Scalable Video Coding

Miska M. Hannuksela

Research Leader, Nokia Research Center

http://research.nokia.com/people/miska_hannuksela/

NOKIA
Connecting People

About the Lecturer

- Research Leader, Media Systems and Transport
- Interests
  - Scalable, error-resilient and multi-view video coding
  - Real-time multimedia transport (RTP, H.324, MPEG-2 Systems)
  - Multimedia file formats (especially derivates of ISO base media file format)
  - Multimedia communication systems
  - Subjective quality of audio-visual services
  - Multimedia applications and APIs for handheld devices
- Standardization:
  - H.263++, H.264/AVC, Scalable Video Coding (SVC), Multiview Video Coding (MVC)
  - IETF RTP payload formats for H.264/AVC and SVC
  - ISO base media file format, 3GP file format, DVB file format
  - 3GPP multimedia specifications, DLNA RTP profile, DVB IP data casting
- Read more: http://research.nokia.com/people/miska_hannuksela/

NOKIA
Connecting People

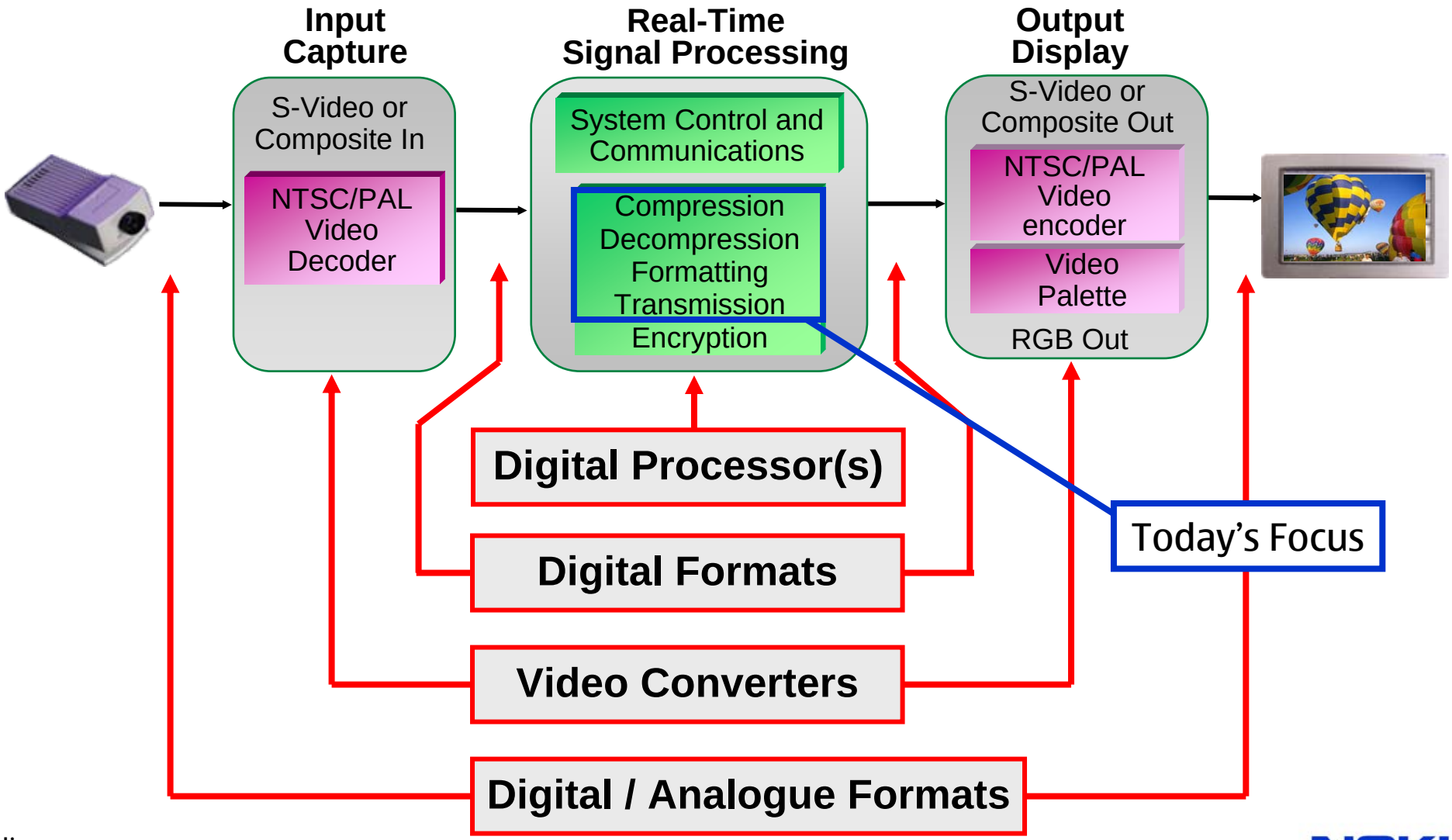
Outline

1. Scalable video coding:
basics, history, and motivation

2. Features of
the Scalable Video Coding (SVC) standard

3. Integration of SVC into services

Typical Digital Video System



NOKIA
Connecting People

Scalable Video Coding

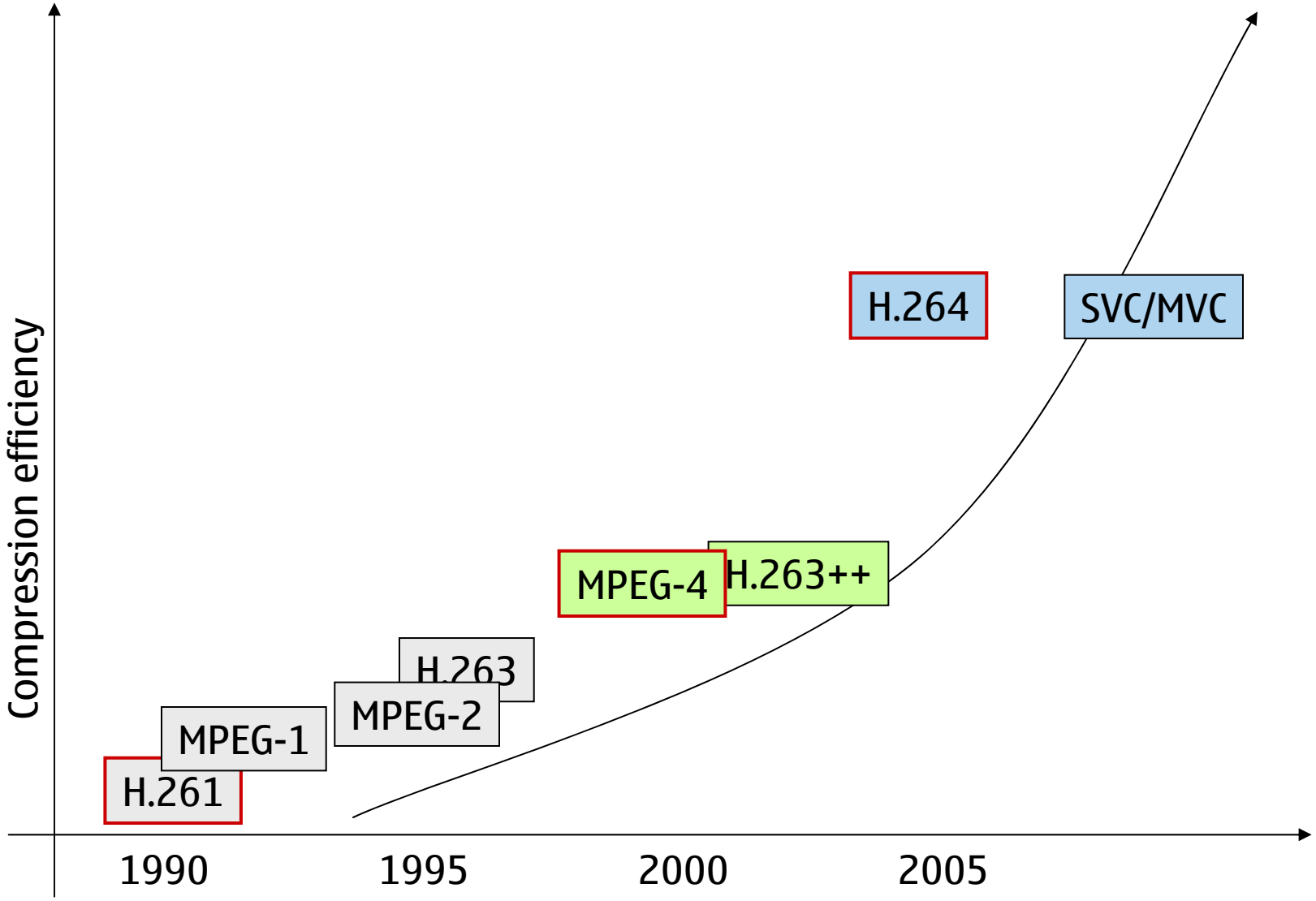
- Video is compressed once and played back at the optimal picture size for a display or optimal bit rate for a network
- Unified video content and services for mobile and wired use; anything from mobile to high-definition television

NOKIA
Connecting People

Types of Scalability

- Temporal Scalability
- Spatial Scalability
- Quality Scalability

Brief History of Video Coding Standards



Public

NOKIA
Connecting People

Applications of H.264/AVC

- High Definition Video
  - Blu-Ray Discs
  - HD Broadcasting
- Internet Video
  - YouTube high quality videos
- Mobile Multimedia
  - Mandatory or recommended in 3GPP and DVB-H
  - iPod Video

History of Scalable Video Coding Standards

- Temporal, spatial, and quality scalability have been included in all codecs since MPEG-2
- Only temporal scalability is used commonly
  - Improves compression efficiency
  - Fast forward and rewind functionality
- Spatial and quality scalability before SVC not used
  - Higher computational complexity and lower compression efficiency than in non-scalable codecs
- SVC has better chances to become widely used
  - The same compression efficiency as H.264/AVC with 10% bit rate increase (source: MPEG verification tests)
  - Single-loop decoding – moderate computational complexity increase in decoding
  - Built on top of H.264/AVC
  - Being adopted in all DVB services and ATSC M/H mobile television

NOKIA
Connecting People

Why Scalable Video?

Internet and mobile transmission are
best-effort/shared-resource and
becoming primary distribution mechanisms
→ Need for graceful degradation, bitrate adaptation

Lot of mobile video applications
→ Need for power adaptation

Variety of terminals and display sizes: QCIF, QVGA, VGA, SD, HD
→ Need for resolution adaptation

Alternatives to Scalable Video Coding

- In unicast streaming
  - Multiple streams encoded for different bit rates
  - Switching between the streams according to network throughput and/or 3GPP PSS rate adaptation signaling
- In multicast/broadcast streaming
  - Simulcast = simultaneous transmission of multiple independent streams
- In few other services, such as multiparty conferencing
  - Transcoding = (partial) decoding and re-encoding

Benefits of Scalability – One Video Fits All

- A single scalable service can cover wide range of devices and networks
  - Today's widely deployed AVC decoders can always decode the base layer
  - No need to tailor services specifically for mobile use
  - Scalable video is a key enabler for Internet services suitable for both wireline and mobile use

- A single video fits all devices and environments
  - Same content can be played and shared among low-end and high-end devices
  - Playback is optimized for the available display resolution
  - Allows low-power playback in battery-constrained cases

NOKIA
Connecting People

Benefits of Scalability – Improved User Experience

- When the same content is broadcast using multiple picture sizes, SVC brings considerable bit rate saving
  - E.g. 17% bit rate saving compared to H.264/AVC simulcast of QVGA and VGA [MPEG verification tests]
- Improved service continuity and resiliency against transmission errors
  - Unequal error protection in broadcasting
  - Unequal error protection in video conferencing
  - Resiliency against unexpected network throughput drops in point-to-point streaming [Schierl et al., ICIP 2005]
- Lower end-to-end delay in multiparty video conferencing
  - >100 msec one-way delay reduction [Eleftheriadis et al., 2006]
- Better picture quality in many services, as no transcoding required

NOKIA
Connecting People

Outline

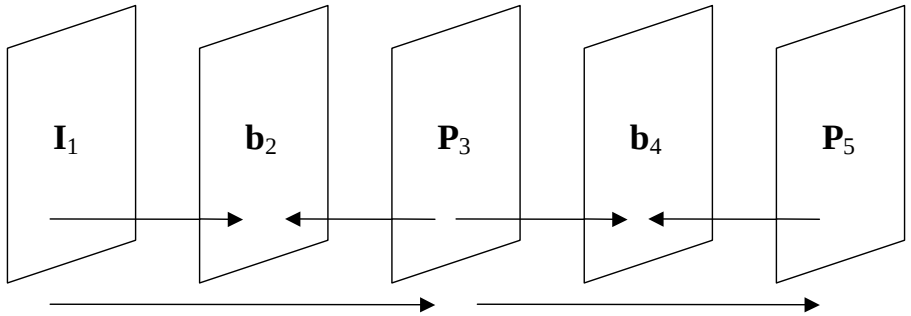
1. Scalable video coding:
basics, history, and motivation

2. Features of
the Scalable Video Coding (SVC) standard

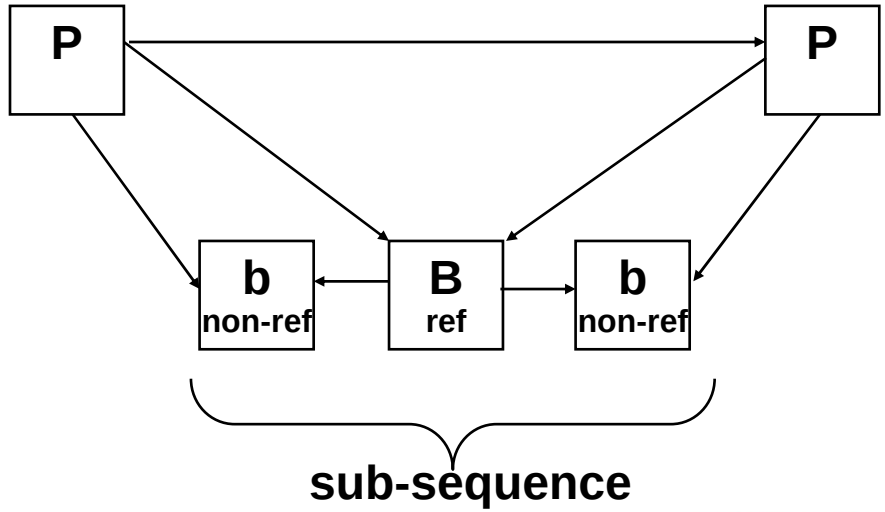
3. Integration of SVC into services

Temporal Scalability

1. Non-reference pictures

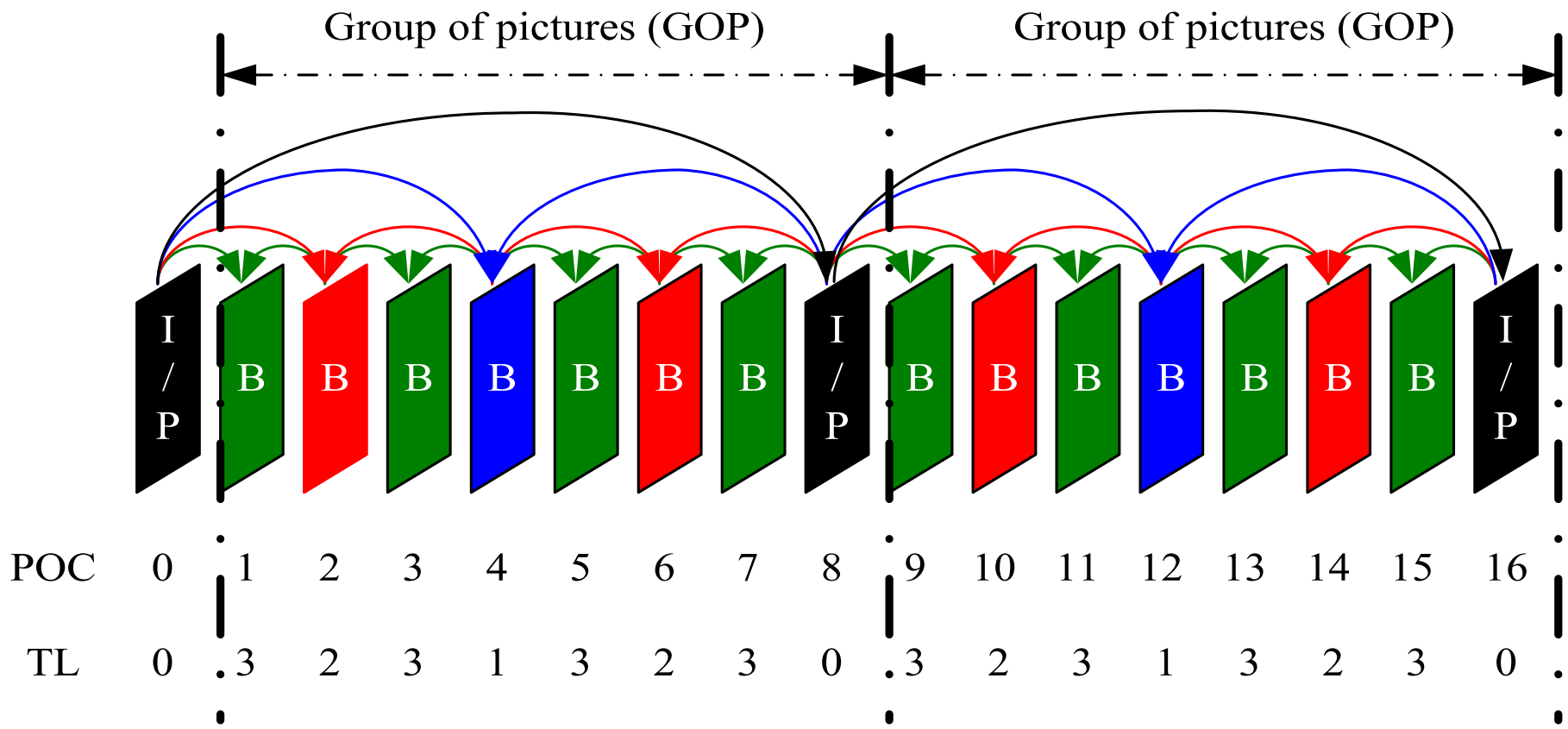


2. Hierarchical temporal scalability / disposable sub-sequences



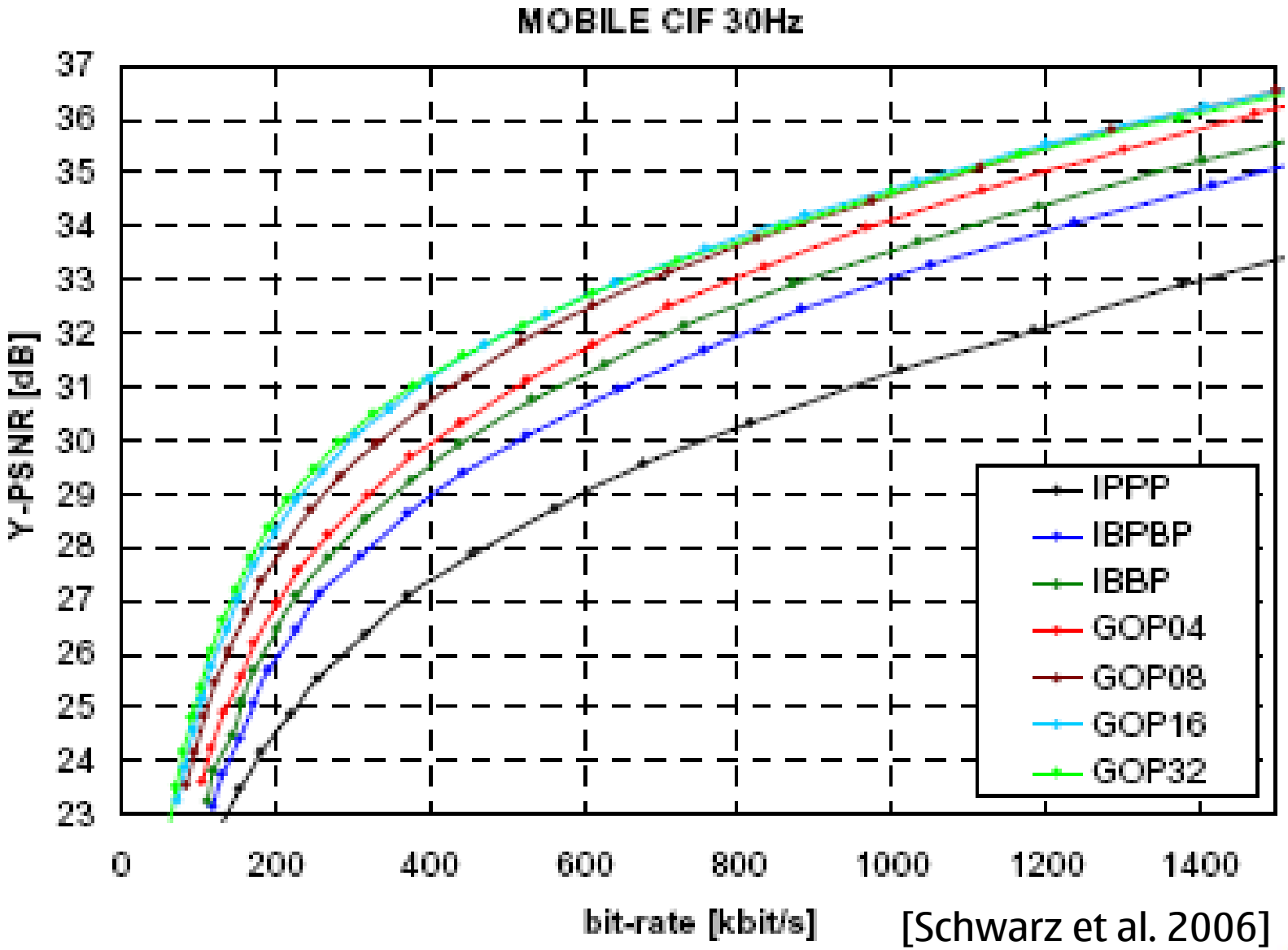
NOKIA
Connecting People

Temporal Scalability



NOKIA
Connecting People

Compression Efficiency of Temporal Scalability



[Schwarz et al. 2006]

NOKIA
Connecting People

Compression Efficiency of Temporal Scalability
(B pictures not in use)

Sequence	Compared to IPPP	
	PSNR Gain	***Bit-rate saving***
container	***1.388***	***7.19%***
foreman	***1.306***	***19.75%***
irene	***1.184***	***19.97%***
mobile	***3.163***	***42.95%***
news	***1.18***	***22.58%***
paris	***2.2***	***28.61%***
silent	***2.141***	***29.71%***
tempete	***2.128***	***34.14%***
Average	***1.836***	***25.61%***

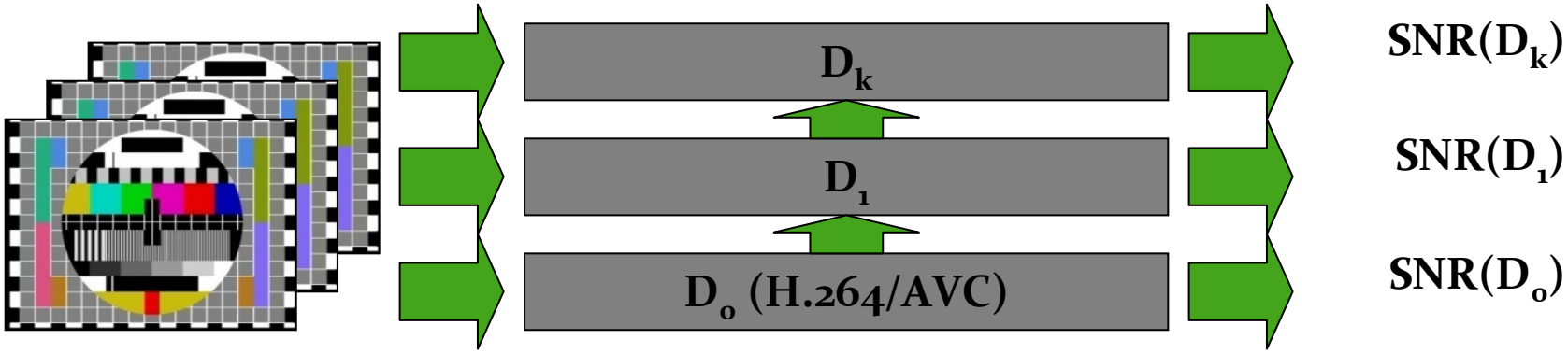
[Wen et al., submitted to ISCAS 09]

NOKIA
Connecting People

Spatial scalability

Quality Scalability

- Coding tools the same as in spatial scalability
- Two types:
  - Coarse Grain Scalability (CGS)
    - Switching between layers at IDR pictures
  - Medium Grain Scalability (MGS)
    - Finer quantization step deltas
    - Switching between layers at any position → controlled drift with base representation



NOKIA
Connecting People

Outline

1. Scalable video coding:
basics, history, and motivation

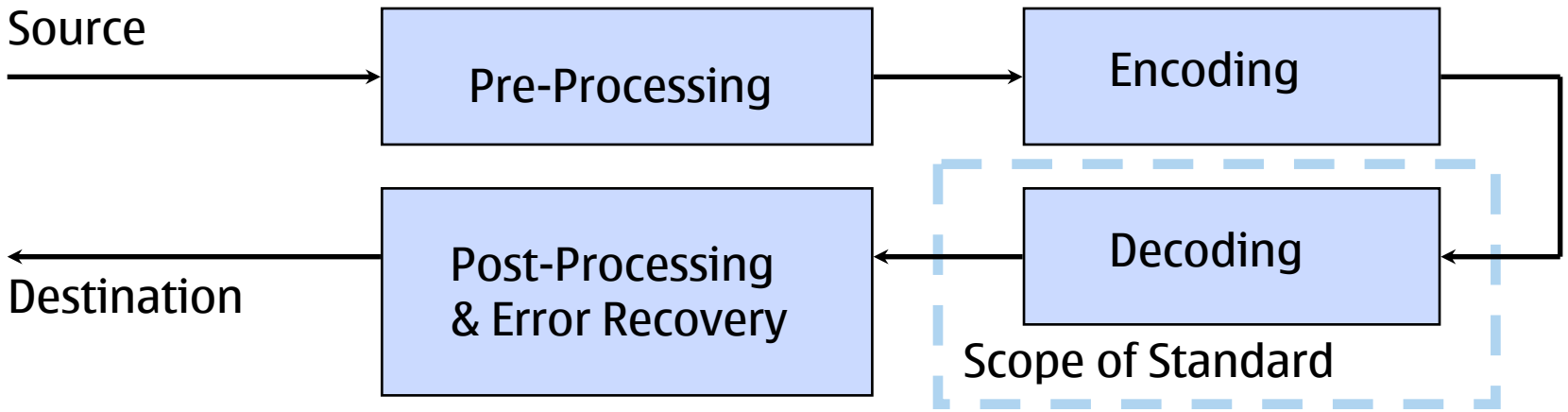
2. Features of
the Scalable Video Coding (SVC) standard

3. Integration of SVC into services

The Scope of Video Coding Standardization

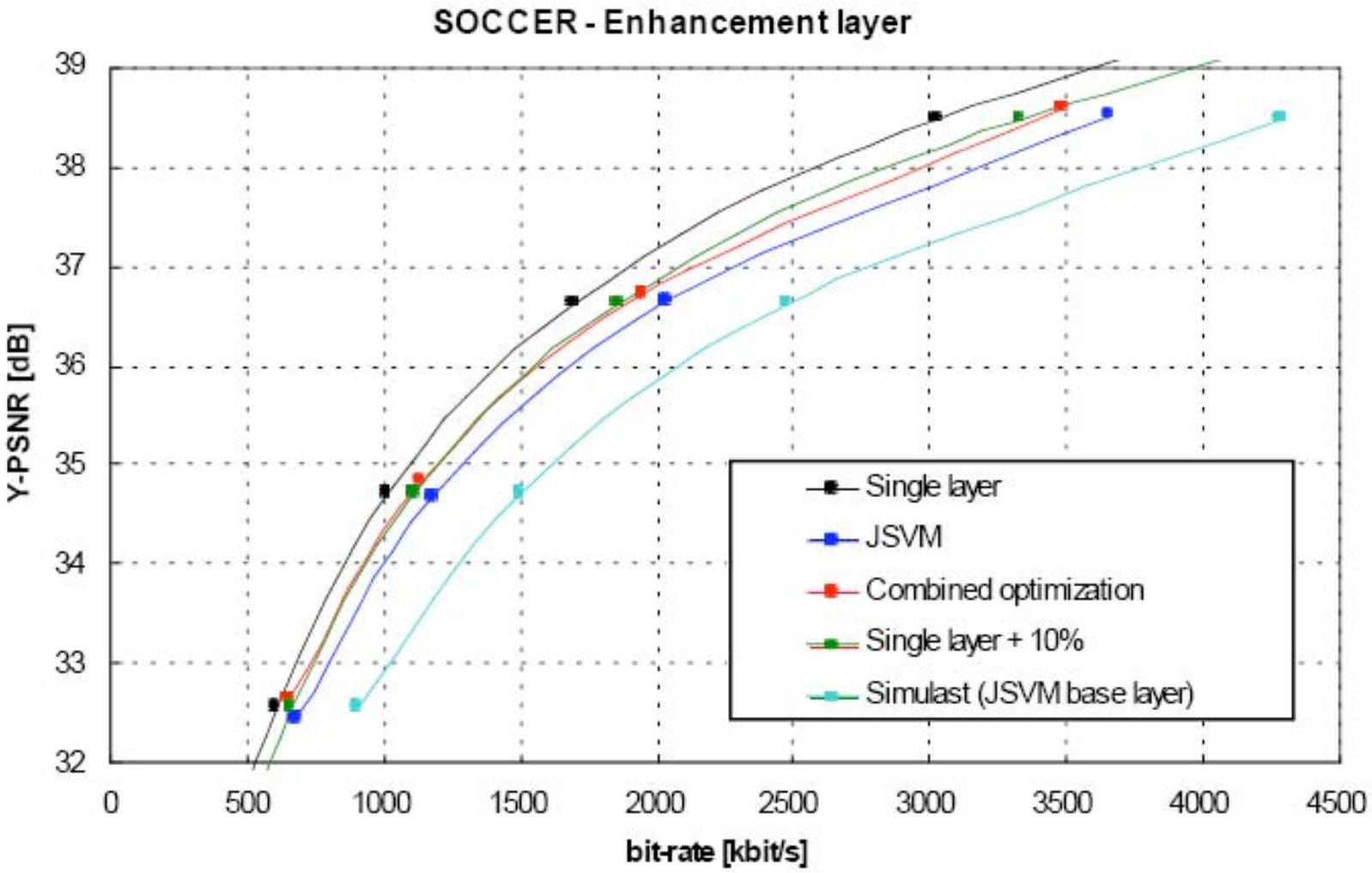
Only restrictions on the bitstream, bitstream syntax, and decoder operation are standardized:

- Permits optimization beyond the obvious
- Permits complexity vs. compression efficiency trade-offs in encoders
- Provides no guarantees of quality



NOKIA
Connecting People

Joint Optimization Results: Spatial: CIF30Hz/4CIF30Hz



NOKIA
Connecting People

Video Coding Standards - Profiles and Levels

- Profile
  - A subset of algorithmic features of the coding standard and constraints on the features
  - Decoders conforming to a profile shall be capable of supporting the entire subset of the algorithmic features of that profile
  - Encoders are not required to use a particular subset of a profile
  - A typically profile is targeted for a set of applications that share a similar trade-off between memory, processing, latency, and error resiliency requirements
- Level
  - A set of limits mainly on memory and computation performance parameters
  - Gives minimum limits for decoders
- Profile and level
  - Indicate characteristics of bitstreams. Can be used in session/stream description.
  - Indicate capabilities of decoders. Can be used in capability exchange process.

NOKIA
Connecting People

SVC Profiles

- Scalable Baseline Profile
  - Resolution ratios of 1.5 and 2 between successive spatial layers in both horizontal and vertical direction and macroblock-aligned cropping
  - Progressive sources
  - Enhancement layers: B slices, weighted prediction, CABAC, 8x8 luma transform
  - Base layer conforms to the H.264/AVC constrained baseline profile
- Scalable High Profile
  - Restrictions of Scalable Baseline Profile are removed
  - Base layer conforms to the H.264/AVC high profile
- Scalable High Intra Profile
  - Professional applications
  - Only IDR pictures (for all layers)
  - Scalable High Profile is supported

25

Accompanying Specifications

- SVC file format
  - Specifies how SVC streams are stored in MP4, 3GP, DVB, and other similar file formats
  - File metadata helping in adapting the stream
- SVC transport over MPEG-2 transport stream
  - For most digital television systems
- RTP payload format for SVC
  - For real-time SVC transport over IP networks
  - Technically stable, last call to be issued soon

Summary

- Types of scalability
  - Temporal
  - Spatial
  - Quality
- Benefits of scalability
  - One video fits all
  - Improved user experience
- SVC Performance
  - Temporal scalability improves compression efficiency significantly
  - Spatial and quality scalability: The same compression efficiency as H.264/AVC with 10% additional bitrate
  - ~15% bitrate saving for dyadic spatial scalability compared to simulcast
  - Cross-layer-optimized encoder needed for achieving the best compression gain
  - Single-loop decoding keeps decoding complexity reasonable