Introduction to Multiple-Description Coding

a Joint Source-Channel Coding Paradigm with a Twist Towards Digital TV

Jan Østergaard

Multimedia Information and Signal Processing
Department of Electronic Systems
Aalborg University

Universidad Técnica Federico Santa María, April 2010

Appetizer - Candy for the Ear & Sweets for the Mind

Appetizer - Candy for the Ear & Sweets for the Mind

- It's about time we all enjoy Eric Clapton!

Appetizer - Candy for the Ear & Sweets for the Mind

- It's about time we all enjoy Eric Clapton!
- Let's enjoy the music when using conventional audio compression techniques and streaming over a lossy packet-switched network and well — say the packet loss rate is 30%

Appetizer - Candy for the Ear & Sweets for the Mind

- It's about time we all enjoy Eric Clapton!
- Let's enjoy the music when using conventional audio compression techniques and streaming over a lossy packet-switched network and well — say the packet loss rate is 30%
- It is not particular pleasent for the ear

Appetizer - Candy for the Ear & Sweets for the Mind

- It's about time we all enjoy Eric Clapton!
- Let's enjoy the music when using conventional audio compression techniques and streaming over a lossy packet-switched network and well — say the packet loss rate is 30%
- It is not particular pleasent for the ear
- Thus, we should definitely not try to do near *real-time* audio streaming over lossy networks

Appetizer - Candy for the Ear & Sweets for the Mind

- It's about time we all enjoy Eric Clapton!
- Let's enjoy the music when using conventional audio compression techniques and streaming over a lossy packet-switched network and well — say the packet loss rate is 30%
- It is not particular pleasent for the ear
- Thus, we should definitely not try to do near *real-time* audio streaming over lossy networks
- Or should we???

Outline

Outline

The Digital Switch

- All over the world today, terrestrial broadcast of analogue TV is being replaced by broadcast of digital TV

The Digital Switch

- All over the world today, terrestrial broadcast of analogue TV is being replaced by broadcast of digital TV
- The main motivation is the promise of more TV channels

The Digital Switch

- All over the world today, terrestrial broadcast of analogue TV is being replaced by broadcast of digital TV
- The main motivation is the promise of more TV channels
- One analogue TV channel occupies approximately 7-8 MHz bandwidth of the available (UHF) frequency spectrum

The Digital Switch

- All over the world today, terrestrial broadcast of analogue TV is being replaced by broadcast of digital TV
- The main motivation is the promise of more TV channels
- One analogue TV channel occupies approximately 7-8 MHz bandwidth of the available (UHF) frequency spectrum
- In comparison, when converted into a “digital” channel, this bandwidth corresponds to approximately 20 Mbits per second

The Digital Switch

- All over the world today, terrestrial broadcast of analogue TV is being replaced by broadcast of digital TV
- The main motivation is the promise of more TV channels
- One analogue TV channel occupies approximately 7-8 MHz bandwidth of the available (UHF) frequency spectrum
- In comparison, when converted into a “digital” channel, this bandwidth corresponds to approximately 20 Mbits per second
- A raw standard TV signal (PAL) has a resolution of 720×576 pixels and a frame rate of 25 frames per second

The Digital Switch

- All over the world today, terrestrial broadcast of analogue TV is being replaced by broadcast of digital TV
- The main motivation is the promise of more TV channels
- One analogue TV channel occupies approximately 7-8 MHz bandwidth of the available (UHF) frequency spectrum
- In comparison, when converted into a "digital" channel, this bandwidth corresponds to approximately 20 Mbits per second
- A raw standard TV signal (PAL) has a resolution of 720×576 pixels and a frame rate of 25 frames per second
- Each pixel color is represented by at least 8 bits, and there are three color components Red, Green, and Blue

$$\text{SDTV} \quad 720 \times 576 \times 25 \times 3 \times 8 \approx 248 \text{ Mbits/sec}$$

$$\text{HDTV} \quad 1920 \times 1080 \times 25 \times 3 \times 8 \approx 1.2 \text{ Gbits/sec}$$

The Digital Switch

- All over the world today, terrestrial broadcast of analogue TV is being replaced by broadcast of digital TV
- The main motivation is the promise of more TV channels
- One analogue TV channel occupies approximately 7-8 MHz bandwidth of the available (UHF) frequency spectrum
- In comparison, when converted into a “digital” channel, this bandwidth corresponds to approximately 20 Mbits per second
- A raw standard TV signal (PAL) has a resolution of 720×576 pixels and a frame rate of 25 frames per second
- Each pixel color is represented by at least 8 bits, and there are three color components Red, Green, and Blue

$$\text{SDTV} \quad 720 \times 576 \times 25 \times 3 \times 8 \approx 248 \text{ Mbits/sec}$$

$$\text{HDTV} \quad 1920 \times 1080 \times 25 \times 3 \times 8 \approx 1.2 \text{ Gbits/sec}$$

- Let’s face it — we need video compression!

Lossy Image Compression - Source Coding

- 8 bit gray-scale image 512 $\times$ 512 pixel: Bit-rate is 2 Mbit

Lossy Image Compression - Source Coding

- Represent the image by the MSB: Bit-rate is 0.25 Mbit

Lossy Image Compression - Source Coding

- Represent the image by the two MSB's: Bit-rate is 0.5 Mbit

Lossy Image Compression - Source Coding

- Coded using JPEG: Bit-rate is 0.1 Mbit $\approx$ 0.38 bits/pixel

Outline

Compression for Storage and Retrieval not Streaming

- MP3, JPEG, MPEG2, MPEG4, etc. are meant for storage and retrieval of audio, images, and video
- Long delay (allowed to work on the entire audio or video file)
- Are these compression formats suitable for real-time streaming?
- Not if there's gonna be varying delay, varying bandwidth, packet losses

(recall poor Clapton?)

Outline

1. Introduction to Compression
   - Motivation for Compression
2. Compression Meets the Network
   - Coding without Network Awareness
   - Coding with Network Awareness
3. Multiple-Description Coding
   - Theory
   - Practice
4. Conclusions

What if the Network is not Ideal?

- Packet loss concealment
- Use back-channel (feedback) and request lost packets
- Channel Coding
- Hybrid ARQ
- Link Layer Adaptation
- Multi Resolution Coding

- *Analogue Coding, Space-Time Coding, Super Position Coding (Onion Peeling), Coset Coding*

- Multiple-Description Coding

Packet loss concealment

- Replace the lost packet by “something clever”

1) Repeat previous packet (GSM)
2) Extrapolate from previous packet (hard to do)
3) Wait for next packet and interpolate

Packet loss concealment

- Replace the lost packet by “something clever”

1) Repeat previous packet (GSM)
2) Extrapolate from previous packet (hard to do)
3) Wait for next packet and interpolate

In digital tv in Europe: Mute sound and freeze picture!

Packet loss concealment

- Replace the lost packet by “something clever”

1) Repeat previous packet (GSM)
2) Extrapolate from previous packet (hard to do)
3) Wait for next packet and interpolate

In digital tv in Europe: Mute sound and freeze picture!

- Pros & Cons:

+ Does not cost any additional bits (done at the decoder)
- Usually requires a delay to work well
- Often hard to make a good substitution
- Doesn’t work for consecutive packet droppings

Feedback Channel

- Makes it possible to request lost or broken packets

- Pros & Cons:

+ Only need additional bits when there is a packet loss
- Requires at least a round trip delay (before arrival of the retransmitted packet)
- The retransmitted packet and the request may be lost
- In many “real-time” applications, delayed packets are often treated as being lost
- Does not work well in broadcast environments due to the feedback implosion problem
(loss of a single packet – leads to multiple requests)

Channel Coding

- In packet-switched networks (e.g., the Internet), packets are generally received error-free or not received at all
- Packet losses are mainly due to router congestions

- Forward Error Correction (FEC) Coding:
  Adds redundancy to combat channel “noise”
  —Parity Check Codes, (Hamming and Reed-Solomon)

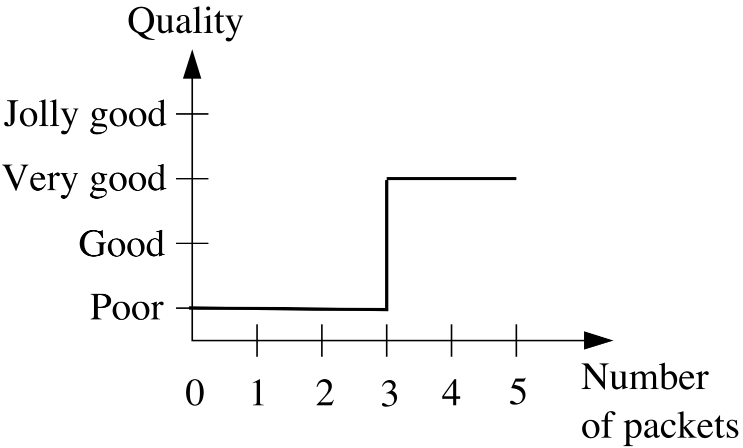
- Pros & Cons:

+ Does not need a back channel
- Severely increases the overall bit rate: Conservative
  —sender does not know the current channel conditions
- Usually requires a few packets delay to be efficient

Channel Coding (how does it work)

- (n, k) FEC erasure codes: $(n > k)$
- Takes k packets as input and outputs n packets
- An $(n, 1)$ erasure code is also called a repetition code
- On reception of less than k packets, the receiver cannot decode at all
- On reception of any k packets, the receiver can completely recover the original k packets
- On reception of more than k packets, there is no quality improvement
- Cliff-wall effect

Cliff-Wall Effect of a $(5,3)$ Erasure Code



- This is todays digital TV in Denmark!

Hybrid ARQ

- Partial use of error correcting codes and partial use of feedback
- Automatic Repeat reQuest (ARQ), re-transmit only lost packets, which cannot be recovered

- Pros & Cons:

\+ Works better than ARQ and ECC alone

+/- Smaller increase in bit-rate than for ECC

+/- Requires a feedback channel but less often than ARQ

\- Hard to find optimal trade-off

Link Layer Adaptation

- Utilizes erasure codes and a back-channel
- Predicts channel conditions for future transmissions
- Trade off bit-budget between source and channel coding
- If the channel will be poor, a strong (high bit-rate) channel coder and a weak (low bit-rate) source coder are used
- If the channel will be good, a high bit-rate source coder and a weak channel coder are used
- Example: AMR Speech Coder for GSM has bit rates of 12.2, 10.2, 7.95, 7.40, 6.70, 5.90, 5.15, and 4.75 kbit/sec. and the total bit rate after channel coding is 22.8 kbit/sec.

Link Layer Adaptation

- Pros & Cons:

+ The overhead is reduced a lot compared to FEC
—the channel condition is (partially) known at the sender
+ Possible to optimize the trade-off between source and channel coding (joint source-channel coding)
- Requires a system that is capable of (accurately) predicting the future channel conditions
- Requires a back channel (perhaps not from receiver)
- The channel must be slowly varying
- Does not work for multicast or broadcast since it is unlikely that all end-users experience identical channel conditions

Multi-Resolution Coding

- Falls under the category of scalable coding
- The signal is encoded into a base layer and a refinement layer
- The base layer yields tolerable audio/video quality
- The refinement layer is no good by itself
- Base + refinement layers ⇒ quality improvement
- Pros & Cons:

+ Allows for “online” quality scalability in a network

\- Requires prioritized networks, “smart” routers, or dedicated “refinement” channels
—everyone needs the base layer
—and only some the refinement layer!

Outline

Multiple-Description Coding

- We split the source into several descriptions:
  —the individual descriptions yield tolerable quality
  —and are able to refine each other!

Multiple-Description Coding

- We split the source into several descriptions:
  —the individual descriptions yield tolerable quality
  —and are able to refine each other!
- This is different from multi-resolution coding

Multiple-Description Coding

- We split the source into several descriptions:
  —the individual descriptions yield tolerable quality
  —and are able to refine each other!
- This is different from multi-resolution coding
- It is not straight-forward since:

Multiple-Description Coding

- We split the source into several descriptions:
  —the individual descriptions yield tolerable quality
  —and are able to refine each other!
- This is different from multi-resolution coding
- It is not straight-forward since:

1) a description is good if it “looks” like the source

Multiple-Description Coding

- We split the source into several descriptions:
  —the individual descriptions yield tolerable quality
  —and are able to refine each other!
- This is different from multi-resolution coding
- It is not straight-forward since:

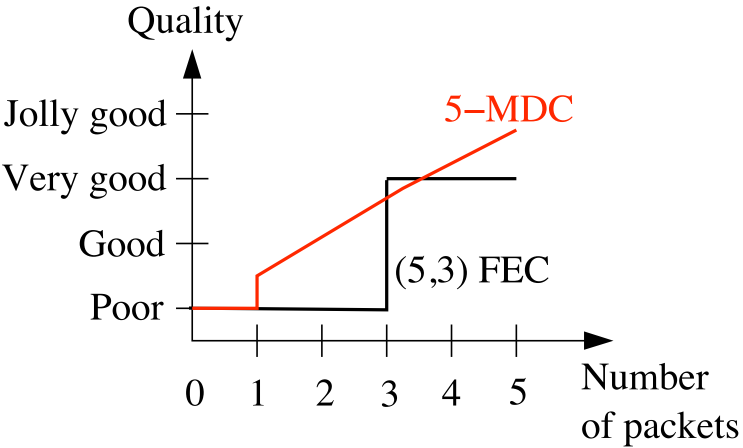
1) a description is good if it “looks” like the source
2) if two individual descriptions are both good
   —they must both “look” a lot like the source
   —thus, they “look” a lot like each other
   —hence, they cannot refine each other very well!

Multiple-Description Coding

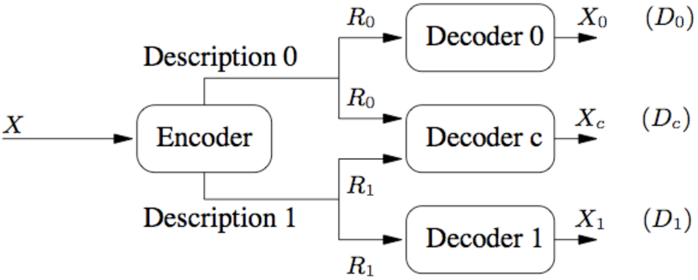
- We split the source into several descriptions:
  —the individual descriptions yield tolerable quality
  —and are able to refine each other!
- This is different from multi-resolution coding
- It is not straight-forward since:

1) a description is good if it “looks” like the source
2) if two individual descriptions are both good
   —they must both “look” a lot like the source
   —thus, they “look” a lot like each other
   —hence, they cannot refine each other very well!
3) Vice versa, if their joint description is good
   —at least one of the individual descriptions must be poor!

Motivation for MDC: No Cliff-Wall Effect!



Two-Channel Multiple-Description Coding



- With MDC — we ask the following questions:

1) What are the information theoretic R-D limits?
2) How do we approach these limits in practice?

The MD RDF Quadratic Gaussian (Memoryless)

- $R_0 \geq I(X;X_0) = R(D_0) = \frac{1}{2}\log_2(\sigma_X^2/D_0)$

The MD RDF Quadratic Gaussian (Memoryless)

- $R_0 \geq I(X; X_0) = R(D_0) = \frac{1}{2}\log_2(\sigma_X^2/D_0)$
- $R_1 \geq I(X; X_1) = R(D_1) = \frac{1}{2}\log_2(\sigma_X^2/D_1)$

The MD RDF Quadratic Gaussian (Memoryless)

- $R_0 \geq I(X; X_0) = R(D_0) = \frac{1}{2}\log_2(\sigma_X^2/D_0)$
- $R_1 \geq I(X; X_1) = R(D_1) = \frac{1}{2}\log_2(\sigma_X^2/D_1)$
- $R_0 + R_1 \geq I(X; X_0) + I(X; X_1) + I(X_0; X_1)$

The MD RDF Quadratic Gaussian (Memoryless)

- $R_0 \geq I(X;X_0) = R(D_0) = \frac{1}{2}\log_2(\sigma_X^2/D_0)$
- $R_1 \geq I(X;X_1) = R(D_1) = \frac{1}{2}\log_2(\sigma_X^2/D_1)$
- $R_0 + R_1 \geq I(X;X_0) + I(X;X_1) + I(X_0;X_1)$
  $= R(D_c) + \frac{1}{2}\log_2(\gamma) = \frac{1}{2}\log_2(\sigma_X^2/D_c) + \frac{1}{2}\log_2(\gamma)$

where

$$\gamma = \begin{cases} \mathbf{1}, & \boldsymbol{D_c < D_0 + D_1 - \sigma_X^2} \\ \frac{\sigma_X^2 D_c}{D_0 D_1}, & D_c > \left(\frac{1}{D_0} + \frac{1}{D_1} - \frac{1}{\sigma_X^2}\right)^{-1} \\ \frac{(\sigma_X^2 - D_c)^2}{(\sigma_X^2 - D_c)^2 - \left(\sqrt{(\sigma_X^2 - D_0)(\sigma_X^2 - D_1)} - \sqrt{(D_0 - D_c)(D_1 - D_c)}\right)^2}, & \text{o.w.,} \end{cases}$$

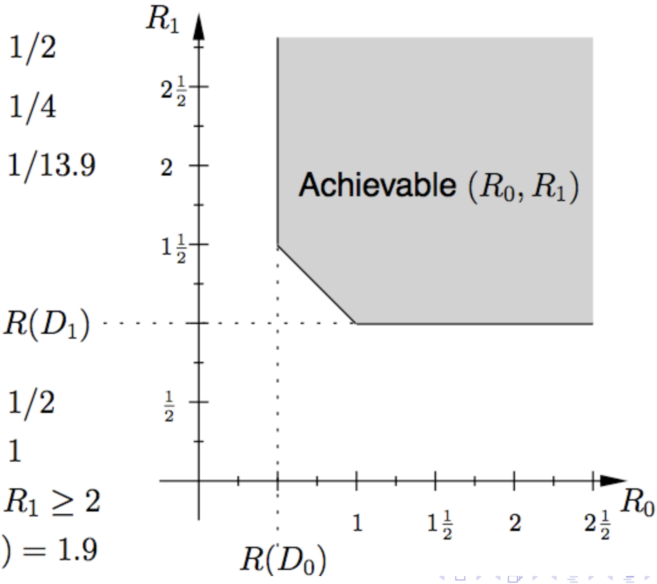
[Ozarow'80, Chen et al.'06]

MD Rate-Rate Region (Quadratic Gaussian)

$D_0 = 1/2$

$D_1 = 1/4$

$D_c = 1/13.9$



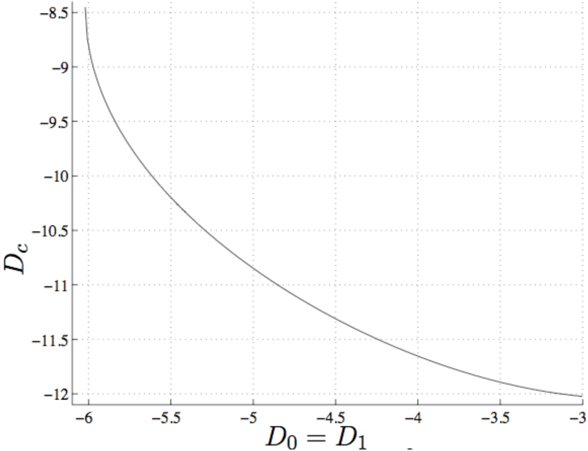
$R_0 \geq 1/2$

$R_1 \geq 1$

$R_0 + R_1 \geq 2$

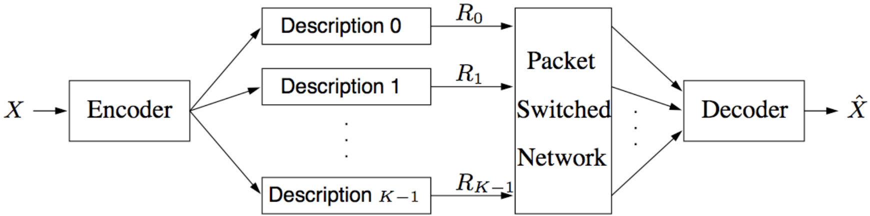
$R(D_c) = 1.9$

MD Distortion-Distortion Function (Q. Gaussian)



$R_0 = R_1 = 1$ bit/dimension. X is unit-variance and i.i.d. Gaussian.

Multiple-Description Coding for Many Descriptions



- Encode X into K descriptions
- Send descriptions over lossy network
- Distortion depends upon which descriptions are received
- Descriptions refine each other
- Very much an open problem in general

Outline

Practical Constructions of Multiple Description Codes

- The simple engineering approach:

Practical Constructions of Multiple Description Codes

- The simple engineering approach:

1) **Take any multi-resolution coder**

Practical Constructions of Multiple Description Codes

- The simple engineering approach:

1) **Take any multi-resolution coder**

Base layer	Refinement

2) **Duplicate the base layer**

Description 1 | Base layer

Description 2 | Base layer

Practical Constructions of Multiple Description Codes

- The simple engineering approach:

1) **Take any multi-resolution coder**

Base layer	Refinement

2) **Duplicate the base layer**

Description 1 | Base layer

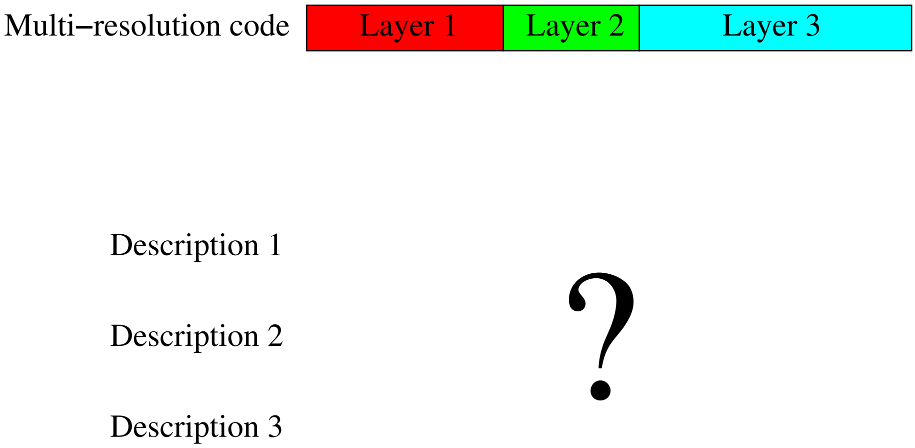
Description 2 | Base layer

3) **Split the refinement layer into two parts**

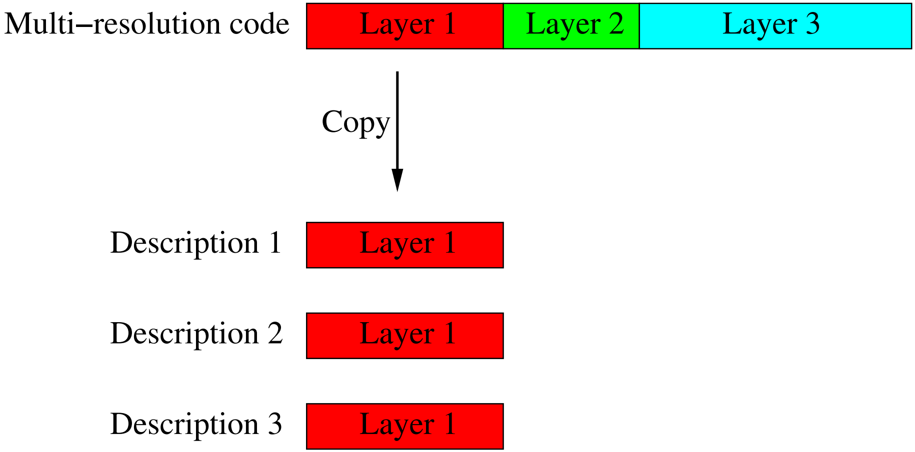
Description 1 | Base layer | Refine

Description 2 | Base layer | ment

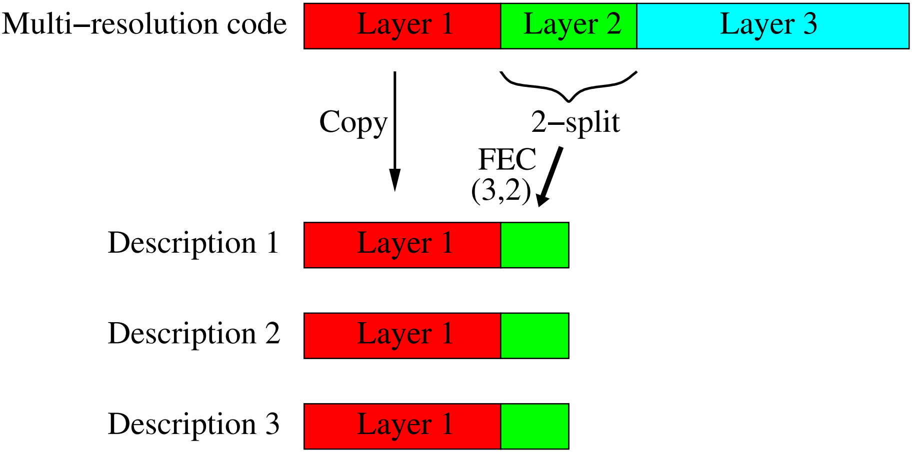
Practical Constructions of Multiple Description Codes



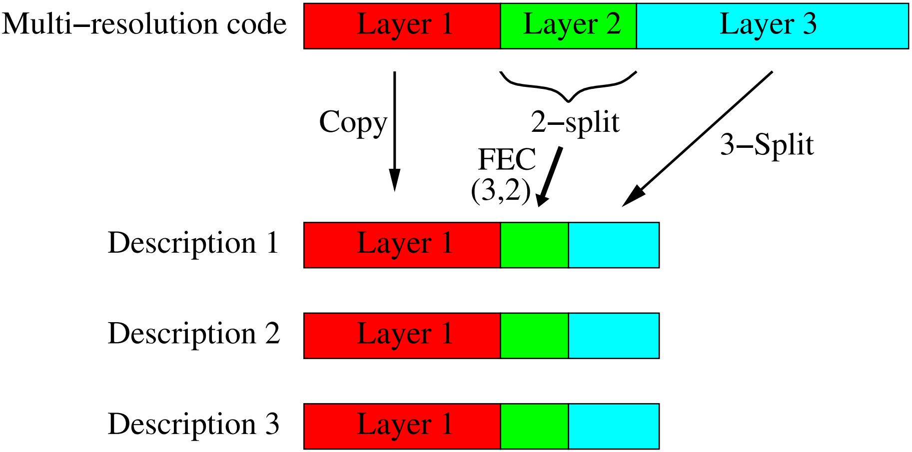
Practical Constructions of Multiple Description Codes



Practical Constructions of Multiple Description Codes



Practical Constructions of Multiple Description Codes



Practical Constructions of Multiple Description Codes

- The h.264 MPEG AVC (advanced video coding) standard used in digital TV, has recently been generalized to include scalable coding: h.264 MPEG SVC (scalable video coding)
- Thus, one can create a *prioritized* layered video stream, which by use of FEC's can be turned into an *unprioritized* multiple-description stream
- The h.264 AVC and h.264 SVC are both available for free online (C++ implementation)
- Erasure codes are also available standard components

Practical Constructions of Multiple Description Codes

- Erasure codes (FEC)
- Index-assignments (Ph.D. work)
- Quantization splitting
- Correlating transforms
- Frame expansions
- Redundant filterbanks
- Compressed Sensing
- ℓ_1-Compression (Tobias L. Jensen)
- Oversampling and Delta-Sigma Quantization

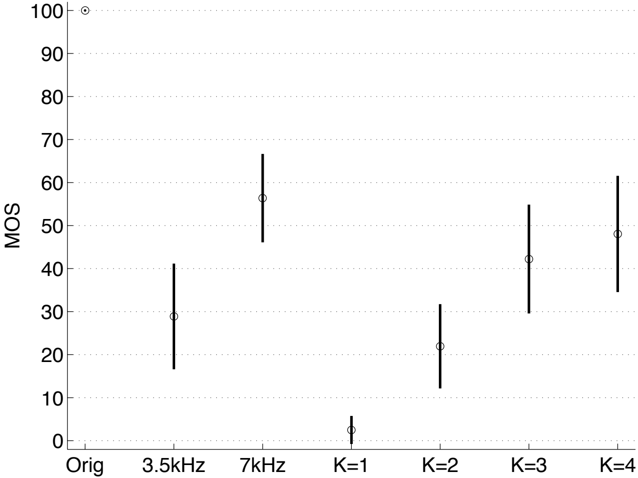
Back to the real-time streaming question

- We developed a transform coder, which resemble mp3 (MPEG-2 AAC)
- Sampling frequency of music is 48 kHz (mono)
- The coder is block based, each block beeing approx 20 ms. (1024 samples)
- We use multiple description coding: 1 to 4 descriptions
- Total target bit-rate is 96 kbit/sec. (2 bits/sample)

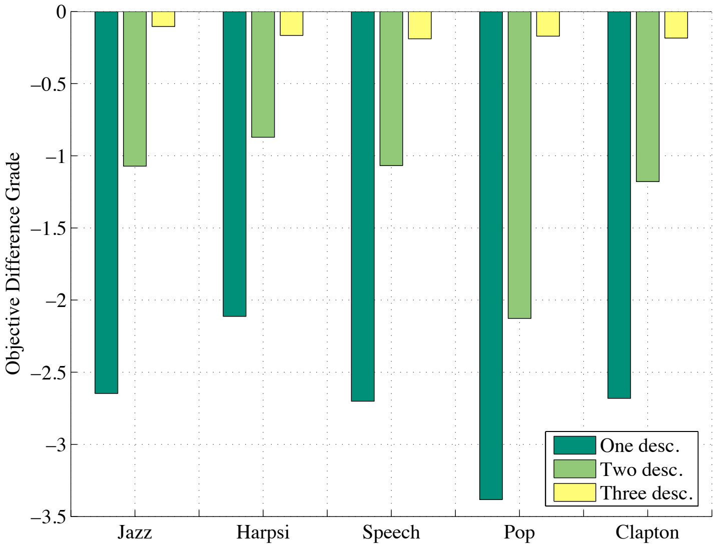
1 desc.	2 desc.	3 desc.	4 desc.
96 kbps	48 kpbs	32 kpbs	24 kbps

- The packet loss probability is 30%
- The losses are i.i.d.

30 % i.i.d. Packet losses (Eric Clapton) - MUSHRA



Objective Difference Grade



Conclusions

- **Yesterday**, we used to compress for storage and retrieval

Conclusions

- **Yesterday**, we used to compress for storage and retrieval
- **Today**, we compress to satisfy channel/network requirements

Conclusions

- **Yesterday**, we used to compress for storage and retrieval
- **Today**, we compress to satisfy channel/network requirements
- Message to take home:
- joint source-channel coding is getting more and more important
- multiple-description coding is one such technique, which achieves a certain robustness without compromising delay

Conclusions

- **Yesterday**, we used to compress for storage and retrieval
- **Today**, we compress to satisfy channel/network requirements
- Message to take home:
- joint source-channel coding is getting more and more important
- multiple-description coding is one such technique, which achieves a certain robustness without compromising delay
- **Tomorrow:**

Conclusions

- **Yesterday**, we used to compress for storage and retrieval
- **Today**, we compress to satisfy channel/network requirements
- Message to take home:
- joint source-channel coding is getting more and more important
- multiple-description coding is one such technique, which achieves a certain robustness without compromising delay

- **Tomorrow:**

— The Anillo Project —

joint source-channel coding and ***control***