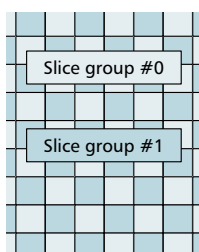


H.264/AVC VIDEO FOR WIRELESS TRANSMISSION

THOMAS STOCKHAMMER, NOMOR RESEARCH
MISKA M. HANNUKSELA, NOKIA RESEARCH CENTER



The authors introduce the features of the H.264/AVC coding standard that make it suitable for wireless video applications, including features for error resilience, bit rate adaptation, integration into packet networks, interoperability, and buffering considerations.

ABSTRACT

H.264/AVC will be an essential component in emerging wireless video applications thanks to its excellent compression efficiency and network-friendly design. However, a video coding standard itself is only one component within the application and transmission environment. Its effectiveness strongly depends on the selection of appropriate modes and parameters at the encoder and decoder, as well as in the network. In this article we introduce the features of the H.264/AVC coding standard that make it suitable for wireless video applications, including features for error resilience, bit rate adaptation, integration into packet networks, interoperability, and buffering considerations. Modern wireless networks provide many different means to adapt quality of service, such as forward error correction methods on different layers and end-to-end or link layer retransmission protocols. The applicability of all these encoding and network features depends on application constraints such as the maximum tolerable delay, possibility of online encoding, and availability of feedback and cross-layer information. We discuss the use of different coding and transport related features for different applications: video telephony and conferencing, video streaming, download-and-play, and video broadcasting. Guidelines for the selection of appropriate video coding tools, video encoder and decoder settings, and transport and network parameters are provided and justified. References to relevant research publications and standards contributions are given.

INTRODUCTION

Most emerging and future mobile client devices will significantly differ from those used for speech communications only: handheld devices will be equipped with a color display and a camera, and have sufficient processing power to allow presentation, recording, and encoding/decoding of video sequences. In addition, emerging and future wireless systems will provide sufficient bit rates to support video communication applications. Nevertheless, bit rates will always be scarce in wireless transmission environments due to physical bandwidth and power limitations; thus, efficient video compression is required. Nowadays H.263 and MPEG-4 Visual Simple

Profile are commonly used in handheld products, but it is foreseen that H.264/AVC [1] will be the video codec of choice for many video applications in the near future. The compression efficiency of the new standard outdoes prior standards roughly by at least a factor of two. Although compression efficiency is the major feature for a video codec to be successful in wireless transmission environments, it is also necessary that a standard provide means to be integrated easily into existing and future networks as well as address the needs of different applications.

This article is organized as follows. Several required features for a video codec to be used in wireless video applications are extracted. We introduce standard components in H.264/AVC that are relevant for wireless communication. We discuss the application of H.264/AVC for bit rate adaptation and error resilience, respectively. Finally, we conclude the article. For space reasons we have decided to exclude the explanation of all acronyms, the provision of extensive references, and the integration of simulation results from the printed version of the article. A supplemental Web page¹ has been set up addressing these issues.

VIDEO OVER WIRELESS

END-TO-END VIDEO TRANSMISSION

Figure 1 attempts to provide a suitable abstraction level of a video transmission system. In order to keep this article focused, we have excluded capturing and display devices, user interfaces, and security issues; also, most computational complexity issues are ignored.

The video encoder generates data units containing the compressed video stream, possibly stored in an encoder buffer before transmission. A wireless transmission system might delay, lose, or corrupt individual data units. The unavailability of a single data unit usually has significant impact on perceived quality due to spatio-temporal error propagation. In modern wireless system designs, data transmission is usually supplemented by additional information between the sender and the receivers, and within the respective entities. Abstract versions of available messages are included in Fig. 1; specific syntax and semantics as well the exploitation in video

¹ <http://www.nomor.de/H264/wcm.html>

transmission systems are discussed in more detail. Furthermore, each processing and transmission step adds some delay, which can be fixed, deterministic, or random. The encoder and decoder buffers allow compensating for variable bit rates produced by the encoder as well as channel delay variations to keep the end-to-end delay constant and maintain the timeline at the decoder. Nevertheless, if the *initial playout delay* is not or cannot be too extensive, late data units are commonly treated to be lost.

WIRELESS VIDEO APPLICATIONS

Ideally, high-quality video transmission would require high transmission bit rates, error-free delivery, as well as low and constant channel delays. Obviously, not all of the requests of the video application can be fulfilled; one has to live with the features and limitations of wireless systems as discussed in detail in [2]. Wireless transmission systems provide different transmission modes resulting in different quality of service (QoS) in terms of supported bit rates, bit rate variations, delay variations, as well as reliable delivery. The appropriate selection of transmission modes, adapted to the considered video application, is discussed. Furthermore, in Table 1 we also categorize video applications with respect to their maximum tolerable end-to-end delay, the availability and usefulness of different feedback messages, the availability and accurateness of channel state information at the transmitter, and the possibility of online encoding in contrast to pre-encoded content. Typical Third Generation Partnership Project (3GPP) applications within each category are mentioned. Especially real-time services streaming and conversational, but also broadcast, services provide challenges in wireless environments, as in general reliable delivery cannot be guaranteed. The suitability of H.264/AVC for these services is discussed in the following. For a review of the application standards and used protocols please refer to [2].

H.264/AVC IN WIRELESS SYSTEMS

Similar to previous video coding standards, the H.264/AVC standard specifies the *decoder operation* for error-free bitstreams as well as the *syntax*

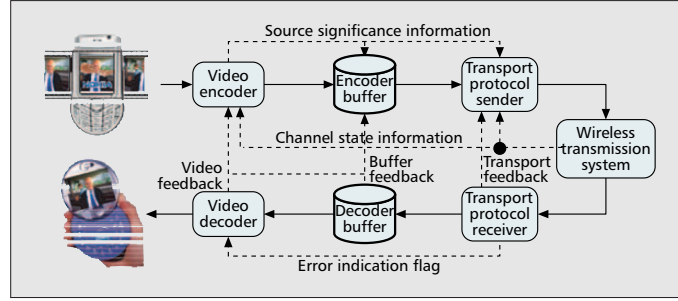


Figure 1. Abstraction of end-to-end video transmission systems.

and *semantics* of the bitstream. Consequently, the deployment of H.264/AVC still provides a significant amount of freedom for encoders and decoding of erroneous bitstreams. In the following subsections we introduce the essential features of H.264/AVC for wireless systems, categorized as network integration, compression efficiency, error resilience, and bit rate adaptivity. It is important to understand that most features are general enough to be used for multiple purposes rather than assigned to a specific application.

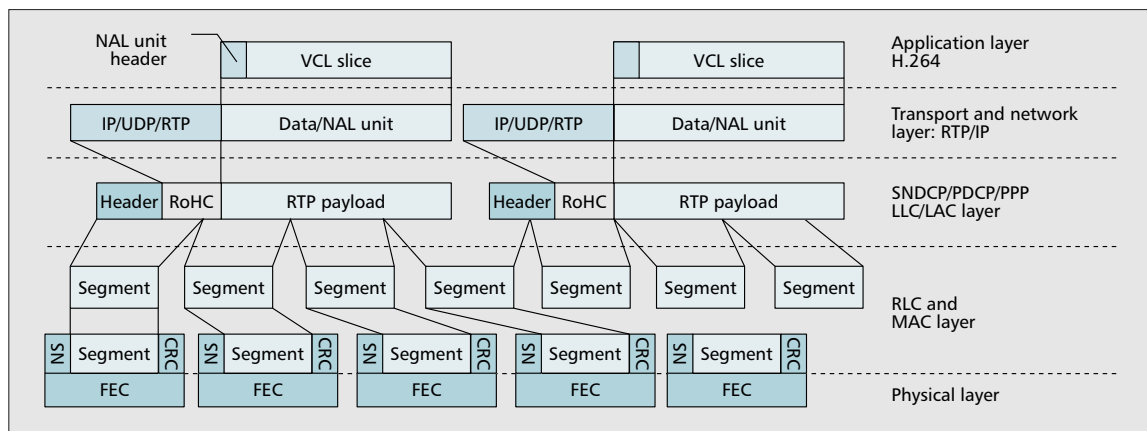
NETWORK INTEGRATION

NAL Units — The elementary unit processed by an H.264/AVC codec is called the network abstraction layer (NAL) unit, which can easily be encapsulated into different transport protocols and file formats, such as MPEG-2 transport stream, Real-Time Transfer Protocol (RTP), and MPEG4 file format. There are two types of NAL units, video coding layer (VCL) NAL units and non-VCL NAL units. VCL NAL units contain data that represent the values and samples of video pictures in the form of slices or slice data partitions. One VCL NAL unit type is dedicated for a slice in an instantaneous decoding refresh (IDR) picture. A non-VCL NAL unit contains supplemental enhancement information (SEI), parameter sets, picture delimiter, or filler data.

Each NAL unit consists of a one-byte header and the payload byte string. The header indi-

Video application	3GPP	Max. delay	Video/buffer feedback		Transport feedback		CSI	Encoding
			Available?	Useful?	Available?	Useful?		
Download-and-play	MMS	N/A	No	—	Yes	Yes	—	Offline
On-demand streaming (pre-encoded content)	PSS	≥ 1 s	Yes	Yes	Yes	Yes	Partly	Offline
Live streaming	PSS	≥ 200 ms	Yes	Yes	Partly	Yes	Partly	Online
Multicast	MBMS	≥ 1 s	Limited	Partly	Limited	Partly	Limited	Both
Broadcast	MBMS	≥ 2 s	No	—	No	—	No	Both
Conferencing	PSC	≤ 250 ms	Limited	Yes	No	—	Limited	Online
Telephony	PSC	≤ 200 ms	Yes	Yes	Limited	Yes	Partly	Online

Table 1. Characteristics of typical wireless video applications.



■ **Figure 2.** Integration example of an VCL slice in the RTP payload and 3GPP framework.

icates the type of NAL unit and whether a VCL NAL unit is part of a reference or non-reference picture. Furthermore, syntax violations in the NAL unit and the relative importance of the NAL unit for the decoding process can be signaled in the NAL unit header.

Parameter Set Concept — H.264/AVC allows sending of sequence and picture level information reliably, asynchronously, and in advance of the media stream that contains the VCL NAL units by the use of parameter sets. Sequence and picture level data are organized into sequence parameter sets (SPS) and picture parameter sets (PPS), respectively. An active SPS remains unchanged throughout a coded video sequence (i.e., until the next IDR picture), and an active PPS remains unchanged within a coded picture. The parameter set structures contain information such as picture size, optional coding modes employed, and macroblock to slice a group map. In order to be able to change picture parameters such as picture size without the need to transmit parameter set updates synchronously to the slice packet stream, the encoder and decoder can maintain a list of more than one SPS and PPS. Each slice header contains a codeword that indicates the SPS and PPS in use.

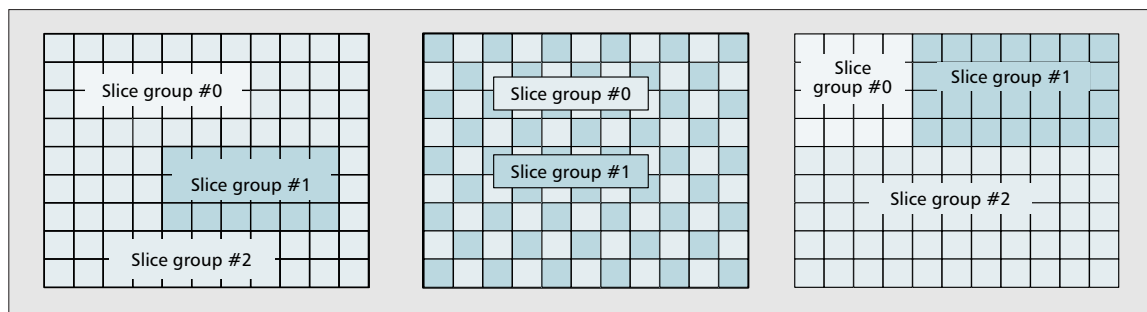
Integration in RTP and 3GPP Multimedia Services — The integration of multimedia services in 3G wireless systems has been addressed in the recommendations of 3GPP (for details see [2]). In the following we concentrate on packet-based real-time video services. H.264/AVC was lately adopted as a recommended codec for all 3GPP video services. Figure 2 shows the basic processing of a VCL slice within the RTP and 3GPP framework. The slice is packetized in a NAL unit, which itself is encapsulated in RTP/UDP/IP according to [3] and finally transported through the protocol stack of a wireless system such as General Packet Radio Service (GPRS), Enhanced GPRS (EGPRS), Universal Mobile Telecommunications System (UMTS), or code-division multiple access (i.e., CDMA2000). The RTP payload specification supports different packetization modes. In the simplest mode a sin-

gle NAL unit is transported in a single RTP packet, and the NAL unit header coserves as an RTP payload header. In noninterleaved mode, several NAL units of the same picture can be packetized into the same RTP packet. In interleaved mode several NAL units of potentially different pictures can be packetized into the same RTP packet, and NAL units do not have to be sent in their decoding order. Both the non-interleaved and interleaved modes also allow fragmentation of a single NAL unit into several RTP packets.

In the following we concentrate on UMTS terminology; the corresponding layers for other systems are shown in Fig. 2. After potential RoHC, the generated IP/UDP/RTP packet is encapsulated into a single PDCP packet that becomes an RLC-SDU. As a typical RLC-SDU has a larger size than a RLC-PDU, it is then segmented into smaller units. The length of the RLC-PDU depends on the selected bearer as well as the coding and modulation scheme in use. The RLC layer in wireless systems can operate in unacknowledged mode (UM) and acknowledged mode (AM), both providing RLC-PDU loss detection. However, whereas UM is unidirectional and data delivery is not guaranteed, in AM an automatic repeat request (ARQ) mechanism is used for error correction. The physical layer generally adds forward error correction (FEC) to RLC-PDUs depending on the coding scheme in use so that a constant length channel-coded and modulated block is obtained. This channel-coded block is further processed in the physical layer before it is sent to the far-end receiver. The receiver performs error correction and detection, and possibly requests retransmissions. It is important to understand that in general the detection of a lost segment results in the loss of an entire PDCP packet, so the encapsulated IP and RTP packet as well as the NAL unit is lost.

COMPRESSION EFFICIENCY

Compression efficiency is the major attribute for a video codec to be successful in wireless environments. Although the design of the VCL of H.264/AVC basically follows the design of prior



■ **Figure 3.** Macroblock allocation maps: foreground slice groups with one left-over background slice group, checkerboard-like pattern with two slice groups, and sub-pictures within a picture.

video coding standards, it contains many new details that enable significant improvement in terms of compression efficiency; for details the interested reader is referred to [1, 4]. The gains do not come from a single new technique, but from an ensemble of advanced prediction, quantization, and entropy coding schemes.

The encoder implementation is responsible for appropriately selecting a combination of different encoding parameters, so-called *operational coder control*. When using a standard with a completely specified decoder, parameters in the encoder should be selected such that good rate-distortion performance is achieved. For a video coder like H.264/AVC, the encoder must select parameters, such as motion vectors, macroblock modes, quantization parameters, reference frames, and spatial and temporal resolution, to provide good quality under given rate and delay constraints. Two simplify matters in deciding on good selections of the coding parameters, commonly this task is divided into three levels [5]:

- *Encoder control* performs local decisions, such as the selection of macroblock modes, reference frames, or motion vectors, on the macroblock level and below most appropriately based on a rate-distortion optimized mode selection applying Lagrangian techniques.
- *Rate control* mainly controls the timing and bit rate constraints of the application by adjusting the QP or Lagrange parameter and is usually applied to achieve a constant bit rate (CBR) encoded video suitable for transmission over CBR channels. The aggressiveness of the quantization/Lagrangian parameter change allows a trade-off between quality and instantaneous bit rate characteristic of the video stream.
- *Global parameter selection* selects the appropriate temporal and spatial resolution of the video based on application, profile, and level constraints. Also, packetization modes, like slice sizes, are usually fixed for the entire session. The parameters are mainly determined by general application constraints.

ERROR RESILIENCE AND BIT RATE ADAPTIVITY FEATURES IN H.264/AVC VCL

In the following we introduce different error resilience and bit rate adaptivity features included in H.264/AVC VCL with respect to their

functionality. For more details we refer to [4, 6, 7, references therein].

Slice Structured Coding — Slices provide spatially distinct resynchronization points within the video data for a single frame. This is accomplished by introducing a slice header, which contains syntactical and semantical resynchronization information. In addition, intra-prediction and motion vector prediction are not allowed over slice boundaries. The encoder can select the location of the synchronization points at any macroblock boundary.

Flexible Macroblock Ordering — FMO allows mapping of macroblocks to slice groups, where a slice group itself may contain several slices. Therefore, macroblocks might be transmitted out of raster scan order in a flexible and efficient way. Some examples of macroblock allocation maps for different applications are shown in Fig. 3. Dispersed macroblock allocations are especially powerful in conjunction with appropriate error concealment (i.e., when the samples of a missing slice are surrounded by many samples of correctly decoded slices).

Arbitrary Slice Ordering — ASO allows that the decoding order of slices within a picture may not follow the constraint that the address of the first macroblock within a slice is monotonically increasing within the NAL unit stream for a picture. This permits, for example, reduction of decoding delay in case of out-of-order delivery of NAL units.

Slice Data Partitioning — In data partitioning mode, each slice can be separated in a header and motion information, intra information, and intertexture information by simply distributing the syntax elements to individual NAL units. Due to this reordering on syntax level, coding efficiency is not reduced, but obviously the loss of individual partitions still results error propagation.

Intra-coding — H.264/AVC distinguishes IDR pictures and regular intra-pictures whereby the latter do not necessarily provide the random access property as pictures before the intra pictures may be used as reference for succeeding predictively coded pictures. H.264/AVC also allows intracoding of single macroblocks for regions

For many use cases it is necessary to adapt the bit rate dynamically in the application to larger bit rates and timescales larger than the initial playout delay allows. In wireless streaming environments bitstream switching provides a simple but powerful means to support bit rate adaptivity.

that cannot be predicted efficiently or due to any other case where the encoder decides for non-predictive mode. The intra mode can be modified such that intra-prediction from predictively coded macroblocks is disallowed. The corresponding constraint intra flag is signaled in the PPS.

Redundant Slices — A redundant coded slice is a coded slice that is a part of a redundant coded picture, which itself is a coded representation of a picture that is not used in the decoding process if the corresponding primary coded picture is correctly decoded. The redundant slice should be coded such that there is no noticeable difference between any area of the decoded primary picture and a decoded redundant picture.

Flexible Reference Frame Concept — H.264/AVC allows reference frames to be selected in a flexible way on a macroblock basis, which provides the possibility to use two weighted reference signals for macroblock interprediction, allows frames to be kept in short-term and long-term memory buffers for future reference, and finally provides temporal scalability. The classical I, P, B frame concept is replaced by a highly flexible and general concept that can be exploited by the encoder for different purposes. However, this concept also requires that not only is the HRD [8] specified in the bitstream domain, but it is also necessary that the encoder be constrained in the amount of frames to be stored in the decoded picture buffer.

Switching Pictures — H.264/AVC allows applying mismatch-free predictive coding even where there are different reference signals. So-called primary SP-frames are introduced in the encoded bitstream, which are in general slightly less efficient than regular P-frames but significantly more efficient than regular I-frames. The major benefit results from the fact that this quantized reference signal can be generated mismatch-free using any other prediction signal. If this prediction signal is generated by predictive coding, the frame is referred to as secondary SP-pictures, which are usually significantly less efficient than P-frames as an exact reconstruction is necessary. To also generate this reference signal without any predictive signal, so-called switching-intra (SI) pictures can be used. SI pictures are only slightly less inefficient than common I pictures and can also be used for adaptive error resilience purposes. For more details on this unique feature within H.264/AVC the interested reader is referred to [9].

BIT RATE ADAPTIVITY PROVISION IN H.264/AVC

Bit rate adaptivity is one of the most important features for applications in wireless systems to react to the dynamics due to statistical traffic, variable receiving conditions, as well as handovers and random user activity. Due to the applied error control features, these variations mainly result in varying bit rates in different

timescales. For applications where online encoding is performed and the encoder has sufficient feedback on the expected bit rate on the channel by some channel state or decoder buffer fullness information, *rate control for VBR channels* can be applied. H.264/AVC obviously supports these features, mainly by the possibility of changing QPs dynamically, but also by the changing temporal resolution.

When channel bit rate fluctuations are not a priori known at the transmitter, or there is no sufficient means or necessity to change the bit rate frequently, *playout buffering* at the receiver can compensate for bit rate fluctuations to some extent. In addition, for anticipated buffer underrun, techniques such as *adaptive media playout* allow a streaming media client, without involvement of the server, to control the rate at which data is consumed by the playout process.

However, these techniques might not be sufficient to compensate for bit rate variations in wireless applications. In this case rate adaptation has to be performed by modifying the encoded bitstream. In today's systems rate adaptation is typically carried out in streaming servers. It is well known that intelligent decisions to drop less important packets rather than dropping random packets — this is treated under the framework of error resilience — can significantly enhance the overall quality. A formalized framework called *rate-distortion optimized packet scheduling* has been introduced [10] and serves as the basis for several subsequent publications. Applying the framework is easiest when important and less important packets are identified in the encoding process. H.264/AVC provides different approaches to support packets with different importance for bit rate adaptivity. First, the temporal scalability features [11] of H.264/AVC relying on the reference frame concept can be used. Second, if frame dropping is not sufficient, one might apply *data partitioning* which can be viewed as a very coarse but efficient method for SNR scalability. Third, flexible macroblock ordering may also be used for prioritization of regions of interest. For example, a background slice group can be dropped in favor of a more important foreground slice group.

For many use cases it is necessary to adapt the bit rate dynamically in the application to larger bit rates and timescales larger than the initial playout delay allows. In wireless streaming environments *bitstream switching* provides a simple but powerful means to support bit rate adaptivity. In this case the streaming server stores the same content encoded with different versions in terms of rate and quality. In addition, each version provides a means to randomly switch into it. IDR pictures provide this feature, but they are generally costly in terms of compression efficiency. The SP-frame concept in H.264/AVC can be used to reduce the loss of compression efficiency in stream switching. In this case the streaming server not only stores different versions of the same content, but also secondary SP pictures as well as SI pictures. When switching streams, the server sends an appropriate secondary SP picture or SI picture [9, 12].

ERROR ROBUSTNESS SUPPORT USING H.264/AVC

This section discusses the endpoint operation in a wireless H.264/AVC video system. The provided H.264/AVC features can be used exclusively or jointly for error robustness purposes, depending on the application. It is necessary to understand that most codec-level error resilience tools decrease compression efficiency. Therefore, the main goal when transmitting video goes along the spirit of Shannon's famous *separation principle* [13]: Combine compression efficiency with link layer features that completely avoid losses such that the two aspects, compression and transport, can be completely separated. Nevertheless, if errors cannot be avoided, the following system design principles are essential:

- *Loss correction below the codec layer*: Minimize the amount of losses in the wireless channel without completely sacrificing the video bit rate.
- *Error detection*: If errors are unavoidable, detect and localize erroneous video data.
- *Prioritization methods*: If losses are unavoidable, at least minimize loss rates for very important data (e.g., control).
- *Error recovery and concealment*: In case of losses, minimize the visual impact of losses on the actual distorted image.
- *Encoder-decoder mismatch avoidance*: Limit or completely avoid encoder and decoder mismatches resulting in annoying error propagation. Use cases of the error resilience features for specific applications are discussed.

Error Control Methods — Error control such as FEC and retransmission protocols are the primary tool to provide QoS in mobile systems, especially on the radio access part. QoS methods are essential in good system designs as minimizing or vanishing transmission errors has many advantages for applications. However, usually the trade-offs between reliability vs. delay have to be considered. Nevertheless, to compensate for the shortcomings of non-QoS-controlled networks (e.g., the Internet or some mobile systems) as well as address total blackout periods caused by, say, network buffer overflow or a handover of transmission cells, advanced transport protocols provide features that allow error control to be introduced at the application layer. For example, MBMS services make use of an application-layer FEC scheme. For point-to-point services, selective application layer retransmission schemes can be used to retransmit RTP packets. For many applications it can be assumed that at least a low-bit-rate feedback channel from the receiver to the transmitter exists that allows general back-channel messages to be sent. For example, RTP is accompanied by Real-Time Transport Control Protocol (RTCP) providing control and management messages. Media receivers can send *receiver reports* including instantaneous and cumulative loss rates as well as delay and jitter information. RTCP has recently been extended with the *extended report* packet type, which allows the loss or reception of each RTP packet to be indicated by the receiver to the sender.

Resynchronization and Error Concealment — Despite error control techniques, error resilience in the

video is still necessary whenever the video decoder observes residual losses. According to [2], these problems mainly occur in conversational applications due to the delay constraints as well as in multicast/broadcast situations due to the missing feedback link. Slice structured coding typically allows the encoder to choose between two slice coding options, one with a constant number of macroblocks within one slice but arbitrary size of bytes, and one with the slice size bounded to some maximum $S_{\max}$ in bytes, resulting in an arbitrary number of macroblocks per slice. The latter is especially useful to introduce some QoS as commonly the slice size determines loss probability in wireless systems due to the processing shown in Fig. 2.

H.264/AVC decoders should detect losses of slices by keeping a record of which slices of a picture have been received and decoded. Entirely lost reference pictures should be detected based on gaps in the sequence number for reference pictures (the `frame_num` syntax element of the H.264/AVC bitstream) or prediction from missing pictures in the reference picture buffer (when a bitstream may include subsequences).

As soon as the erroneous macroblocks are detected, *error concealment* for all of them should be invoked. For example, in the H.264/AVC test model software two types of error concealment algorithms have been introduced [7], one exploiting spatial information only, suitable mainly for intra frames, and one exploiting temporal information. It is important to select the appropriate error concealment technique, spatial or temporal, adaptively to obtain a reasonably good visual quality. This selection can be concluded from a coded slice (e.g., the macroblock mode information of reliable neighbors), or encoders can assist decoders in the decision-making, say, by including spare picture and scene information SEI messages.

Limitation of Temporal Error Propagation — Despite the presented error control and concealment techniques, packet losses still result in imperfect reconstruction of pictures. Thus, the effects of spatio-temporal error propagation resulting from the motion compensated prediction can be severe. Therefore, the decoder has to be provided with other means that allow error propagation to be reduced or completely stopped.

The most common way to accomplish this task is the reduction of temporal prediction in the encoding process by encoding image regions in intra mode. The straightforward way of inserting IDR frames is quite common for broadcast and streaming applications as these frames are also necessary to randomly access the video sequences. However, when transported over CBR channels, the latency caused by IDR pictures can be undesirable, especially in conversational applications. Therefore, more subtle methods are frequently used to synchronize encoder and decoder reference frames.

In early work it was proposed to introduce intra-coded macroblocks using a constant update pattern, randomly, or adaptively based on a cost function. The selection of an appropriate update ratio depends on different parameters such as

The effects of spatio-temporal error propagation resulting from motion compensated prediction can be severe. Therefore, the decoder has to be provided with other means that allow error propagation to be reduced or completely stopped.

Although the standardization process is finalized, the freedom at the encoder as well as the combination with transport modes such as FEC and retransmission strategies promises optimization potentials.

the sequence characteristics, transmission bit rate, and, most important, channel characteristics. Most suitably, the selection of coding modes can be incorporated in the operational encoder control taking into account the influence of the lossy channel. The encoder control is modified such that the expected decoder distortion is used instead of the encoding distortion. For details on the computation of the expected decoder distortion see [7]. In addition to limiting the error propagation with macroblock intra updates, encoders can also guarantee that macroblock intra updates result in gradual decoding refresh (GDR), that is, entirely correct output pictures after a certain period of time. GDR can be signaled with the recovery point SEI message of H.264/AVC and implemented in H.264/AVC encoders using the isolated regions coding technique [14].

The availability of the feedback channel in conversational and unicast streaming applications has led to different standardization and research activities on interactive error control (IEC) in recent years. If *online encoding* is performed, the slice loss information of the decoder can be directly incorporated in the encoding process to reduce, eliminate, or even completely avoid error propagation. The basics of these methods have been founded under the term *error tracking* [15]. The syntax of H.264/AVC permits incorporating methods for reduced or limited error propagation in a straightforward manner. Similar to operational encoder control for error-prone channels, the delayed decoder state can also be integrated in modified encoder control. By the use of SP-pictures IEC can even be extended to applications with offline encoding.

Prioritization and Data Differentiation — The inherent property of video and multimedia data compared to file download results from the fact that some data might be more important than others. The optimized combination of different video features with different transmission features is also referred to as *cross-layer design*; a similar rate-distortion framework as presented in [10] can be used. Some work in this area shows promising potential to be exploited in future system designs, but it is also important to note that some designs do not provide sufficient gains to replace conventional transmission modes in practical systems. Examples for the provision of different priority modes in end-to-end as well as transmission systems include unequal error protection, unequal erasure protection, selective retransmission, proxies, or different priority queues when accessing shared channels. These systems can be combined with H.264/AVC bit rate adaptivity features, such as different frame types, subsequences, and data partitioning.

Timing constraints, ruling out, for example, retransmissions, only relate to data that is live generated such as conversational video, live streaming or live broadcasting, wherein the sending time of the data is usually closely coupled to the display time, referred to as *time-stamp-based* streaming. If pre-encoded data is transmitted and the decoder buffer is sufficiently large, one can transmit data earlier than its nom-

inal sending time, so-called *ahead-of-time* streaming, which allows better exploitation of the channel. This strategy can be even extended by transmitting more important data earlier, allowing more retransmissions for this important data [12]. H.264/AVC even extends this concept to live encoding by the provision of parameter sets and long-term multiple reference frames.

SUMMARY AND OUTLOOK

In this work the benefits of H.264/AVC in wireless transmission environments have been shown. In addition to excellent compression efficiency, H.264/AVC provides features that can be used in one or several application scenarios, and also allows easy integration in most networks. The selection and combination of different features strongly depends on the system and application constraints, namely bit-rates, maximum tolerable playout delays, error characteristics, online encoding possibility, as well as availability of feedback and cross-layer information. Although the standardization process is finalized, the freedom at the encoder as well as combination with transport modes such as FEC and retransmission strategies promise optimization potentials. Therefore, further research in the areas of optimization, cross-layer design, feedback exploitation, and error concealment is necessary to fully understand the potential of H.264/AVC in wireless environments. However, researchers are especially encouraged to integrate transport protocols as well as wireless system options into their considerations rather than assuming QoS-unaware link and transport layers.

ACKNOWLEDGMENTS

The authors would like to thank Thomas Wiegand, Stephan Wenger, Ye-Kui Wang, Günther Liebl, and Ingo Viering for useful discussions on the subject of this work.

REFERENCES

- [1] ITU Rec. H.264/ISO IEC 14996-10 AVC, "Advanced Video Coding for Generic Audiovisual Services," 2003.
- [2] M. Etoh and T. Yoshimura, "Wireless Video Applications in 3G and Beyond," *IEEE Wireless Commun.*, this issue.
- [3] S. Wenger et al., "RTP payload Format for H.264 Video," IETF RFC 3984, Feb. 2005.
- [4] G. Sullivan and T. Wiegand, "Video Compression — From Concepts to H.264/AVC Standard," *Proc. IEEE, Special Issue on Advances in Video Coding and Delivery*, vol. 93, no. 1, Jan. 2005, pp. 18–31.
- [5] T. Wiegand et al., "Rate-Constrained Coder Control and Comparison of Video Coding Standards," *IEEE Trans. Circuits and Sys. for Video Tech.*, vol. 13, July 2003, pp. 688–703.
- [6] S. Wenger, "H.264/AVC over IP," *IEEE Trans. Circuits and Sys.*, vol. 13, no. 7, July 2003, pp. 645–56.
- [7] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in Wireless Environments," *IEEE Trans. Circuits and Sys.*, vol. 13, no. 7, July 2003, pp. 657–73.
- [8] J. Ribas-Corbera, P. A. Chou, and S. Regunathan, "A Generalized Hypothetical Reference Decoder for H.264/AVC," *IEEE Trans. Circuits and Sys.*, vol. 13, no. 7, July 2003, pp. 674–87.
- [9] M. Karczewicz and R. Kurçeren, "The SP and SI Frames Design for H.264/AVC," *IEEE Trans. Circuits and Sys.*, vol. 13, no. 7, July 2003.
- [10] P. A. Chou and Z. Miao, "Rate-distortion Optimized Streaming of Packetized Media," *IEEE Trans. Multimedia*, vol. Feb. 2001, <http://research.microsoft.com/pachou>
- [11] D. Tian, M. M. Hannuksela, and M. Gabbouj, "Subsequence Video Coding for Improved Temporal Scalability," *Proc. 2005 IEEE Int'l. Symp. Circuits and Sys.*, May 2005.

- [12] T. Stockhammer, M. Walter, and G. Liebl, "Optimized H.264-Based Bitstream Switching for Wireless Video Streaming," *Proc. ICME*, July 2005, Amsterdam, The Netherlands.
- [13] C. E. Shannon, *The Mathematical Theory of Communication*, University of Illinois Press, 1948.
- [14] M. M. Hannuksela, Y.-K. Wang, and M. Gabbouj, "Isolated Regions in Video Coding," *IEEE Trans. Multimedia*, vol. 6, no. 2, Apr. 2004, pp. 250–67.
- [15] B. Girod and N. Färber, "Feedback-Based Error Control for Mobile Video Transmission," *Proc. IEEE*, vol. 97, no. 10, Oct. 1999, pp. 1707–23.

BIOGRAPHIES

MISKA M. HANNUKSELA (miska.hannuksela@nokia.com) is a research manager in the Multimedia Technologies Laboratory, Nokia Research Center. He has been an active participant in the ITU-T Video Coding Experts Group since 1999 and in the Joint Video Team of ITU-T and ISO/IEC since its

foundation in 2001. He has co-authored more than 80 technical contributions to these standardization groups. His research interests include video error resilience and video communication systems.

THOMAS STOCKHAMMER (stockhammer@nomor.de) worked at the Munich University of Technology, Germany, and has been a visiting researcher at Rensselaer Polytechnic Institute, Troy, New York, and the University of California at San Diego (UCSD). He has published more than 60 conference and journal papers and holds several patents. He regularly participates and contributes to different standardization activities, such as JVT, IETF, and 3GPP and has co-authored more than 70 technical contributions. He is co-founder of Novel Mobile Radio (NoMoR) Research. Since 2004 he has been working as a research and development consultant for Siemens Mobile Devices. His research interests include video transmission, cross-layer and system design, rate-distortion optimization, information theory, and mobile communications.

IEEE NETWORK MAGAZINE

CALL FOR PAPERS

SPECIAL ISSUE: WIRELESS SENSOR NETWORKING

Wireless Sensor Networks (WSNs) recently received tremendous attention from both academia and industry because of its promise of a wide range of potential applications in both civil and military areas. A WSN consists of a large number of small sensor nodes with sensing, data processing, and communication capabilities, which are deployed in a region of interest and collaborate to accomplish a common task, such as environmental monitoring, military surveillance, and industry process control. Distinguished from traditional wireless networks, WSNs are characterized of dense node deployment, unreliable sensor node, frequent topology change, and severe power, computation, and memory constraints. These unique characteristics and constraints present many new challenges to the design and implementation of WSNs, such as energy conservation, self-organization, efficient data dissemination, and fault tolerance. For example, energy efficiency is the key to prolonging the network lifetime and is thus of primary importance in WSNs. It must be considered not only at the physical layer but also at the link layer and the network layer in sensor network design. Although many networking protocols and algorithms have been developed for traditional wireless ad hoc networks, they cannot effectively address the unique characteristics and constraints and application requirements of sensor networks. To meet the new challenges, innovative protocols and algorithms are needed to achieve energy efficiency, flexible scalability and adaptability, and good network performance. For example, it is highly desirable to develop new energy-efficient protocols for topology discovery, self-organization, medium access control, route discovery, and data dissemination. An efficient query processing and data aggregation algorithm can significantly reduce the number of transmissions of sensor nodes and thus provide substantial energy savings and prolong the lifetime of the network. In addition, open standards are important and imperative to facilitate and improve the development of WSNs. To realize the vision of WSNs, a large amount of research and development activities are going on in recent years. The purpose of this special issue is to expose the readership of IEEE Network to the latest research and development progress in this hot and exciting area.

SCOPE OF CONTRIBUTIONS

This special issue aims to publish a collection of research and survey articles that focus on the latest research and development results in all networking aspects of WSNs. Original research and survey articles are solicited from all researchers and practitioners. Articles should be tutorial in nature and should be written in a style comprehensible to the readers outside the specialty of the article. As applicable to WSNs, topics of interest include but are not limited to:

- Network architectures and protocols
- Energy-efficient medium access control (MAC)
- Query processing and data aggregation
- Field trials and standardization activities
- Topology discovery and self-organization
- Energy-efficient routing and data dissemination
- Fault tolerance and self-healing

MANUSCRIPT SUBMISSION

Authors should submit their manuscripts electronically in PDF format via email to one of the guest editors. With regard to both the content and formatting style of the submissions, prospective contributors should follow the IEEE Network guidelines for authors that can be found at <<http://www.comsoc.org/pubs/net/ntwrk/authors.html>>.

SCHEDULE FOR SUBMISSIONS

Submission deadline: September 1, 2005
 Acceptance notification: January 1, 2006
 Final manuscript due: March 1, 2006
 Publication date: 2nd Quarter, 2006

GUEST EDITORS

Dr. Jun Zheng
 University of Ottawa
 E-mail: jzheng@ieee.org

Dr. Pascal Lorenz
 University of Haute Alsace
 E-mail: lorenz@ieee.org

Dr. Petre Dini
 Cisco Systems, Inc.
 E-mail: pdini@cisco.com