A Cognitive Approach for Effective Coding and Transmission of 3D Video

Simone Milani and Giancarlo Calvagno

E-mail: {simone.milani,calvagno}@dei.unipd.it

Florence, Italy - Oct. 27th, 2010

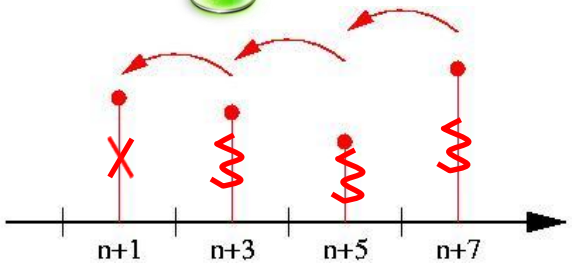
Outline

- Needs of modern video systems
- The problem of orchestrating different units
- Cognitive source coding
- The proposed solution
- Experimental results
- Conclusions and future work

Needs of modern video systems

- Limited power supply to enhance mobility and autonomy

- Robustness to losses, errors, delays, congestions



- Capability of interoperating
- Possibility to adapt to video services and applications

- Ability to adapt to the transmitted signal (video, depth) and grant different levels of Quality-of-Service (QoS)
- Capability of understanding environment and situation

- Possibility to integrate information from other sensors (positioning, movement, etc...)

Characteristics of 3D Video

In Depth Image Based Rendering,
3D video signal is made of

- Texture stream (=strandard video signal)
- Depth stream

Texture

Depth

Previous requirements are even more urgent for 3D video communications

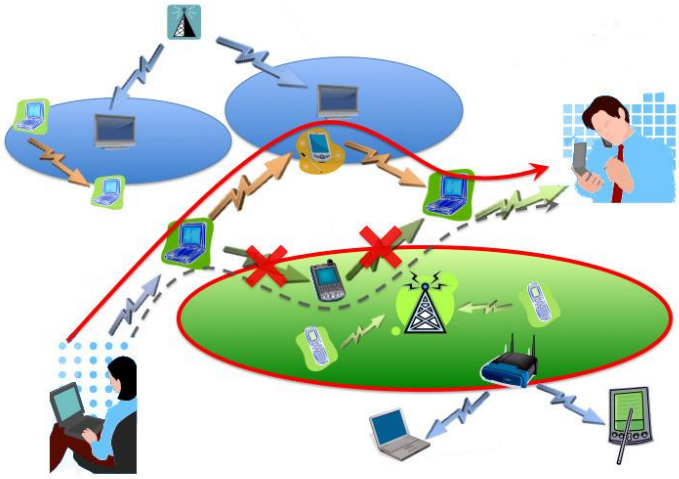
- Different types of signals
- Need to synchronize different streams

Cognitive Source Schemes are CL solutions where the very architecture of the source coder changes implementing different coding strategies.

It can be considered a subset of CL solutions where source coder structure usually does not change.

Cognitive Source Coding

Cognitive Source Coding (CSC) system
... "is an intelligent source coding system that is aware of scene and transmission environment (i.e., outside world), and uses the methodology of understanding-by-building to learn from the outside world and adapt its internal states."



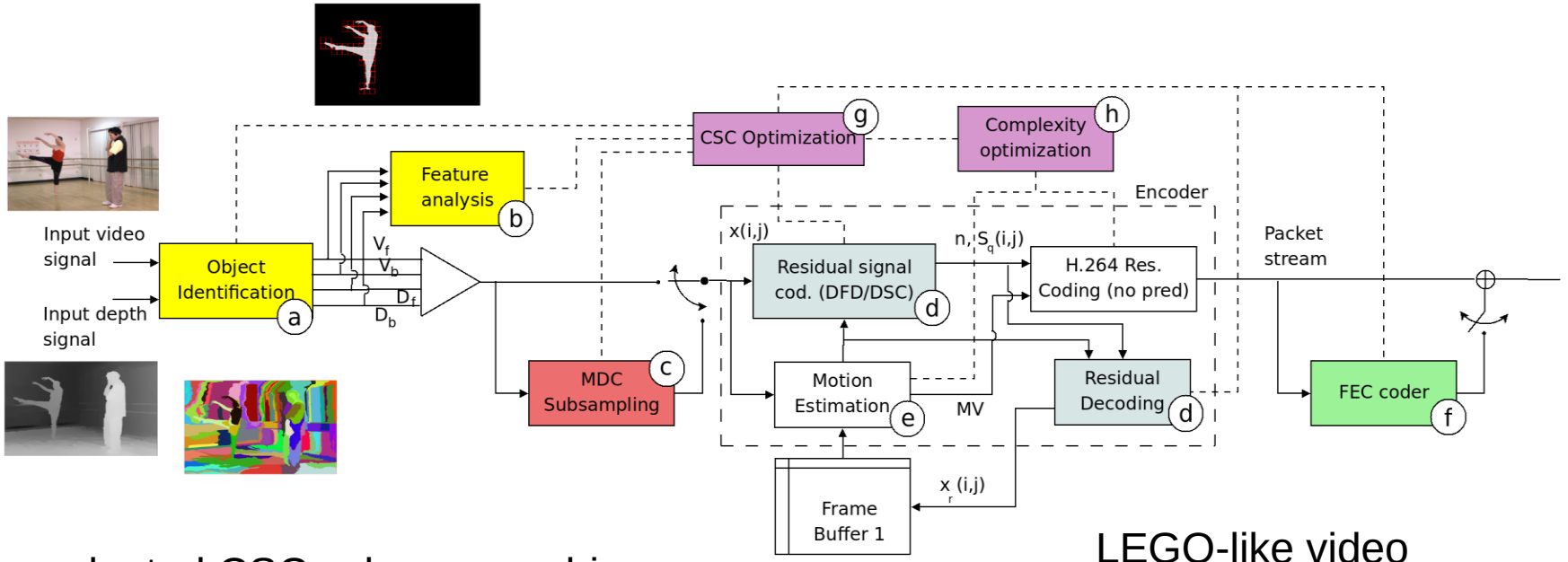
CR		CSC
Different modulation schemes	→	Different source coding schemes
Sensing the channel and adapting	→	Estimating channel state and changing the coding scheme

flexibility

reconfigurability

The adopted scheme for 3D video transmission

The signal is partitioned into different subsignals by segmentation unit followed by a classification unit



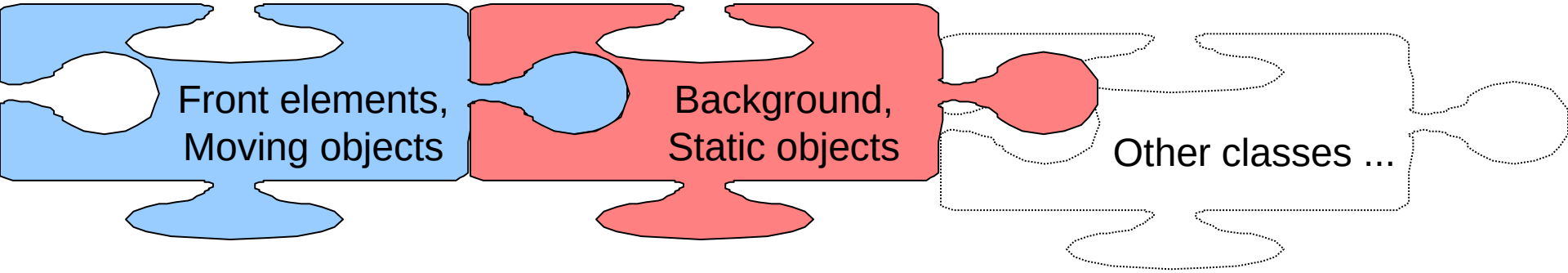
The adopted CSC scheme combines

- Single Description Coder (i.e. H.264/AVC)
- Multiple Description Coder
- Distributed Source Coder
- FEC cross-packet Coder

LEGO-like video source coder

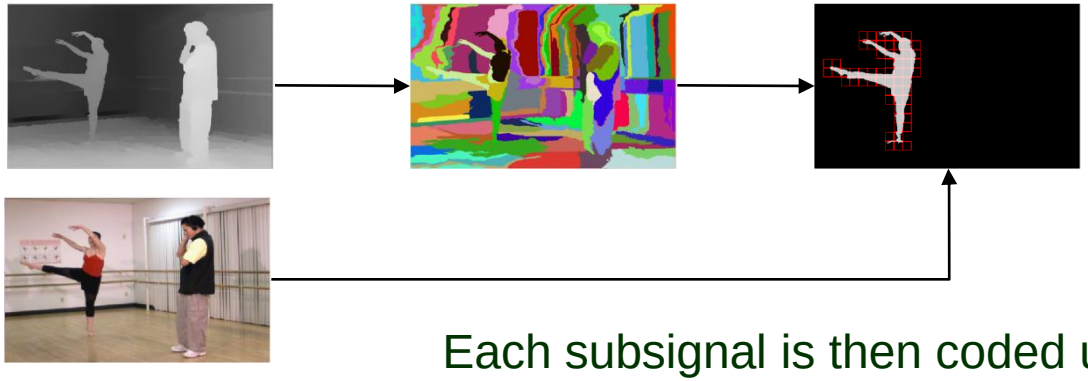
Classification

At he beginning, the signal is splitted into subsignals that include



The input depth signal is oversegmented into “superpixels”

Superpixels are classified according to depth and texture information into segments belonging tomoving and static objects



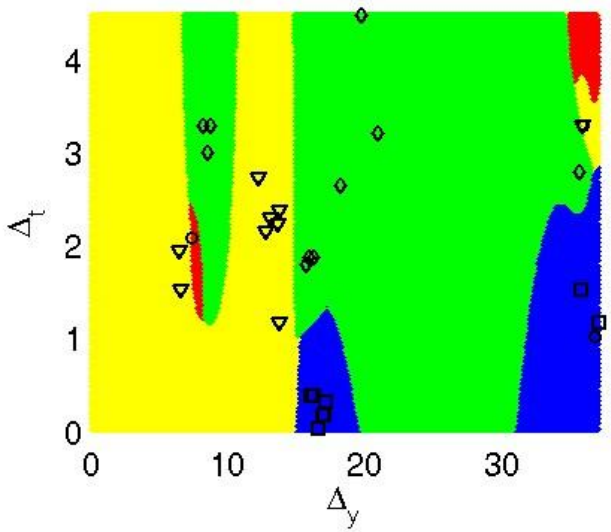
Subsignals	Front	Back
Texture	Vf	Vb
Depth	Df	Db

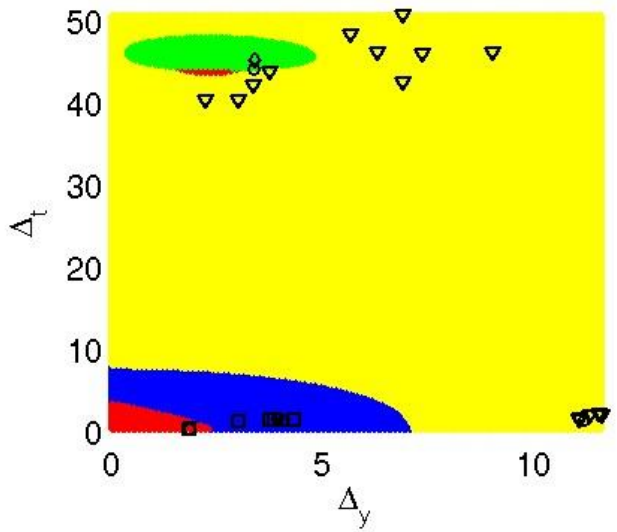
Each subsignal is then coded using the most appropriate source coding configuration.

SVM-based classification strategy

4 configurations (C)
- SD + DFD + FEC
- SD + DVC + FEC
- MD + DFD
- MD + DVC

Performance can be improved by classifying subsignals feature via an SVM partitioning.

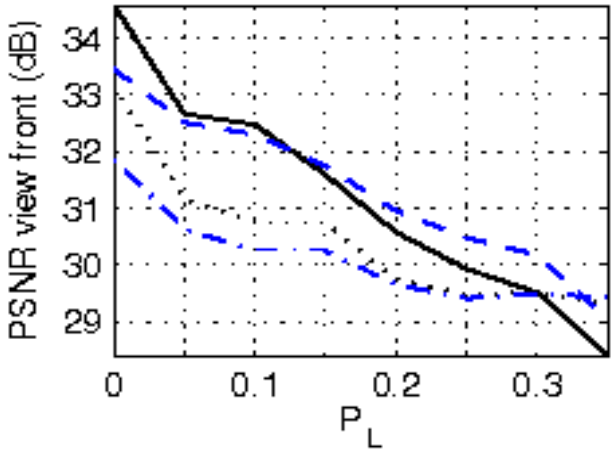


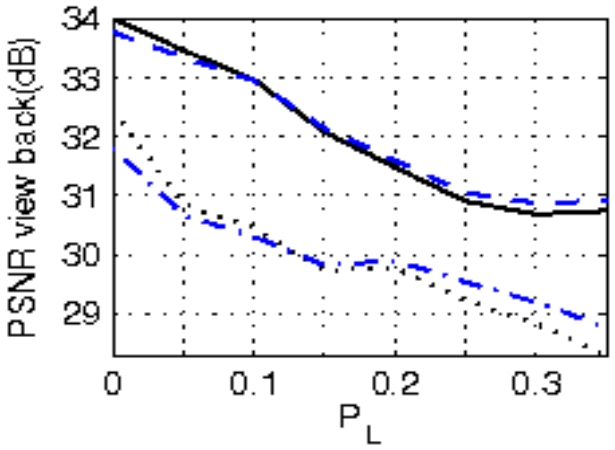


Computational complexity can be reduced by changing
- Motion search window
- Macroblock partitioning modes

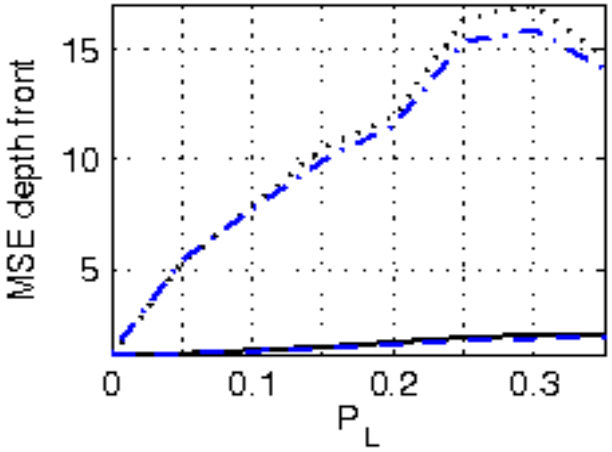
according to the characteristics of the subsignal.

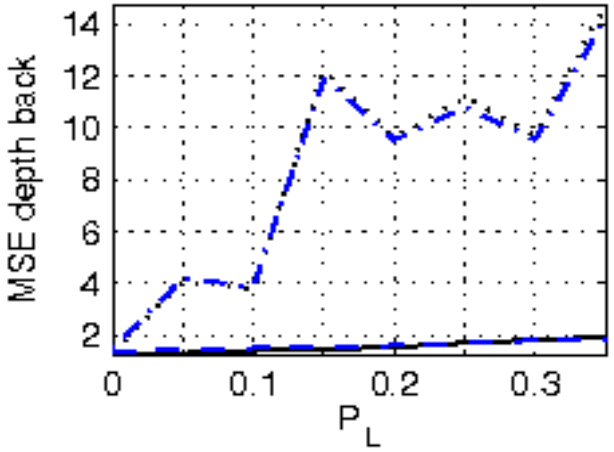
Experimental results (1/3)





Horse sequence
(video + depth)

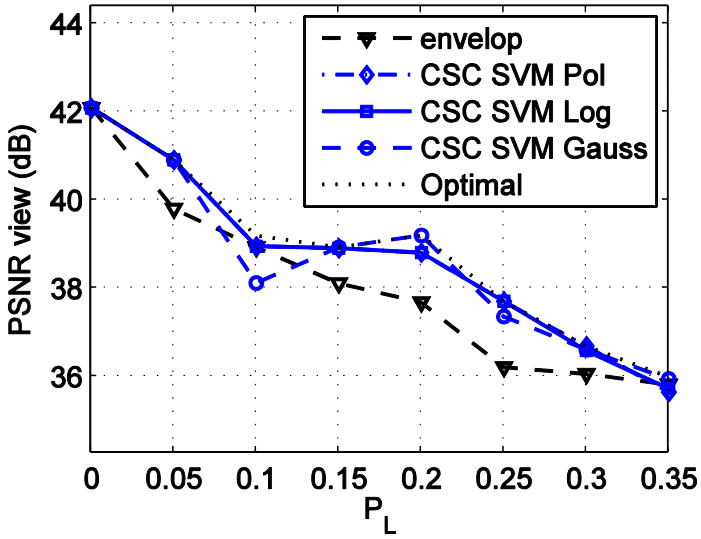




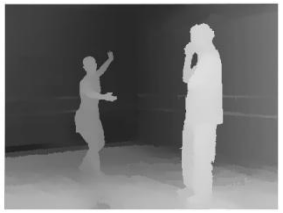
Performance of different configurations for the sequence Horse

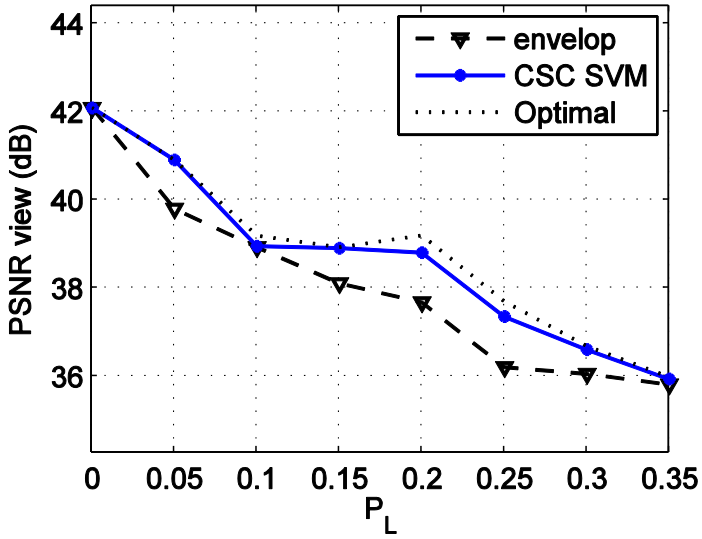
—— SD + DFD + FEC
… MD +DFD
_._ SD + DVC + FEC
_ _ MD + DVC

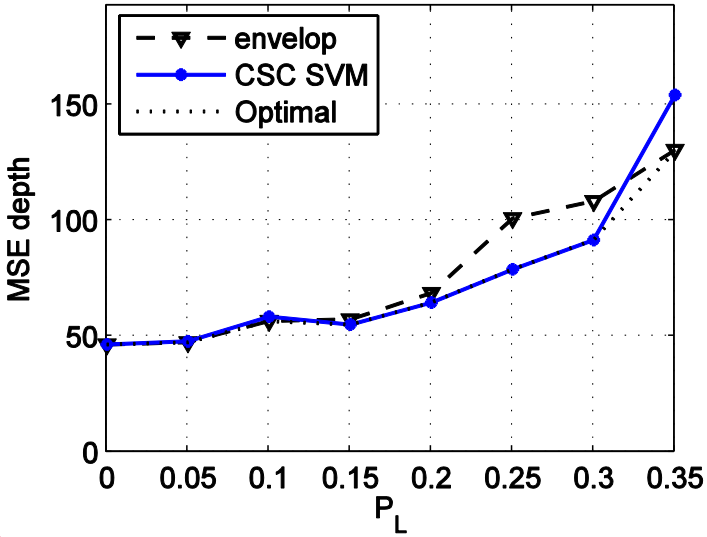
Experimental results (2/3)

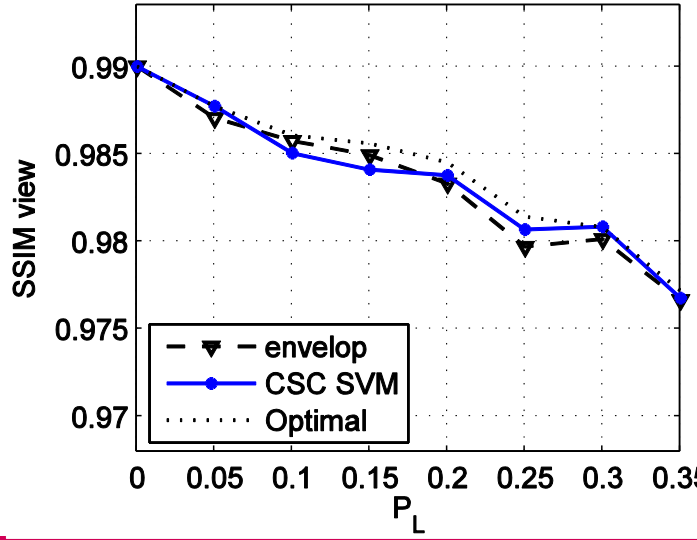


Sequence ballet







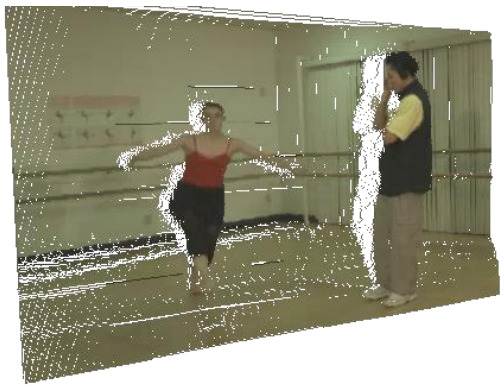


Experimental results (3/3)

Ballet sequence (frame 10)

H.264 + FEC

CSC with SVM



Ballet sequence (frame 4)

Conclusions and future work

- adapting the source coding scheme to the channel characteristics can significantly improve the quality of the reconstructed sequence;
- the adaptation needs to be suited to the different objects;
- a CSC scheme has been defined based on a SVM classification of the elements in the sequence.

Ongoing work ...

LTTM lab setting

- Optimization of the computational complexity
- Optimization of the coded bit rate for the different subsignals
- Building a complete real-time system (acquisition + coding + visualization).

Thank you for the attention !!

Any question ?

More results and documentation can be found at

http://www.dei.unipd.it/~sim1mil