CENIC - QoS Workshop 1/24/02

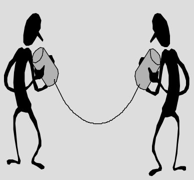
Video Coding and Quality Issues

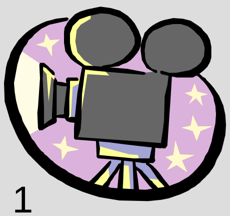
Magda El Zarki (elzarki@uci.edu)

VIP Lab. (www.ics.uci.edu/~vip)

CAL-(IT)^2 & Dept. of ICS

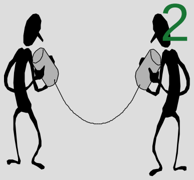
UC, Irvine

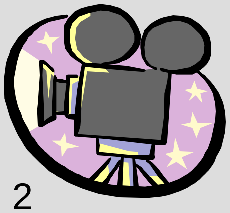




Outline of Talk

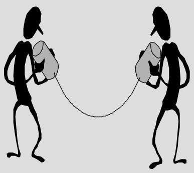
1. Part 1:
    1. Overview of Video Compression
    2. The MPEG suite
    3. Video Quality
    4. Losses
2. Part 2:
    1. Delivery over IP Networks
    2. Feedback Control
3. Part 3:
    1. MPEG-4
    2. Layered and Object based Coding
4. Part 4:
    1. QoS Issues
    2. Conclusions

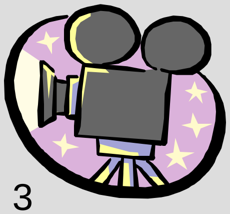




Part 1

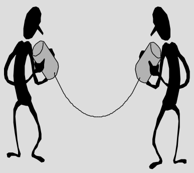
1. Overview of Video Compression
2. The MPEG suite
3. Video Quality
4. Losses

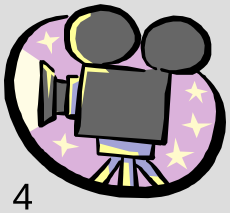




1. Video Compression: Goal

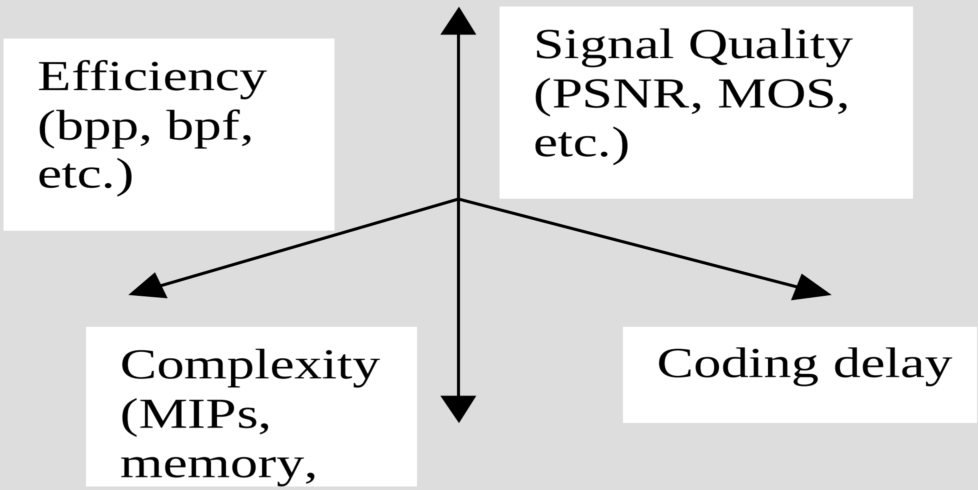
- Goal of video compression is to minimize the bit rate in the digital representation of the video signal while:
  - – Maintaining required levels of signal quality
  - – Minimizing the complexity of the codec
  - – Containing the delay





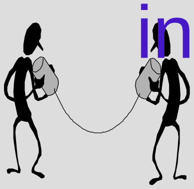
1. Video Compression: Tradeoffs

- The choice of a compression method involves a tradeoff along the following 4 dimensions:



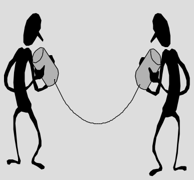
1. Video Compression: Why?

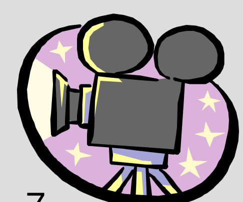
- Video signals are amenable to compression due to the following factors:
  - Spatial correlation: correlation among neighboring pixels
  - Spectral correlation: color images
  - Temporal correlation: correlation among pixels in different frames
- There is considerable irrelevant (from a perceptual viewpoint) information contained in video data.



1. Video Compression: Lossless Coding

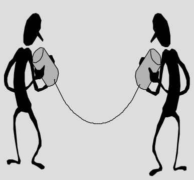
- Lossless coding is a reversible process - perfect recovery of data -> before and after are identical in value. Used regardless of media’s specific characteristics. Low compression ratios.
  - – Example: Entropy Coding
    - data taken as a simple digital sequence
    - decompression process regenerates data completely
    - e.g. run-length coding (RLC), Huffman coding, Arithmetic coding

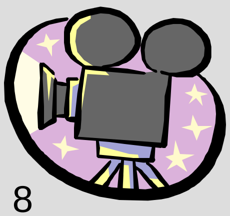




1.Video Compression: Lossy Coding

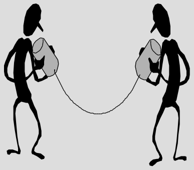
- Lossy coding is an irreversible process - recovered data is degraded -> the reconstructed video is numerically not identical to the original. Takes into account the semantics of the data. Quality is dependent on the compression method and the compression ratio.
  - Example: Source Coding
    - degree of compression depends on data content.
    - E.g. content prediction technique - DPCM, delta modulation

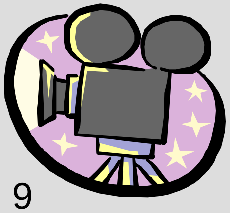




1. Video Compression: Hybrid Coding

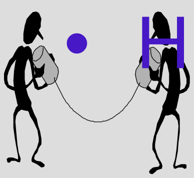
- Used by most multimedia systems
  - combines entropy with source encoding
  - E.g. JPEG, H.263, MPEG-1, MPEG-2, MPEG-4

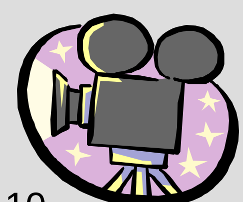




1. Video Compression: Design Choices

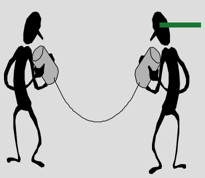
- Lossless or lossy or both
- Compression ratio
- Variability in compression ratio (fixed or variable quality)
- Resilience to transmission errors
- Complexity tradeoffs in codec (memory, processing, etc.)
- Nature of degradations
- Hierarchical representation

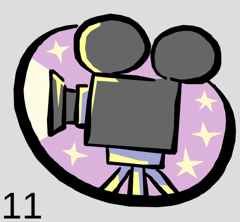




1. Video Compression - Standards

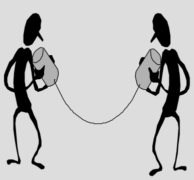
- Broadcast (high bit rate):
  - MPEG-1
  - MPEG-2
- Video Conferencing (low bit rate):
  - H.261
  - H.263
- Interactive (full range of bit rates):
  - MPEG-4

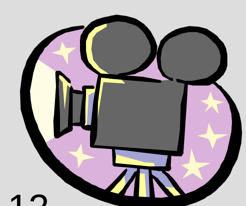




1. Video Compression: Deficiencies of existing standards

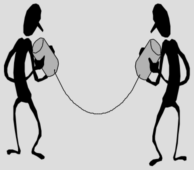
- Designed for specific usage
  - H.263 cannot be stored (no random access)
  - MPEG-1 & MPEG-2: not optimized for IP transport
- No universal file format for both local storage and network streaming
- Output cannot be reused efficiently after composition - encoded once, no versatility

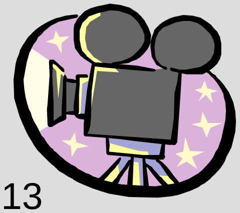




1. Video Compression: Requirements for New Standard

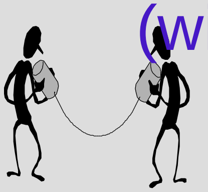
- Efficient coding scheme
  - Code once, use and reuse everywhere
  - optimized for both local access and network streaming
- Works well in both error prone and error free environment
  - Scalable for different bandwidth usage
  - Video format can be changed on the fly
  - Transparent to underlying transport network
- Support efficient interactivity over network

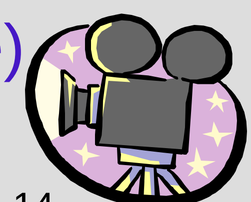




1. Video Compression: A solution - MPEG-4

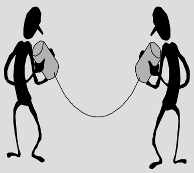
- Internet in the future
  - Not only text and graphics, but also audio and video
- Fast and versatile interactivity
  - Zoom in, zoom out (remote monitoring)
  - Fast forward and fast backward (video on demand)
  - Change viewing point (online shopping, sports)
  - Trigger a series of events (distance learning)
  - On the fly composition
  - Virtual environments
- Support both low bandwidth connections (wireless/mobile) and high bit rates (fixed/wireline)

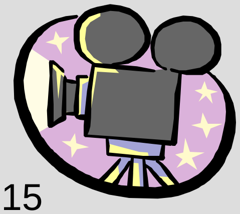




1. Video Compression: What is MPEG-4?

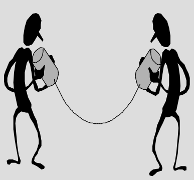
"A coded, streamable representation of audio-visual objects and their associated time-variant data along with a description of how they are combined."

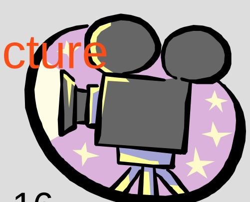




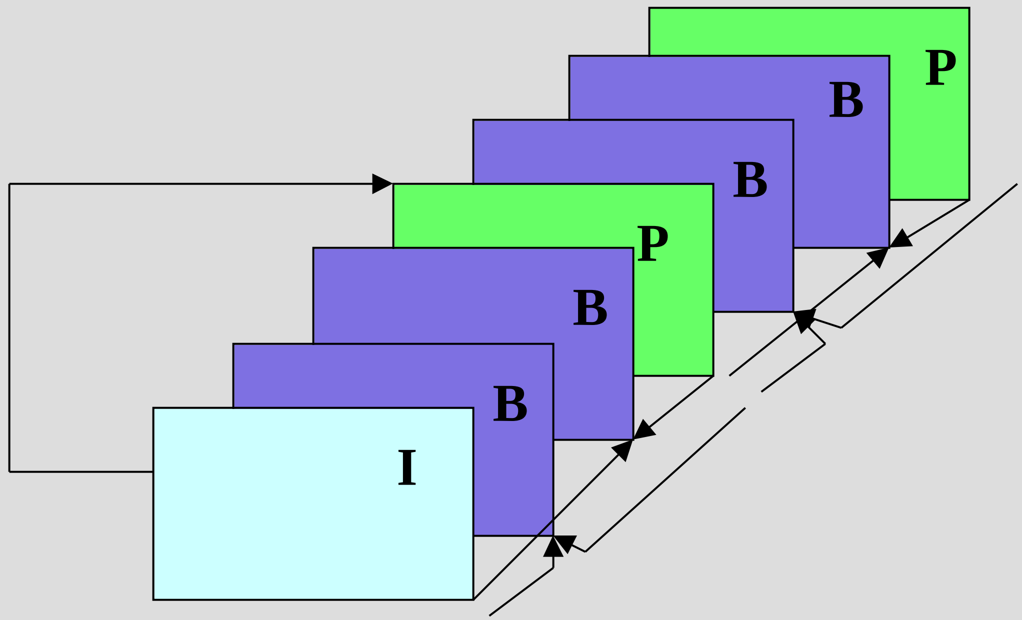
2. MPEG: Overview

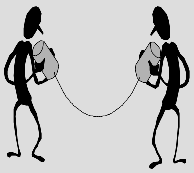
- MPEG exploits not only spatial redundancy in each frame, but also temporal (i.e frame-to-frame) redundancy present in all video sequences.
- Two Categories: intra-frame (spatial) and inter-frame (temporal) encoding
  - Intra: DCT based compression for the reduction of spatial redundancy - I frame
  - Inter: Block-based motion compensation for exploiting temporal redundancy
    - Causal (predictive coding) - current picture is modeled as transformation of picture at some previous time - P frame
    - Non-causal (interpolative coding) - uses past and future picture reference - B frame

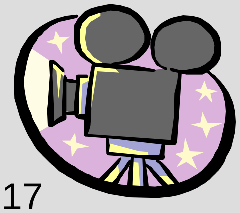




2. MPEG: Stream Components

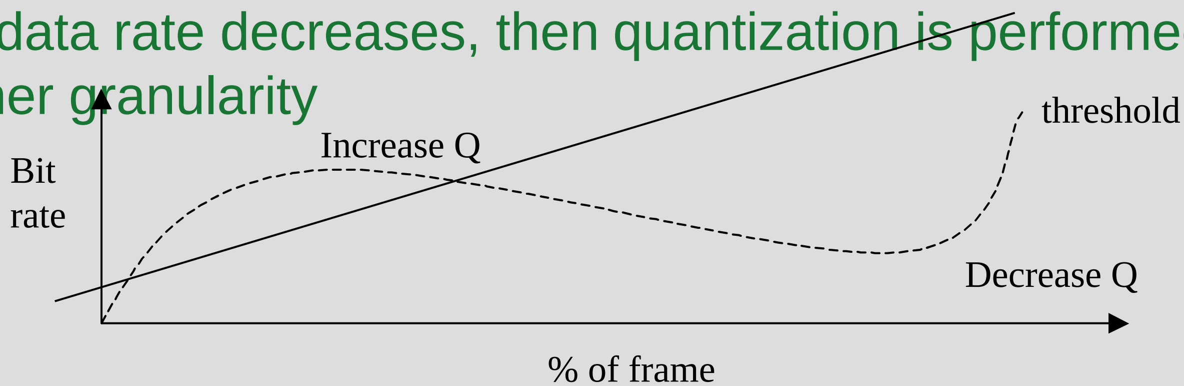


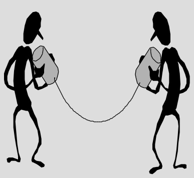


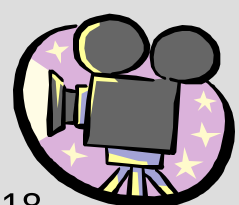


2. MPEG: The Quantization Parameter

- The quantization step is the main knob used to control the output bit rate of MPEG based encoders.
- For CBR encoders MPEG quantization is adjusted as follows:
  - If data rate increases over threshold, then quantization enlarges the step size
  - If data rate decreases, then quantization is performed with finer granularity







3. Quality

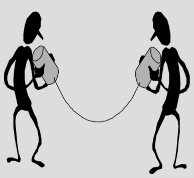
- What is video quality?
  - Generally judged using PSNR
    - Easy to compute

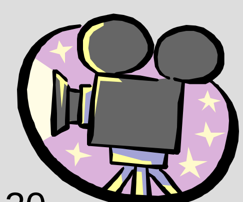
    BUT
    - Not a good estimate of quality
  - New objective quality measurements
    - Hard to compute

    BUT
    - More accurate

3. Quality: Assessment Techniques

- Traditional Objective Assessment - Peak Signal to Noise Ratio (PSNR)
- Subjective Assessment – DSCQS (Double Stimulus Continuous Quality Scale)
- Perceptual Objective Assessment -
  - Human visual perception based
  - Capturing image imperfections



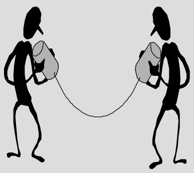


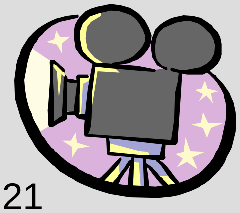
3. Quality: Peak Signal to Noise Ratio

- For a video sequence of K frames of NxM dimension with 8 bit depth:

$$RMSE = \sqrt{\frac{1}{N\,M\,K}\sum_{k=1}^{K}\sum_{n=1}^{N}\sum_{m=1}^{M}\left[x_k(n,m)-\tilde{x}_k(n,m)\right]^2}$$

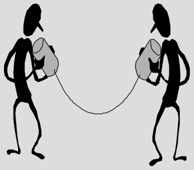
$$P\ldots = 20\log_{10}\frac{255}{\ldots}$$

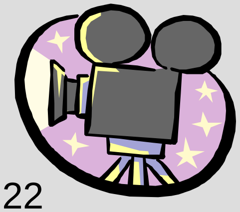




3. Quality: Advantages of PSNR

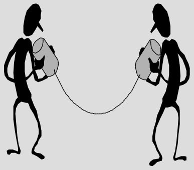
- Very easy to compute
- Well understood by most researchers
- Results are close to DSCQS but they do not translate accurately to human perception

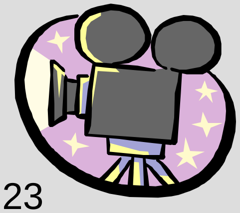




3. Quality: Disadvantages of PSNR

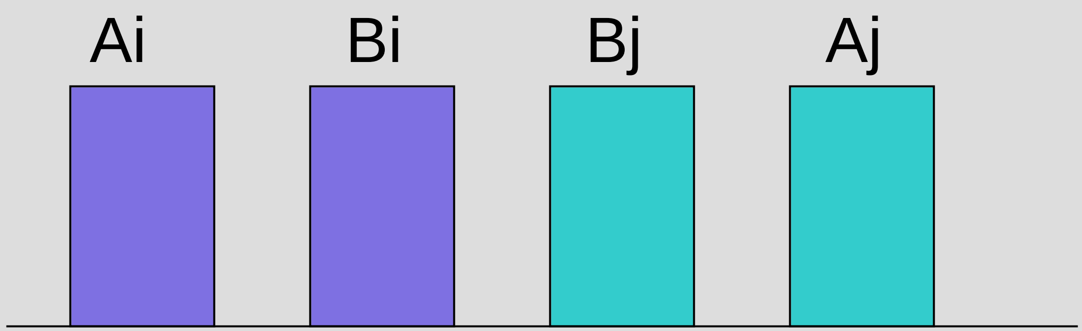
Some reconstructed images with different errors have the same PSNR values

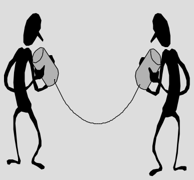


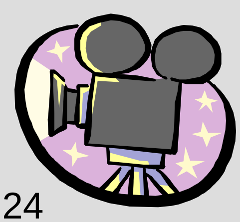


3. Quality: Subjective Assessment: DSCQS

- Source (A) and Processed (B) video clips are presented in pairs
- The video presentation sequences are randomized

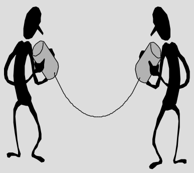


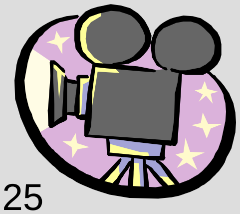




3. Quality:DSCQS scoring

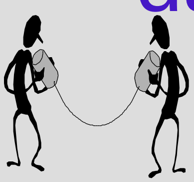
- Viewers grade each clip’s quality
- Data is gathered in pairs

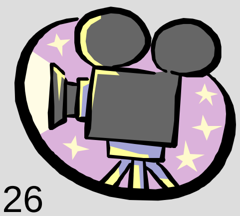




3. Quality: Issues with DSCQS

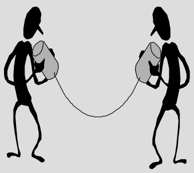
- Until now the most reliable quality measurement method
- Requires special viewing room and equipment
- Needs a large group of people
- Large amount of post processing on data

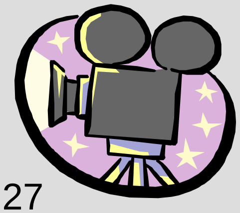




3. Quality: Objective Assessment (OA)

- Establish a good quality assessment model
- The model takes as inputs the source and the processed video clips.
- Compare the model output to DSCQS test score
- If the result is consistent with DSCQS measurement, the model can substitute DSCQS



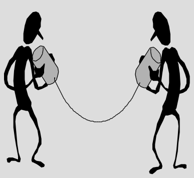


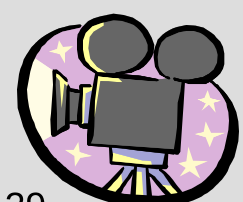
3. Quality: OA Requirements

- Ability to predict subjective quality with low error
- Predictions agree monotonically with the relative magnitudes of subjective quality ratings
- Prediction is robust with respect to a variety of video impairments

3. Quality: OA Models - 2 approaches

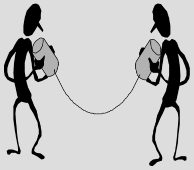
1. Establish a model that simulates the human visual stimulation
2. Find the relationship between measurable parameters and perceptual distortion (blurring, tiling, noise)

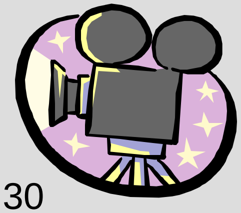




3. Quality: Issues related to Method 1

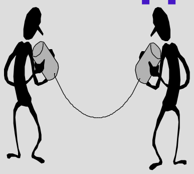
- Advantages:
  - Considers both luminance and chrominance
  - Some methods show very high correlation with DSCQS for some video sequences
- Disadvantages:
  - Not capable of in-service evaluation
  - Not consistent over all video bit rate ranges
  - Computationally complex

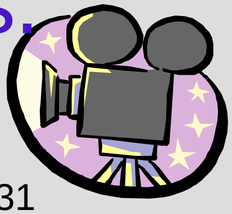




3. Quality: Method 2 - ITS Model

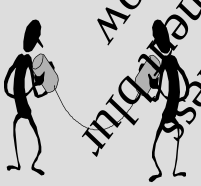
- Institute for Telecommunication Studies (ITS) were the first group to propose an objective measure several years ago.
- They have since refined (or fine tuned) the model to capture more of the image imperfections.
- They map image imperfections onto measurable mathematical parameters.

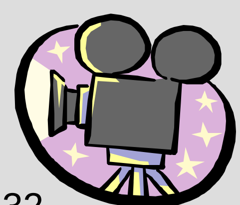




3. Quality: Perceptual Impairment Factor Vs. A Measurable Parameter

Image blur	Step response rise time
Edge busyness	Step response jitter width
False contouring	Sp-p to minimum quantizing
Granular noise	Equivalent analog signal SNR expiress as Sp-p/ Nrms
Dirty Window	Maximum noise amplitude
Movement blur	Rise time of a moving edge
Jerkiness	Field or frame difference of moving edge position





3. Quality: Perceptual Impairment Factor

a

b

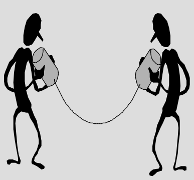
c

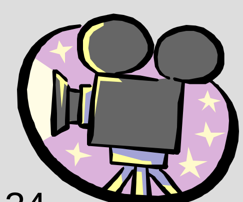
d

a) Original, b) Blur, c) Tiling, d) Noise

3. Quality: Advantages of ITS Model

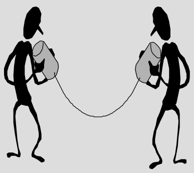
- Works well for a wide range of bit rates
- Produces results that are consistent with subjective tests (DSCQS)
- Computationally efficient
- Bandwidth efficient (384:1)
- In service quality monitoring

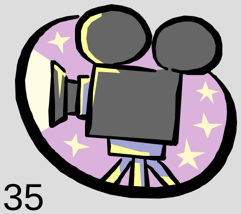




3. Quality: Disadvantages of ITS Model

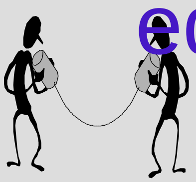
- Based on no visual model (vs.method 1)
- Only considers luminance value

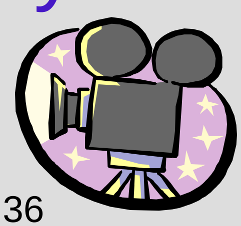




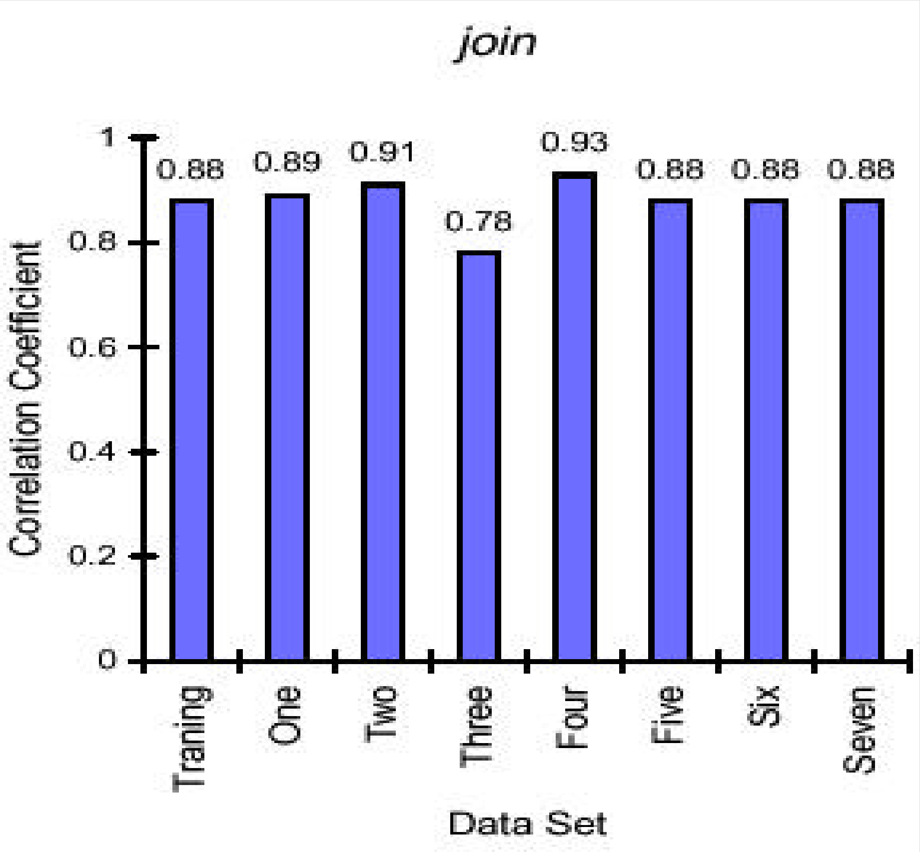
3. Quality: Video Quality Experts Group (VQEG)

- Several models have been under evaluation
- Tested video bit rate from 768 kbps to 50Mbps (4:2:0 - 4:2:2 MPEG-2)
- Both NTSC and PAL signals tested
- Viewing Distance limit to 6:1
- Carefully calibrated and aligned display equipment



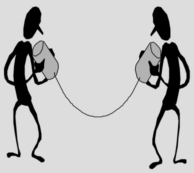


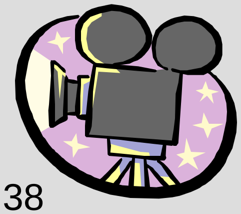
3. Quality: DSCQS vs ITS



3. Quality: Conclusions

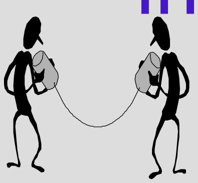
- All models have strengths and weaknesses, not one can substitute DSCQS
- Some display fairly consistent behavior for different video resources
- Developed to judge encoder quality not to assess damage caused by packet losses
- No quality measures developed yet for shape coding

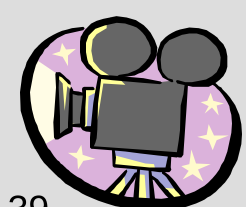




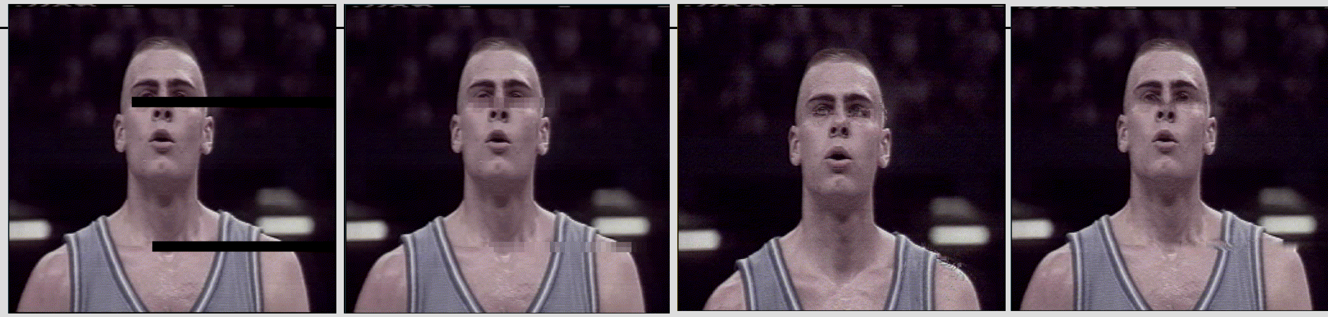
4. Losses

- Packet losses may cause the quality of the video to degrade to unacceptable viewing levels
- It is not always easy to assess the degree of degradation - highly dependent on what portion is lost
- Error concealment techniques can improve quality substantially





4. Losses: Error Concealment

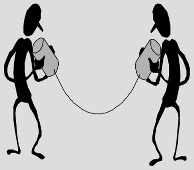


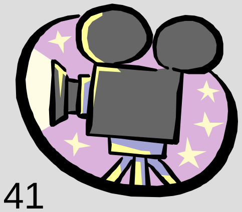
(a). Unconcealed Image; (b). Frequency concealment (FC); (c). The 16th frame after initial FC on the first image; (d). Spatial Concealment

Left: Unconcealed Image. Middle: Concealed by simple motion vector estimation, simply averaging the top and bottom mvs. Right: Concealed with more motion vectors. All the adjacent mvs are used directly or indirectly.

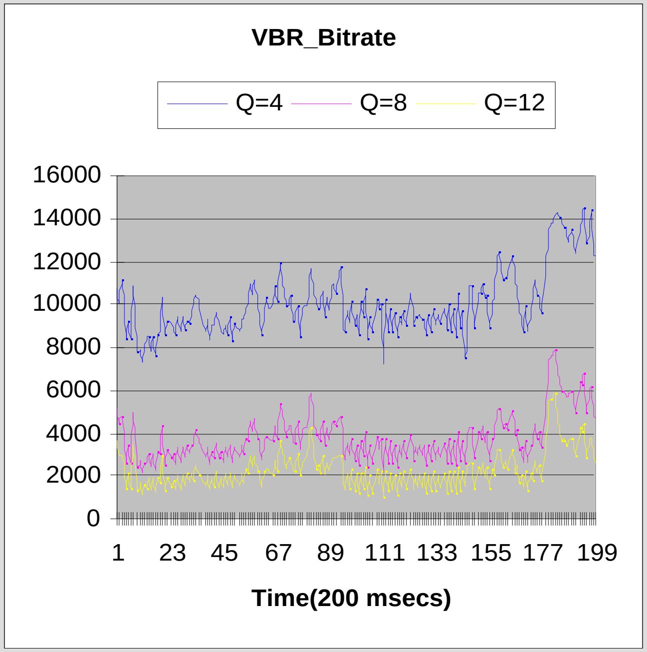
Part 2

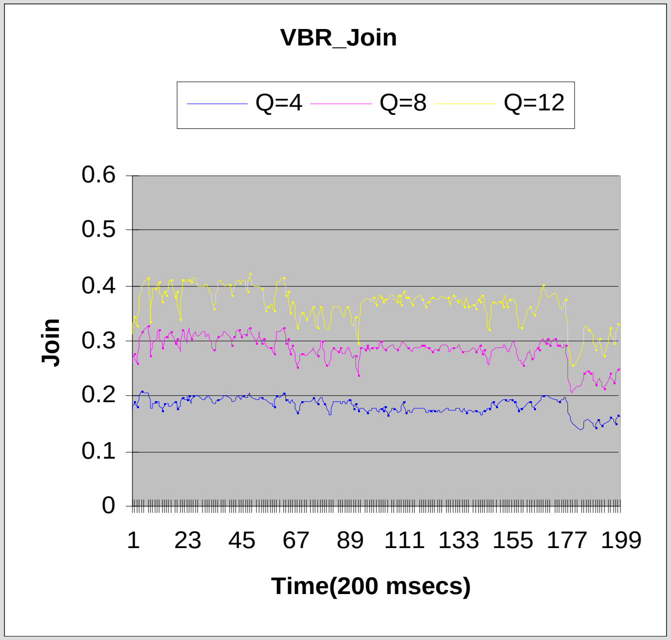
1. Delivery over IP Networks
2. Feedback Control

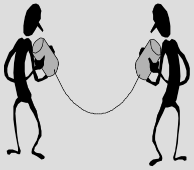


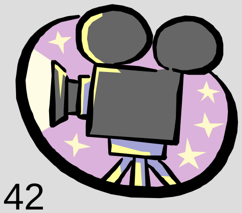


1. Delivery: Bit rate & Quality (VBR)

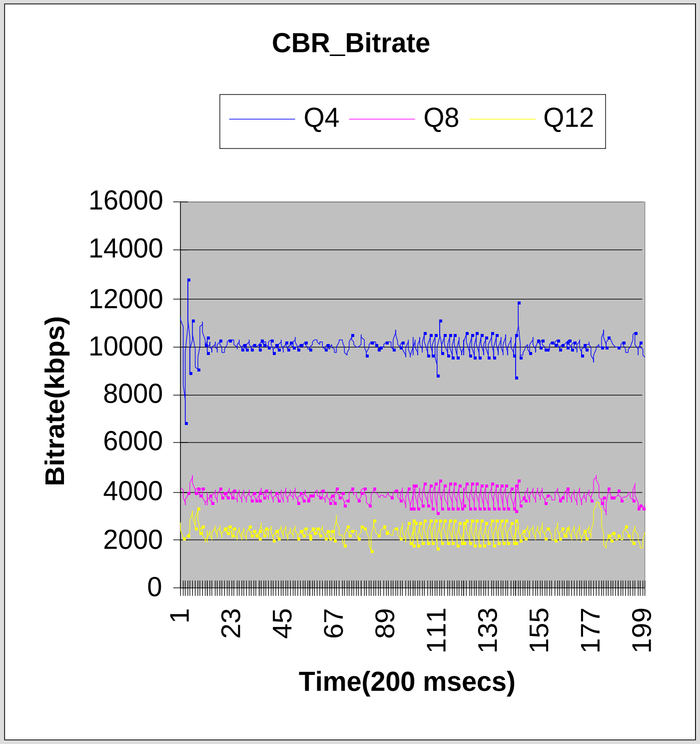


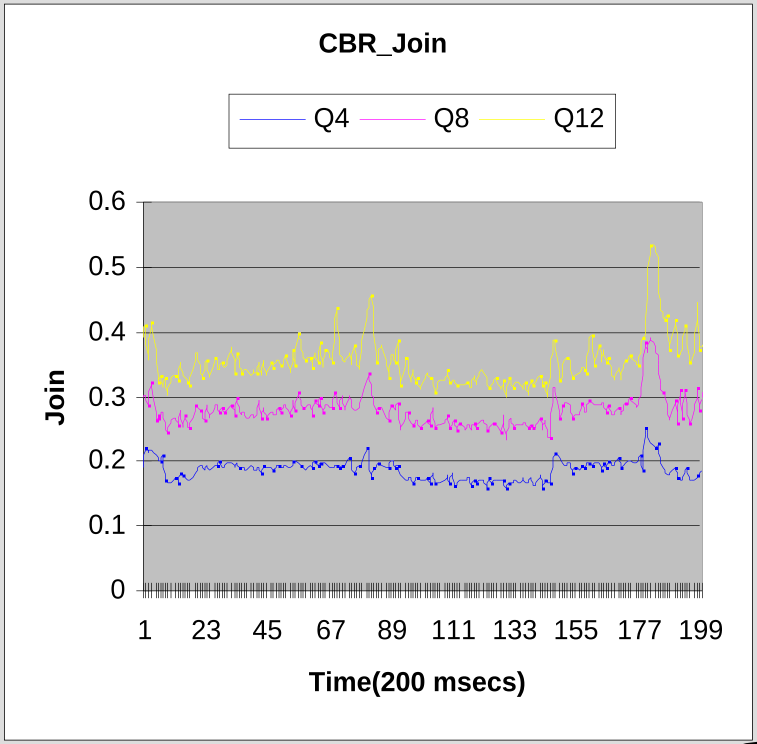


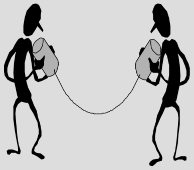


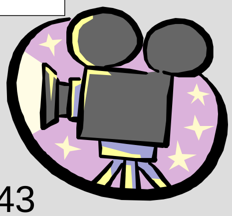


1. Delivery: Bit Rate & Quality (CBR)

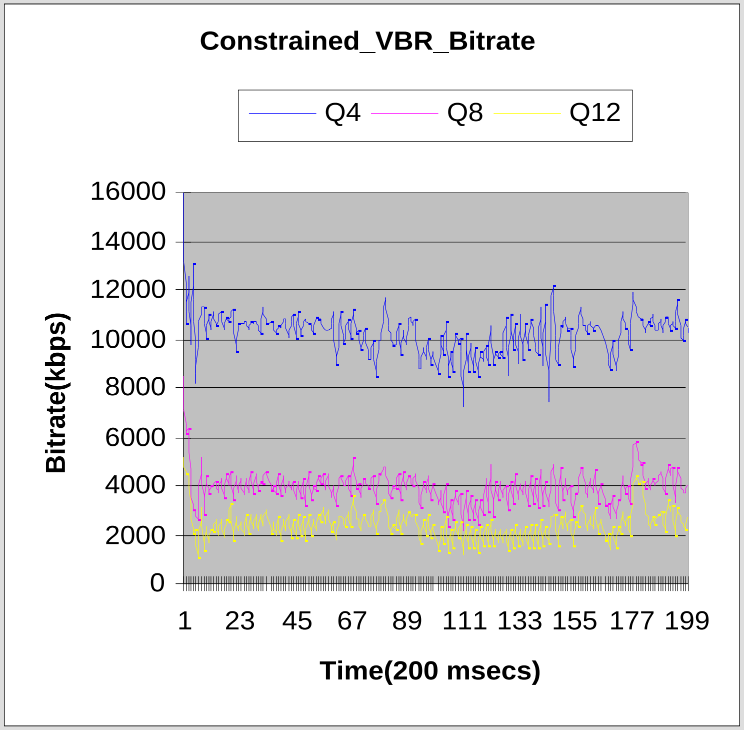


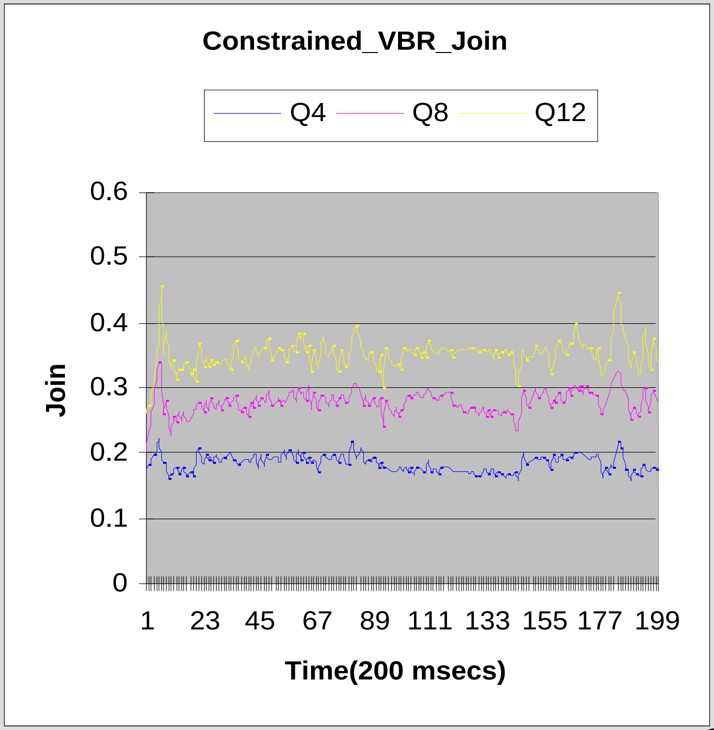


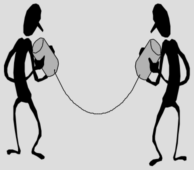


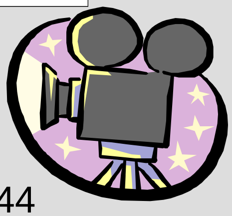


1. Delivery: Bit Rate & Quality (Constrained VBR)





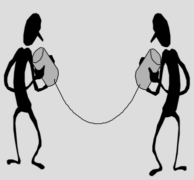


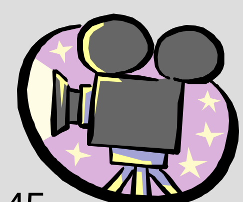


1. Delivery: Comparison Table

	Join		PSNR		Bit Rate	
	Average	Std_dev	Average	Std_dev	Average	Std_dev
VBR_Q4	0.202004	0.002655	35.91496	0.44465	10053	1545
VBR_Q8	0.317597	0.004489	32.96177	0.52643	3813	1127
VBR_Q12	0.409873	0.004677	31.42474	0.67014	2302	889
CBR_Q4	0.204929	0.003747	35.72485	0.98738	10054	483
CBR_Q8	0.329049	0.024604	32.80154	1.15316	3815	408
CBR_Q12	0.458771	0.058536	31.36981	1.56755	2307	374
Constrained_VBR_Q4	0.206998	0.003203	35.78648	0.83835	10066	947
Constrained_VBR_Q8	0.320873	0.010033	32.87081	0.87554	3829	736
Constrained_VBR_Q12	0.412427	0.016802	31.38822	0.92129	2326	647

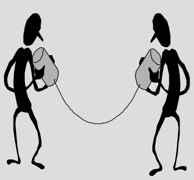
Table 1. Quality and Bit rate: average and standard deviation comparison

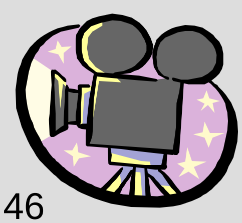




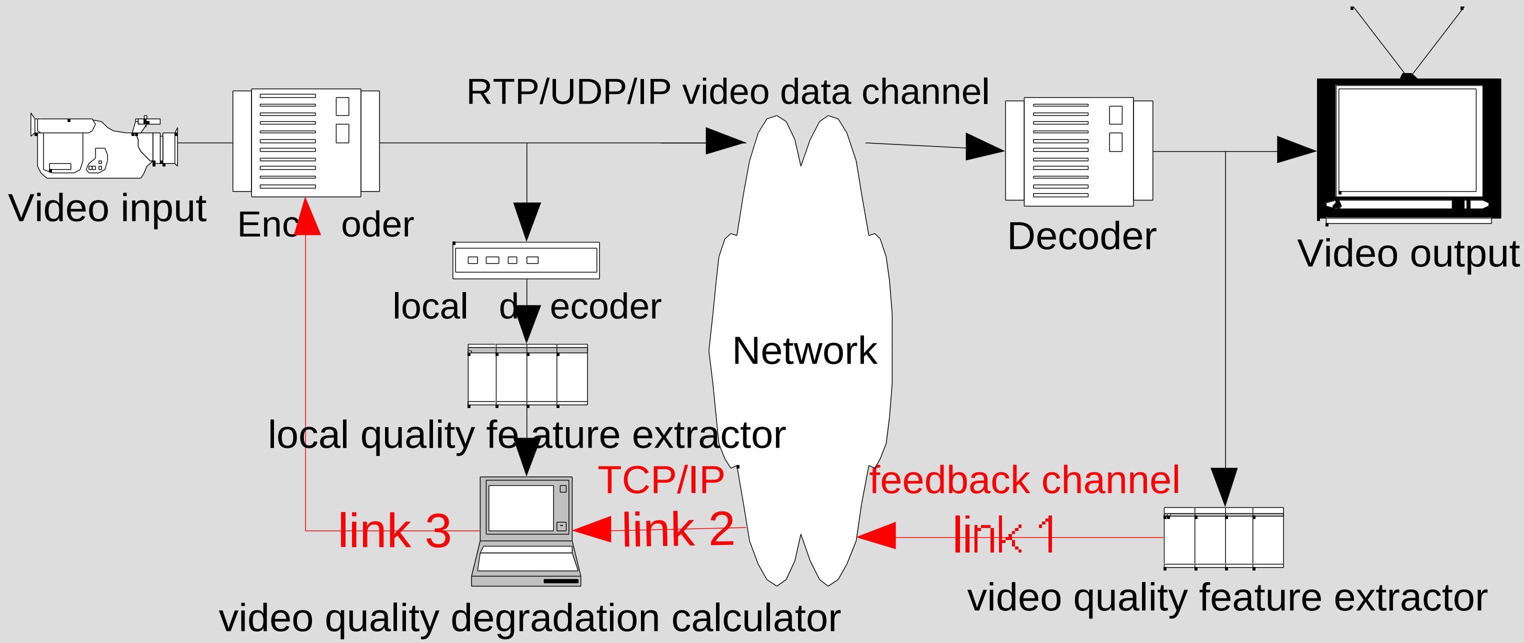
2. Feedback Control

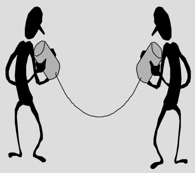
- Feedback control can be used to control the source encoder - change the output bit rate by changing the quantization parameter (Q) based on some performance metrics
  - Use packet loss rates (RTCP error reports)
  - Use TCP congestion information
  - Use the perceived quality at the receiver

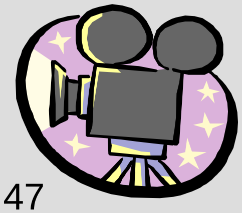




1. Feedback Control - Implementation







2. Feedback Control: Impact

Picture-1

Picture-2

Quality degradation caused by 1) packet loss and 2) compression algorithm

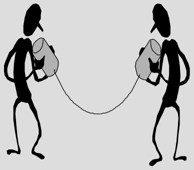
2. Feedback Control: Issues

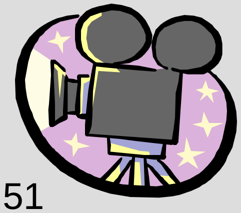
- Fine tuning of feedback control:
  - Error Concealment
  - Evaluation window
  - Degree of correction
  - Thresholds for increasing or decreasing “Q”
    - QoS issues
    - Pricing
  - Impact on perception - variability in quality
  - Etc.

2. Feedback Control: Example

Part 3

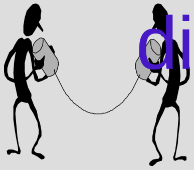
1. MPEG-4
2. Layered and Object based Coding

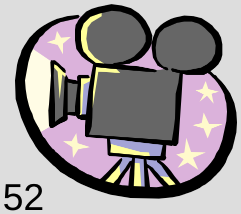




1. MPEG-4: Overview

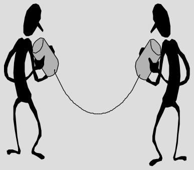
- MPEG-4 aims to pave the way towards a uniform, high quality encoding and decoding standard, that would replace the many proprietary streaming technologies in use on the Internet today
- MPEG-4 is object-based, multi stream
- Can accommodate a wide range of bit rates including very low bit-rate communication for mobile receivers or wristwatches that can display video.

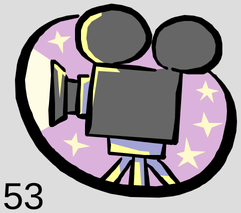




1. MPEG-4: What's new?

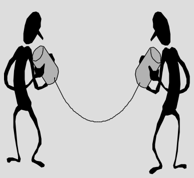
- Improved Coding Efficiency
  - Hybrid data coding: mixing of synthetic and natural
  - Arbitrary shape coding (as opposed to rectangular)
  - Coding of multiple concurrent data streams
- Content-based Interactivity
  - Does not deem video frame as a whole anymore
  - Code each audio/video/text/graphics object into separate stream
  - User can interact with each object in the scene
- Universal Access
  - Robustness, independent of environment
  - Content-based scalability based on client's request
  - Dynamically adaptive to available network bandwidth

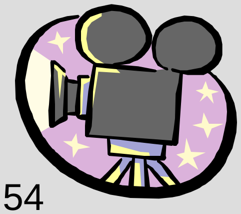




1. MPEG-4: Object Coding

- Improves reusability and coding efficiency of individual components
  - Apply different coding algorithms on different objects
- Allows content-based scalability
  - High resolution only on interesting part
  - Streaming object, pre downloaded object and local object can work together
  - Object based QoS support
  - Allows more flexible user interactivity - each object can be paused, FF, removed, etc.

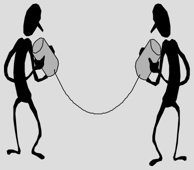


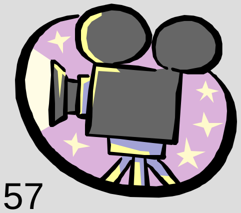


Integration of Natural and Synthetic Content

Augmented Reality

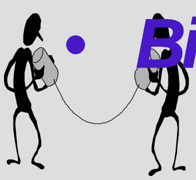
Tele Presence

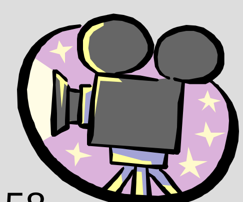




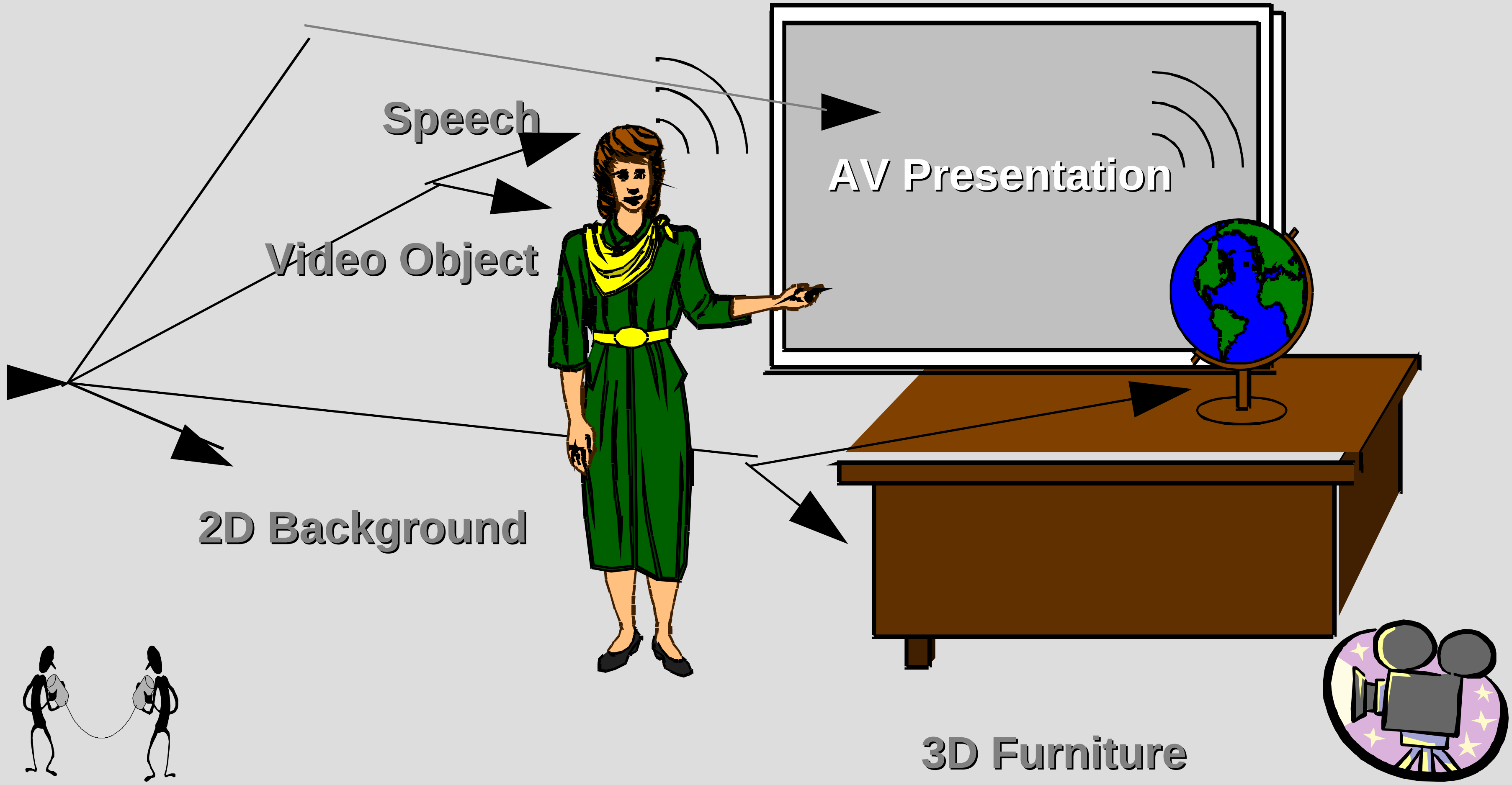
1. MPEG-4: Scene Description

- A '**compositor**' composes objects in a scene (A&V, 2&3D) creating a composite scene
- A scene description defines how objects appear on End User screen (composition view)
- With the scene description an End User can
  - change the position of individual video object
  - Zoom in/out interesting object
  - Choose different audio track (language, music)
  - Turn on/off individual object
  - Change resolution of an object, etc.
- ***Bi***nary ***F***ormat for ***S***cene Description : '**BIFS**'

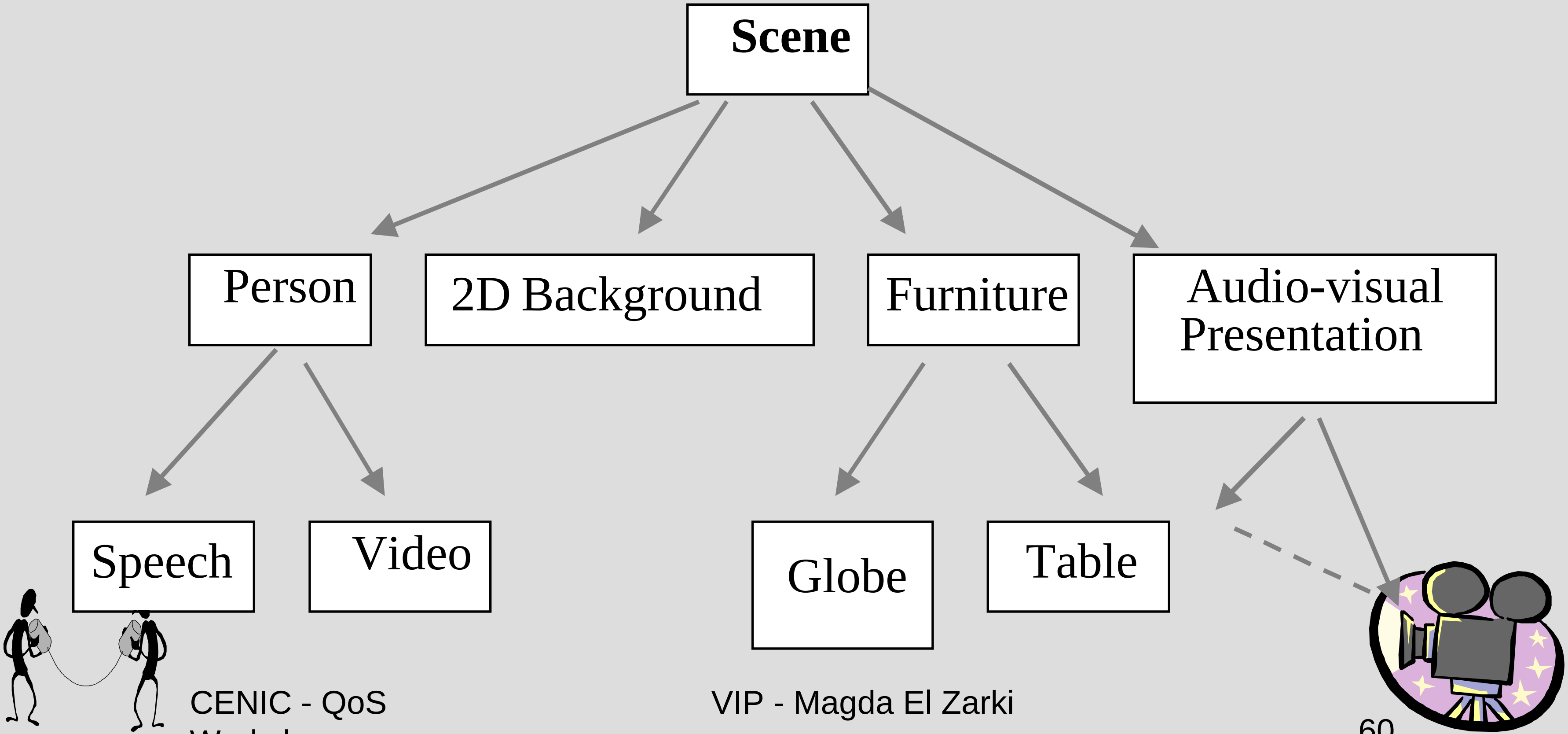




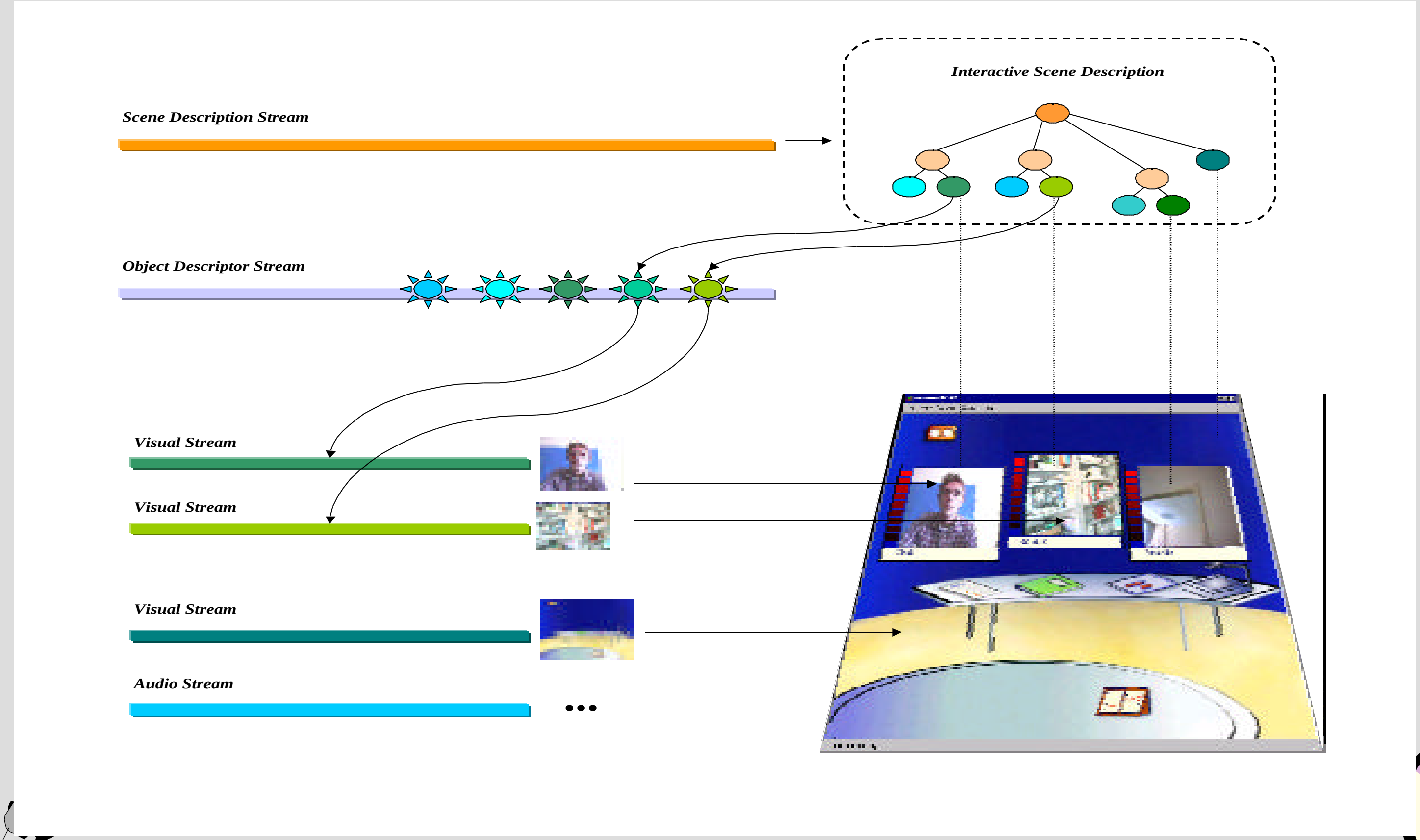
1. MPEG-4: Scene Composition



1. MPEG-4: Composition of Scenes

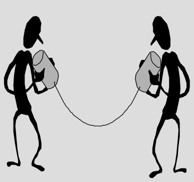


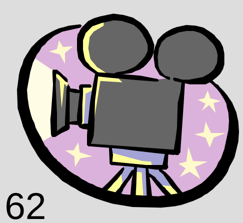
1. MPEG-4: Object Streams



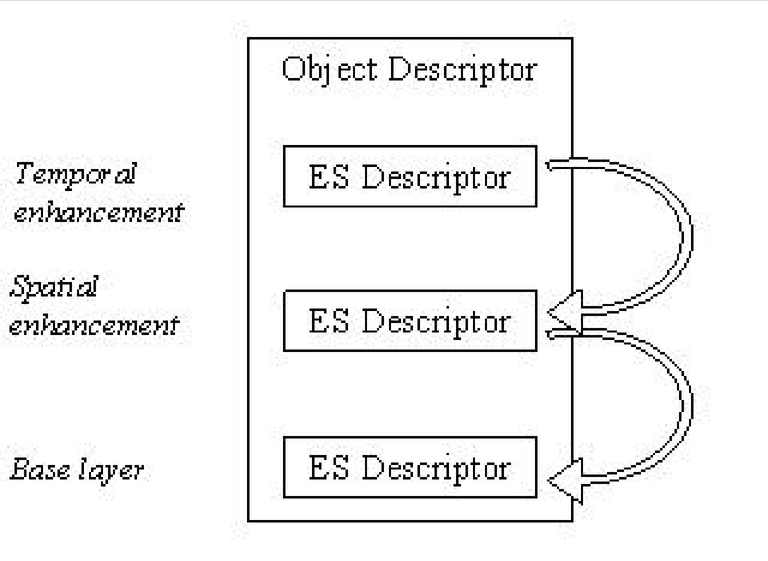
2. Object Coding: Object Descriptor (OD)

- Groups a set of Elementary Streams (ESs) associated with a particular object as a single entity (e.g. base and enhancement layers)
- Transported in object descriptor stream
- Object descriptors can be updated dynamically over time

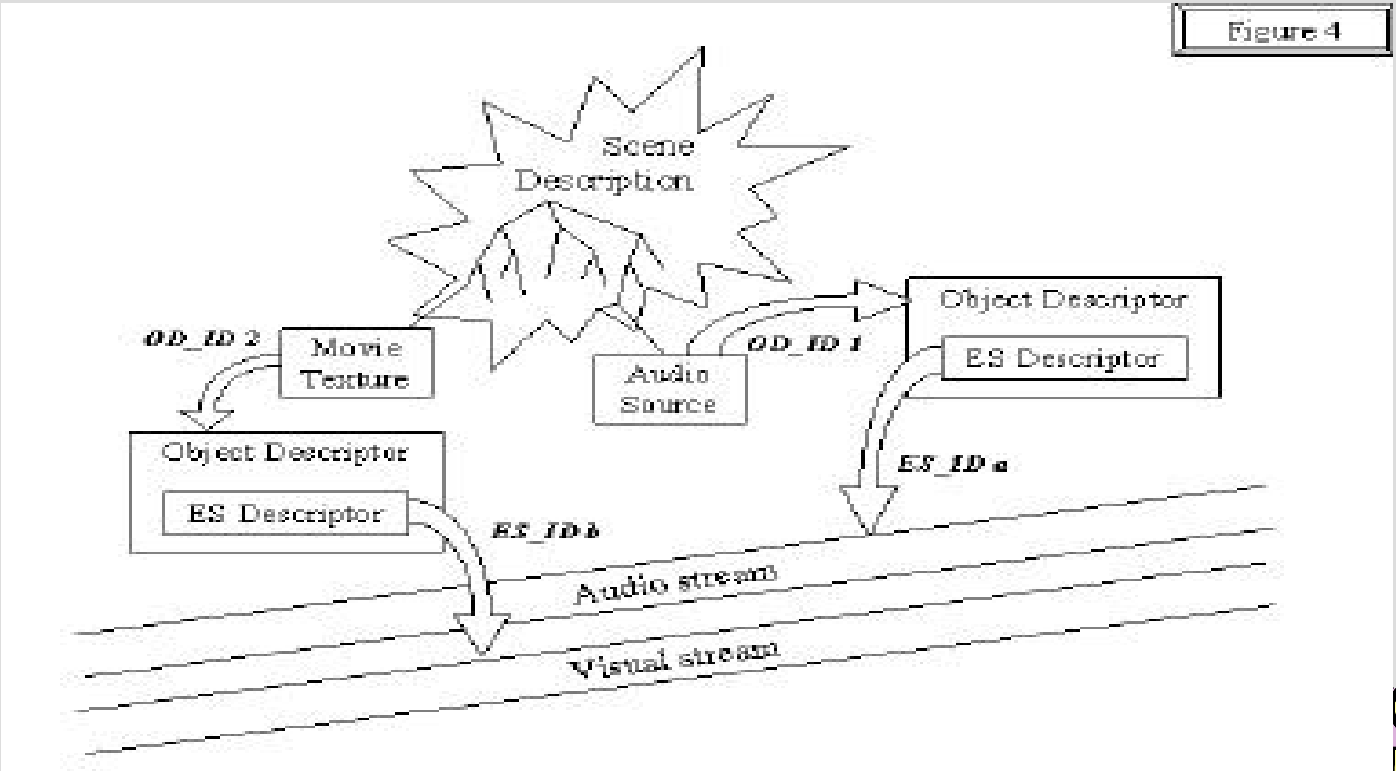




Generic Sample of an Object Descriptor

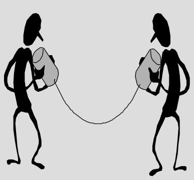


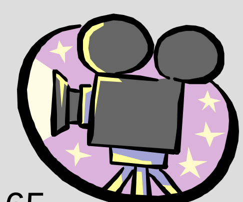
Examples of Object Descriptors



ES-Descriptors

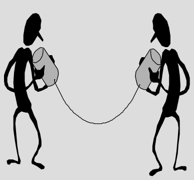
- Each describes one Elementary Stream (ES) (audio stream, video stream, etc.)
- Includes configuration information for dedicated stream decoder (DecoderConfig)
- Contains sync layer configuration information for this stream (SLConfig)
- Conveys QoS Requirements to transport channel (***optional*** QoS descriptor)

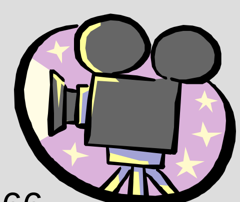




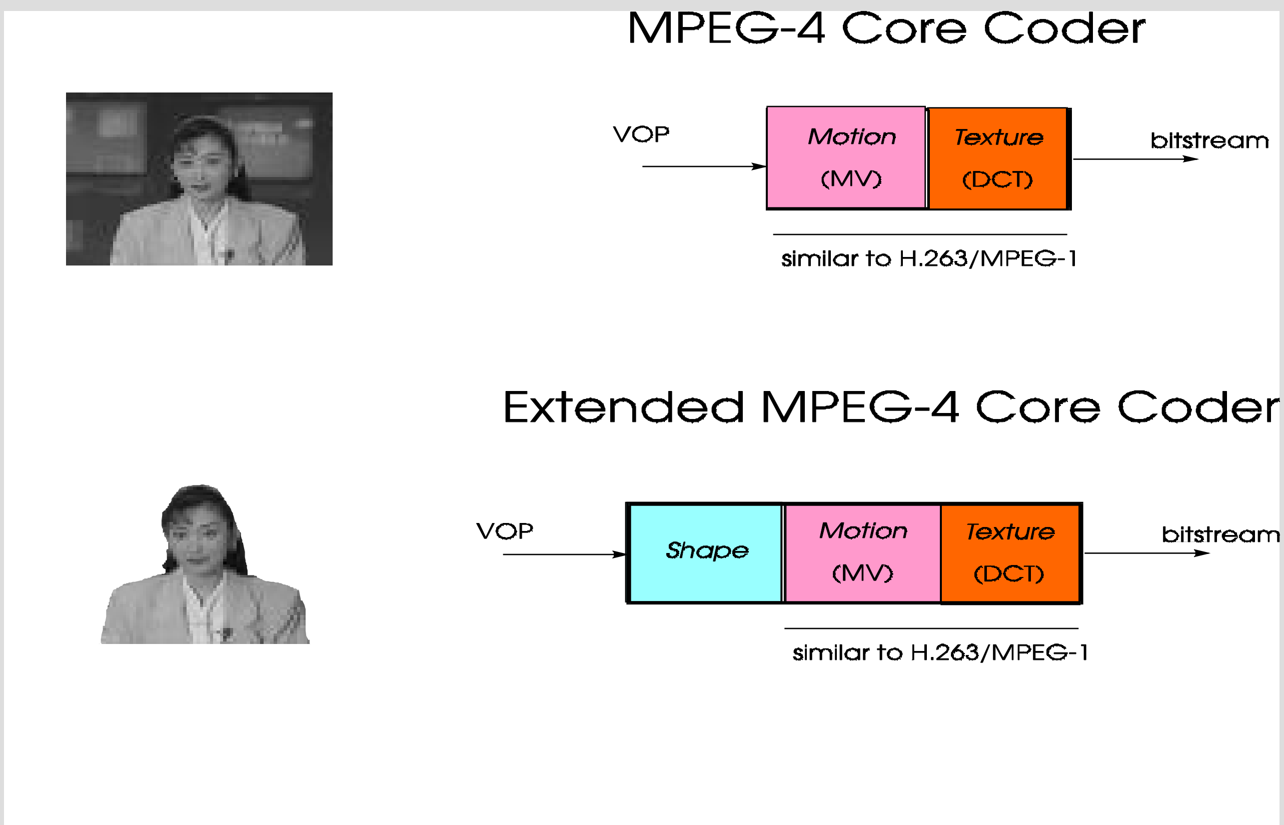
Coding Modes of MPEG-4

- Baseline
  (***Conventional rectangular coding***)
  - Compression
  - Error Resilience
  - Scalability
- Extended
  (***Object (shape) coding***)
  - Content-based Coding
  - Still Texture Coding



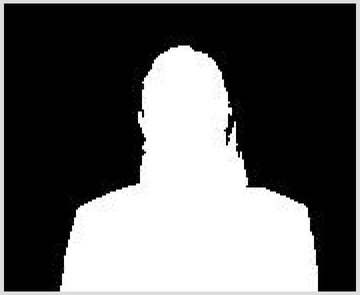


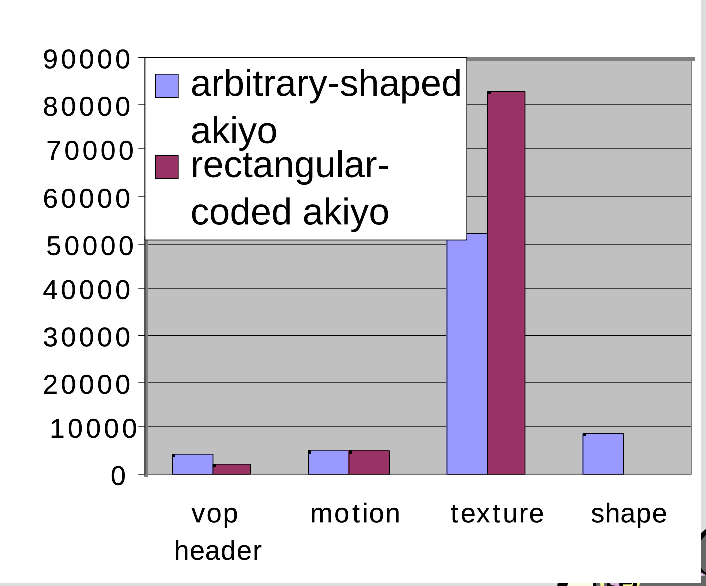
Comparison of Modes



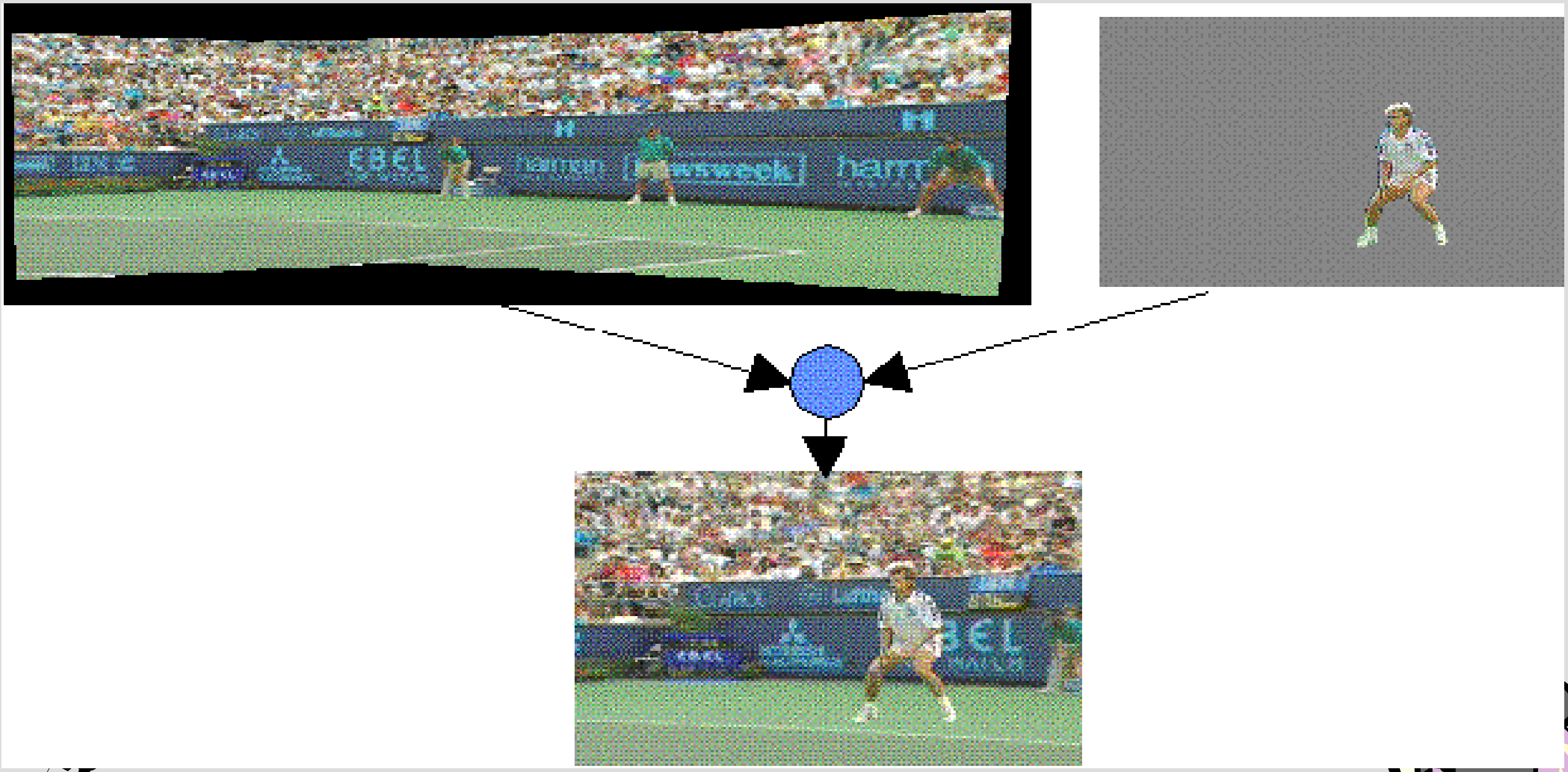
Shape Coding

Comparison of Arbitrary-shape coding and Rectangular-based coding (Q = 6). Total no. of frames: 40, format: qcif



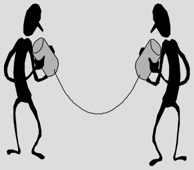


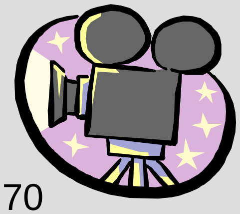
Combining Objects



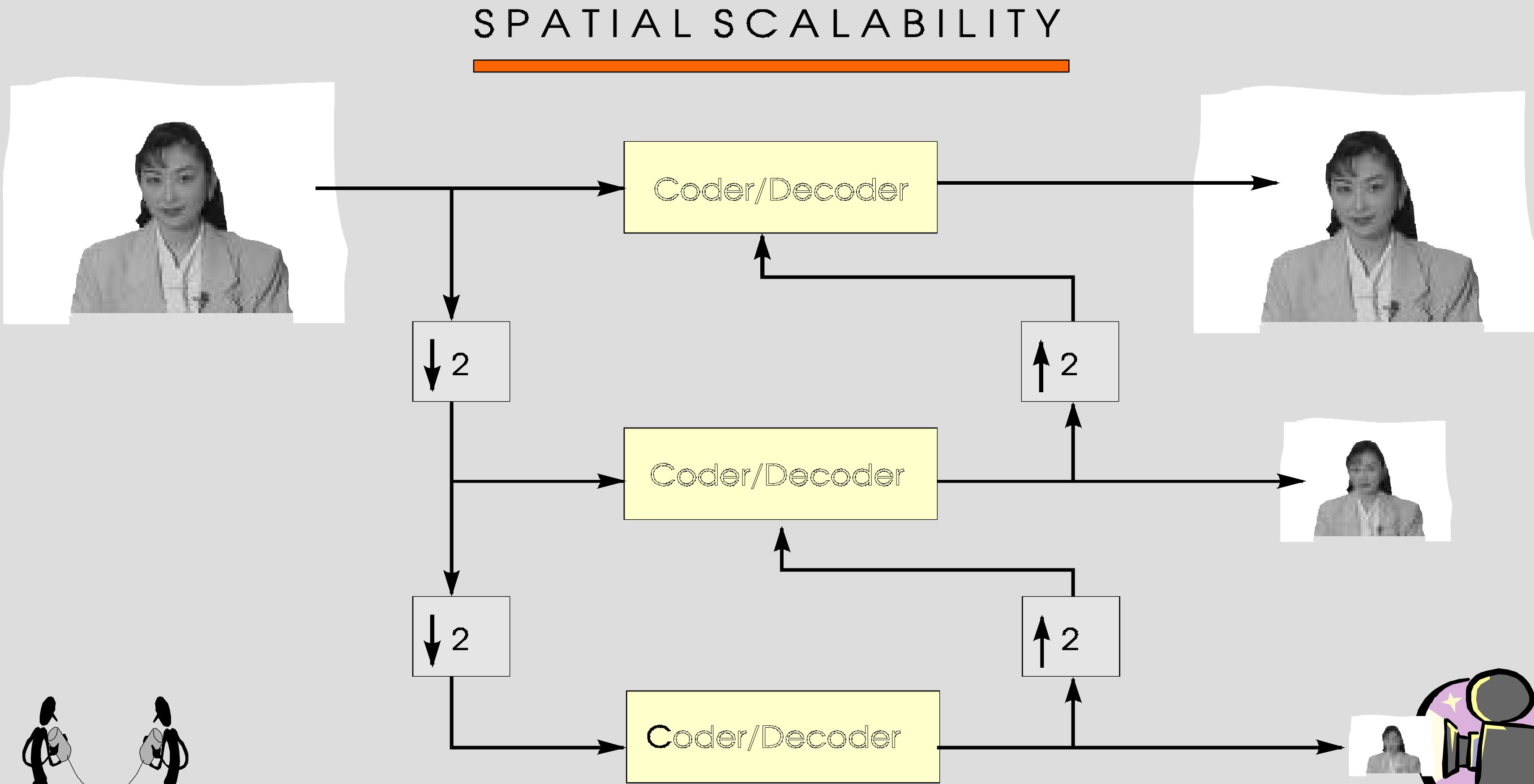
Scalable Coding

- Object based spatial scalability
- Object based temporal scalability
- Both provide resilience to transmission errors

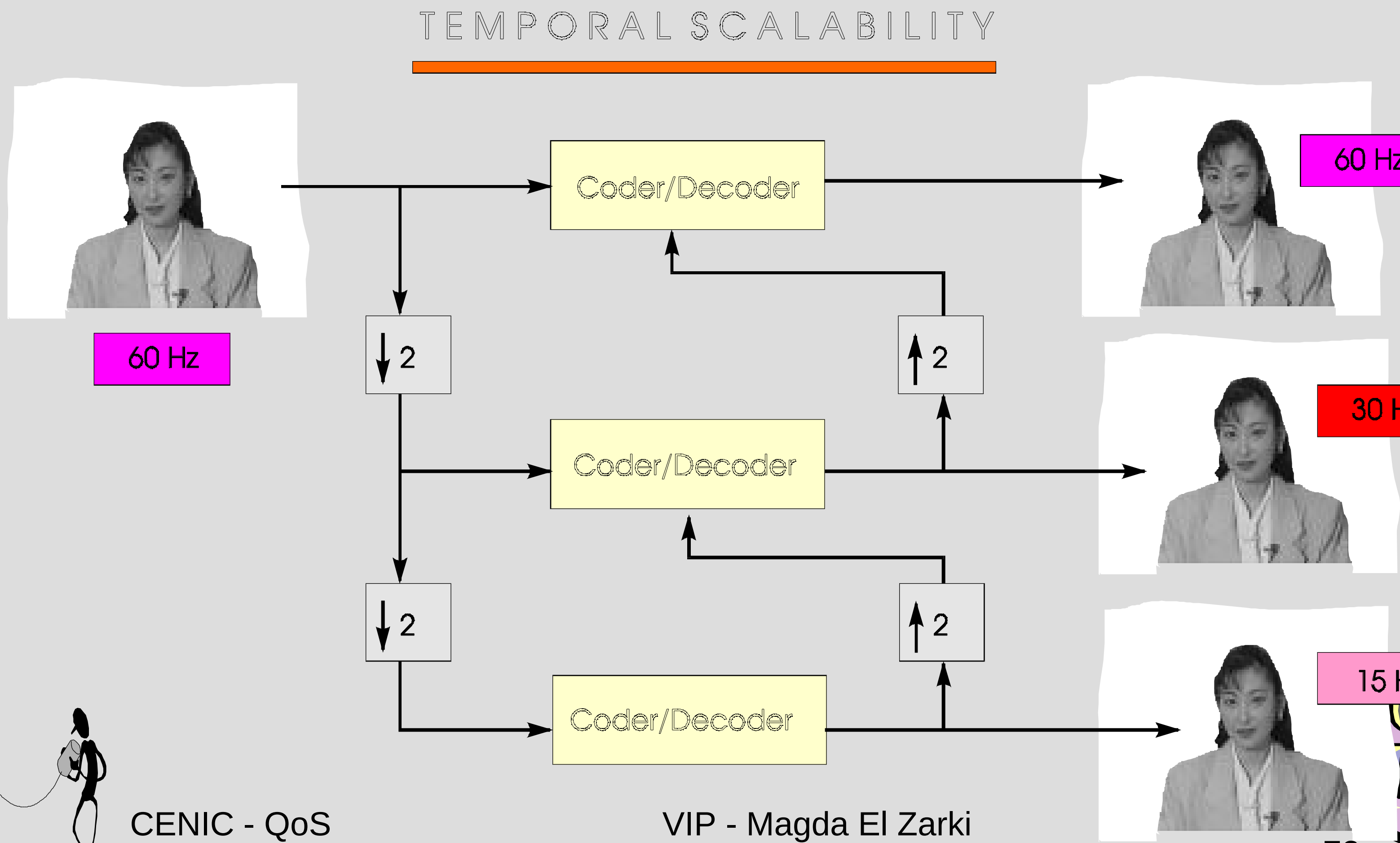




Spatial Scalability

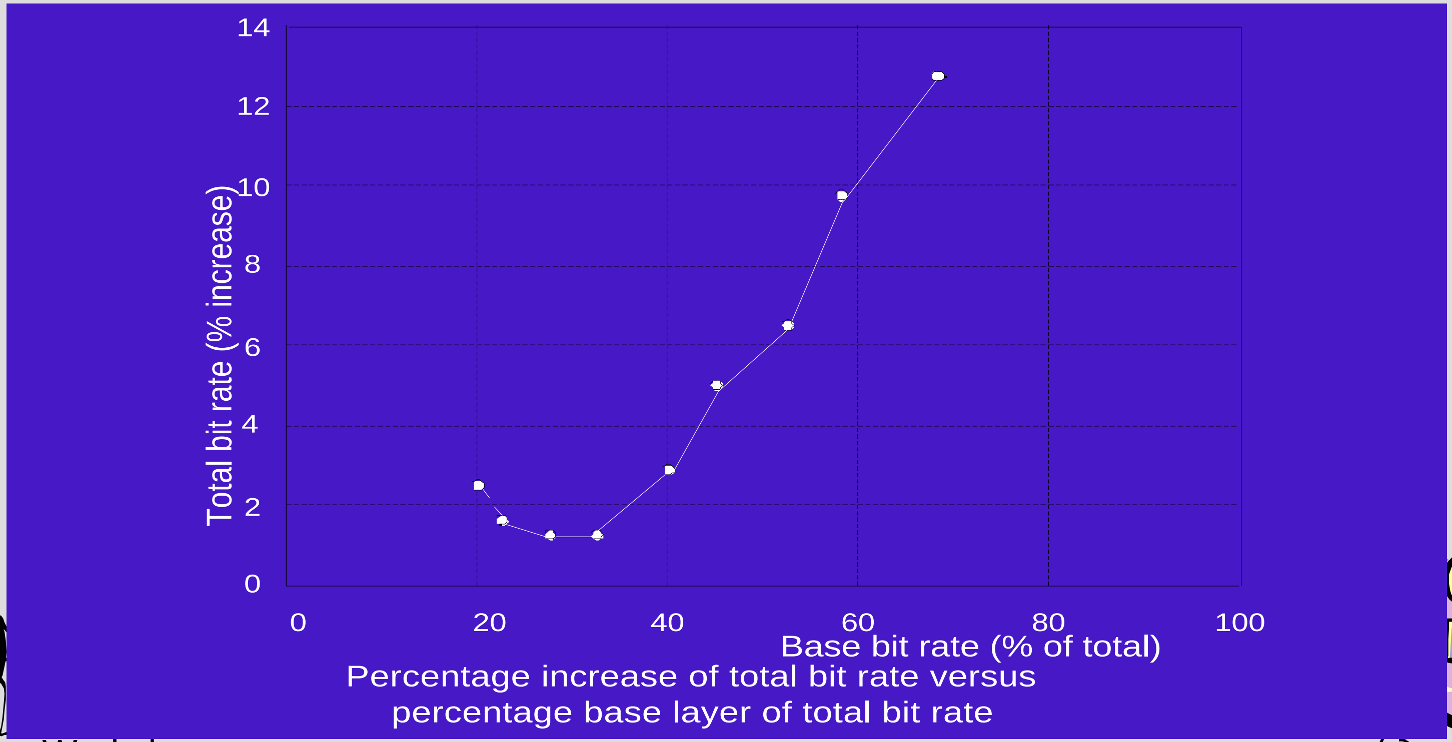


Temporal Scalability

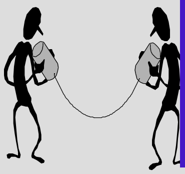


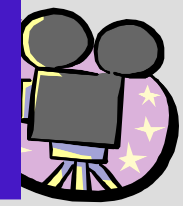
Tradeoff of Layered Coding

Layered coding schemes incur an increase in bit rate or decrease in video quality in comparison to a single-layer codec of equivalent quality.



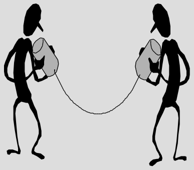
Percentage increase of total bit rate versus percentage base layer of total bit rate

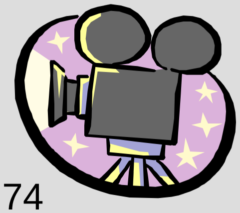




Scalable Coding & ESs

- Each layer is coded into an individual ES with unique ES_ID
- All layers belonging to the same object (i.e,. all ES that refer to the same object) are placed in the same Object Descriptor with its unique OD_ID

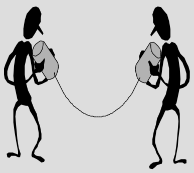


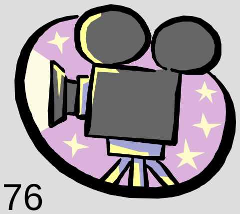


Example of Layered Spatial and Shape Coding

Part 4

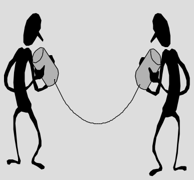
1. QoS Issues
2. Conclusions

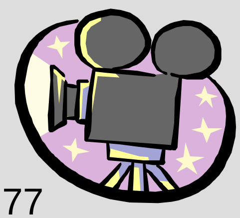




1. QoS Issues

- Need to understand the application
- Need to understand its usage
- Need to understand its content
- Need to understand its versatility
- Need to understand quality tradeoffs
- Need to familiarize ourselves with resilience of the data, recovery and control mechanisms
- Finally we can discuss QoS





2. Conclusions

We have still a long way to go -

- Layered coding combined with shape coding shows promise
- Multi streaming not supported over IP
- Quality tools not there yet
- Pricing/Quality trade offs have to be defined
- Finally: Guarantees of Service are required as Best Effort does not work well!