Method of Video Transcoding from MPEG-2 to HEVC with Picture Structure Information

Mitsuo Ikeda, Takeshi Kumaki and Takeshi Ogura

Ritsumeikan University
1-1-1 Noji-Higashi, Kusatsu, Shiga 525-8577, Japan
E-mail: gr0166hp@ed.ritsumei.ac.jp, kumaki@fc.ritsumei.ac.jp, togura@se.ritsumei.ac.jp

Abstract: A method is presented for transcoding from MPEG-2 to HEVC by using picture structure information. One of the main issues in the process of transcoding from MPEG-2 to HEVC is the efficiency of HEVC encoding. For HEVC encoding, a hierarchy picture structure called a random-access mode is usually used for high efficiency coding; however, the random-access mode may not always be the best structure for transcoding from MPEG-2 to HEVC. We propose some picture structures using MPEG-2 stream information. The experimental results show that the information on the picture coding structure in the MPEG-2 video stream is useful for transcoding MPEG-2 to HEVC. The proposed method with hierarchy picture structure using the MPEG-2 structure information reduced bitrate by 7.1% in the transcoding experiments.

Keywords: video coding, MPEG-2, H.265/HEVC, transcoding, video quality, picture coding information

1. Introduction

Video coding technologies have been developed and improved rapidly. MPEG-2[1][2] is one of the most common technologies for video coding, and the technologies have spread worldwide. A lot of video content has been coded by MPEG-2 encoders in many applications such as digital broadcasting, digital video discs, and video transmission. After the MPEG-2 standard was established in 1994, some new video coding technologies have been developed, such as MPEG-4[3], H.264/MPEG-4 AVC[4][5] and H.265/HEVC[6][7]. The coding efficiency of H.264 is double that of MPEG-2. Also, the coding efficiency of HEVC is said to be double that of H.264. Video transcoding technologies are becoming more important to utilize these coding efficiencies for existing MPEG-2 contents. Several methods were studied about the transcoding of MPEG-2 to H.264[8-11]. These studies mainly include methods of mode decision, motion estimation and block coding for transcoding from MPEG-2 to H.264. In the aspect of transcoding to HEVC, some methods of transcoding from MPEG-2 or H.264 to HEVC were proposed[12][13]. One of the main purposes of those studies is to reduce the computational complexity while decreasing the loss of coding efficiency compared to that of HEVC re-coding. On the other hand, the main purpose of this paper is to describe how we improve the efficiency of transcoding beyond that of the conventional HEVC re-coding.

In general, how to determine the picture coding structure is one of the most important factors in improving the coding efficiency. Although the picture coding structures used for HEVC coding are different from those for MPEG-2 coding in usual, the improvement of the picture coding structure for transcoding from MPEG-2 to HEVC has not been discussed. This paper describes our study of the improvement in the picture coding structure for transcoding from MPEG-2 to HEVC, and the method of transcoding by using picture structure information is proposed.

Section 2 presents the transcoding method from MPEG-2 to HEVC and its improvement. The experimental results of the proposed transcoding methods are presented in Section 3. Section 4 concludes the paper.

2. Transcoding Method

2.1 Transcoding from MPEG-2 to HEVC

Figure 1 shows the outline of the video transcoding from MPEG-2 to HEVC. First, the original video is coded by an MPEG-2 encoder. The result is an MPEG-2 video stream. After the MPEG-2 video stream is decoded by an MPEG-2 decoder, the video is recoded by an HEVC encoder. The process of transcoding from MPEG-2 to HEVC is a combination of MPEG-2 decoding and HEVC encoding.
Transcoding from MPEG-2 to HEVC

Figure 2 presents an example of the rate-distortion curves for MPEG-2, HEVC, and transcoding from MPEG-2 to HEVC. In this graph, the X-axis indicates the bitrate of the coded stream. The Y-axis is the peak signal-to-noise ratio (PSNR) of the decoded video compared with the original video. The curves of MPEG-2 and HEVC show that the difference in the PSNR is about 4 dB between MPEG-2 and HEVC at the same bitrate. On the other hand, in order to satisfy the PSNR of 32 dB, the HEVC needs the bitrate of 30 Mbps, and the HEVC needs about 10 Mbps. In Fig. 2, the rate-distortion curve of transcoding from MPEG-2 stream at 30.5 Mbps is also shown. The PSNR is calculated for the HEVC decoded video compared with the original video before MPEG-2 encoding.

In order to increase the coding efficiency of transcoding, we utilized picture coding information of MPEG-2 for HEVC recoding[14]. Figure 3 presents a coding structure of MPEG-2 with M=3, which is used most widely for MPEG-2 encoding. In the figure, “I”, “P” and “B” indicate the picture coding type of “intra coded picture”, “predictive coded picture” and “bi-predictive coded picture”, respectively[1]. The arrows in the figure indicate the relationship of the reference pictures for motion prediction. The value “M” is the interval of P pictures. For example, there are two successive B-pictures between P-pictures which are referred by the B-pictures where the value “M” equals to three, as shown in Fig. 3.

Figure 4 shows a coding structure of the HEVC random-access mode, which is used most commonly for HEVC encoding. Although the picture structure of M=3 like the one in Fig. 3 is permitted in the HEVC standard, the structure of the random-access mode usually produces better coding efficiency than that of the structure with M=3 for encoding. In the case of transcoding, however, the random-access mode may not always be the best. One of the issues is the reference of pictures. Pictures which are often referred in MPEG-2 coding are not always referred in HEVC coding. For example, Picture “3” in Fig. 3 is referred by Picture “1”, “2”, “4”, and “5” in MPEG-2 coding. However, the Picture “3” in Fig. 4 is not referred by any other pictures in HEVC recoding. On the other hand, pictures which are not referred by any other pictures in MPEG-2 coding may be often referred in HEVC coding, such as Picture “4” and “8” in Fig. 3 and Fig. 4. These things may act on transcoding disadvantageously for the following reasons. In general, the quality of a coding picture is low when the quality of the picture referred by the coding picture is low. Therefore, making the quality of pictures better is desirable if the picture is referred by other pictures. If the relationship of the reference picture in HEVC coding is different from that in MPEG-2 coding, the quality of the pictures referred by other pictures in HEVC coding may not always be coded with good quality in MPEG-2 coding.
2.2 Improvement in picture structure for transcoding

The information of picture coding structure in the MPEG-2 video stream content is utilized to increase the transcoding efficiency. Figure 5 presents a picture coding structure of HEVC for transcoding from a MPEG-2 stream with M=3. It is intended to be similar to the picture coding structure of the MPEG-2 stream to be transcoded. The information of the picture coding structure for MPEG-2 can be obtained from the MPEG-2 video stream.

In addition, we propose the picture coding structure shown in Fig. 6. In the structure, two successive B-pictures such as Picture “1” and “2” are placed between the pictures referred by the B-pictures. The structure of pictures besides the successive B-pictures is same as the structure of random access mode. Consequently, this structure has not only the feature of hierarchy coding like random-access mode for HEVC coding, but also the picture structures with M=3 inherited from MPEG-2 coding. The number of the successive B-pictures is correspond to the value of M-1. The concept of the method is applicable to another value of M.

3. Experimental Results

3.1 HEVC encoding

The coding efficiency of each method for HEVC encoding was evaluated before experiments of the transcoding from MPEG-2 to HEVC. The encoder in the experiments is based on the HEVC reference software HM-16.6. Main conditions of encoding are shown in Table 1. Table 2 shows the list of test sequences. These conditions are commonly used in the following experiments of MPEG-2 encoding and HEVC re-coding.

| Table 1 Coding conditions |
|---------------------------|
| Profile  | Main |
| Bit depth | 8 bits |
| Chroma format | 4:2:0 |

| Table 2 Test sequences |
|------------------------|
| Video sequence | “Crowd Run” | “Into Tree” | “Park Joy” |
| Image size | 1920 x 1080 | 1920 x 1080 | 1920 x 1080 |
| Frame rate | 50 fps | 50 fps | 50 fps |
| Frames | 97 frames | 97 frames | 97 frames |

Thumbnails of the test video sequences used in the experiments are shown in Figs. 7-9.
Figures 10–12 present the rate-distortion curves of the encoding for the test sequences. The rate-distortion curves in these figures show that the random-access structure had better coding-efficiency than that of the structure with M=3.

Table 3 presents the BD-bitrate [15] of each method in encoding. In Table 3, the value “N” indicates the interval of intra coded pictures. The “anchor” of the BD-bitrate is the random-access method. Table 3 shows that the random-access method had the best coding-efficiency in most cases of encoding. Only the case of the hierarchy structure with M=3 and N=36 had a better result for the sequence of “Into Tree” than that of the random-access structure with N=32. In general, the higher frequency of intra-picture reduces the efficiency of coding for the following reasons. Because only the redundancy in the picture can be used for coding in intra coded pictures, the coding efficiency in these pictures is less than that of inter coded pictures that can also use the redundancy of inter pictures. Thus, if the value “N” is large, the ratio of intra pictures in the sequence of pictures decreases and the coding efficiency usually improves.
### Table 3 BD-bitrate of encoding

| Sequence     | Method             | N   | BD-bitrate   |
|--------------|--------------------|-----|--------------|
| “Crowd Run” | Random Access      | 32  | -            |
|              | M=3 Structure (A)  | 30  | +9.0%        |
|              | M=3 Hierarchy      | 24  | +5.8%        |
|              | Structure (B)      | 36  | +3.6%        |
| “Into Tree” | Random Access      | 32  | -            |
|              | M=3 Structure (A)  | 30  | +8.0%        |
|              | M=3 Hierarchy      | 24  | +1.7%        |
|              | Structure (B)      | 36  | -4.4%        |
| “Park Joy”  | Random Access      | 32  | -            |
|              | M=3 Structure (A)  | 30  | +12.8%       |
|              | M=3 Hierarchy      | 24  | +6.5%        |
|              | Structure (B)      | 36  | +3.6%        |
| Average      | Random Access      | 32  | -            |
|              | M=3 Structure (A)  | 30  | +9.9%        |
|              | M=3 Hierarchy      | 24  | +4.7%        |
|              | Structure (B)      | 36  | +0.9%        |

### 3.2 Transcoding from MPEG-2 to HEVC

MPEG-2 streams were prepared for the test sequences to evaluate the transcoding from MPEG-2 to HEVC. Table 4 shows the coding conditions of MPEG-2 for the evaluation.

### Table 4 Coding parameters of MPEG-2

| Video sequence | “Crowd Run” | “Into Tree” | “Park Joy” |
|----------------|-------------|-------------|------------|
| M              | 3           | 3           | 3          |
| N              | 30          | 30          | 30         |
| Quantization   | I, P        | 22          | 10         | 20         |
| parameter      | B           | 32          | 14         | 28         |

Figures 13-15 present the rate-distortion curves of the transcoding for the test sequences. In these graphs, the X-axis indicates the bitrate. The Y-axis is the PSNR of the transcoded video that was compared with the original video before MPEG-2 encoding. In these figures, the value “N” for the curve with M=3 hierarchy structure equals to 24. The rate-distortion curves in these figures show that the proposed method with M=3 hierarchy structure had better coding efficiency than the method with the random-access structure, although the value “N” is smaller.
Table 5 presents the BD-bitrate of each method in transcoding. The “anchor” of the BD-bitrate is the random-access method. Table 3 shows that the transcoding by the proposed method had better coding efficiency than the transcoding with the random-access mode in the measurement of the BD-bitrate for the test sequences.

Table 5 BD-bitrate of transcoding

| Sequence      | Method            | N   | BD-bitrate |
|---------------|-------------------|-----|------------|
| “Crowd Run”   | Random Access     | 32  | -          |
|               | M=3 Structure (A) | 30  | -2.7%      |
|               | M=3 Hierarchy     | 24  | -5.5%      |
|               | Structure (B)     | 36  | -7.6%      |
| “Into Tree”   | Random Access     | 32  | -          |
|               | M=3 Structure (A) | 30  | +1.1%      |
|               | M=3 Hierarchy     | 24  | -9.2%      |
|               | Structure (B)     | 36  | -13.4%     |
| “Park Joy”    | Random Access     | 32  | -          |
|               | M=3 Structure (A) | 30  | -4.7%      |
|               | M=3 Hierarchy     | 24  | -6.6%      |
|               | Structure (B)     | 36  | -8.2%      |
| Average       | Random Access     | 32  | -          |
|               | M=3 Structure (A) | 30  | -2.1%      |
|               | M=3 Hierarchy     | 24  | -7.1%      |
|               | Structure (B)     | 36  | -9.7%      |

The structure with M=3 and N=30 shown in Fig. 5 reduced BD-bitrate by 2.1% in the transcoding experiments. The hierarchy structure with M=3 and N=24 shown in Fig. 6 reduced BD-bitrate by 7.1%, even though the value “N” was smaller than the structure with the random-access mode. In a comparison between the structure with M=3 and the hierarchy structure with M=3, the hierarchy structure had better results in two sequences of “Crowd Run” and “Park Joy.” Only in the sequence of “Into Tree,” the structure with M=3 and N=30 had better results than the hierarchy structure with M=3 and N=24. One of the reasons is that the value N in the structure is bigger than that in the hierarchy structure.

A comparison of each method for encoding and transcoding is shown in Fig. 16. The PSNR presented in the Fig.16 is the average value of PSNR for the test sequences. The BD-bitrate of transcoding using the MPEG-2 coding structure was better than that of random access structure. This means the information on the picture coding structure in the MPEG-2 video stream is useful for the transcoding of MPEG-2 to HEVC.

From the view point of computational complexity, the average time of transcoding for each method is shown in Tab. 6.

Table 6 Average time for HEVC transcoding

| Method          | N   | Time of encoding |
|-----------------|-----|------------------|
| Random Access   | 32  | 10,310 sec.      |
| M=3 Structure (A) | 30  | 10,456 sec.      |
| M=3 Hierarchy   | 24  | 9,060 sec.       |
| Structure (B)   | 36  | 9,523 sec.       |

The encoding time in the experiments is mainly depend on the bitrate, as shown in Fig. 17. The method with M=3 structure often generates more number of bits than other methods with same quantization parameter. That is one of the reasons why the average time for the method with M=3 structure is longer than other methods in Tab. 6.

One of the features of proposed methods is that the method does not need high-throughput processing in coding blocks such as coding unit(CU) and prediction unit(PU). The additional processing for handling the picture coding information occurs once in a picture. Since the picture rate is usually less than 100 Hz., the time for additional processing is negligible.
Fig. 17 HEVC encoding time for the test sequence “Park Joy”

4. Conclusion

We presented methods for transcoding from MPEG-2 to HEVC by using picture structure information. Although the random-access structure usually had better coding efficiency than other coding structures for HEVC encoding, it was not always the best structure for transcoding from MPEG-2 to HEVC. The method with the hierarchy picture structure by using the MPEG-2 coding structure was proposed in order to improve the efficiency of transcoding. The structure has not only the feature of hierarchy coding like random-access mode for HEVC coding, but also the picture structures with the value M inherited from MPEG-2 coding. The proposed method achieved the bitrate saving by 7.1% as the result of transcoding experiments.

References

[1] Test Model Editing Committee: MPEG-2 Video Test Model 5, ISO/IEC JTC1/SC29/WG 11 Doc. N0400, April 1993.
[2] Recommendation ITU-T H.262 | ISO/IEC 13818-2: Generic Coding of Moving Pictures and Associated Audio Information - Part 2 Video, 1995.
[3] ISO/IEC 14496-2: Information Technology - Coding of Audio-Visual Objects - Part 2: Visual, 1999.
[4] T. Wiegand, G. J. Sullivan, G. Bjontegaard and A. Luthra: Overview of the H.264/AVC video coding standard, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No. 7, pp.560 - 576, July 2003.
[5] Recommendation ITU-T H.264 | ISO/IEC 14496-10: Advanced Video Coding for Generic Audiovisual Services, 2003.
[6] G. J. Sullivan, J.-R. Ohm, W.-J. Han and T. Wiegand: Overview of the high efficiency video coding(HEVC) standard, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 22, No. 12, pp. 1649-1668, Dec. 2012.
[7] Recommendation ITU-T H.265 | ISO/IEC 23008-2: High Efficiency Video Coding, 2013.
[8] Z. Zhou, S. Sun, S. Lei and M. T. Sun: Motion information and coding mode reuse for MPEG-2 to H.264 transcoding, Proceedings of 2005 IEEE International Symposium on Circuits and Systems, pp. 1230-1233, Vol. 2, May 2005.
[9] X. Lu, A. M. Tourapis, P. Yin and J. Boyce: Fast mode decision and motion estimation for H.264 with a focus on MPEG-2/H.264 transcoding, Proceedings of 2005 IEEE International Symposium on Circuits and Systems, pp. 1246–1249, Vol. 2, May 2005.
[10] J. Xin, A. Vetro, S. Sekiguchi and K. Sugimoto: Motion and mode mapping for MPEG-2 to H.264/AVC transcoding, 2006 IEEE International Conference on Multimedia and Expo, pp.313-316, July 2006.
[11] T. Qiang and P. Nasiopoulos: Efficient motion re-estimation with rate-distortion optimization for MPEG-2 to H.264/AVC transcoding, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 20, No. 2, pp. 262–274, Feb. 2010.
[12] E. Peixoto and E. Izquierdo: A complexity-scalable transcoder from H.264/AVC to the new HEVC codec, Proceedings of 2012 19th IEEE International Conference on Image Processing, pp. 737-740, Sept. 2012.
[13] T. Shanableh, E. Peixoto and E. Izquierdo: MPEG-2 to HEVC video transcoding with content-based modeling, IEEE Transactions on Circuits and Systems for Video Technology, pp. 1191-1196 , Vol. 23, No. 7, July 2013.
[14] M. Ikeda, T. Kumaki and T. Ogura: A transcoding method of MPEG-2 to HEVC with picture structure information, Proceedings of RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing 2016, pp. 493-496, March 2016.
[15] G. Bjontegaard: Calculation of average PSNR differences between RD-curves, Technical Report VCEG-M33, ITU-T SG16/Q6, 2001.
Mitsuo Ikeda received his B.E. and M.E. degrees from Kyoto University, Kyoto, Japan, in 1989 and 1991, respectively. He joined Nippon Telegraph and Telephone Corporation (NTT), Japan in 1991. He has been engaged in research on VLSI design and video coding. Currently he is a doctoral candidate in the Department of Electronic and Computer Engineering, Ritsumeikan University. Mr. Ikeda is a member of the Institute of Electronics, Information and Communication Engineers (IEICE) and the Institute of Image Information and Television Engineers (ITE).

Takeshi Kumaki received his B.S. degree from the Department of Mathematics, Faculty of Science and completed the first half of the M.E. program in information mathematics from National Defence Academy, Kanagawa, Japan in 1998 and 2003, respectively, and Ph.D., degree in electric engineering from Hiroshima University, Hiroshima, Japan in 2006. From 2003 to 2004, he was affiliated with the Japan Air Self-Defence Force Electric Experimentation Group. From 2005 to 2009, he joined the Research Center for Nanodevices and Systems (RCNS) and the Research Institute for Nanodevice and Bio Systems (RNBS), Hiroshima University, Japan, where he has engaged in the system design and architecture research. From 2010 to 2012, he became an Assistant Professor in the Department of VLSI System Design, Ritsumeikan University. Since 2013, he has been a Lecturer in the Department of Electronic and Computer Engineering, Ritsumeikan University. He is interested in content addressable memory, SIMD processing architecture and these applications. Dr. Kumaki is a senior member of the Institute of Electrical and Electronics Engineers (IEEE), the Institute of Electronics, Information and Communication Engineers of Japan (IEICE) and the Research Institute of Signal Processing (RISP).

Takeshi Ogura received his B.S., M.S., and Ph.D. degrees in electrical engineering from Osaka University, Osaka, Japan, in 1976, 1978, and 1991, respectively. In 1978, he joined Nippon Telegraph and Telephone Corp.(NTT). In NTT, he engaged in the research and development of CAM LSIs and image encoding LSIs, and their applications. Since 2004 he has been a Professor in the Department of VLSI System Design, Ritsumeikan University, Shiga, Japan. Dr. Ogura is a member of the IEEE, IEICE, IPSJ and RISP.

(Received May 16, 2016; revised August 21, 2016)