SPIE Proceedings Vol. CR60, Standards and Common Interfaces for Video
Information Systems, October 25-26, 1995, Philadelphia, USA

Comparison of the H.263 and H.261 Video Compression Standards

Bernd Girod, Eckehard Steinbach, Niko Färber

Telecommunications Institute
University of Erlangen-Nuremberg
Cauerstrasse 7, 91058 Erlangen, Germany
girod@nt.e-technik.uni-erlangen.de

Invited Paper

Dedicated to Prof. Dr.-Ing. H. G. Musmann on the occasion of his 60th birthday

ABSTRACT

The draft international standard ITU-T H.263 is closely related to the well known and widely used ITU-T Recommendation H.261. However, H.263 does provide the same subjective image quality at less than half the bit-rate. In this paper we investigate to what extend single enhancements of H.263 contribute to this performance gain, and consider the trade-off quality vs. complexity. Based on the test sequence "Foreman", H.263 in its default- and optional coding-modes is compared to H.261 on the basis of rate distortion curves at bit-rates up to 128 kbps. At 64 kbps, the performance gain of H.263 in its default mode compared to H.261 is approximately 2 dB PSNR. This improvement is achieved with only little increase of complexity, and is mainly due to more accurate motion compensation with half-pel accuracy. Considering the trade-off quality vs. complexity, the combination of the optional coding-modes "Advanced prediction mode" and "PB-frames mode" seems to be a good compromise, resulting in an additional performance gain of 1 dB PSNR at 64 kbps. The "Syntax-based arithmetic coding mode" on the other hand, offers only a very small performance gain (0.2 dB PSNR at 64 kbps) for its increased computational complexity. Results from profiling a H.263 software codec are presented in order to support complexity considerations of the optional coding-modes.

1 INTRODUCTION

The demand for videophone applications over communication channels with low bandwidth like PSTN or mobile links requires new standards for the compression of image sequences at very low bit-rates. Though video transmission at 64 kbps was already reported in 1979 [1], reasonable picture quality at bit-rates below 64 kbps is still challenging. However, the ITU-T Study Group XV has now drafted such a standard suitable for video transmission below 64 kbps.

The ITU-T draft international standard H.263 [3] is closely related to the well known and widely used ITU-T recommendation H.261 [2], which has been designed by the same Study Group. This close relationship helped to arrive at the new standard in a short period of time, including not only the video coding algorithm but also the corresponding audio (G.723), multiplex (H.223), control (H.245) and system (H.324) aspects. Though H.261 and H.263 share the same basic codec structure, there is a significant improvement in performance. Side-by-side comparisons show that the same subjective image quality can be achieved with less than half the bit-rate. This performance gain is due to improved and optimized coding techniques, as well as optional coding-modes ("options") which may be switched on by the coder. For a given application, however, the trade-off quality vs. complexity has to be considered as well. Therefore, a comparison of the H.263 and H.261 video compression standards regarding quality and complexity will be investigated in this paper. We assume that the reader is familiar with H.261.

The paper is organized as follows. In section 2 we briefly describe the hybrid coding structure common to H.261 and H.263. In section 3 the differences between the two standards are pointed out, followed by a description of the additional options of H.263 in section 4. Then we compare the performance of H.261 and H.263 with several options in section 5. The increase in complexity due to the H.263 options is discussed in section 6.

2 CODER STRUCTURE

Both H.261 [2] and H.263 [3] use the same basic structure of the encoder (Fig. 1). It is a hybrid of interframe prediction exploiting temporal redundancy and transform coding of the residual prediction error exploiting spatial redundancy and adaptively reducing spatial resolution. Temporal prediction is based on a block-based motion estimation (ME) and compensation (MC), while a discrete cosine transform (DCT) is used for spatial redundancy reduction.

The motion compensated prediction error is subdivided into 8×8 blocks and each block is transformed by the DCT. The coefficients are quantized (Q) and pairs of zero-runs and quantizer levels are combined, resulting in an variable length codeword (VLC) for each block. One of 31 uniform threshold quantizers may be selected adaptively. Finally, run-level pairs, motion vectors (MV), and quantization parameters (Qp) are entropy coded along with other side information and multiplexed to the bitstream.

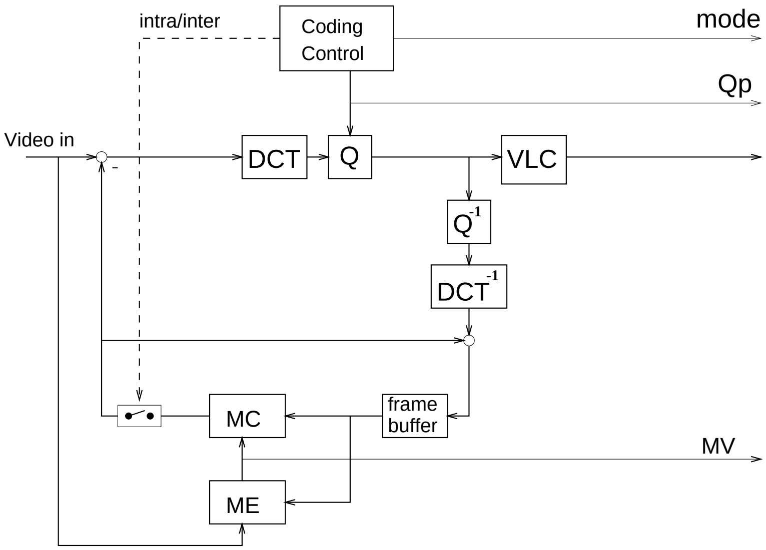

Figure 1: The hybrid coding structure of H.261 and H.263

3 DIFFERENCES

Though H.261 and H.263 share the same basic coding structure, there are both minor and major differences. These are covered in this section. A special section is dedicated to the H.263-options, which are further enhancements not included in H.261.

Target bit-rate

The target bit-rate of H.261 is p×64 kbps (p = 1,2,. . .,30) whereas H.263 aims at bit-rates below 64 kbps. Those bit-rates reflect the typical applications envisioned when the standards were designed. In the case of H.261, this was visual telephony over ISDN, whereas H.263 shall enable the same service over PSTN (e.g., using V.34).

Picture formats

H.261 operates on two picture formats, CIF (Common Intermediate Format) and quarter-CIF (QCIF). As the name implies, the spatial resolution of QCIF is one quarter of the CIF-resolution. In addition to CIF and QCIF, H.263 supports a third format called sub-QCIF. As can be seen in Table 1, the spatial resolution is further reduced by a factor of approximately 2. Sub-QCIF resolution does further limit the image quality, but it does provide a better subjective impression at extremely low bit-rates, e.g., below 10 kbps. In addition, the reduced number of samples is attractive for inexpensive low-end terminals.

CIF	352×288	pels
QCIF	176×144	pels
sub-QCIF	128×96	pels

Table 1: Spatial resolution of luminance components

Motion compensation accuracy

While H.261 is limited to motion compensation (MC) with integer-pel accuracy, H.263 provides half-pel accuracy. We will see in section 5, that this extension improves performance significantly. The improvement due to half-pel MC is a well established fact [6] and has already been utilized successfully in ITU-T H.262 (MPEG-2) [7].

Filter in the loop

H.261 utilizes a spatial lowpass filter in the predictor, the "filter in the loop", which can be switched on a macroblock basis. H.263 does not include such a loop filter. This is in fact the only feature from H.261 not adopted by H.263. Since there is a significant performance gain due to loop filtering in H.261, this may seem surprising. The reason, however, is that the bilinear interpolation used in H.263 for half-pel MC introduces spatial lowpass filtering as a side affect. In addition, one of the H.263 options includes overlapped block MC (see section 4), which has an inherent filtering effect as well.

GOB layer

Both standards use a hierarchical syntax decomposing a sequence into pictures, group of blocks (GOB), and macroblocks (MB). In H.263, the size of a GOB has been reduced to a single MB-row (11 MBs for QCIF, 8 MBs for sub-QCIF) compared to three MB-rows (11×3 MBs) in H.261. The second difference in the GOB layer is that H.263 allows to insert header information optionally. This enables the coder, e.g., to insert extra synchronization words, or to reduce the overhead for applications requiring a lower robustness.

Error correction

The ITU-T Recommendation H.261 includes forward error correction using a BCH (511,493) code, while no particular error protection scheme is recommended for H.263. However, ITU-T Study Group XV continues to work on error protection for applications of H.263, e.g. , in mobile environments.

Other

Besides the differences mentioned above H.263 contains several minor improvements compared to H.261, which shall be explained briefly. Though each single improvement only makes a small difference, they all add up and contribute to the overall performance.

- VLC tables
  According to the new syntax, there are several new tables for variable length codes (VLC). Other tables are optimized for better performance. E. g., the events coded in the VLC table for DCT coefficients now consist of triplets (run, level, eob) rather than pairs (run, level).
- Motion vector prediction
  Motion vectors are coded differentially as in H.261, but with a more sophisticated predictor. Not only the preceding MB is used for prediction, but also MBs in the previous MB-row.
- Adaptation of quantization parameter
  While H.261 allows any quantizer for any MB, H.263 restricts the transitions from MB to MB within a GOB to the two next coarser or finer quantizers. This yields a smaller and more effective VLC table.
- Macroblock addressing
  For skipped MBs, H.261 uses a MB address (MBA) indicating how many MBs are skipped. In H.263, for every skipped MB a single bit (COD) is transmitted.

4 H.263 OPTIONS

An H.263-coder may use optional coding techniques ("options") to further improve its performance. Options have to be negotiated with the decoder via external means (for example according to ITU-T H.246). There are four options available in H.263, which will be briefly explained in the following.

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming inter-face) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.

WHAT WILL YOU BUILD? | **sales@docketalarm.com** | **1-866-77-FASTCASE**

fastcase®
Smarter legal research.