LORD: LOw-complexity, Rate-controlled, Distributed video coding system

Rami Cohen and David Malah

Signal and Image Processing Lab
Department of Electrical Engineering
Technion - Israel Institute of Technology

26 November, 2012

SITIS 2012
Naples, Italy
Outline

1 Distributed Video Coding (DVC)
   - Why DVC?
   - DVC Systems - Overview
1 Distributed Video Coding (DVC)
   - Why DVC?
   - DVC Systems - Overview

2 LORD: LOw-complexity, Rate-controlled, Distributed video coding system
   - Motivation
   - Encoder
   - Decoder
   - Adaptation to Endoscopy Video Compression
Outline

1. Distributed Video Coding (DVC)
   - Why DVC?
   - DVC Systems - Overview

2. LORD: LOw-complexity, Rate-controlled, Distributed video coding system
   - Motivation
   - Encoder
   - Decoder
   - Adaptation to Endoscopy Video Compression

3. Experimental Results
Outline

1. Distributed Video Coding (DVC)
   - Why DVC?
   - DVC Systems - Overview

2. LORD: LOw-complexity, Rate-controlled, Distributed video coding system
   - Motivation
   - Encoder
   - Decoder
   - Adaptation to Endoscopy Video Compression

3. Experimental Results

4. Conclusion
Why DVC?

- There are cases in which standard (complex) encoders are impractical
- DVC paradigm offers low complexity encoders with good performance

Limited-complexity video encoders: Examples
DVC Systems - Overview

Main Parts

- Usually, the input is separated into key (intra-coded) frames and Wyner–Ziv (WZ) frames (Group of Pictures (GOP) of size 2)
- Side information creation: prediction ($Y$) of the frame to be encoded ($X$), created at the decoder
- Noise prediction model: estimating $X$ from $Y$
  - Probability distribution models for $N = X - Y$
Motivation for developing LORD

Adaptation to the Video statistics
- On-line estimation of the parameters of the noise model

Rate control
- Not affected by the decoder
- Suitable for channels with constant rate constraint

Low delay
- No feedback-channel

Medical application
- Adaptation to the compression of endoscopy videos
Encoder

Block Classification
**COSET mode**

- The differences between the first 15 AC coeffs. are quantized uniformly to symmetric $2^m$ levels.
- The quantization indices are sent using Huffman code.
- The remaining coeffs. are INTRA-coded.

![Diagram showing COSET mode with quantization indices and codewords]
Rate-Distortion Optimization

- Assuming $P$ DCT bands in the GOP:

$$
\min_{b_i} D = \sum_{i=1}^{P} m_i h_i \sigma_i^2 2^{-2b_i}, \quad \text{s.t.} \quad \sum_{i=1}^{P} b_i \leq B
$$

- The solution:

$$
b_i = \bar{b} + \frac{1}{2} \log_2 \frac{\sigma_i^2}{\sigma^2} + \frac{1}{2} \log_2 \frac{h_i}{H} + \frac{1}{2} \log_2 \frac{m_i}{M}
$$

$$
\bar{b} = \frac{B}{P}, \quad \sigma^2 = \left( \prod_{i=1}^{P} \sigma_i^2 \right)^{1/P}, \quad H = \left( \prod_{i=1}^{P} h_i \right)^{1/P}, \quad M = \left( \prod_{i=1}^{P} m_i \right)^{1/P}
$$
Rate Control

- Rate control (RC) algorithm is used for enforcing the optimal bit distribution.
- The following relationship between the coding bit rate \( R \) and the fraction \( \rho \) of zeros among the quantized intra-coded coefficients is employed [He & Mitra, 2002] :

\[
R(\rho) = \theta (1 - \rho)
\]

- \( \theta \) is a constant related to the image content.
- The number of zeros is controlled by the parameter \( q \) used in JPEG.
RC algorithm: Example

- Low complexity, one-pass algorithm
LORD: Decoder

Scheme

- Intra-coded Key Frame: $\hat{X}_{2k-1}$
- Decoded Key Frame: $\hat{X}_{2k-1}$
- Previously decoded frame: $\hat{X}_{2k-1}, \hat{X}_{2k-2}$
- Current Quantized WZ frame: $\hat{X}_{2k}$
- MMSE Reconstruction
- Estimated WZ Frame in the DCT Domain
- IDCT

Steps:
1. JPEG-Like Decoder
2. Side Information Creation
3. Noise Correlation Modelling
4. MMSE Reconstruction
5. IDCT
Decoder

SI Creation: Motion Extrapolation

- Qpel (quarter-pixel) full search motion estimation is performed between two already decoded frames, $\hat{X}_{2k-2}$ and $\hat{X}_{2k-1}$
- Assuming linear motion, the pixels from $\hat{X}_{2k-1}$ are projected to the next (extrapolated) frame, which is used as side information
Noise Prediction Model

- The noise model is used for ”improving” the basic estimate provided by the SI
  - Estimated on-line, for each WZ frame
- The noise (\( N \)) between \( \hat{X}_{2k-2} \) and \( \hat{X}_{2k-1} \) serves as an estimate of the noise between the SI (\( Y \)) and the WZ frame (\( X \))
- \( N \) is assumed to be Laplace-distributed:
  \[
  f_{X|Y}(x) = f_N(x - y) = \frac{\alpha}{2} e^{-\alpha |x - y|}
  \]
- \( \alpha \) is calculated for each band, using the ML estimator
Decoder

MMSE Reconstruction

- The boundaries of the quantization interval (denoted $z_i$ and $z_{i+1}$) of the COSET coeffs. are provided by encoder.
- An MMSE estimate of the source $X$, using both these boundaries and the side information, is obtained:

$$\hat{x} = \mathbb{E} [x | x \in [z_i, z_{i+1}), y] = \frac{\int_{z_i}^{z_{i+1}} x f_{X|y} (x) \, dx}{\int_{z_i}^{z_{i+1}} f_{X|y} (x) \, dx}$$
Endoscopy videos

Endoscopy

- Endoscopy refers to looking inside the body for medical reasons using an endoscope
- An endoscope is consisted of a long, thin, flexible tube that has a light source and an attached camera
- Recently, a shift towards transmission of endoscopy videos over a wireless channel - limited power resources
Bayer Filter

- Bayer color filter array (CFA) is composed of filter blocks of size 2x2, which are 50% green, 25% red and 25% blue.
- Almost universal on consumer digital cameras, used in endoscopes.
Endoscopy Videos

Rate-Distortion Optimization

- The distortion is calculated separately for each color component.
- Each distortion is weighted according to $w_C^2$ ($C = R, G, B$):

$$D = w_R^2 \cdot \sum_{DCT\ bands} m_{Ri} h_i \sigma_i^2 2^{-2b_i} + w_G^2 \cdot \sum_{DCT\ bands} m_{Gi} h_i \sigma_i^2 2^{-2b_i}$$

- Distortion from R component

$$+ w_B^2 \cdot \sum_{DCT\ bands} m_{Bi} h_i \sigma_i^2 2^{-2b_i}$$

- Distortion from G component

- Distortion from B component

- Considering the available bits $B$, we get an optimization problem.
- The solution is a simple extension of the previous one.
Experimental Results

Standard Videos

PSNR Results

Foreman, QCIF, 15Fps

Coastguard, QCIF, 15Fps
Endoscopy Videos

PSNR Results

Simulation - chicken, 368x480, 28Fps

Simulation - gastrointestinal tract, 240x320, 28Fps
Summary

- New DVC codec was developed
- On-line estimation of the parameters of the noise model
- Rate-distortion model and rate control algorithm are used, at the encoder
- No feedback channel is used
- Adaptation to endoscopy videos (Bayer format)
- Improvement over standard intra coding, for both standard videos and endoscopy videos
Thank You