ITU-T H.263 AND H.263 + STANDARDS AS TOOLS FOR CORRELATION BETWEEN OBJECTIVE MEASURE AND SUBJECTIVE SCORE IN VIDEO COMPRESSION

Andreja Samčović[1], Ionel Stanču[2], Zoran Bojković[1]

[1] University of Belgrade, Yugoslavia

[2] «Politehnica» University of Timisoara, Romania

Abstract: Some performances of the ITU-T H.263 and H.263 + standards are compared in the mean of the objective measure for picture quality PSNR, and subjective score in video compression. It is shown that the using of prediction achieves PSNR gain of 1 dB, with little loss in picture quality. The picture formats are considered and QCIF is proposed, taking into account the quality of pictures. The block predictor structures of H.263 and H.263 + are analyzed and compared. H.263 + brings features and capabilities to H.263.

1. INTRODUCTION

Several separate international standards for digital video compression have been set for different applications: storage and retrieval, digital TV and HDTV, videoconference and videophony. The tendency towards integration of video services and the increasing importance of multimedia services ask for new functionality of the coding algorithms in future standards.

This paper describes the standardization of very low bitrate video coding H.263 algorithms, by the International Telecommunications Union (ITU) [1]. After the introduction, the configuration and algorithms of H.263 are given in the second part of the paper. The third part explain the design H.263 coder and invokes the picture formats. Finally, the performances of H.263 options are analyzed and compared.

2. H.263: CONFIGURATION AND ALGORITHMS

Recommendation H.263 is the first one for very low bitrate video compression, allowing transmission of audio-visual information on narrow channels down to 9,6 kb/s. The compression scheme used in H.263 is similar to that of the Recommendation H.261, to which few enhancements have been added. The major enhancements are: a new type of coded frame called PB (Predicted Bidirectional), in which both predicted and bidirectional macroblocks can reside; an advanced motion estimation with a motion field resolution down to 8 x 8 pixels; half-sample motion estimation, as well as overlapping motion compensation.

H.263 is followed by ITU-T H.263 + standard.

An uncompressed video sequence for very low bitrate applications typically requires a bitrate of up to 10 Mb/s. To achieve very low data rates, compression ratios on the order 1000:1 are required to meet the needs of the consumer market, which has access only to very narrow communication channels for the few coming years. Intensive research has been carried out in the last decade to attain this objective. However, severe blocking artifacts may still occur at very low data rates.

The coding structure of H.263 is based on hybrid motion-compensated discrete cosine transform (DCT), variable length coding (VLC), scalar quantization algorithm and macroblock structure. Block-based coding algorithm has been used in the H.263 standard, where the coder has been optimized basically in motion prediction and frame formats.

3. PICTURE FORMATS FOR H.263 CODER

Five picture formats were adapted for H.263: CIF (Common Intermediate Format, 352 x 288 pixels), QCIF (Quaternal CIF, 176 x 144), sub-QCIF (128 x 96), 4 CIF (704 x 576), as well as 16 CIF (1408 x 1152).

QCIF is preferable, because of the need for a better picture quality and interworking with ITU-T H.320 terminals. However, there is also a need for lower cost, smaller picture format, picture quality and feasibility for future trends.

H.263 has the optional coding methods: UMV - Unrestricted Motion Vector mode, SAC - Syntax-based Arithmetic Coding mode, AP - Advanced Prediction mode, and PB - frames mode,

In UMV mode, motion vectors are allowed to point outside the picture. The referenced pixels outside the picture area are replaced by the nearest edge pixel. Therefore, it is possible good prediction for all pixels except those coming in as new information. This psrt of the option is useful when the camera is moving.

In SAC mode, arithmetic coding is used instead of VLC coding.

The AP mode includes 8 x 8 block motion vectors. In this option, four 8 x 8 vectors are used instead of one 16 x 16 vector for some of the macroblocks in the picture, and motion vectors are allowed to point outside the picture. The encoder has to decide which type of vectors should be used. However, four vectors instead one use more bits, but can obtain better resolution for prediction.

In PB mode, the frames are grouping of a P- and B-frame. There are some differences between B- and PB-frames. Namely, in PB-frame mode there can be only one B-frame among P-frames. Secondly, no extra data is used for the B-frame in PB-block. The vectors for the B-frame are obtained from scaling down vectors for the relevant P-block. Finally, P- and B-blocks are coded as one unit.

4. PERFORMANCE OF H.263 OPTIONS

Performance of the H.263 AP and SAC options are compared in the peak signal-to-noise ratio (PSNR) of the luminance component for the «Foreman» test-sequence in QCIF picture format (frames 0-200). The frame rate is 12,5 frames/s. The results are shown in Fig.1 [2].

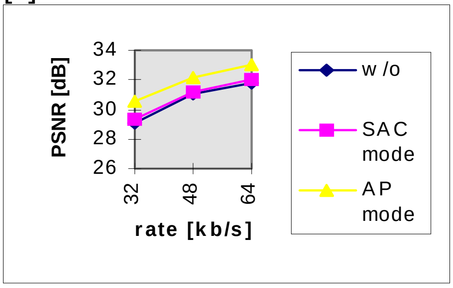

Fig.1 Performance of the H.263 AP and SAC options

The H.263 options are compared to the baseline H.263 coder (w/o options), providing the performance gain due to single options.

The performance gain due to the AP mode is illustrated. It results 1,2 dB at 64 kb/s. The improvement due to the SAC mode is relatively small, 0,2 dB at 64 kb/s. The reason is entropy coding, when the PSNR gain is not significant, but fewer bits are produced, on the other side [3].

The main benefit of the PB mode is increasing frame rate with little additional bit usage. The frame rate is typically

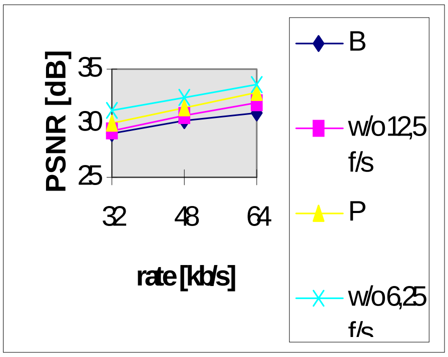

Fig.2 Performance of the H.263 PB-frames mode

doubled when the PB mode is used. Fig.2 shows the sequence coded at 6,25 frames/s. The PSNR of the upper curve is good (33,5 dB at 64 kb/s), but the low temporal resolution results in jerky motion. Doubling the frame rate to 12,5 frames/s results in significant loss of PSNR (1,7 dB at the same rate). A better trade-off between temporal and spatial resolution is possible with the PB mode. The PSNR is calculated separately for P- and B-frames, and is shown on dashed curves at the rate of 12,5 frames/s in Fig.2. One can see that the same number of P-frames/s is transmitted with little loss of PSNR (0,6 dB). However, with the use of the B-frames, the frame rate is doubled. The B-frames provide the subjective impression of smooth motion.

The motion vectors for the B-frame would be interpolated from the P-frame motion vectors and a small offset to the interpolated vectors could also be transmitted in the bitstream. Although in most cases, a PB pair of frames would be more efficient than a pair of P-frames, where this arrangement would perform worse. The PB-frames mode is improved (IPB-mode) in the ITU-T standard H.263 +. An improvement to the original mode allows a complete motion vector to be transmitted for the B-frames, thus a PB-frame pair can do no worse than a pair of P-frames.

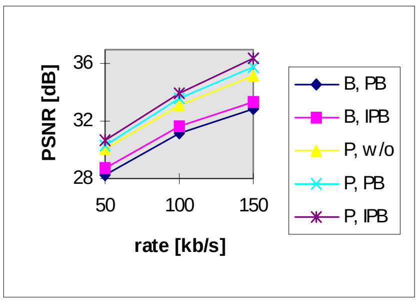

Fig.3 IPB-frames mode

Fig.3 illustrates the performance of the IPB-mode, for the luminance component of the video sequence «Foreman» at 10 frames/s. One can see a significant improvement in PSNR comparing to the H.263 PB-frames mode. Also, the Figure shows the PSNR gains achieved with the H.263 PB mode and H.263 + IPB mode, compared to the H.263 baseline coder (w/o options). Also, the PSNR of B-frames is increased in the IPB mode by approximately 0,5 dB for a given bitrate, to the PB mode.

However, the PB-frames and improved PB-frames modes increase the encoder's complexity and computational requirements. The complexity of the IPB-frames mode is slightly larger than that of the PB-frames, because it involvs backward and forward prediction. Namely, forward prediction is usually very computationally intensive, and requires more memory.

The prediction process in PB-frames mode is illustrated in Fig.4.

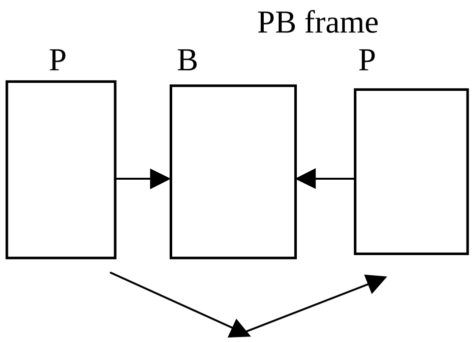

Fig.4 Prediction in PB-frames mode

The prediction of a B-macroblock in a PB-frame, for luminance and chrominance blocks, is performed on a block basis (8 x 8 pixels). The P-macroblock has to be decoded and reconstructed. This block is called PR. From the motion vector for each P-block, a backward and forward vector are derived. Pixels where the backward points inside PR, use bidirectional prediction. This is obtained as the average of the forward prediction using the forward vector relative to the previously decoded P-frame, and the backward prediction using the backward vector relative to the just decoded P-block. For all other pixels, forward prediction using the forward vector relative to the previously decoded P-frame is taken into account.

One of the modifications of the H.263 + to the original H.263 standard concerns motion vector range. In H.263, the motion vector range was for the most part [-16,-15,5]. When H.263 + is invoked, the range is generally larger and depends on the frame size, as shown in Table 1 [4].

Table 1. Motion vector ranges in H.263 +

FRAME SIZES UP TO ...	MOTION VECTOR RANGE
352 X 288	[-32,31,5]
704 X 576	[-64,63,5]
1408 X 1152	[-128,127,5]
WIDTHS UP TO 2048	HORIZONTAL RANGE [-256,255,5]

5. CONCLUSION

This paper has compared the performances of H.263 and H.263 + in the PSNR, as objective measure of picture quality, and the subjective score of perceived pictures. It is shown that the prediction, which includes motion block compensation, yields gain of 1 dB. The PB-frames mode allows to almost double the frame rate with only little loss of picture quality. Very small gain can be realized with arithmetic entropy coding. Also, the picture formats are considered and among them QCIF is preferable, taking into account the picture quality. The block predictor structures of the H.263 and H.263 + standards are analyzed and compared. H.263 + brings new features and capabilities to H.263.

REFERENCES

[1] K.Rijkse: «ITU standardization of very low bitrate video coding algorithms», Signal Processing: Image Communication, Vol.7, 1995, pp 553-565

[2] B.Girod, N.Faerber, E.Steinbad: «Standards based video communications at low bitrates», EUSIPCO’96, Trieste, Italy, September 1996.

[3] T.Gardos: «H.263 +: the new ITU-T recommendation for video coding at low bitrates», ICASSP’98, Seattle, WA, USA, 1998.

[4] G.Cote, B.Erol, M.Gallant, F.Kossentini: «H.263 +: video coding at low bitrates», IEEE Transactions on Circuits and Systems for Video Technology CSVT-8, No.7, November 1998, pp 849-865