Distributed Video Coding Applications for Scalable and Flexible Decoding

Ngai-Man (Man) CHEUNG
Huisheng WANG
Antonio ORTEGA

Signal and Image Processing Institute
University of Southern California
Los Angeles, CA, USA

Content

This talk will cover two DSC applications:

- **Scalable video coding**
  - Error robust scalable coding
  - Low complexity encoding for multiple layers coding
- **Flexible video decoding**
  - Viewpoint switching in multiview video
  - Robust video transmission

USC Viterbi School of Engineering

DSC Application to Scalable Video Coding

USC

Some Work on Scalable Coding Based on DSC

USC Viterbi School of Engineering

- [Xu, Xiong; VCIP 04, IEEE Trans. IP 06]
  - Embedded enhancement layer similar to MPEG-4/H.26L FGS
  - Robust to error in base layer
  - Do not suffer performance loss due to layering
- [Tagliasacchi, Majumdar, Ramchandran, Tubaro; PCS 04, Eurasip SP 06]
  - Spatial/temporal/SNR scalability
  - Based on PRISM [Puri, Ramchandran; Allerton 02]
  - Robust to channel losses
  - Flexible distribution of complexity
- [Sehgal, Jagmohan, Ahuja; PCS 04]
  - Multiple Wyner-Ziv encoded versions for different possible predictors
  - Encoder streams an appropriate encoded version based on knowledge of predictors available at decoder
- [Wang, Cheung, Ortega; PCS 04, Eurasip JASP 06]
  - Improvement of MPEG4-FGS: Utilizing EL reconstruction for prediction
  - Low complexity overhead: Avoid replicating EL reconstruction at encoder

USC

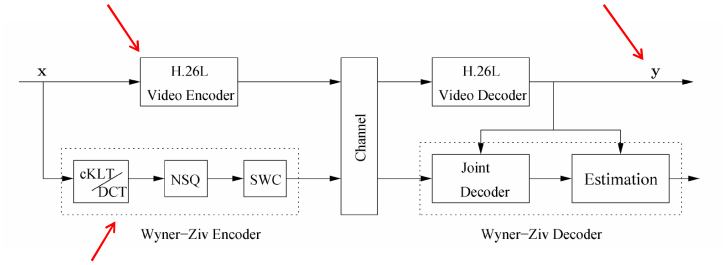
Error Robust Scalable Coding

USC Viterbi School of Engineering

[Xu, Xiong; VCIP 04, IEEE Trans. IP 06]

Base layer (BL): Standard video codec

Side Information (SI): BL reconstruction



Enhancement layer (EL) based on bit-plane coding and DSC

Error corrupted base layer (SI) may still be used for DSC decoding

USC

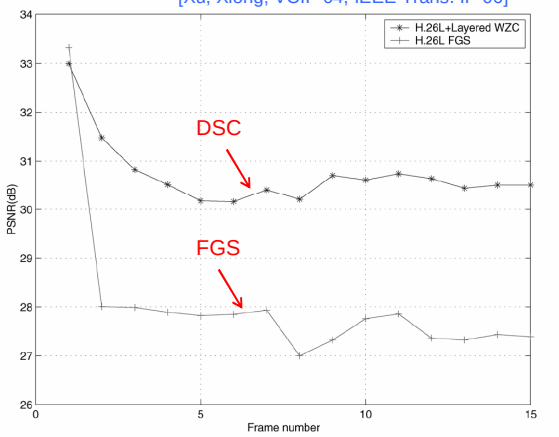
Robust to Transmission Error

USC Viterbi School of Engineering

DSC-based system can achieve substantial improvement in situations with transmission error

[Xu, Xiong; VCIP 04, IEEE Trans. IP 06]

Football
1% macroblock loss at BL
BL: 1450 kbps
EL: 200 kbps



In addition, the system does not suffer performance loss due to layering, i.e., same performance as the monolithic WZ coding

USC

Some Work on Scalable Coding Based on DSC

USC Viterbi School of Engineering

- [Xu, Xiong; VCIP 04, IEEE Trans. IP 06]
  - Embedded enhancement layer similar to MPEG-4/H.26L FGS
  - Robust to error in base layer
  - Do not suffer performance loss due to layering
- [Tagliasacchi, Majumdar, Ramchandran, Tubaro; PCS 04, Eurasip SP 06]
  - Spatial/temporal/SNR scalability
  - Based on PRISM [Puri, Ramchandran; Allerton 02]
  - Robust to channel losses
  - Flexible distribution of complexity
- [Sehgal, Jagmohan, Ahuja; PCS 04]
  - Multiple Wyner-Ziv encoded versions for different possible predictors
  - Encoder streams an appropriate encoded version based on knowledge of predictors available at decoder
- [Wang, Cheung, Ortega; PCS 04, Eurasip JASP 06]
  - Improvement of MPEG4-FGS: Utilizing EL reconstruction for prediction
  - Low complexity overhead: Avoid replicating EL reconstruction at encoder

USC

Spatial/Temporal/SNR Scalability

USC Viterbi School of Engineering

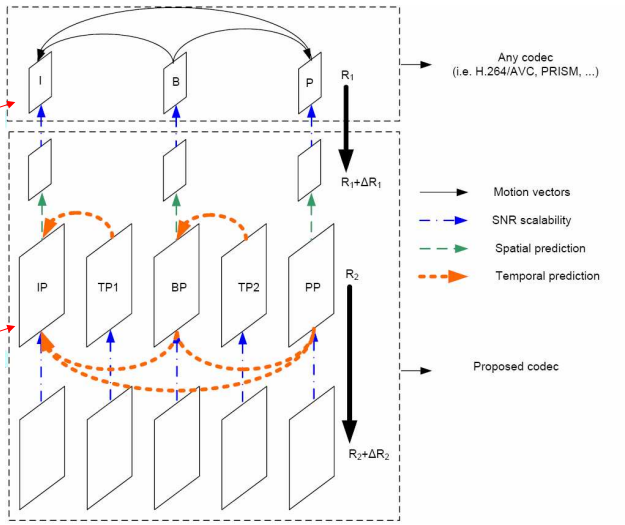
[Tagliasacchi, Majumdar, Ramchandran, Tubaro; PCS 04, Eurasip SP 06]

BL: Standard video codec/PRISM

Motion estimation (ME) at BL resolution: low complexity

EL: PRISM

Encoder does not perform ME at EL: Use BL motion information for correlation estimation



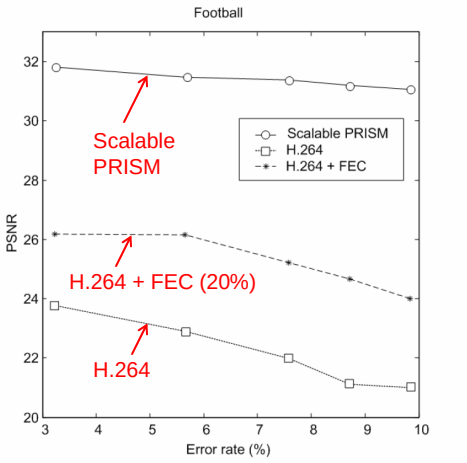
USC

Robust to Transmission Error

USC Viterbi School of Engineering

[Tagliasacchi, Majumdar, Ramchandran; PCS 04]

Football, CIF, 15fps, 1800 kbps
Packet dropped at both EL and BL



USC

Some Work on Scalable Coding Based on DSC

USC Viterbi School of Engineering

- [Xu, Xiong; VCIP 04, IEEE Trans. IP 06]
  - Embedded enhancement layer similar to MPEG-4/H.26L FGS
  - Robust to error in base layer
  - Do not suffer performance loss due to layering
- [Tagliasacchi, Majumdar, Ramchandran, Tubaro; PCS 04, Eurasip SP 06]
  - Spatial/temporal/SNR scalability
  - Based on PRISM [Puri, Ramchandran; Allerton 02]
  - Robust to channel losses
  - Flexible distribution of complexity
- [Sehgal, Jagmohan, Ahuja; PCS 04]
  - Multiple Wyner-Ziv encoded versions for different possible predictors
  - Encoder streams an appropriate encoded version based on knowledge of predictors available at decoder
- [Wang, Cheung, Ortega; PCS 04, Eurasip JASP 06]
  - Improvement of MPEG4-FGS: Utilizing EL reconstruction for prediction
  - Low complexity overhead: Avoid replicating EL reconstruction at encoder

USC

Some Work on Scalable Coding Based on DSC

USC Viterbi School of Engineering

- [Xu, Xiong; VCIP 04, IEEE Trans. IP 06]
  - Embedded enhancement layer similar to MPEG-4/H.26L FGS
  - Robust to error in base layer
  - Do not suffer performance loss due to layering
- [Tagliasacchi, Majumdar, Ramchandran, Tubaro; PCS 04, Eurasip SP 06]
  - Spatial/temporal/SNR scalability
  - Based on PRISM [Puri, Ramchandran; Allerton 02]
  - Robust to channel losses
  - Flexible distribution of complexity
- [Sehgal, Jagmohan, Ahuja; PCS 04]
  - Multiple Wyner-Ziv encoded versions for different possible predictors
  - Encoder streams an appropriate encoded version based on knowledge of predictors available at decoder
- [Wang, Cheung, Ortega; PCS 04, Eurasip JASP 06]
  - Improvement of MPEG4-FGS: Utilizing EL reconstruction for prediction
  - Low complexity overhead: Avoid replicating EL reconstruction at encoder

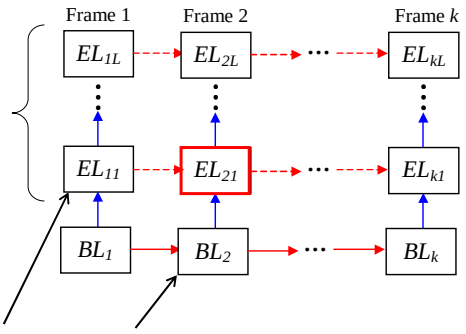
USC

Scalable Coding with Multiple Motion-Compensated Prediction (MCP) Loops

USC Viterbi School of Engineering

Multiple MCP loops approach
[Rose, Regunathan; IEEE Trans. IP 2001]

Encoder replicates all the EL reconstructions



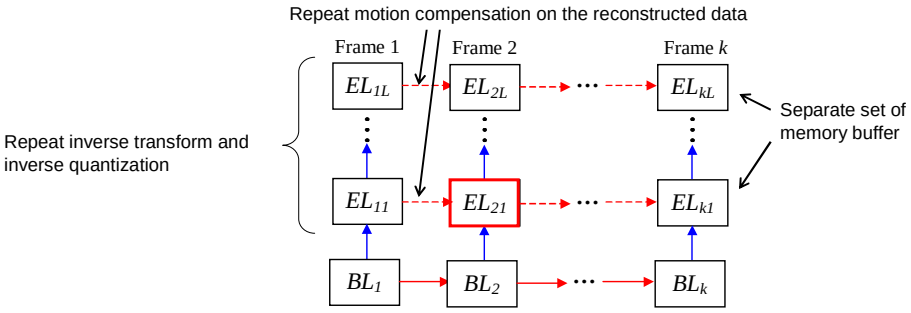
Use both BL and EL for prediction: good coding efficiency

USC

Multiple MCP Loops Approach: Non-trivial Complexity

USC Viterbi School of Engineering

In multiple loops approach, the complexity of replicating all the EL reconstructions could be non-trivial in multiple layers coding



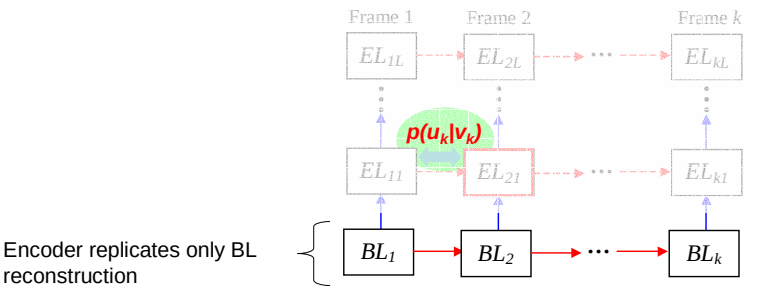
Our contribution: DSC counter-part to the multiple loops approach to achieve low complexity encoding

USC

Wyner-Ziv Scalable Video Coding (WZS) – Overview

USC Viterbi School of Engineering

WZ scalable coding: Encoder does not replicate EL reconstruction
- EL predictor available only at decoder
- Side Information: EL predictor



Key point: DSC requires only the correlation information instead of the exact reconstructed data in encoding

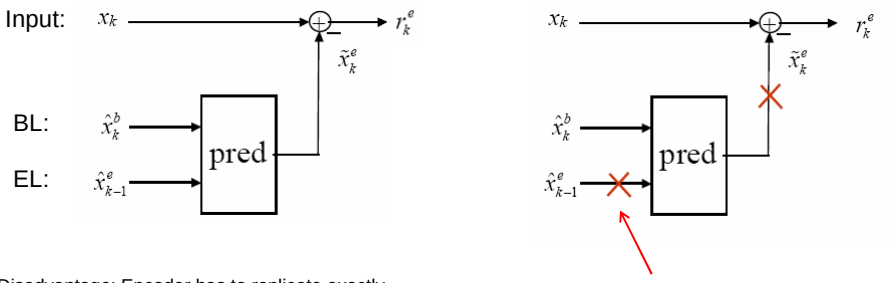
USC

WZS: Encoder does not replicate EL reconstruction

USC Viterbi School of Engineering

Multiple MCP loops

Wyner-Ziv Scalable



Disadvantage: Encoder has to replicate exactly all the possible EL reconstructions, so that the prediction residue can be computed

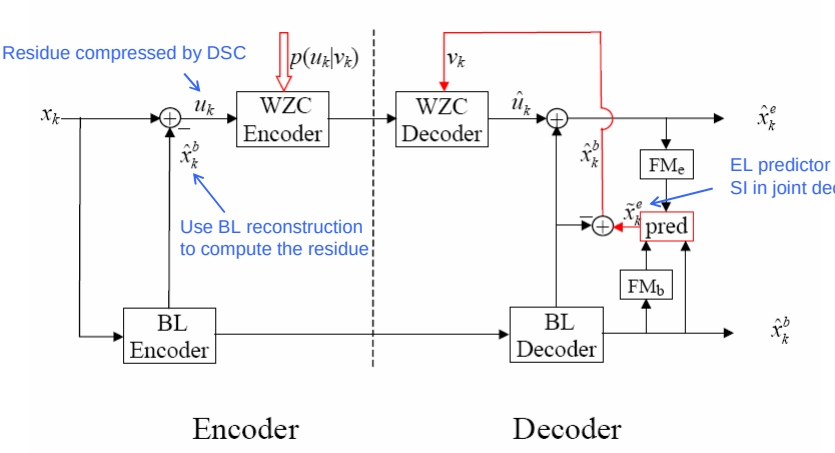
Encoder does not replicate EL reconstruction

USC

15

Cast scalable coding as a Wyner-Ziv problem

USC Viterbi School of Engineering



How to estimate $p(u_k|v_k)$?

USC

16

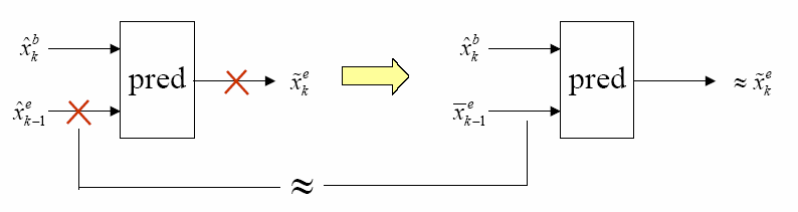
Correlation Estimation at Encoder

USC Viterbi School of Engineering

- Problem: How to estimate the correlation?
- Solution
  - Computing an approximation of the EL reconstruction at the encoder
  - Use approximation for correlation estimation

$$\overline{x}^e_{k-1} = Q_e(x_{k-1}) \approx \hat{x}^e_{k-1}$$

Compute the approximation by quantizing original frame to a quality level similar to EL



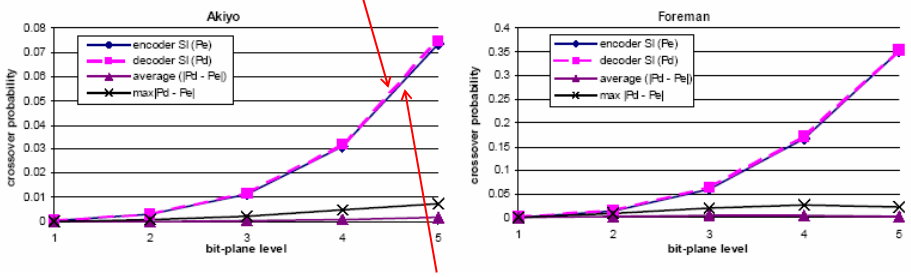
USC

Correlation information can be estimated with low complexity

USC Viterbi School of Engineering

Our approximation can provide accurate statistics for DSC encoding

Correlation computed from EL reconstruction (available only at decoder)



Correlation estimated from the approximation (at encoder)

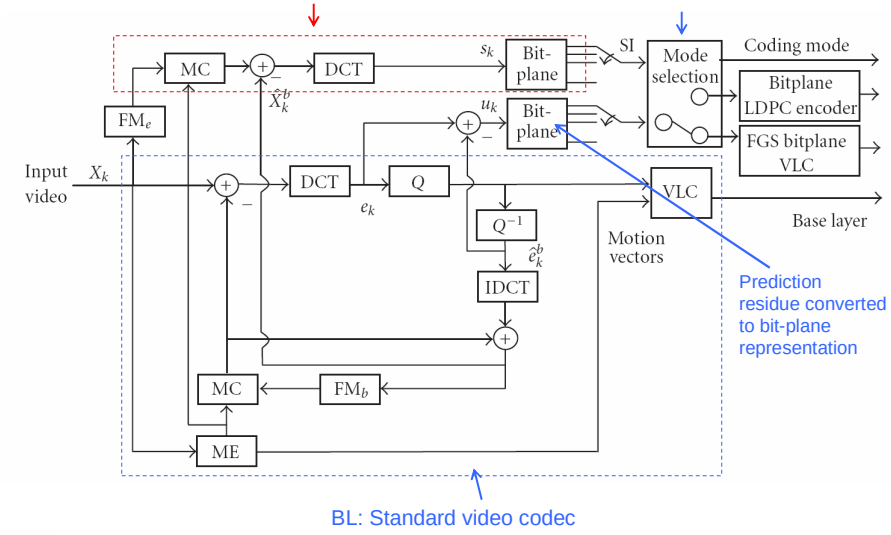
USC

WZS Encoder Block Diagram

USC Viterbi School of Engineering

Process original frame to estimate the correlation
(Encoder has to only do this once, even for multiple layers)

Bit-plane compressed by DSC
or FGS bit-plane coding

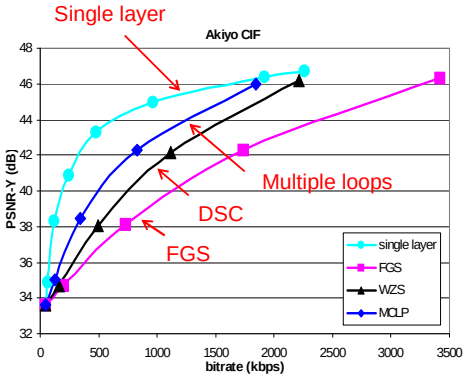


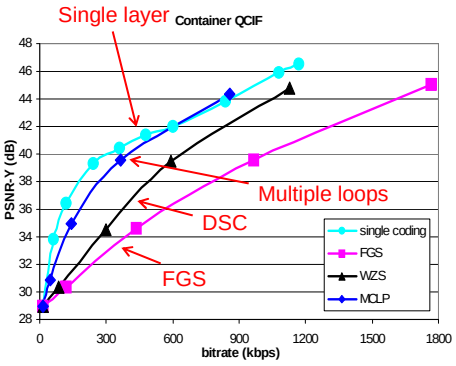
BL: Standard video codec

Results

USC Viterbi School of Engineering

Our proposed algorithm out-performs MPEG-4 FGS





Summary: DSC Application to Scalable Coding

USC Viterbi School of Engineering

- Robust to transmission errors
  - Corrupted SI (BL or spatial/temporal predictor) may still be used for DSC decoding
- Low complexity scalable encoding
  - Motion estimation at BL resolution (spatially down-sampled)
- Improvement to MPEG-4 FGS
  - Exploring temporal correlation at EL
  - Apply DSC techniques to reduce complexity overhead

USC

USC Viterbi School of Engineering

DSC Application to Flexible Video Decoding

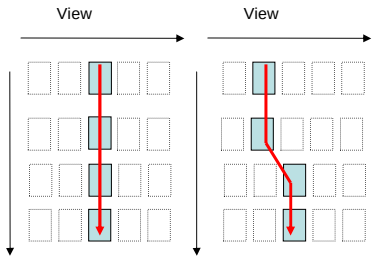
USC

Flexible Video Decoding

USC Viterbi School of Engineering

Apply DSC to generate a *single* bitstream that can be decoded in *several* different ways

- Free viewpoint switching in multiview video compression
- Forward and backward video playback
- Robust video transmission



USC

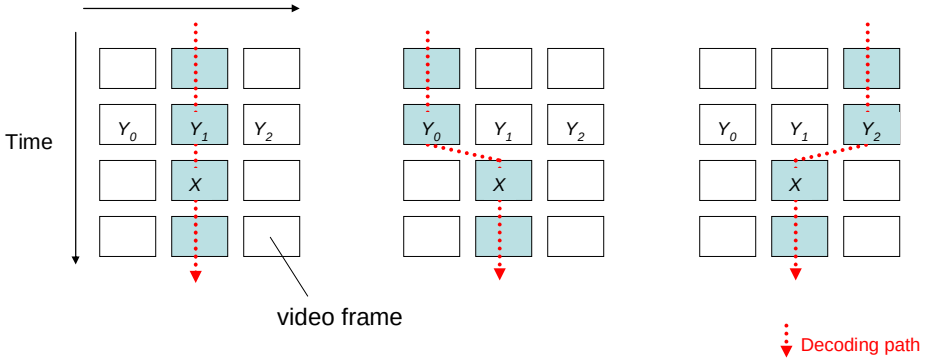
Flexible Decoding Example – Free Viewpoint Switching in Multiview Video

USC Viterbi School of Engineering

[Cheung, Ortega; PCS 07]

User selects a particular view: View

User switches between different views during playback:



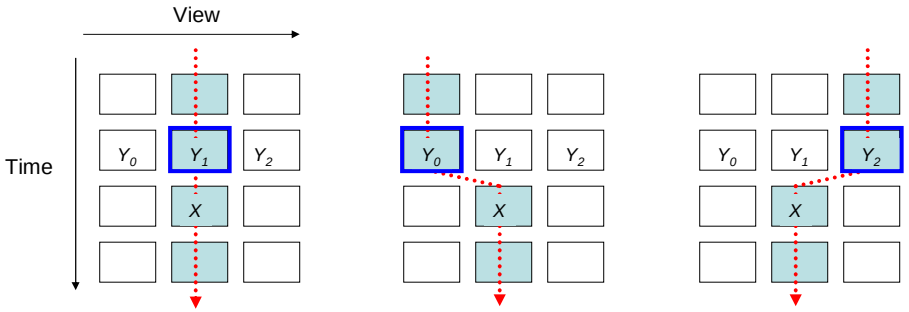
In order to address viewpoint switching, compression schemes have to support **multiple decoding paths**

USC

Free viewpoint switching poses challenges to multiview compression

USC Viterbi School of Engineering

When users can choose among different decoding paths, it is not clear which previous reconstructed frame will be available to use in the decoding



Multiple decoding paths ⇒ Either Y_0 or Y_1 or Y_2 will be available at decoder ⇒ Uncertainty on predictor status at decoder!!!!

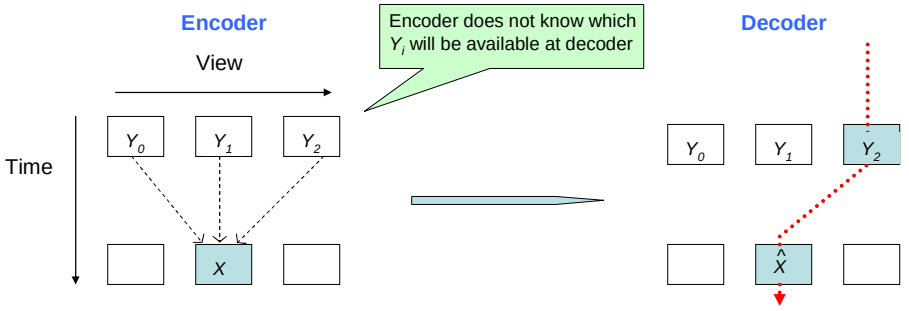
USC

Problem Formulation

USC Viterbi School of Engineering

To support low-delay free viewpoint switching (flexible decoding), encoder needs to operate under uncertainty on decoder predictor

[Cheung, Ortega; MMSP 07]



- Assume feedback is not available
  - Low-delay, interactive application
  - Offline encoding of multiview data

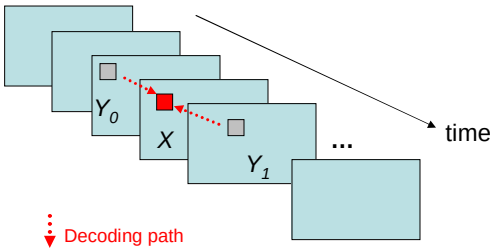
USC

Other Flexible Decoding Examples – Forward/Backward Frame-by-frame Video Playback

USC Viterbi School of Engineering

[Cheung, Wang, Ortega; VCIP 06]

User can choose to play back in either direction:
Either past **or** future reconstructed frame will be available at decoder



Y_0, Y_1: best motion-compensated predictor for X

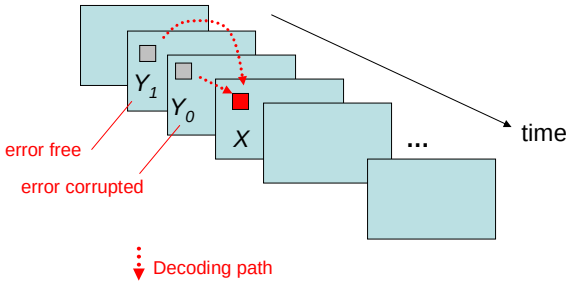
- Not B-frame (as in video coding standards)
- Not multiple reference frames

USC

Other Flexible Decoding Examples – Robust Video Transmission

USC Viterbi School of Engineering

[Wang, Prabhakaran, Ramchandran; ICIP 06]



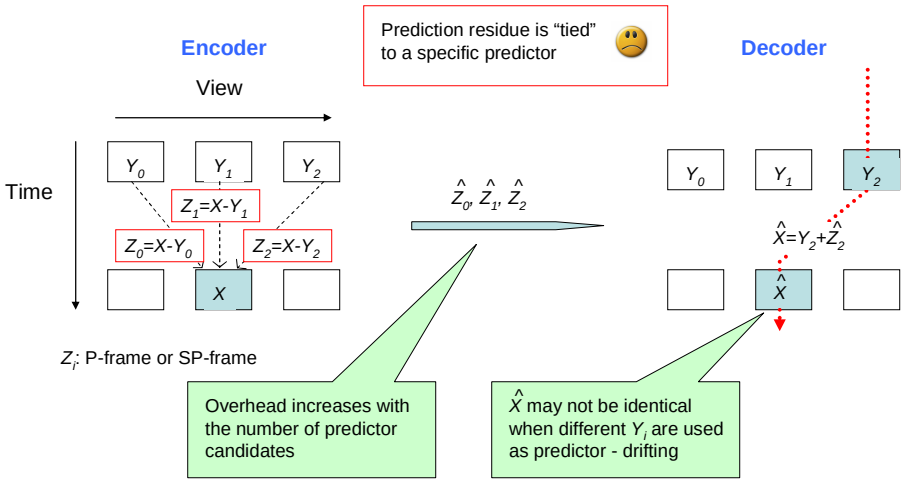
Some reference frames have error, but encoder does not know which one (assume feedback is not available)

USC

Address Viewpoint Switching/Flexible Decoding Within Closed-Loop Predictive (CLP) Coding Framework

USC Viterbi School of Engineering

Encoder has to send **multiple** prediction residues to the decoder



USC

Our Contributions

USC Viterbi School of Engineering

- Propose distributed source coding (DSC) based coding algorithm to address viewpoint switching (flexible decoding)
- **Apply DSC to generate a *single* bitstream that can be decoded in *several* different ways**
- Propose an efficient encoding algorithm
  - Macroblock mode
  - Significance coding
  - Correlation model
  - Minimum MSE dequantization

USC

Related Work

USC Viterbi School of Engineering

Differences between our work on viewpoint switching and distributed multiview coding

(E.g., [Zhu, Aaron, Girod; SSP 03], [Toffetti, Tagliasacchi, Marcon, Sarti, Tubaro, Ramchandran; EUSIPCO 05])

	Distributed multiview image/video coding	**Free viewpoint switching**
Key objective	Distributed, independent encoding at spatially-separated sensors	Centralized encoding to generate a single bitstream to support multiple decoding paths
Encoder complexity	Low-complexity encoding	Not primary issue. Applications use off-line encoding
Encoder access to SI	SI not accessible	Encoder has access to all SI candidates

USC

Related Work

USC Viterbi School of Engineering

- Provide random access in compression of image-based rendering data/multiview video
  - [Jagmohan, Sehgal, Ahuja; Asilomar 03]
  - [Aaron, Ramanathan, Girod; MMSP 04]
  - [Guo, Lu, Wu, Gao, Li; VCIP 06]
  - **All assume encoder has knowledge on predictor status, e.g., using feedback**
- Robust video transmission [Wang, Prabhakaran, Ramchandran; ICIP 06]
  - Specific algorithms used are different
- Improve decoding flexibility and accessibility [Naman, Taubman; ICIP 07]
  - Feedback and conditional replenishment
- Forward/backward video playback [Cheung, Wang, Ortega; VCIP 06]
  - Considerably different algorithm to achieve better coding performance

USC

USC Viterbi School of Engineering

Solution Based on Distributed Source Coding

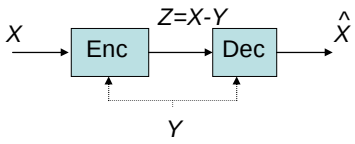
USC

DSC - Virtual Communication Channel Perspective

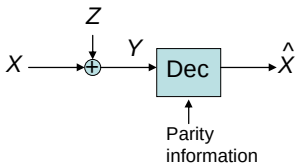
USC Viterbi School of Engineering

In DSC, encoder can communicate *X* by sending **parity information**
(E.g., [Girod, Aaron, Rane, Rebollo-Mondero; Proc. IEEE 04])

CLP:



DSC:

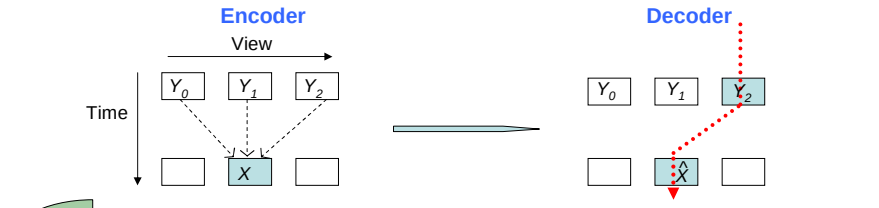


Parity information is **independent** of a specific predictor
- What matters is the **amount** of parity information

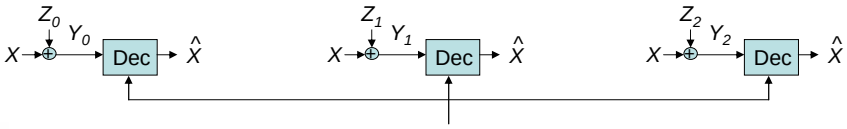
USC

Address Viewpoint Switching (Flexible Decoding) Using DSC

Under predictor uncertainty, encoder can communicate X by sending an amount of parity corresponding to the **worst case correlation**



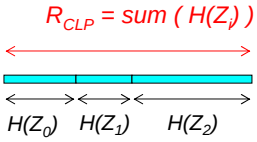
N virtual channels:

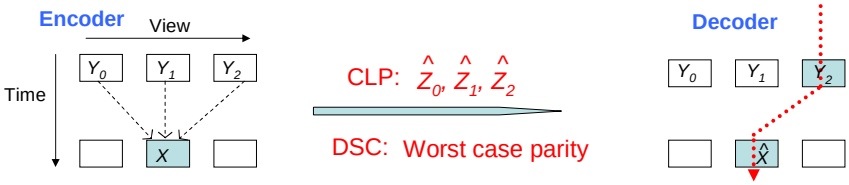


Parity corresponding to the worst case correlation noise

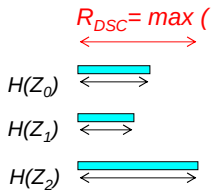
Viewpoint Switching (Flexible Decoding) – CLP vs. DSC

CLP





DSC

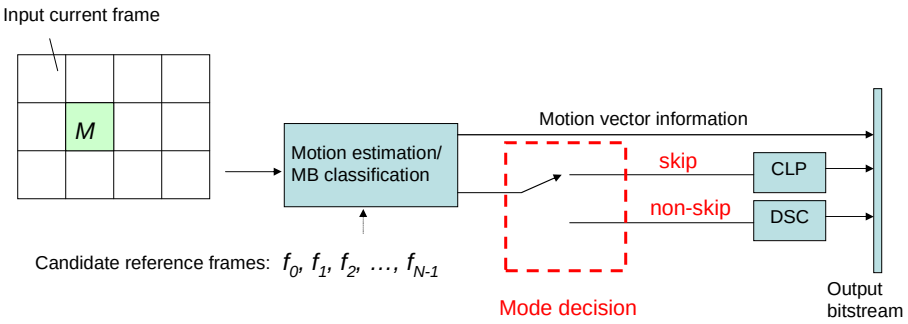


Bits required to communicate X with Y_i at decoder : $H(X \mid Y_i) = H(Z_i)$

Encoding Algorithm – Motion Estimation and Macroblock Classification

USC Viterbi School of Engineering

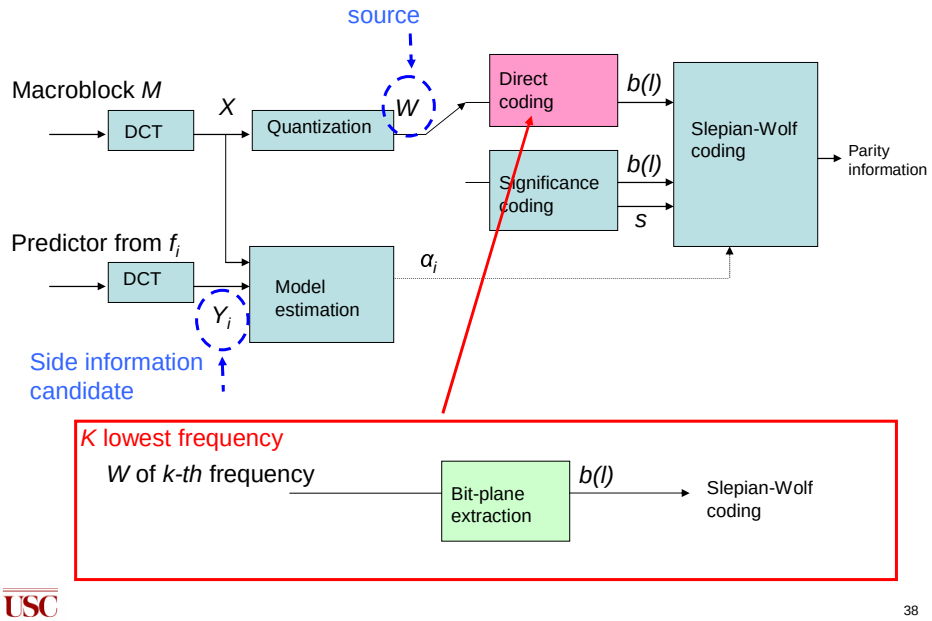
M may be classified to be in a *skip* mode if the difference between *M* and predictors from some f_i is small



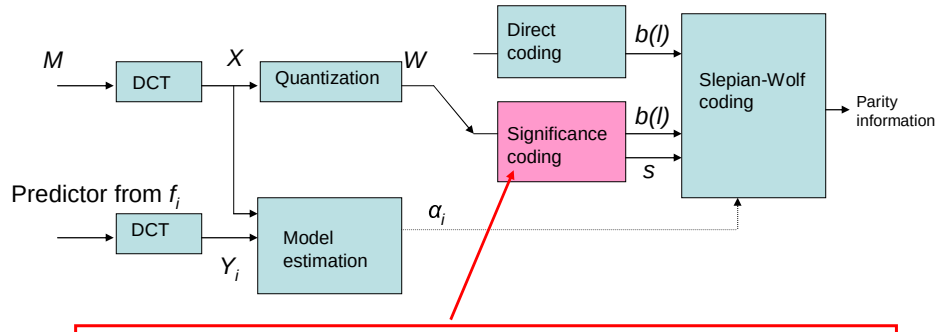
Majority: using DSC

Encoding Algorithm – DSC Coded MB

USC Viterbi School of Engineering

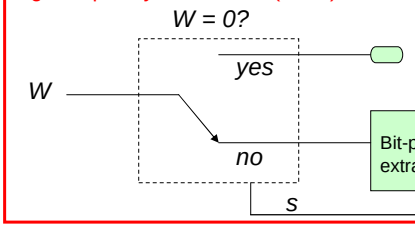


Encoding Algorithm – Significance Coding



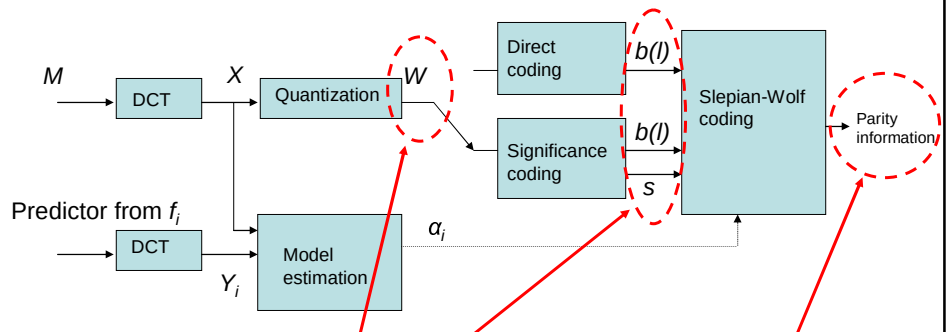
High frequency coefficients ($k>=K$)

- Expected number of source bits = $1*p_k+(1+L_k)*(1-p_k)$
- Lead to source bits saving when $p_k>1/L_k$



39

Drift-free: *W* in DSC coded MB are exactly recovered



Exactly recovered in decoding

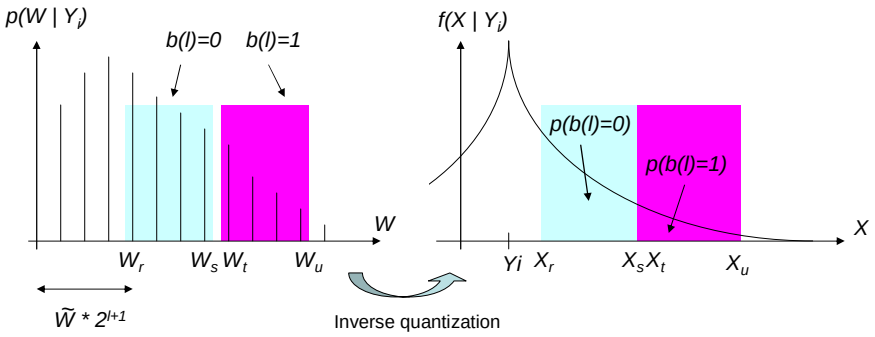
Amount of output parity determined by the worst case correlation

40

Models and Conditional Probability Estimation

USC Viterbi School of Engineering

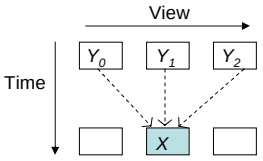
Estimate $p(b(l) \mid Y_i, b(l+1), b(l+2), \ldots)$ by integrating $f(X \mid Y_i)$ as follow:

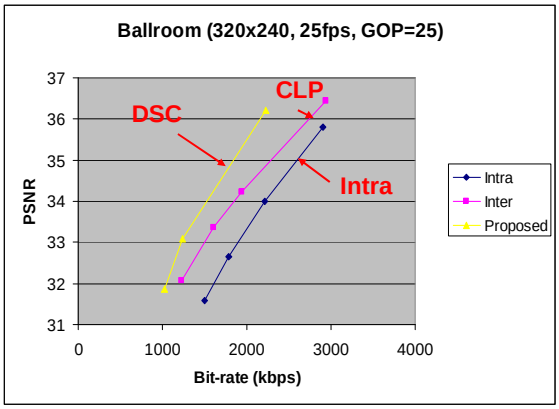


Experimental Results - Multiview Video Coding

USC Viterbi School of Engineering

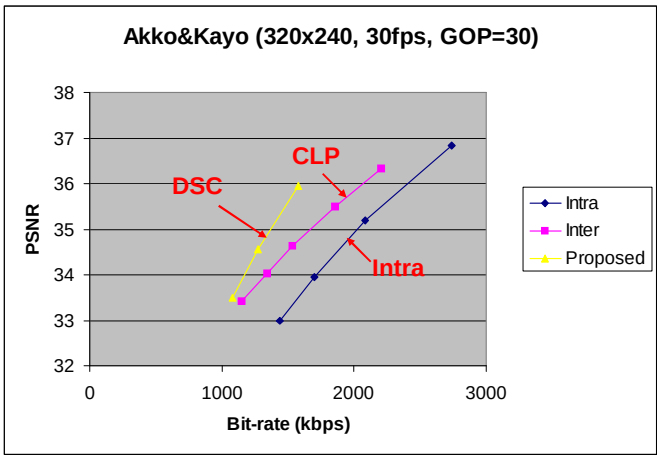
Allow switching from adjacent views: three predictor candidates





Our proposed algorithm out-performs CLP and intra coding

Experimental Results

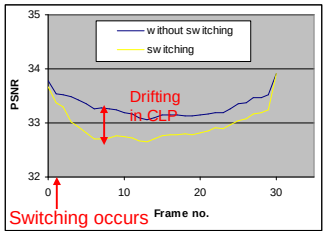


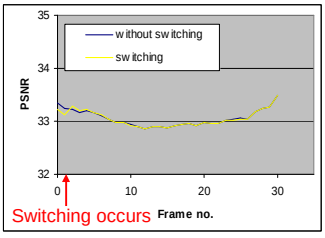
Drifting Experiments

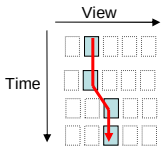
Our proposed algorithm is almost drift-free, since quantized coefficients in DSC coded MB are identically reconstructed

CLP DSC

Akko&Kayo (GOP=30)





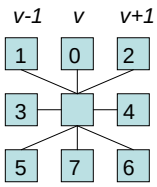


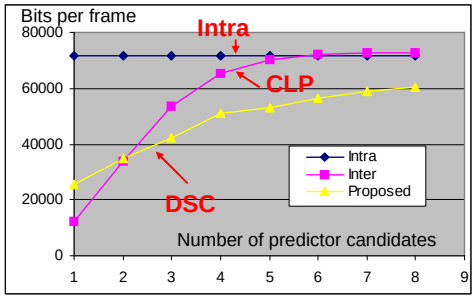
View switching occurs at frame number 2

Scaling Experiments

USC Viterbi School of Engineering

- Number of coded bits vs. number of predictor candidates





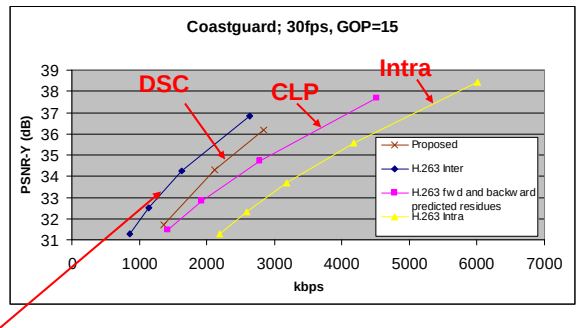
- Bit-rate of DSC-based approach increases at a slower rate compared with CLP
  - An additional candidate incurs more bits only if it has the worst correlation among all candidates

USC

Experimental Results – Forward/backward video playback

USC Viterbi School of Engineering

Forward/backward playback: **two** predictor candidates



Coastguard CIF

Inter-frame coding with one prediction residue: cannot support flexible decoding

USC

Summary: DSC Application to Viewpoint Switching/Flexible Decoding

USC Viterbi School of Engineering

- DSC-based coding algorithm to address viewpoint switching/flexible decoding
  - **Single** bitstream to support **multiple** decoding paths
  - Parity information **independent** of a specific predictor
  - Overhead depends on the **worst correlation** rather than the number of decoding paths
  - Outperform CLP and intra coding in terms of coding performance
  - Our proposed system is almost drift-free

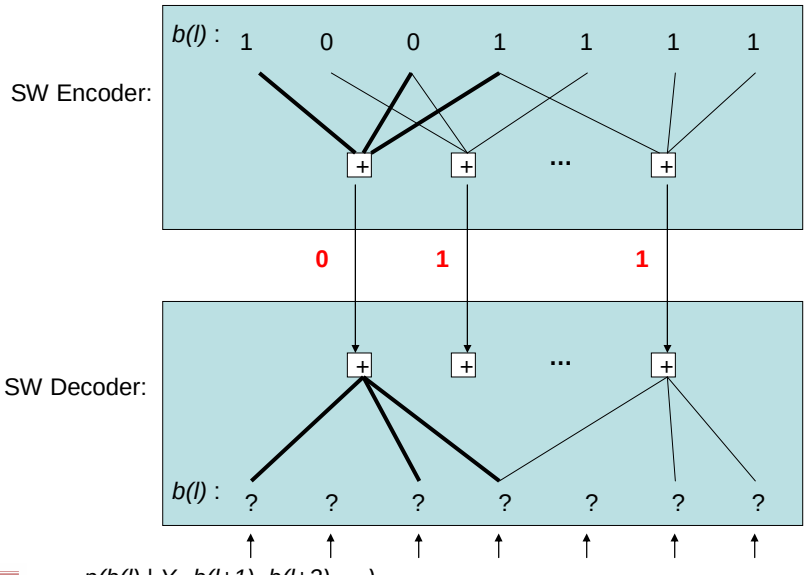
Comments or Questions?

USC

47

Parity information is independent of a specific predictor

USC Viterbi School of Engineering



USC

$p(b(l) \mid Y_i, b(l+1), b(l+2), \ldots)$

48