PERFORMANCE COMPARISON OF THE EMERGING H.264 VIDEO CODING STANDARD WITH THE EXISTING STANDARDS

Nejat Kamaci, Yucel Altunbasak *

Center for Signal and Image Processing
Georgia Institute of Technology, Atlanta, GA, USA
kamaci,yucel@ece.gatech.edu

ABSTRACT

The emerging ITU-T H.264 video coding standard has been developed to achieve significant improvements over the existing standards in the compression performance. Although the basic coding framework of the standard is similar to that of the existing standards, H.264 introduces many new features. The standard is close to completion, and preliminary tests were reported to give satisfactory results, however the developing standard requires further tests for conformance. In this paper, we provide an independent analysis of the coding performance improvement of H.264 over various other existing standards. Our experiments demonstrated that the future H.264 standard can achieve 50% coding gain over MPEG-2, 47% coding gain over H.263 baseline, and 24% coding gain over H.263 high profile encoders.

1. INTRODUCTION

Development of the international video coding standards such as MPEG-1, MPEG-2 and H.261 boosted a diverse range of multimedia applications, including digital video recording, and teleconferencing [1, 2, 3]. As a result of ongoing demand for better compression performance, advanced standards such as MPEG-4 and H.263 were also introduced [4, 5]. H.264 is developed by the ITU-T - ISO/IEC Joint Video Team (JVT) and is projected to become a standard in the year 2003. Although the technical content of the standard is mature, there is still ongoing development, and this paper refers to the reference software joint model 4.2 (JM-4.2) [6].

The H.264 standard is designed in two distinct layers: a video coding layer (VCL), and a network adaptation layer (NAL). Our work is concentrated on the performance of the VCL. The overall scheme of the VCL is superficially similar to the encoding scheme of the previous video coding standards. As in H.263 and MPEG-2, the VCL of H.264 uses translational block-based motion compensation and transform based residual coding. It also features scalar quantization with adjustable quantization step size for output bit-rate control, zigzag scanning and run-length VLC coding of the quantized transform coefficients. However, there are significant differences in the details. H.264 uses a more flexible and efficient model for motion compensation. Use of multiple reference pictures and seven different block sizes are supported for performing motion compensation. Motion vectors can be specified with $\frac{1}{8}$-pixel accuracy. The standard also specifies use of an improved deblocking filter within the motion compensation loop in order to reduce visual artifacts and improve prediction.

In addition to advances in traditional coding functions H.264 also features new coding functions. It features an integer transform instead of the DCT. This transform reduces blocking and ringing artifacts, and also eliminates encoder-decoder mismatches in the reverse transform which was present in the earlier floating-point DCT. It supports spatial prediction within the frames that helps reduce the residual energy of motion compensation. The standard also features a more complex and efficient context-based arithmetic coding (CABAC) for entropy coding of the quantized transform coefficients.

With the newly-introduced features and advancements to the pre-existing features, the emerging H.264 standard was expected to achieve significant improvements in coding performance to all existing standards, in a wide variety of applications. Although initial performance tests by the development committee reported significant coding gains, the standard requires further testing for conformance. In this study we present our performance test results with the future H.264 encoder in terms of coding gains obtained against some of the existing video encoders. We compare H.264 standard with MPEG-2, H.263 Baseline and H.263 Conversational High Compression (CHC) Profile standards.

---

* This work was supported in part by the National Science Foundation under CAREER Award CCR-0133221. The authors are all with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA.

2. TEST ENVIRONMENT

In this work, we used three sets of video sequences for comparison. These sets of sequences represent a range of typical video content for low and high latency applications. We used 200 frames of each video sequence for encoding. The first set consists of three sequences: CAROUSEL, FLOWER GARDEN, and FOOTBALL. These sequences are in CCIR-601 format (720×480) at 30 fps frame rate. The second set consists of five sequences: IRENE, MOBILE AND CALENDAR, PARIS, STUDENTS, TEMPETE. These sequences are in CIF format (352×288), and at 30 fps frame rate. The third set consists of AKIYO, CAR PHONE, CLAIRE, COAST GUARD, CONTAINER, FOREMAN, IRENE, MISS AMERICA, MOBILE AND CALENDAR, NEWS, PARIS, and TEMPETE sequences. In this set, each sequence is in QCIF format (176×144 pixels) and at 15 fps frame rate.

The coding performances are compared based on output bit rate and PSNR of the encoded video sequences. Only the luminance component is used for PSNR computation.

Sequence	Coding Gain
CAROUSEL	41%
FLOWER GARDEN	55%
FOOTBALL	52%
Average Gain	49.3%

Table 1. Coding gain of H.264 over MPEG-2 for various CCIR sequences.

Sequence	Coding Gain
IRENE	53%
MOBILE & CALENDAR	58%
PARIS	49%
STUDENTS	45%
TEMPETE	41%
Average Gain	49.2%

Table 2. Coding gain of H.264 over MPEG-2 for various CIF sequences.

3. COMPARISON OF H.264 VS MPEG-2

MPEG-2 [2] is the most common standard used for high quality video storage and transmission. We compare the coding performance of MPEG-2 and the emerging H.264 standard on first and second set of video sequences.

We used an MPEG-2 encoder based on the codec developed in University of California at Berkeley [8], enhanced with rate-distortion optimized macroblock mode decision and motion estimation and an advanced rate control module to generate the MPEG-2 results. We used the public JM-4.2 test model encoder [6] to generate the H.264 results. In the tests, we used same GOP structure, with two B-type pictures between each P-P or I-P pictures. The GOP size was selected as 15. The H.264 encoder was configured to have five frames for inter motion search, $\frac{1}{4}$-pel motion vector resolution, context-based adaptive binary coding (CABAC) for symbol coding, and rate-distortion optimized mode decision. Both encoders used full search motion estimation with same search range that is 16×16 for both P- and B-type pictures.

In H.264 simulations, constant quantization parameter values of 26, 28, 31, 36, and 41 were used to cover a practical range of low to moderately high bit rates. For MPEG-2, the bit rates were chosen such as the encoded video visual qualities (PSNR) are close to the corresponding H.264 video streams. For each sequence, rate-quality curves are produced. Fig. 1 and 2 show the rate-quality curves for CAROUSEL and FLOWER GARDEN sequences, respectively. For each sequence, the coding gain of H.264 over MPEG-2 is calculated by averaging the bit savings for each simulation point on the rate-quality curves. The coding gains of H.264 over MPEG-2 are summarized in table 1 and 2 for each of the sequences in set-1 (CCIR-601 format) and set-2 (CIF format), respectively. The H.264 encoder achieves an average of 49.3% coding gain over MPEG-2.

4. COMPARISON OF H.264 AND H.263 BASELINE ENCODERS

H.263 [5] is commonly used for low-delay and low to medium bit-rate applications such as video conferencing. We do not use B-type pictures in both the H.263 and the H.264 encoders to satisfy the low delay requirements. We used the public H.263 codec by Telenor [9] to produce the baseline H.263 encoding results. For the H.263 baseline encoder, advanced prediction and syntax-based arithmetic coding options were turned on. For both encoders, the first picture was encoded as I-type, and the remaining frames are encoded as P-type. The H.264 encoder was configured to have five frames for inter motion search, $\frac{1}{4}$-pel motion vector resolution, context-based adaptive binary coding (CABAC) for symbol coding, and rate-distortion optimized mode decision.

The sequences chosen for this comparison are from set two and three. The average sequence bit rates and PSNR values of encoding for each sequence are collected using constant quantization parameter values of 21, 26, 31, and 41 for H.264 and 8, 15, 20, and 33 for H.263. These parameters correspond to a range of low to moderately high bit rates for low latency applications.

Table 3 and 4 show the coding gains obtained by the H.264 encoder over the H.263 encoder for CIF and QCIF se-

Sequence	Coding Gain
IRENE	53%
MOBILE & CALENDAR	58%
PARIS	42%
STUDENTS	45%
TEMPETE	41%
Average Gain	49.2%

Table 3. Coding gain of H.264 over H.263 Baseline Profile for various CIF sequences.

Sequence	Coding Gain
AKIYO	51%
CAR PHONE	40%
CLAIRE	48%
COAST GUARD	38%
CONTAINER	41%
FOREMAN	42%
IRENE	50%
MISS AMERICA	44%
MOBILE & CALENDAR	48%
NEWS	56%
PARIS	52%
TEMPETE	37%
AVERAGE GAIN	45.6%

Table 4. Coding gain of H.264 over H.263 Baseline Profile for various QCIF sequences.

quences, respectively. H.264 achieves an average of 49.2% gain for the selected five CIF sequences at 30 fps, and 45.6% gain for the selected 12 QCIF sequences at 15 fps. As an example, rate-quality curves for CIF format PARIS ad QCIF format FOREMAN sequences are show in Fig. 3 and 4, respectively.

5. COMPARISON OF H.264 VS H.263 CONVERSATIONAL HIGH COMPRESSION PROFILE

Finally, we compare H.264 coding performance with that of the H.263 encoder with Conversational High Compression (CHC) Profile. We used the public H.263 encoder developed at The University of British Columbia [7] to generate the results. For the H.263 CHC encoder, advanced prediction, syntax-based arithmetic coding, advanced intra coding, and unrestricted motion vector modes were turned on. The encoder is also configured to use deblocking filter, and five reference frames for inter coding. For both of the encoders, a GOP size of 15 with one B-type picture between each P-P or I-P pictures was used. The H.264 encoder was configured to have five frames for inter motion search, $\frac{1}{4}$-pel motion vector resolution, context-based adaptive binary coding (CABAC) for symbol coding, and rate-distortion optimized mode decision.

Sequence	Coding Gain
IRENE	29%
MOBILE & CALENDAR	31%
PARIS	22%
STUDENTS	24%
TEMPETE	20%
Average Gain	25.4%

Table 5. Coding gain of H.264 over H.263 Conversational High Compression Profile for various CIF sequences.

Sequence	Coding Gain
AKIYO	25%
CAR PHONE	20%
CLAIRE	34%
COAST GUARD	19%
CONTAINER	22%
FOREMAN	20%
IRENE	26%
MISS AMERICA	25%
MOBILE & CALENDAR	22%
NEWS	28%
PARIS	22%
TEMPETE	17%
AVERAGE GAIN	23.3%

Table 6. Coding gain of H.264 over H.263 Conversational High Compression Profile for various QCIF sequences.

Sequence set two and three were used for comparison. The average sequence bit rates and PSNR values of encoding for each sequence are collected using constant quantization parameter values of 21, 26, 31, and 41 for H.264 and 8, 15, 20, and 33 for H.263 CHC.

Table 5 and 6 show the average coding gains obtained by the H.264 encoder over the H.263 encoder for set two and set three sequences, respectively. H.264 achieves an average of 25.4% gain for the selected five CIF sequences at 30 fps, and 23.3% gain for the selected 12 QCIF sequences at 15 fps.

6. CONCLUSIONS

In this paper, we presented an evaluation of the emerging H.264 video coding standard in terms of coding bit rate savings compared to existing most common video coding standards. We performed encoding tests at a wide range of rates for both low- and high-latency application. According to our test results,the future H.264 standard achieves 50% average coding gain over MPEG-2, 47% average coding gain over H.263 baseline, and 24% average coding gain over H.263 high profile encoders.

7. REFERENCES

[1] D. LeGall, "MPEG: A video compression standard for multimedia application," *Commun., ACM*, vol. 34, pp. 46-58, Apr. 1991

[2] "The MPEG-2 international standard," ISO/IEC, Reference number ISO/IEC 13818-2, 1996.

[3] "Video coding for audiovisual services at $p\times 64$ kbits," ITU-T, ITU-T Recommendation H.261, Mar. 1993.

[4] T. Sikora, "The MPEG-4 video standard verification model," *IEEE Transactions on Circuit and Systems for Video Technology*, vol. 7, pp. 19-31, Feb. 1997.

[5] "Video coding for low bit rate communications," ITU-T, ITU-T Recommendation H.263, ver. 1, 1995.

[6] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Joint Model Reference Software Version 4.2.

[7] "ITU-T H.263 Encoder, version 2", Signal Processing and Multimedia Group, University of British Columbia, Canada.

[8] MPEG-2 video codec, Computer Science Division-EECS, Univ. of Calif. at Berkeley.

[9] Telenor H.263 codec, "ITU-T/SG-15, video codec test model, TMN5", Telenor Research, 1995.

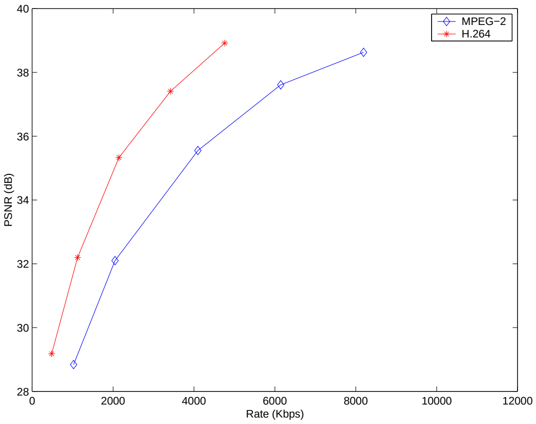

Fig. 1. Rate-PSNR curves of CAROUSEL sequence encoded using H.264 and MPEG-2 encoders.

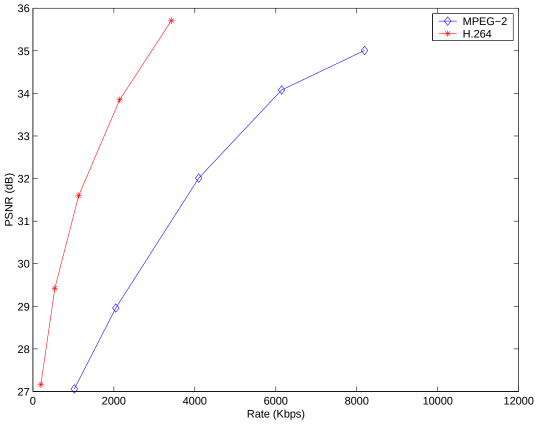

Fig. 2. Rate-PSNR curves of FLOWER GARDEN sequence encoded using H.264 and MPEG-2 encoders.

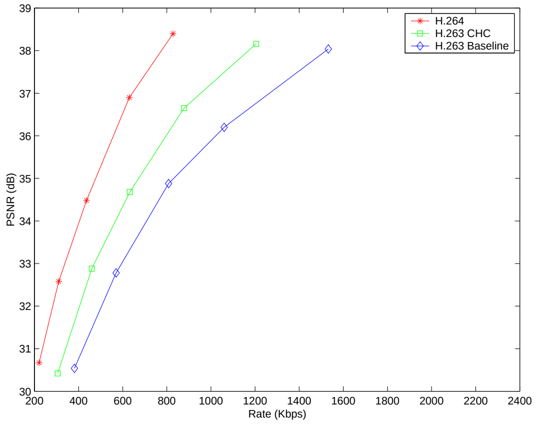

Fig. 3. Rate-PSNR curves of CIF PARIS sequence encoded using H.264, H.263 Baseline and H.263 CHC encoders.

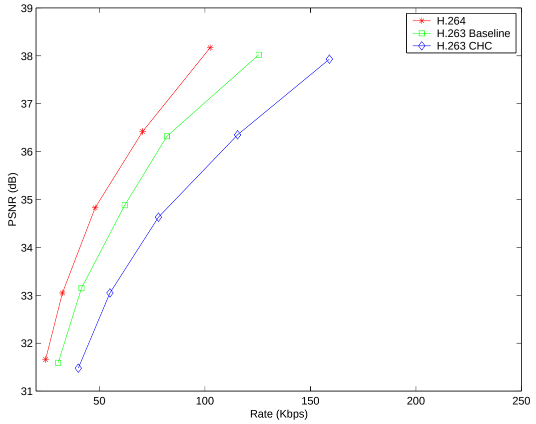

Fig. 4. Rate-PSNR curves of QCIF FOREMAN sequence encoded using H.264, H.263 Baseline and H.263 CHC encoders.