TI Developer Conference

November 29-30, 2007 • Bangalore, India

Minds in Motion

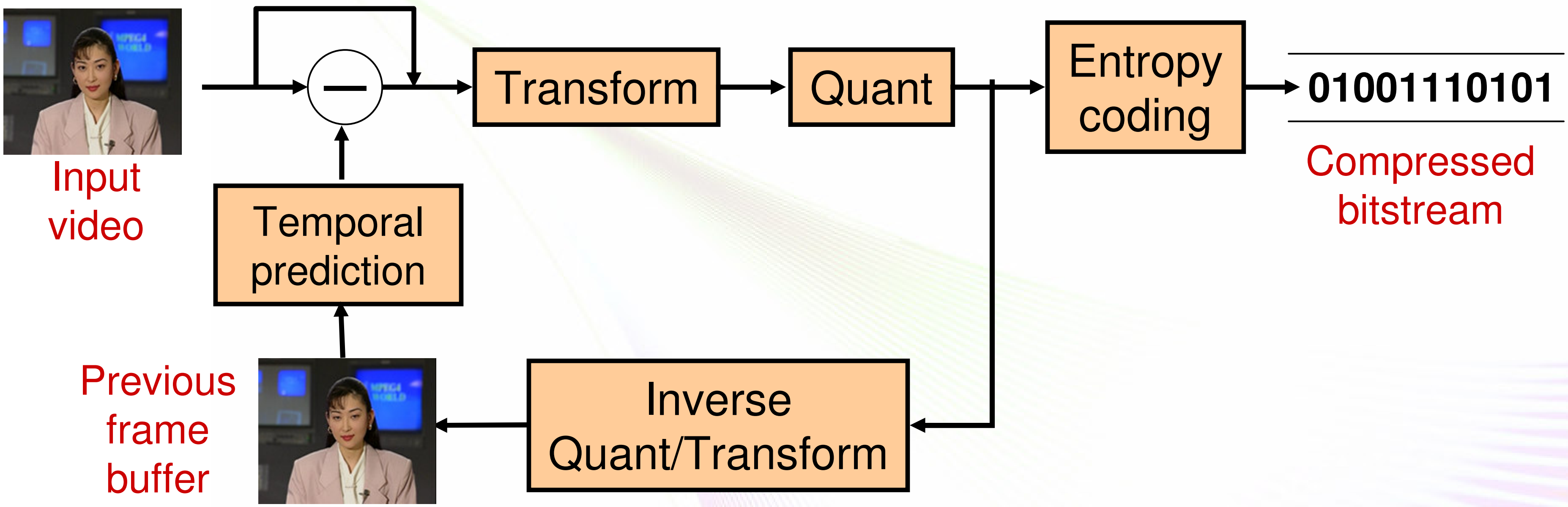
Upcoming Video Standards

Madhukar Budagavi, Ph.D.
DSPS R&D Center, Dallas
Texas Instruments Inc.

Outline

- Brief history of Video Coding standards
- Scalable Video Coding (SVC) standard
- Multiview Video Coding (MVC)
- ITU-T SG16/Q6 activities towards a next-generation video coding standard
- Conclusions

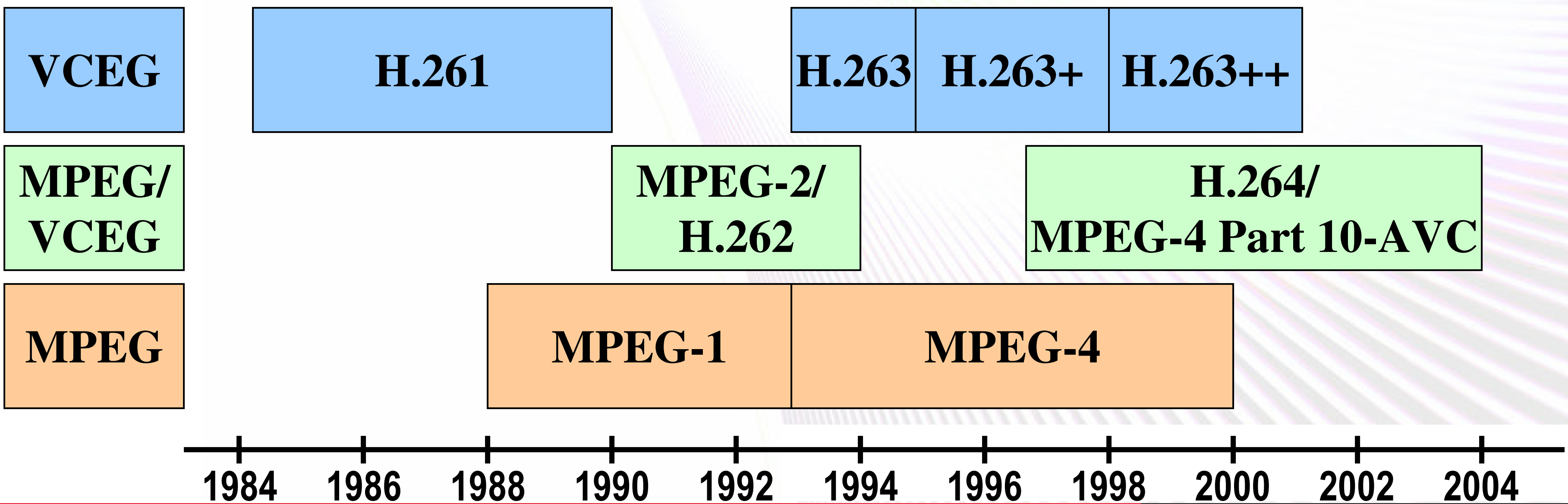
Video Coding Basics



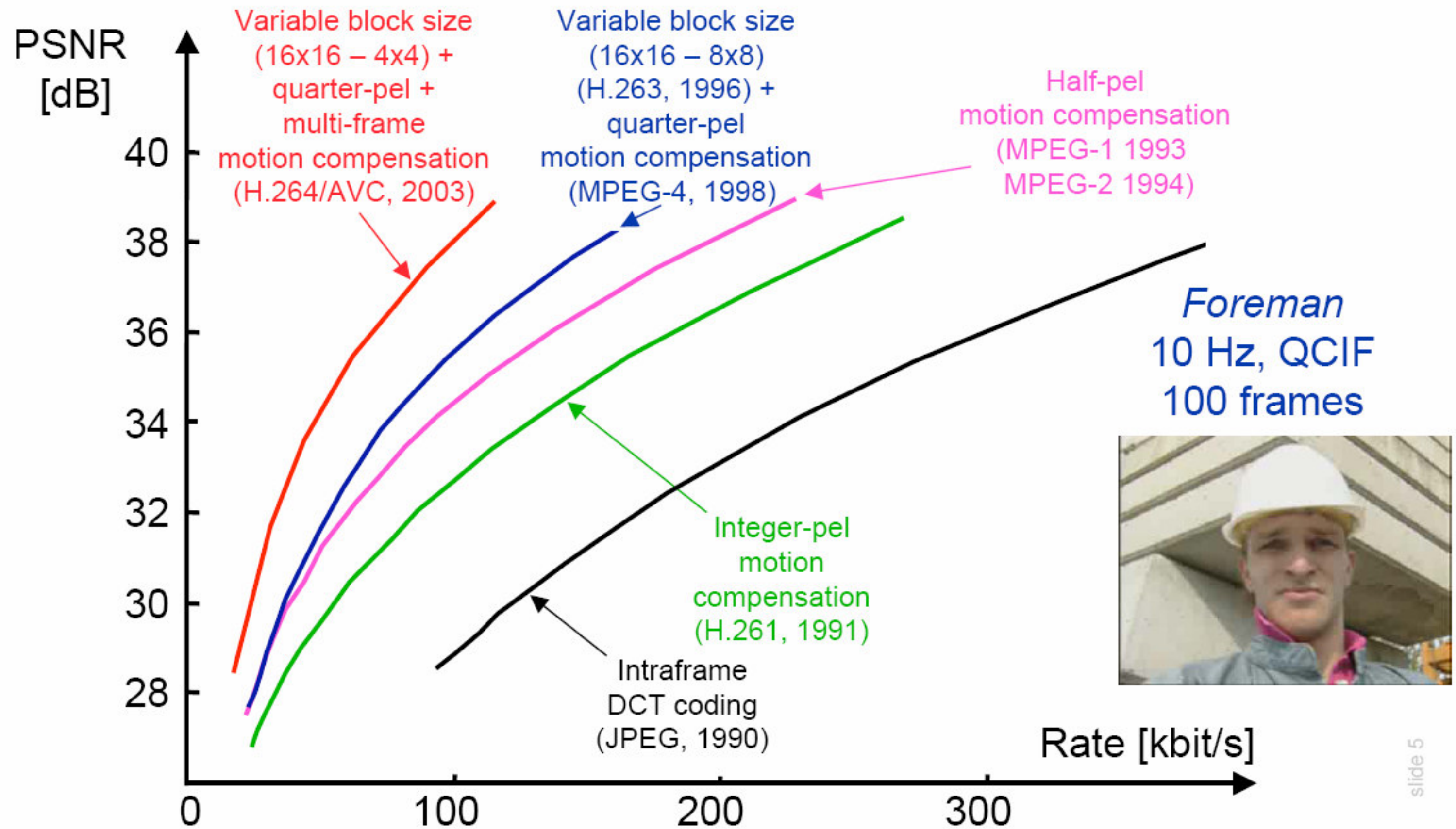
- **Key techniques used in video compression**
  - *Spatial* redundancy → *Transformation*
  - *Perceptual* redundancy → *Quantization*
  - *Statistical* redundancy → *Entropy coding*
  - *Temporal* redundancy → *Temporal prediction*

History of Video Coding Standards

- ISO/IEC JTC1/SC29/WG11
  - MPEG: Moving Picture Experts Group
- ITU-T SG16/Q.6
  - VCEG: Video Coding Experts Group
- Other standards: VC1, RealVideo, China AVS



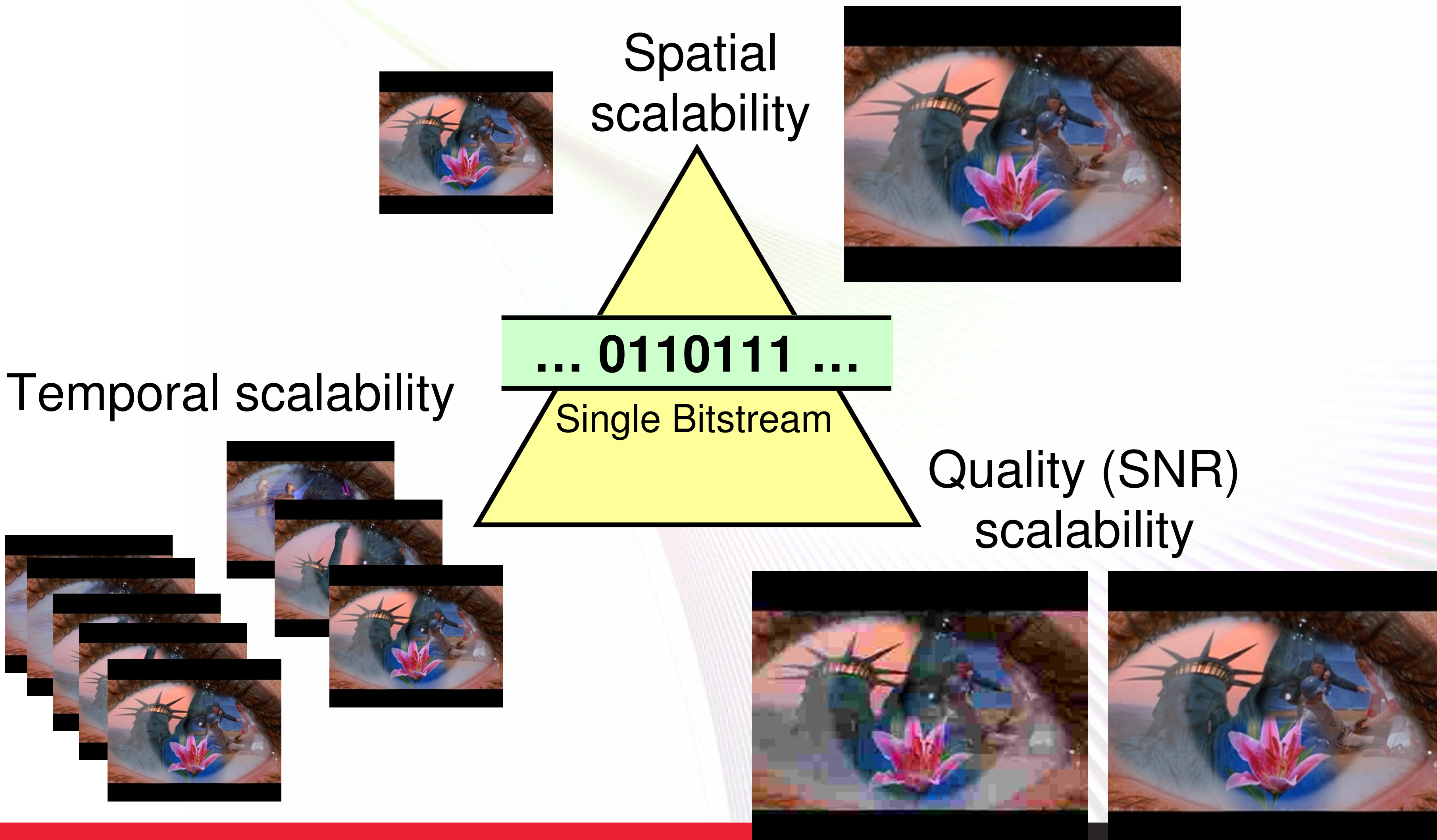
Video Coding Milestones



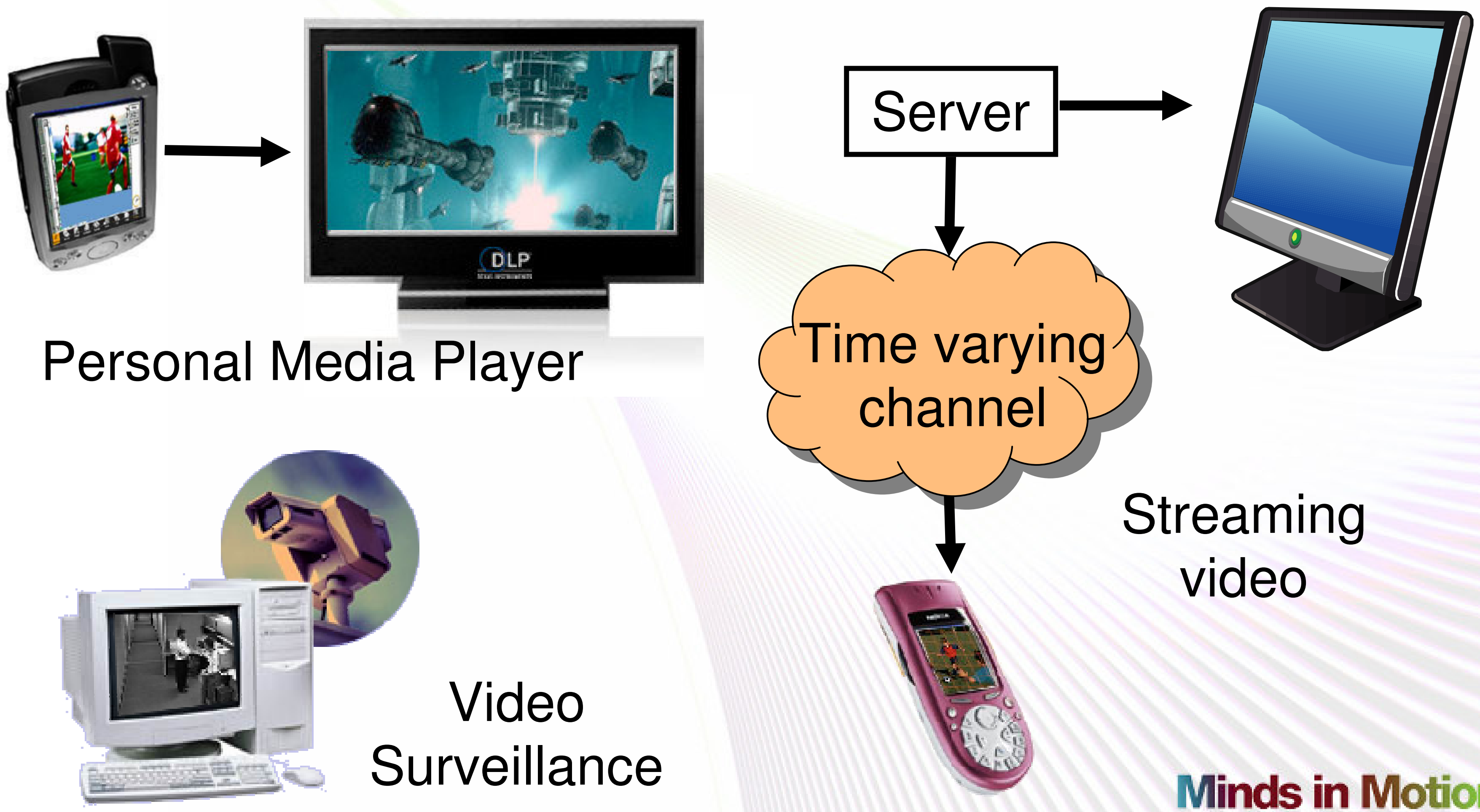
- What's next?

Figure extracted from T. Wiegand, JVT-W132 with permission

Scalable Video Coding (SVC)



Need for scalability

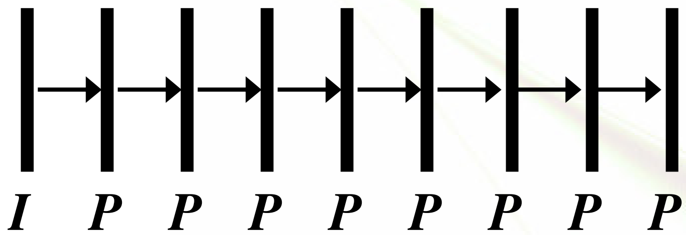


Minds in Motion

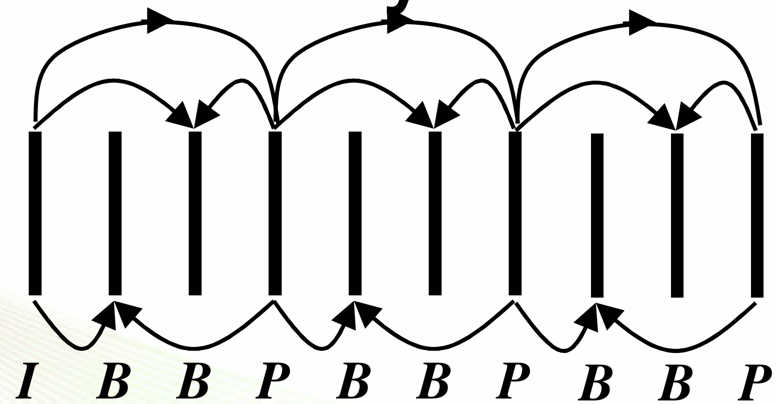
SVC standard

- Extension of H.264/AVC
  - Annex G of specification
  - Version 1 finalized in July 2007
  - Standard by Joint Video Team of ITU-T | ISO/IEC
- Supports
  - Temporal scalability, Spatial scalability, SNR scalability
  - (and a combination of the above)
- Base layer coding is H.264/AVC compatible

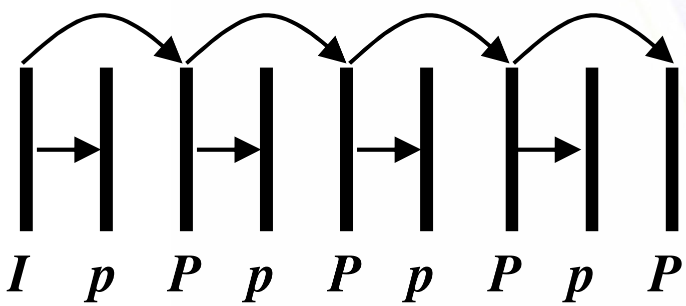
Temporal scalability



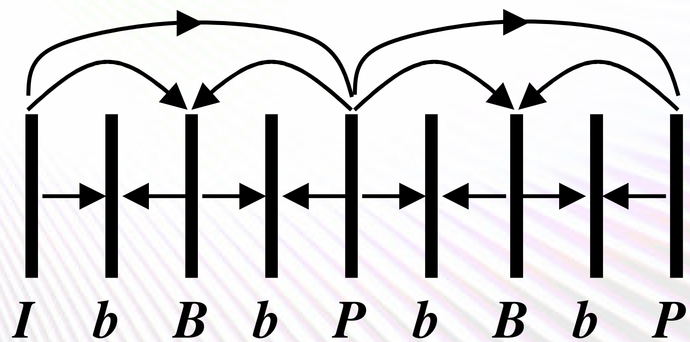
IPPP coding



IBBP coding



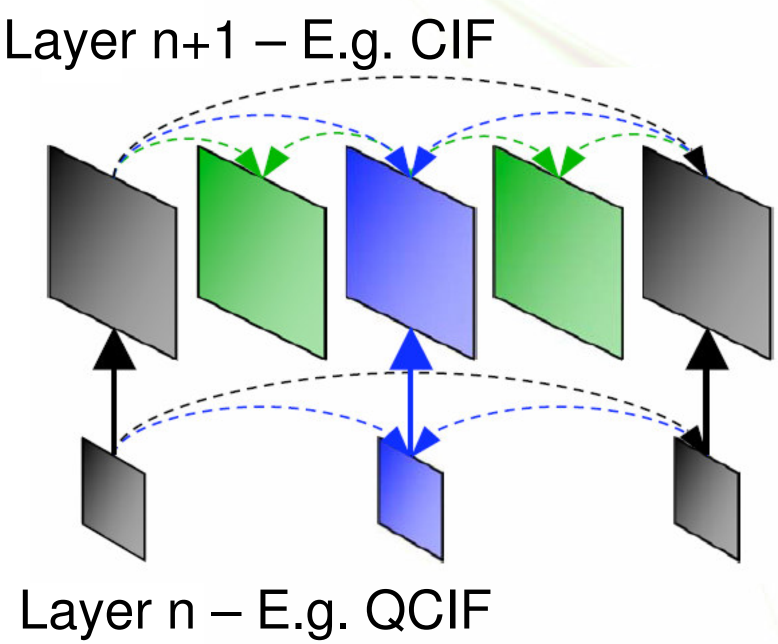
Hierarchical P-frames



Hierarchical B-frames

- p, b – Non-reference frames

Spatial scalability

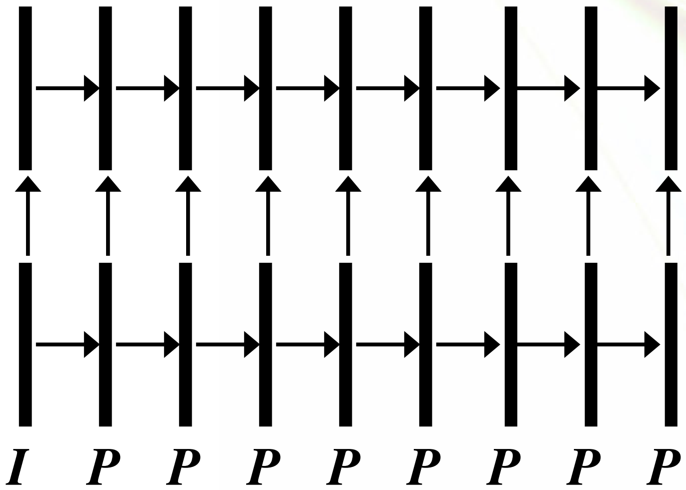


- Layered coding
- Higher layers have higher spatial resolution when compared to lower layers
- Upper layers re-uses data from lower layers
  - Upper layers use motion information, partitioning, and residual and not reconstructed picture from lower layer

Minds in Motion

Figure extracted from T. Wiegand, JVT-W132 [1] with permission

SNR scalability



- Layered coding
- Higher layers have better quality than lower layers
- Re-uses data from lower layers in coding of upper layers
- Two types of SNR scalabilities supported:
  - Coarse-grained
  - Medium-grained

SVC complexity and performance

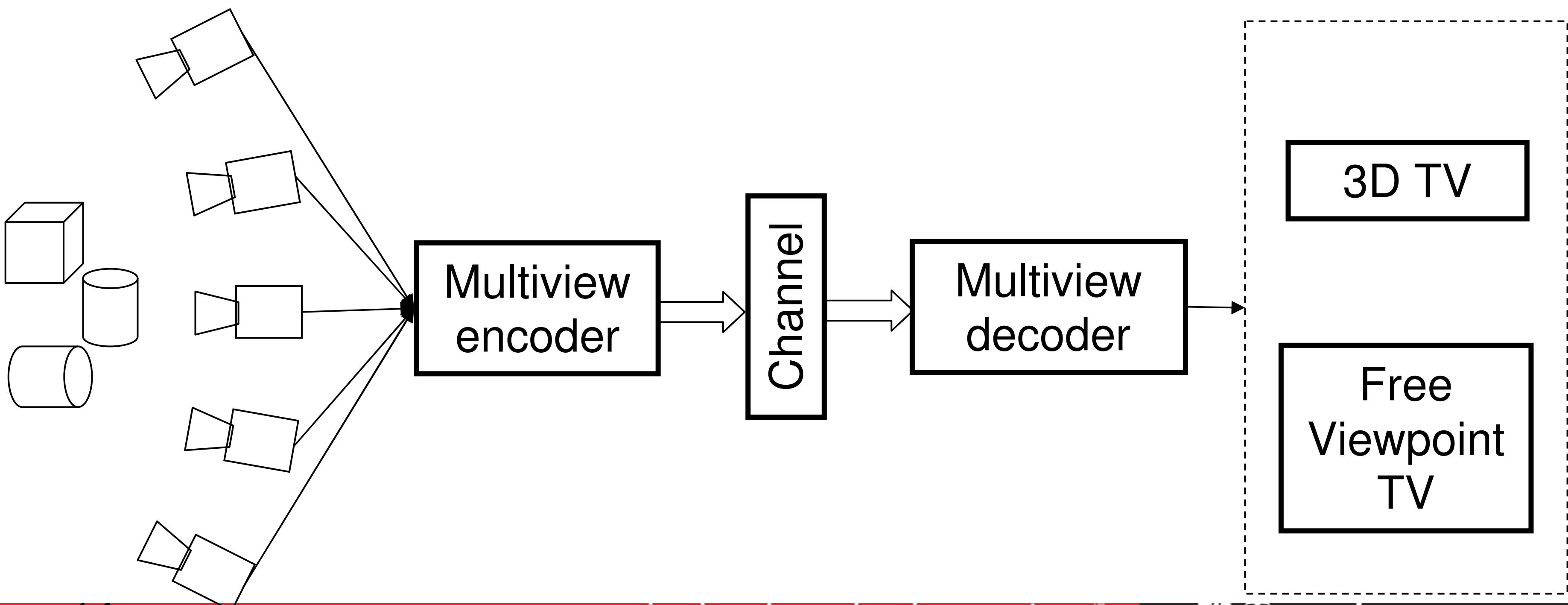
- Decoding complexity:
  - More complex than single-layer decoding because of inter-layer prediction
- Encoding complexity:
  - More complex than independent encoding of all layers because of inter-layer prediction
- Performance of scalability tools:
  - Temporal: 10-20% bitrate savings
  - Spatial:
    - 10% bit-rate penalty compared to single layer
    - 10% bit-rate savings compared to simulcast
  - SNR: 10-40% bit-rate penalty

SVC profiles

- Scalable Baseline
  - Base layer compliant to H.264/AVC Baseline profile with restrictions
  - Spatial scalability ratios of 2 and 1.5
  - Applications: Mobile, Video conferencing
- Scalable High
  - Base layer compliant to H.264/AVC High profile
  - Application: High-delay applications
- Scalable High Intra
  - I slice type only

Multiview coding

- Coding of video from multiple views (multiple cameras) of the same scene
- Application areas
  - 3D TV, Free Viewpoint TV (FTV)



3D TV

- 3D display of scenes
  - Requires two views
  - Stereoscopic displays
    - Require special glasses
  - Auto-stereoscopic displays
    - Require no glasses (e.g. using lenticular lens)

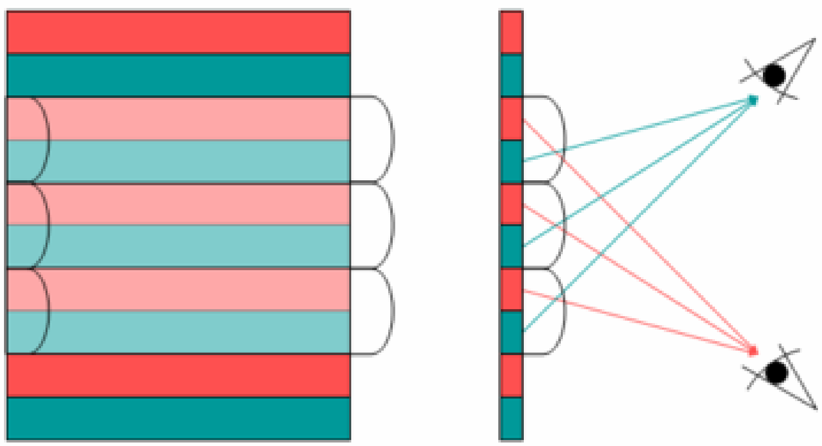


Figure from wikipedia

Minds in Motion

Free Viewpoint TV (FTV)

- User can choose own viewpoint within a scene
- Common in 3D games, but FTV tries to achieve it on natural video scenes
- Uses multiple cameras to record the same scene from different directions
- Appropriate view displayed on 2D screen based on viewer direction
  - Might require view interpolation
- FTV effort is being started in MPEG
  - Technology not mature

Coding of video from multiple views

- Exploits interview redundancy by using interview prediction – framework is based on H.264/AVC
- Need to compensate for:
  - Illumination changes in scene
  - Color balancing differences between cameras
  - Focus changes between views
- Standards work on-going, expected to be finalized April 2008
  - Around 1-1.5 dB gain reported by using interview prediction when compared to independent coding of views
  - Debate on-going on whether an alternate representation of video + depth is better than multiview coding for 3DTV

ITU-T SG16/Q6 activities towards a next-generation video coding standard

- New standards activity not yet launched
- However, improvements to H.264 are being presented in ITU-T Q6/SG16 Video Coding Experts Group (VCEG)
- Two ad-hoc groups (AhGs) in VCEG are carrying out technical work
  - Computational efficiency AhG
    - Focus on low-power and low-complexity work
  - Coding efficiency AhG

Motivation for low-power and low-complexity work

- Described in ITU-T contribution T05-SG16-C-0215: “Desired features in future video coding standards”
  - Texas Instruments, Nokia, Polycom, Samsung AIT, Tandberg, (Covi Technologies)
- Application areas:
  - Mobile phone video applications: Camera phones, Video telephones, Streaming video, Mobile TV
  - Handheld portable devices: Digital still cameras, camcorders, personal media players
  - Videoconferencing, Video surveillance
- Low power consumption and low complexity are important in all these application areas

Desired features in future video coding standards

- Low power consumption
  - Battery size limited in mobile devices because of cost and form factor considerations
  - Power-hungry real-time video encoding is a key requirement on these devices (video telephony and mobile user created content application) – therefore power issues should not be limited to decoder
  - Low power consumption important when encoding is performed in camera - heat negatively impacts video quality (sensor noise increases with temperature)

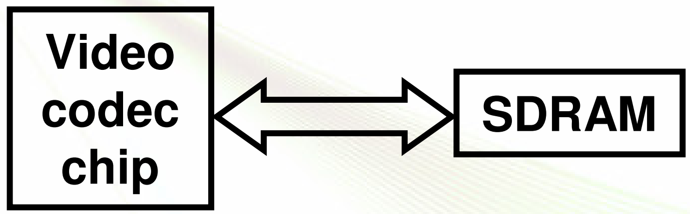
Desired features in future video coding standards (cont.)

- High coding efficiency at low complexity
  - Coding efficiency continues to be very important – because of limited transmission bandwidth and cost of storage
  - Devices are substantially resource constrained (in term of computational power, memory bandwidth, and memory size), hence low complexity coding desirable
- Low cost implementation
- Computational scalability
- Low delay
- Temporal scalability

Technology considerations for future video coding standards

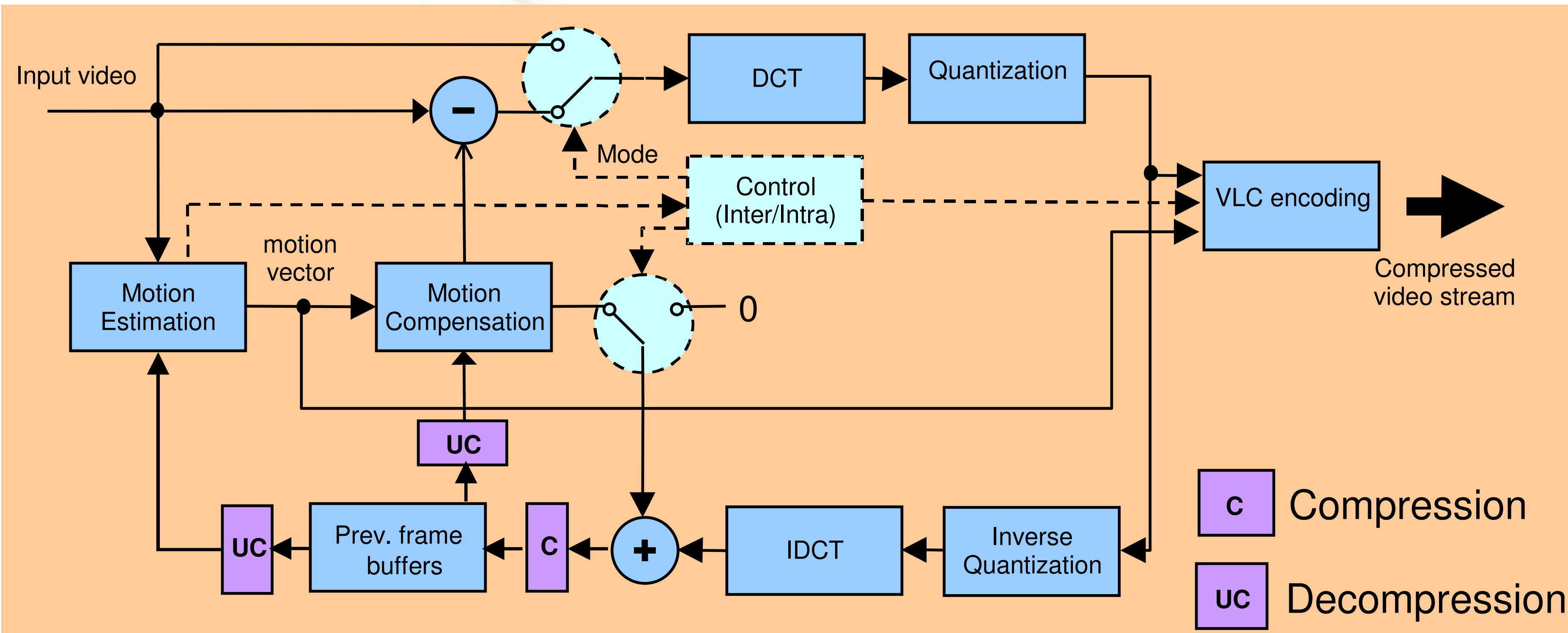
- Video coding methods should be suitable for implementation on several hardware platforms such as FPGAs, DSPs, ASICs, software configurable processors, as well as hybrid combinations of those platforms
- Video coding methods should address the following
  - Memory access bandwidth reduction
  - Enhanced parallelism support
  - Computation cycles reduction
  - Memory size reduction

Memory access bandwidth reduction



- SDRAM memory accesses consume power
  - Power consumed by memory accesses fairly significant when compared to codec power
- Increased memory bandwidth requirements (especially at HD resolution) lead to increased system cost
- Need techniques for reducing memory bandwidth

Video coding with compressed reference frames



- About 25% estimated savings in memory bandwidth and memory size

Other techniques presented in VCEG

- Simplified interpolation (fixed) and deblocking filters
- Adaptive interpolation filtering
- Intra prediction improvements
- Motion compensated prediction with 1/8-pel vector resolution
- Motion vector competition
- Adaptive prediction error coding in spatial and frequential domain
- Adaptive quantization matrix selection

Conclusions

- Scalable Video Coding
  - Extension of H.264/AVC
  - Base layer compliant to H.264/AVC
- Multiview Video Coding
  - Extensions of H.264/AVC being considered
  - Application areas still not mature
- Low-power consumption important goal for next-generation standard

Acknowledgements

- Minhua Zhou, TI, and several other TIers who provided feedback on the slides
- T. Wiegand, HHI, for permission to use material from his presentation JVT-W132 on SVC

Upcoming Video Standards

Madhukar Budagavi, Ph.D.
DSPS R&D Center, Dallas
Texas Instruments Inc.

Minds in Motion