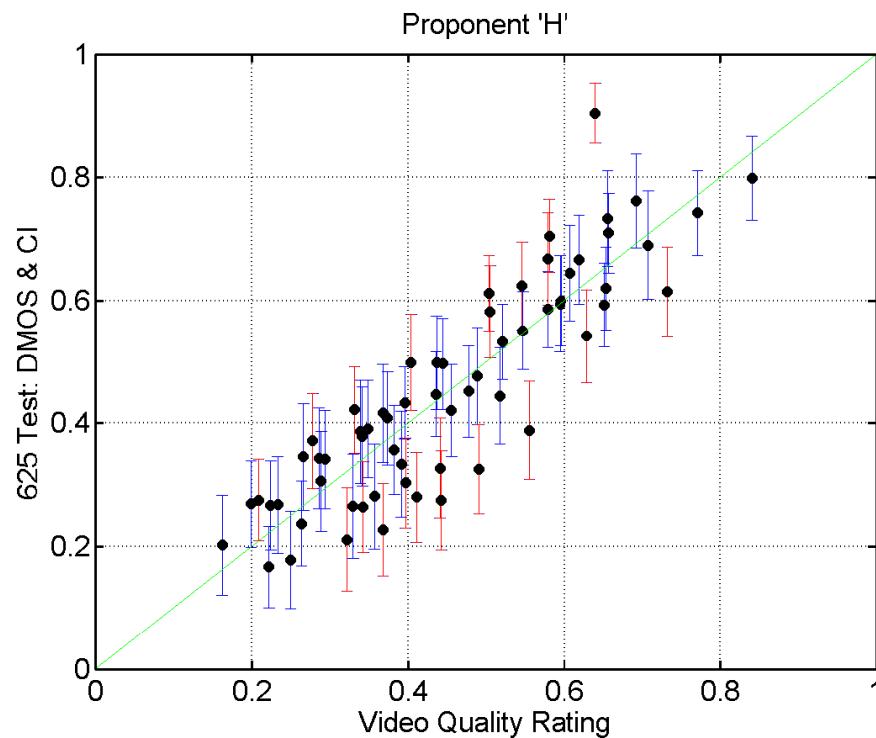
- Tomar una imagen natural
- Dale más bits para las áreas que se mira más
- Dar menos bits a las áreas se miran menos
- Valoración subjetiva será alto, bajo PSNR

Mejores Indicadores



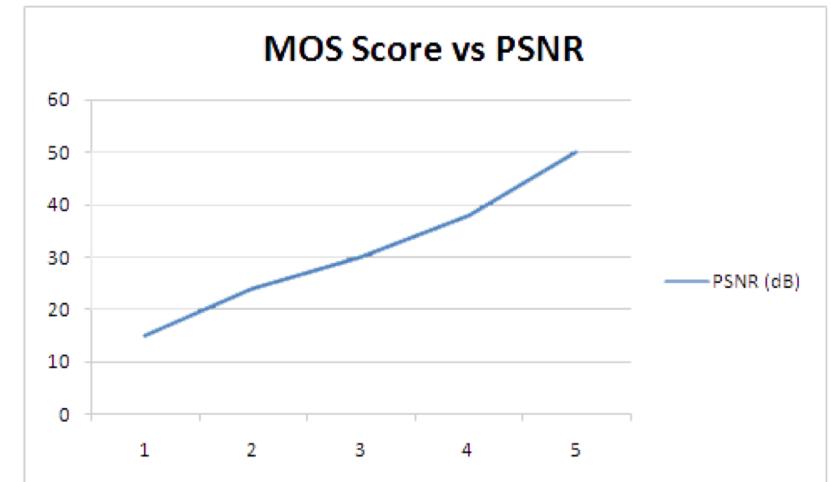
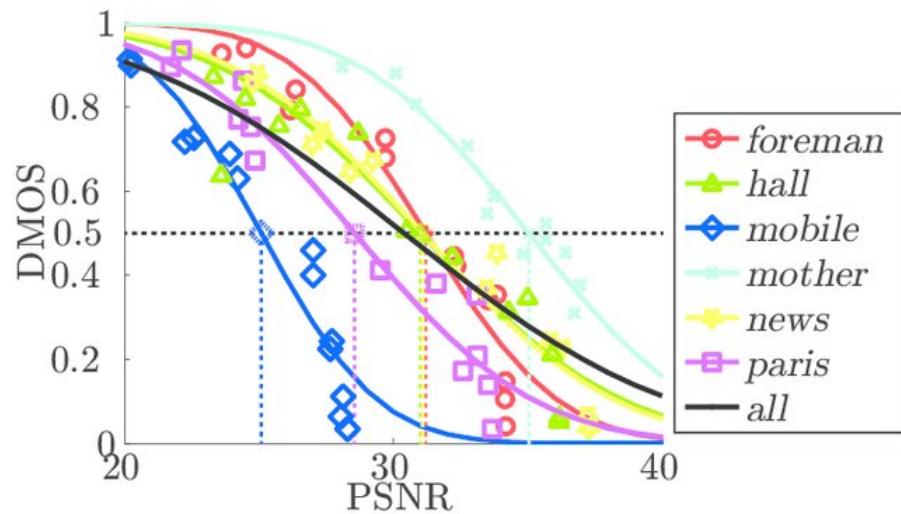
Video Quality Experts Group

■ Evaluación de medidas objetivos



Video Quality Experts Group

■ Evaluación de medidas objetivos



Video Quality Studio

Video Quality Studio :: Objective Evaluation

File Help

Reference Video Stream

File: E:\vq-test-pool\videos\sample04\huff.avi Browse

Stream Information

Size: 720x352
Frames: 254
FourCC: hfyu

Compressed/Impaired Video Stream

File: E:\vq-test-pool\videos\sample04\xvid-1000cbr.avi Browse

Stream Information

Size: 720x352
Frames: 254
FourCC: xvid

A B

Excellent

Good

Fair

Poor

Bad

Video Quality Studio :: Subjective Evaluation

Test Participants Options Help

Subjective Test

Participants

Resize Video
 Keep video always on top
 Allow videos to have different lengths
 Shuffle playlist

Vote!

+ - [x, y] Go!

You can add any number of reference samples with any number of test samples up to depth of 2.
i.e. Any one reference can have any number of children but all test samples cannot have any



Conclusiones

- La compresión de vídeo es importante
- La compresión de vídeo no es fácil
- La compresión de vídeo ha recorrido un largo camino
- No es tan maduro como la compresión de imágenes => Definitivamente hay espacio para mejorar
- Nuevos paradigmas en la computación se dictan las direcciones futuras de investigación