MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

A Study of MPEG-2 to H.264/AVC Transcoding with Half-Horizontal Resolution

Jun Xin, Anthony Vetro, Shun-ichi Sekiguchi

TR2008-003 April 2008

Abstract

This paper analyzes the performance of MPEG-2 to H.264/AVC transcoding to half-horizontal resolution. Both the quality of the resulting video and complexity of the system are examined in the context of next-generation HDD recording systems. Experimental results show the importance of motion mapping to maintain high picture quality and that quality improvement over full-resolution transcoding is possible. It is also shown that transcoding to half-horizontal resolution allows for substantial reduction in complexity.

IEEE Intĺ Conference on Consumer Electronics, January 2008

1.4-1

A Study of MPEG-2 to H.264/AVC Transcoding with Half-Horizontal Resolution

Jun Xin[1], Anthony Vetro[2], Shun-ichi Sekiguchi[3]

[1] Xilient Inc, Cupertino, USA*
[2] Mitsubishi Electric Research Labs (MERL), Cambridge, USA
[3] Mitsubishi Electric Corporation, Kanagawa, Japan

Abstract — *This paper analyzes the performance of MPEG-2 to H.264/AVC transcoding to half-horizontal resolution. Both the quality of the resulting video and complexity of the system are examined in the context of next-generation HDD recording systems. Experimental results show the importance of motion mapping to maintain high picture quality and that quality improvement over full-resolution transcoding is possible. It is also shown that transcoding to half-horizontal resolution allows for substantial reduction in complexity.*

I. INTRODUCTION

The ability to efficiently store high-definition video on consumer products, such as Blu-ray Disc, HD-DVD, and HDD systems, is an important feature and will be supported widely in next-generation video storage systems. Since MPEG-2 is still the primary format for broadcast video, and the next-generation storage devices will be capable of decoding the more efficient H.264/AVC format, conversion from MPEG-2 to H.264/AVC could be utilized for more efficient storage [1].

Transcoding aims to reduce complexity of a full decoding and re-encoding approach, while maintaining picture quality. A substantial amount of complexity is reduced by reusing macroblock level information, such as motion vectors, from the input stream when forming the output stream. Another substantial reduction could also be achieved by scaling the picture to half-horizontal resolution (HHR), thereby only requiring half the number of output blocks. The trade-off between quantization noise and scaling to a lower resolution has been studied in [2]. In this paper, we focus on scaling the video within the context of a transcoding system that aims to store the compressed video at a substantially lower bit-rate (i.e., half the input rate) and utilizing a different set of coding tools (i.e., MPEG-2 versus H.264/AVC).

The rest of this paper is organized as follows. The next section describes the system and motion mapping algorithms used in this study. Next, experimental results that examine both quality and complexity of HHR transcoding compared to the full-resolution approach are examined. Finally, concluding remarks are provided.

II. TRANSCODING SYSTEM & TECHNIQUES

In the following, the transcoding system is briefly outlined and the filtering process employed is given. The motion mapping techniques under consideration are also described.

A. System & Filtering

The overall system is similar to that presented in [1] with the exception that a down-sampling operation is performed on the MPEG-2 decoded pictures and up-sampling is performed prior to display. To perform the down-sampling, the following filter is first applied {2, 0, -4, -3, 5, 19, 26, 19, 5, -3, -4, 0, 2}, then the output is formed by even samples in the horizontal direction of the filtered picture. To perform the up-sampling for each row of a picture, a zero value sample is first inserted to the right of every existing sample, then the following up-sampling filter is applied {1, 0, -5, 0, 20, 32, 20, 0, -5, 0, 1}.

B. Motion Mapping Techniques

Two motion mapping techniques for HHR transcoding are presented. The first is referred to as a baseline mapping technique and represents a simple algorithm that is used for comparison, while the second is a modified version of the novel distance weighted average (DWA) presented in [3] that is suitable for HHR transcoding.

In the baseline mapping approach, the output block is fixed to 16x16. The available motion vectors from the two input macroblocks are scaled by a factor of 2 in the horizontal direction. In the case of field motion, a field-to-frame motion vector conversion is performed prior to the scaling [3]. Finally, the resulting motion vectors are averaged. However, if one of the input blocks is intra, then the scaled motion vector of the other block is selected as the output.

In the modified DWA approach presented in this paper, a direct motion mapping to variable block size partitions including 16x16, 16x8, 8x16 and 8x8 is performed. This approach also utilizes the motion vectors of neighboring macroblocks. The basic idea is to estimate the motion vector

* This work was performed while J. Xin was with MERL.

using a weighted average of motion vectors of candidate macroblocks from the decoded MPEG-2 video, where the weights are determined according to the geometric centers of the input and output blocks as illustrated in Fig. 2. If there is no motion information available for an input macroblock, the weight for that motion vector is set to zero. For the 16x16 partition, the modified DWA algorithm is essentially equivalent to the baseline approach. Improved motion prediction is enabled by deriving accurate motion vectors for other output block partitions.

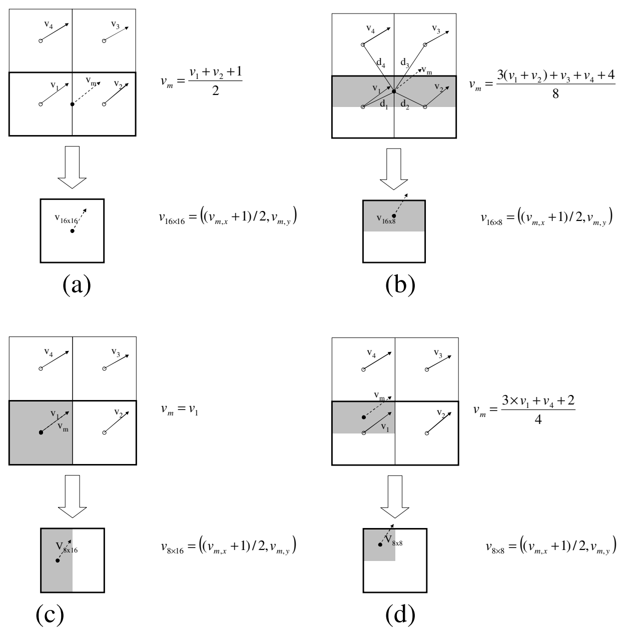

Fig. 1. Motion mapping for various output block partitions using DWA: (a) 16x16, (b) 16x8, (c) 8x16, (d) 8x8.

III. EXPERIMENTAL RESULTS

To validate the effectiveness HHR transcoding and evaluate the effect of the motion mapping techniques, four transcoders are simulated: a full-resolution transcoder using the DWA algorithm (FULL_DWA), the HHR transcoder with the baseline mapping algorithm (HHR_BASELINE), the HHR transcoder using the modified DWA mapping algorithm (HHR_DWA), and the HHR reference transcoder using JM11 (HHR_REF). The simulations were conducted with five 1080i input sequences: Harbor_scene, Horse_Race, Soccer_Action, Street_Car, and Whale_Show. As inputs to the transcoders, the sequences are encoded with typical broadcasting quality and configuration settings. In all simulations, the output streams are coded using RDO and CAVLC for the entropy coding. Also, motion vector refinement of ±1 is used.

A sample plot of the RD performance is shown in Fig. 2, where the line labeled “MPEG-2” shows the PSNR of the input MPEG-2 video, and the line labeled “HHR_UB” shows the PSNR of the down-sampled input MPEG-2 video followed by up-sampling. Intuitively, the “HHR_UB” line is the upper bound of the HHR transcoding quality. A complexity comparison based on CPU time is also plotted in Fig. 3.

Based on the complete set of results (not provided here due to space constraints), the following observations are made regarding transcoding quality. First, HHR_DWA consistently outperforms HHR_BASELINE for all sequences, thereby confirming that the modified DWA mapping provides significant improvement over the baseline algorithm. Also, the performance of HHR_DWA is very close to HHR_REF. Finally, it is noted that HHR_DWA outperforms FULL_DWA for most sequences (both objectively and subjectively); the one exception is StreetCar that contains strong vertical edges and incurs some loss due to down-sampling. For complexity, HHR_DWA is slightly more than half that of FULL_DWA, and incurs some increase in complexity over the baseline method due to motion mapping/refinement and mode decision for an increased number of inter-prediction modes.

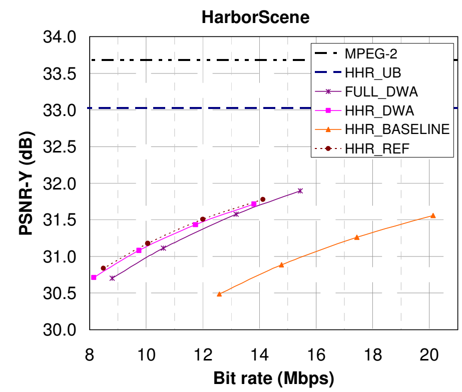

Fig 2. Comparison of RD curves for sample sequence.

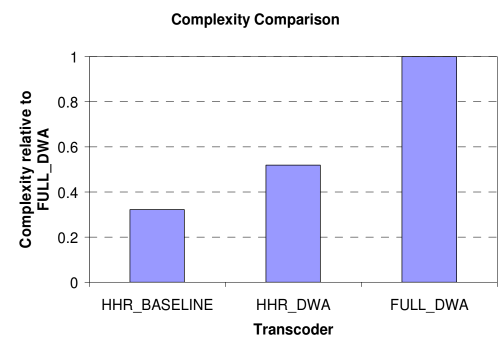

Fig 3. Comparison of complexity.

IV. CONCLUDING REMARKS

The performance of HHR transcoding in the context of consumer video storage systems that utilize MPEG-2 to H.264/AVC conversion has been studied. The importance of efficient motion mapping with the ability to directly map the input motion vectors to various output block partitions have been highlighted. Results demonstrate that HHR transcoding provides an excellent tradeoff between quality and complexity.

REFERENCES

[1] J. Xin, et al., “MPEG-2 to H.264/AVC Transcoding for Efficient Storage of Broadcast Video Bitstreams,” ICCE’06.

[2] A.M. Bruckstein, et al. “Down-Scaling for Better Transform Compression,” *IEEE Trans. Image Processing*, Sept. 2003.

[3] J. Xin, et al., “Motion Mapping for MPEG-2 to H.264/AVC Transcoding,” ISCAS’07.