IEICE TRANS. FUNDAMENTALS, VOL.E85–A, NO.8 AUGUST 2002

LETTER *Special Section on Digital Signal Processing*

QCIF Video Coding Based on JPEG2000 Using Symmetry of Images

Ayuko TAKAGI[†a)], ***Student Member and*** **Hitoshi KIYA**[†], ***Regular Member***

SUMMARY This paper describes an effective technique for coding QCIF video sequences based on a JPEG2000 codec. In the proposed method, multiple frames are combined into one large picture. The larger picture enables images to be coded more efficiently. Image quality is further improved by combining the frames symmetrically. The video sequence is efficiently coded by adapting the time correlation of the video sequences to spatial correlation. We demonstrated the effectiveness of this method by encoding QCIF video sequences using JPEG2000.

key words: *JPEG2000, MotionJPEG2000, symmetric extension, QCIF, video encoding*

1. Introduction

In most multimedia applications, video data is compressed by either of the following two methods: inter-frame coding and intra-frame coding. It is known that the neighboring frames have a high correlation for video sequences. Inter-frame coding, which is used in the MPEG codec, uses motion estimation to eliminate the temporal redundancy of the video frames, thus attaining a high rate of compression. However, motion estimation is very time consuming [1]. In contrast, intra-frame coding is a simpler coding process, and each video frame is encoded independently. Since the temporal correlation of video sequences is not used, intra-frame coding has a lower coding efficiency compared to inter-frame coding.

Codecs using intra-frame coding include JPEG, which are widely used. Recently, JPEG2000 has been standardized and the expectation for this new codec is growing. JPEG2000, which is based on wavelet transform, enables more efficient coding than JPEG, which is based on DCT (discrete cosine transform). However, compared with MPEG which uses inter-frame coding, the quality of images is inferior at the same rate of compression. Moreover, the coding efficiency for small pictures, such as QCIF (Quarter Common Intermediate Format) video sequences, is low for JPEG2000 codec.

In this paper, we propose an effective technique for efficiently coding small video sequences using JPEG2000. In our proposed method, we generate a larger picture by combining multiple frames. Moreover, the frames are combined symmetrically and the coding efficiency increases. This proposed method is performed prior to the encoding, and does not modify the JPEG2000 standard.

Manuscript received December 10, 2001.
Manuscript revised March 1, 2002.
Final manuscript received April 19, 2002.

[†]The authors are with the Department of Electrical Engineering, Tokyo Metropolitan University, Hachioji-shi, 192-0397 Japan.
a) E-mail: ayuko@isys.eei.metro-u.ac.jp

2. Relationship between Image Size and Compression Rate

In JPEG2000, when the picture size is reduced, compression efficiency decreases. This occurs due to the following reasons:

(a) Decrease in frequency resolution
In wavelet transform, which JPEG2000 is based on, a large picture can be decomposed into several frequency components; that is, frequency resolution is high. Conversely, frequency resolution is low for small pictures. Therefore, the coding efficiency of JPEG2000 decreases.

(b) Small number of lowest sub-band coefficients
In JPEG2000, the efficiency of arithmetic coding of the lowest sub-band component of a wavelet coefficient is statistically good compared with that of the components of other sub-bands. Therefore, with a small picture, the size of the lowest sub-band region is also small and coding efficiency decreases relatively.

(c) Increase in header information in a bitstream
The header information used to generate a JPEG2000 bitstream does not depend on picture size. When the picture is small, the amount of header information contained in a bitstream is relatively large. Therefore, the number of bits assigned to coding the image data decreases relatively, reducing coding efficiency.

To demonstrate the effects of picture size on coding efficiency, we show the results of a simulation using JPEG2000. The picture used was one frame of video sequence "mobile & calendar"; the picture size was 704×576 pixels. The simulation conditions were as follows:

a. The whole picture was encoded at once.
b. The picture was divided into four sections (352×288 pixels) and each section was encoded individually.
c. The picture was divided into 16 sections (176×144 pixels) and each section was encoded individually.

The encoding conditions for JPEG2000 were as follows: the wavelet transform was performed in three stages;

the code block size was 64×64. The encoding bit rate varied from 0.4 to 1.0 [bpp]; Fig. 1 shows the average value of PSNR for each bit rate. At the same bit rate, the highest image quality was equivalent to the maximum coding efficiency. The results showed that there was a slight difference between the image quality obtained with a picture size of 704×576 and that with a picture size of 352×288. However, image quality was lower with a picture size of 176×144. The level of image degradation was about 0.5 [dB]. That is, coding efficiency for data for small pictures decreased. To improve coding efficiency for small pictures, we suggest treating larger pictures.

3. Proposed Method for More Efficient Coding

3.1 Method for Efficient Coding

Here, we describe a method for efficiently coding video sequences of QCIF size using JPEG2000. Given the small picture size, it is not efficient to use JPEG2000 to encode each frame of a QCIF video as it is. However, the problem of decreased coding efficiency, outlined in Sect. 2, can be reduced by using a larger picture. Therefore, we generate a larger picture from several frames and use this as an input. Figure 2 shows the coding procedures for both the conventional method and the proposed method. In the conventional method, Fig. 2(a), each frame is encoded individually by the JPEG2000 encoder. Figure 2(b) shows the proposed method which generates the larger picture using four frames. Here, we generate a larger picture by simply placing four frames side by side. The generated picture is then used as an input for the JPEG2000 encoder. Coding is more efficient since a larger picture is used as an input. After decoding, a reverse procedure is performed and decoded pictures of the original size are obtained by dividing the large decoded picture. The resulting pictures have higher image quality than when each frame is coded individually.

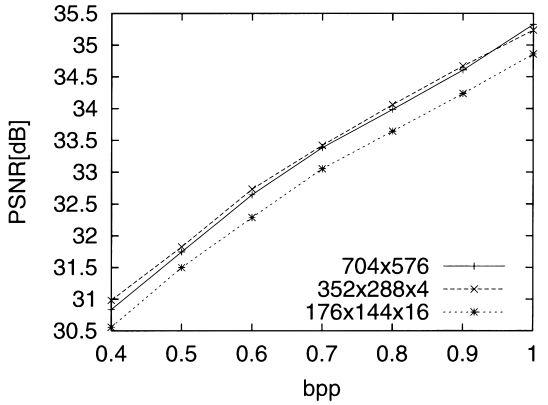
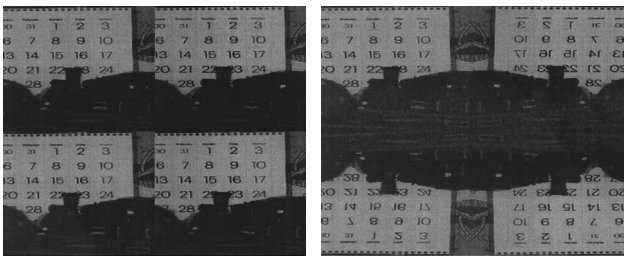


Fig. 1 Influence of picture size on image quality in JPEG2000.

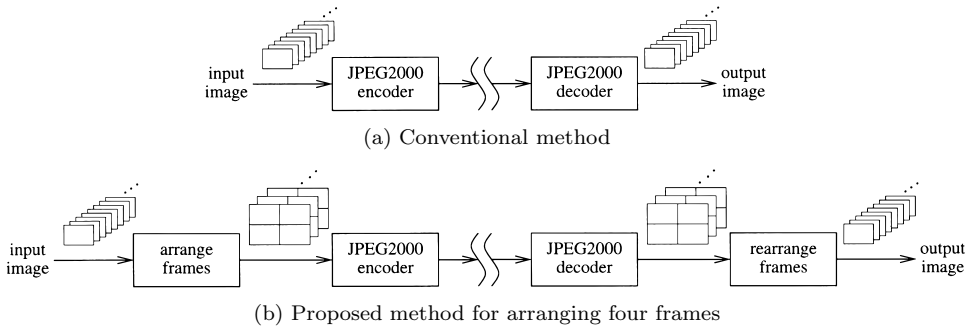
3.2 Proposed Method

Some flexibility is possible in the method used to generate larger pictures from the original frames; the number of frames to be used and the method of arranging them can vary. Image quality is affected by the way these frames are placed in the large picture. Here, we propose a method for arranging the frames in order to obtain higher image quality. Figure 3 shows the proposed method for arranging four frames. In video sequences, time correlation is high and MPEG achieves a high rate of compression by removing the time correlation. Frames placed close to each other have high correlation, and these frames look alike. In a large picture in which adjacent frames are placed side by side, it ap-



(a) Simple extension (b) Symmetric extension (proposed)

Fig. 3 Proposed method for arranging four frames.



(a) Conventional method

(b) Proposed method for arranging four frames

Fig. 2 Coding procedure.

pears as though one frame has simply been extended, as shown in Fig. 3(a). We propose arranging several frames by reversing the four directions of each frame and putting them in order. A large picture generated using this method resembles the result obtained by performing a symmetrical extension of one of the frames. Symmetrical extension is used to smoothen the border of an image during sub-band still image coding, and it is known to improve image quality [4], [5]. The example of arranging four frames symmetrically is shown in Fig. 3(b). When using four frames, the frames at the upper right, lower left, and lower right are reversed horizontally, vertically, and both horizontally and vertically, respectively. In the large picture generated, it appears as though a symmetric extension has been performed on the upper left frame. Figure 3(b) shows that this symmetrical method avoids the discontinuity between frames, which is produced by a simple extension in Fig. 3(a).

When using the proposed method, the arrangement procedure is needed for several frames in advance. This causes a delay in the encoding scheme. Also, when using a larger number of frames, a larger amount of memory is needed for both encoding and decoding. In the decoding procedure, the memory is needed for buffering the decoded large picture, and each frame is rearranged and is used as an output. Although there are disadvantages, the image quality improves on using the proposed method.

4. Simulation

We show the results of a simulation to test the effectiveness of the proposed method. The video sequences used are "carphone," "claire," and "mobile & calendar." They were all grayscale sequences. The picture size was QCIF (176×144) and the number of frames used was 16. For comparison, the video sequences were encoded on both JPEG2000 and MPEG-1 [6] codecs. For JPEG2000, the number of stages used for wavelet transform was three, and the code block size was 64×64. We encoded the video sequences on JPEG2000 under the following conditions:

a. Frames were symmetrically extended to generate a large picture (proposed method).
b. Frames were simply extended to generate a large picture.
c. Each frame was encoded individually.

In conditions (a) and (b), we used four, eight, and sixteen frames and generated the large pictures as in Fig. 4. The simulation conditions for the MPEG-1 codec were as follows:

d. no motion estimation; intra-frame coding only, and
e. using motion estimation for inter-frame coding. I/P frame distance was three.

The frame rate for MPEG-1 is 30 [frames/sec] and the target bit rate is given by bps = bpp × width × height × frames/sec. We encoded 16 frames with bit rates of 0.4, 0.5, 0.7, and 1.0 [bpp]. The header size of JPEG2000 bit stream was 84 bytes for all results. Table 1 gives the results of the simulation, showing the average PSNR of 16 frames at each bit rate. Figure 5 shows the result for carphone with the bit rate of 1.0 [bpp]. Figure 6 shows the PSNR for each frame of a carphone with a target bit rate of 1.0 [bpp], and the result which used only JPEG2000 is shown in Fig. 7.

It is shown that the proposed method, (a), results in the best image quality compared to (b), (c),

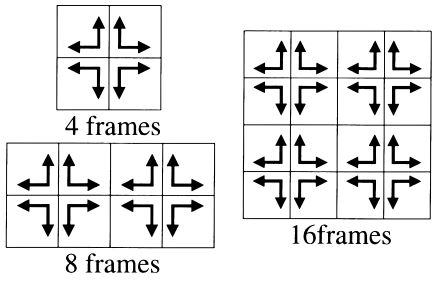


Fig. 4 Example of arranging frames symmetrically.

Table 1 PSNR comparison between different coding conditions [dB]. (a: JPEG2000 with proposed method, b: JPEG2000 with simple extension, c: JPEG2000, d: MPEG without ME, e: MPEG with ME).

	[bpp]	a	b	a	b	a	b	c	d	e
		16 frames		8 frames		4 frames				
carphone	0.4	32.94	32.06	32.81	32.18	32.65	32.07	31.86	28.84	36.16
	0.5	34.63	33.87	34.59	33.97	34.43	33.97	33.67	30.31	37.45
	0.7	37.28	36.39	37.22	36.57	37.06	36.55	36.27	32.70	39.40
	1.0	40.71	39.86	40.65	39.95	40.52	39.94	39.86	35.55	41.50
claire	0.4	39.13	38.02	39.07	38.55	38.87	38.34	37.83	32.79	43.67
	0.5	41.27	40.76	41.25	40.80	41.05	40.54	40.04	34.53	44.60
	0.7	44.58	44.13	44.55	44.02	44.43	43.79	43.49	37.96	45.99
	1.0	48.34	48.05	48.27	48.03	48.22	47.86	47.67	41.61	48.44
mobile & calendar	0.4	31.04	30.72	30.95	30.70	30.81	30.67	30.34	28.11	32.40
	0.5	32.37	31.95	32.29	31.93	32.12	31.88	31.41	28.93	33.47
	0.7	34.69	34.27	34.63	34.28	34.51	34.25	33.92	30.71	34.83
	1.0	37.23	36.96	37.16	36.98	37.09	36.95	36.66	32.90	36.46

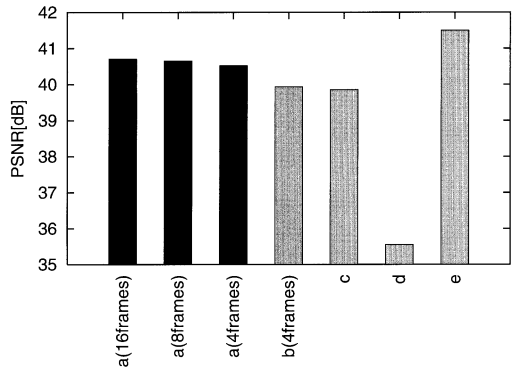


Fig. 5 PSNR result of carphone with target bitrate of 1.0 [bpp]. (a: JPEG2000 using proposed method b: JPEG2000 with simple extension, c: JPEG2000, d: MPEG without ME, e: MPEG with ME).

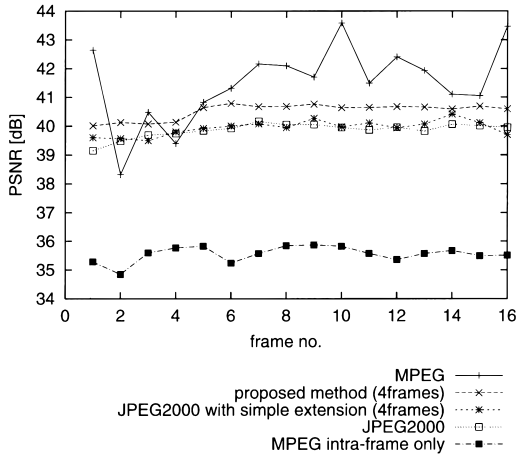


Fig. 6 Dispersion of PSNR result for carphone encoded with target bitrate of 1.0 [bpp].

and (d) which use intra-frame coding. In (d), MPEG without motion estimation was used, which only uses intra-frame coding. A comparison of (d) and the use of JPEG2000, (c), showed that image quality was improved by using JPEG2000. Placing frames side by side by simple extension, (b), resulted in higher image quality than that obtained using (c). Placing frames symmetrically, as proposed, (a), further improved image quality. It is thought that the continuity between frames and the spatial symmetry in the extended picture contributes to improving coding efficiency. By using more frames to generate a large picture leads to an improvement in image quality. In (e), when motion estimation was performed in MPEG, time redundancy in the video sequence was eliminated and it was efficiently coded.

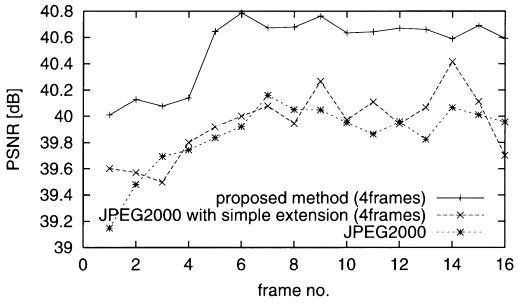


Fig. 7 Dispersion of PSNR result for carphone encoded by JPEG2000 with target bitrate of 1.0 [bpp].

The above result shows that when encoding video sequences with JPEG2000, when the pictures are small, as in QCIF, it is better to encode a larger picture generated by combining several frames that are continuous in time. Furthermore, the use of our proposed method improves image quality by about 1 [dB] compared to when an original picture image is directly coded using JPEG2000.

5. Conclusion

We described a technique for coding QCIF video sequences based on JPEG2000. Multiple frames were combined to form one large picture, and coding efficiency was improved by adapting the time correlation of video sequences to spatial correlation. Image quality was further improved by arranging frames symmetrically. The effectiveness of the method was shown by coding QCIF video sequences using JPEG2000 with the proposed method.

References

[1] B. Furht, J. Greenberg, and R. Westwater, Motion Estimation Algorithms for Video Compression, Kluwer Academic Publishers, 1997.
[2] ISO/IEC 15444-1, "Information technology—JPEG2000 image coding system—Part 1: Core coding system," ISO/IEC JTC 1/SC 29/WG1, Jan. 2001.
[3] ISO/IEC JTC/1/Sc29/WG1 N1987, "Motion JPEG2000 Committee Draft 1.0," Dec. 2000.
[4] H. Kiya, K. Nishikawa, and M. Iwahashi, "A development of symmetric extension method for subband image coding," IEEE Trans. Image Processing, vol.3, no.1, pp.78–81, Jan. 1994.
[5] S.A. Martucci, "Symmetric convolution and the discrete sine and cosine transforms," IEEE Trans. Signal Processing, vol.42, no.5, pp.1038–1051, 1994.
[6] MPEG Software Simulation Group, "MPEG-2 encoder/decoder version 1.2," [Online] Available WWW: http://www.mpeg.org/MPEG/MSSG/