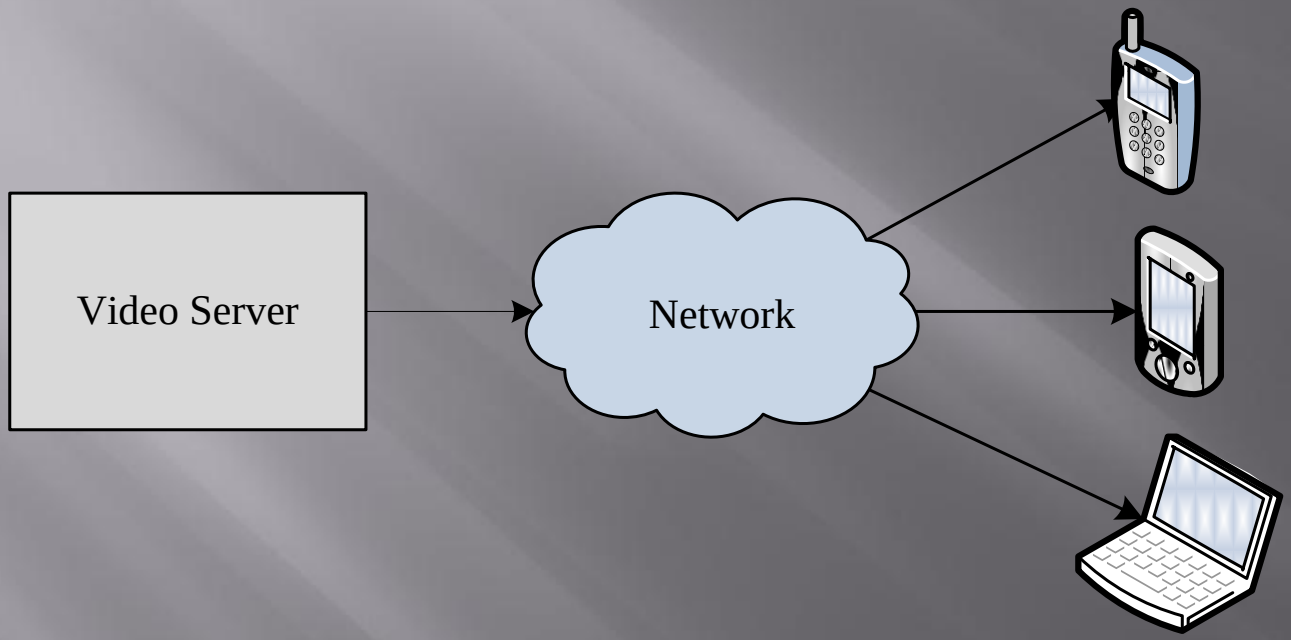
AN APPROACH FOR AN AVC TO SVC TRANSCODER WITH TEMPORAL SCALABILITY

Rosario Garrido-Cantos, José Luis Martínez, Pedro Cuenca, and Antonio Garrido

Outline

- Introduction
- Temporal Scalability
- H.264/AVC-to-SVC Transcoder
  - Proposal
  - Implementation results
- Conclusions

Introduction



Small display (QQVGA)
Low processing power
Low rate network connection

Medium display (QVGA)
Medium processing power
Medium rate network connection

Large display (VGA)
High processing power
Medium rate network connection

Introduction

- SVC (Scalable Video Coding):
  - H.264/AVC extension
- Structure with layers:
  - Base layer
  - Enhancement layers
- Types of scalability:
  - Spatial
  - Temporal
  - Quality (SNR)

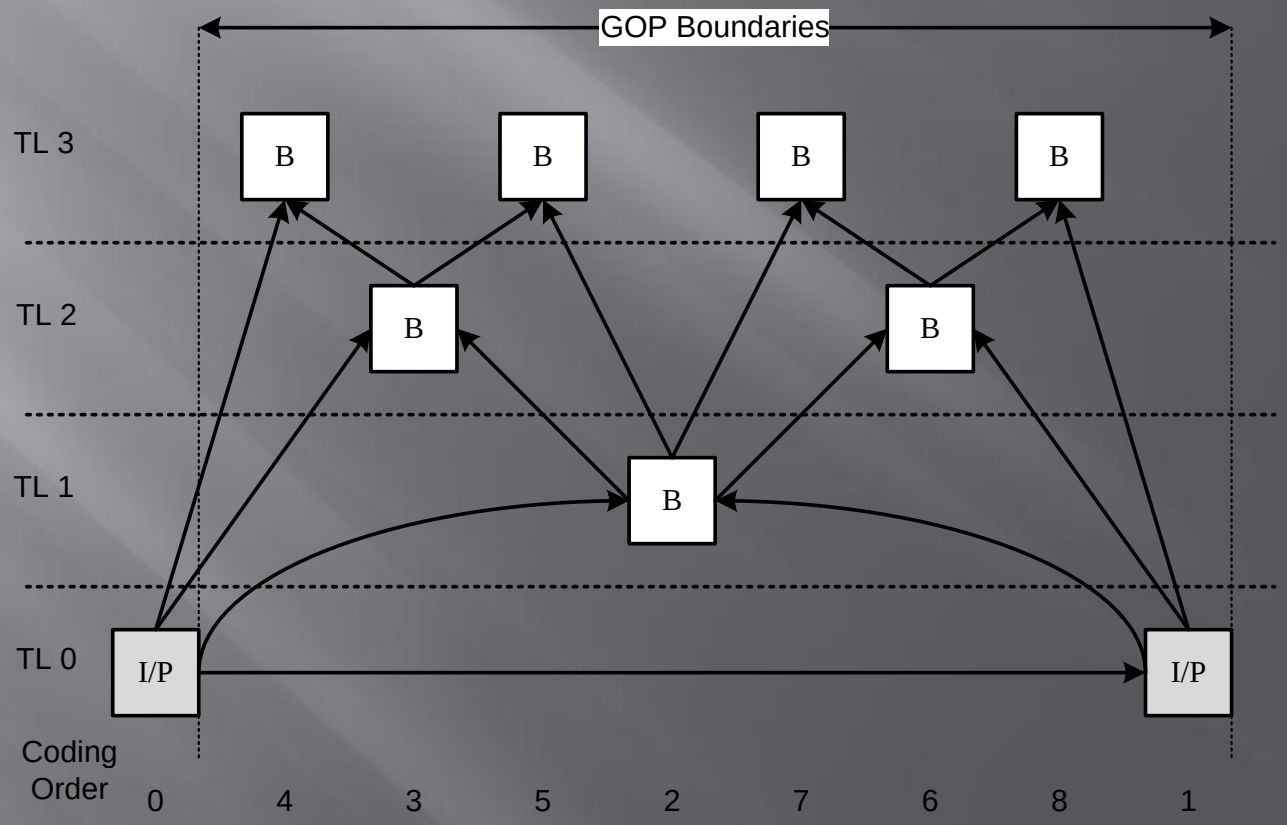
Introduction

- Most of the video contents today are still created in a single-layer format (H.264/AVC video streams).

- To provide these video bitstreams with scalability is necessary to develop efficient adaptation techniques.

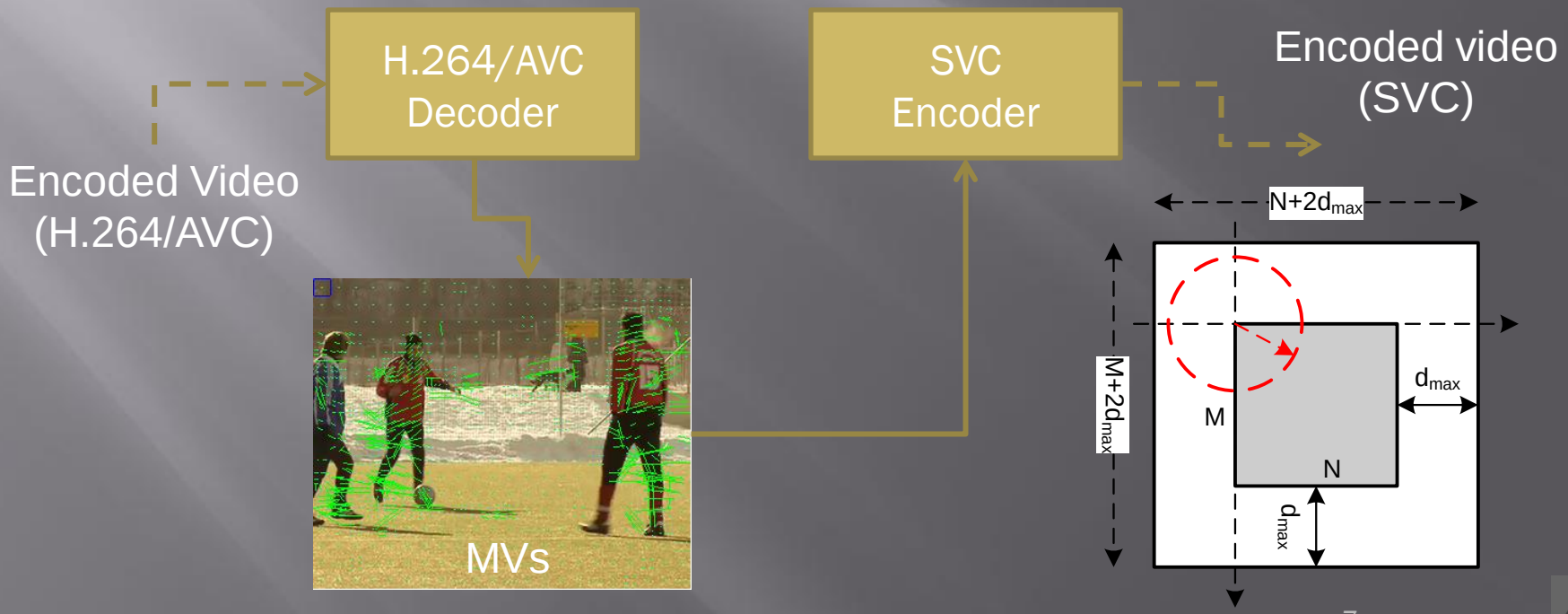
Temporal Scalability

- Different frame rates (15Hz, 30Hz, ...)
- Hierarchical B-pictures



H.264/AVC-to-SVC Transcoder

- Converting H.264/AVC to SVC bitstream
- Reducing computational cost of ME:
  - Reusing motion vectors from decoder.



Implementation Results

- Foreman, Bus, Football, Mobile, Soccer and Hall
  - QCIF and CIF resolution with key pictures every 0.5s.
- H.264/AVC
  - JM 16.2
  - IBBP GOP pattern
  - QP = 28
- SVC
  - JSVM 9.19.3
  - Hierarchical GOP patterns
  - Different QP = 28, 32, 36, 40

Implementation Results

Encoding time (%) of every temporal layer – CIF (30 Hz)									
	GOP = 8				GOP = 16				
Sequence	TL 0	TL1	TL 2	TL 3	TL 0	TL1	TL 2	TL 3	TL4
Foreman	4.72	13.59	27.29	54.40	1.52	6.63	13.11	26.34	52.40
Bus	4.73	13.67	27.29	54.31	1.75	6.47	13.29	26.22	52.27
Football	4.70	13.62	27.41	54.27	1.55	6.58	13.13	26.34	52.40
Mobile	4.72	13.60	27.23	54.45	1.49	6.61	13.09	26.32	52.50
Soccer	4.71	13.59	27.25	54.45	1.54	6.55	13.11	26.35	52.45
Hall	4.68	13.57	27.26	54.49	1.57	6.56	13.15	26.39	52.33
Average	4.71	13.61	27.28	54.40	1.57	6.57	13.15	26.33	52.38

Encoding time (%) of every temporal layer – QCIF (15 Hz)							
	GOP = 4			GOP = 8			
Sequence	TL 0	TL1	TL 2	TL 0	TL1	TL 2	TL 3
Foreman	11.98	29.45	58.57	4.71	13.73	27.32	54.24
Bus	11.71	29.26	59.03	4.84	13.58	27.06	54.52
Football	11.72	29.40	58.87	4.67	13.73	27.26	54.34
Mobile	11.94	29.36	58.70	4.72	13.65	27.37	54.26
Soccer	11.74	29.42	58.84	4.99	13.44	27.13	54.44
Hall	11.69	29.51	58.80	5.08	13.52	27.10	54.30
Average	11.80	29.40	58.80	4.84	13.61	27.21	54.34

Implementation Results

RD performance of AVC/SVC transcoder GOP = 8 - QCIF (15 Hz)			
Sequence	ΔPSNR (dB)	ΔBitrate (%)	ΔTime (%)
Foreman	-0.1109	4.44	-51.95
Bus	-0.1845	12.51	-46.21
Football	-0.1095	1.82	-26.12
Mobile	-0.1727	6.45	-85.31
Soccer	-0.2873	9.26	-40.46
Hall	-0.0877	3.12	-87.30
Average	-0.1 588	6.27	-56.23

RD performance of AVC/SVC transcoder GOP = 16 - CIF (30 Hz)			
Sequence	ΔPSNR (dB)	ΔBitrate (%)	ΔTime (%)
Foreman	-0.1027	3.04	-50.34
Bus	-0.1774	11.37	-45.56
Football	-0.0997	1.63	-25.21
Mobile	-0.1653	5.96	-84.78
Soccer	-0.2774	8.78	-39.36
Hall	-0.0337	2.95	-86.49
Average	-0.1427	5.62	-55.29

Conclusions

- Motion estimation has been accelerated.
  - Reusing motion vectors to build a dynamic motion window.
- Applied to the last two enhancement temporal layers.
- Achieved a Time Saving of around 55%.

Thank you!

Q&A