B-frame Transcoding in H.264 to MPEG-2 Transcoders

Philip Kunzelmann
Florida Atlantic University
Boca Raton, FL
pkunzelm@fau.edu

Hari Kalva
Florida Atlantic University
Boca Raton, FL
hari@cse.fau.edu

Abstract

In this paper we review fast transcoding techniques for H.264/MPEG-2 B pictures. B pictures are a bi-directional prediction mode used by the H.264/AVC and MPEG-2 video compression standards. B pictures allow motion estimation (ME) to be performed using both backward and forward reference frames. This type of inter-prediction is highly effective but also highly complex making reduced complexity methods highly desirable for the transcoding process. Results show an application of reduced complexity techniques may reduce overall transcoding time by up to 70%.

Keywords
Transcoding, MPEG-2, H.264.

INTRODUCTION

The H.264 specification represents a significant advance in the state of video coding technology by providing video of MPEG-2 comparable quality at an average of half the required bandwidth. Although widespread adaptation to H.264 is anticipated, many legacy systems including virtually all existing digital TVs and home receivers use MPEG-2. This motivates the need for an architecture that both efficiently leverages the lower cost of H.264 video and does not require a significant investment in additional video coding hardware. Figure 1 shows the configuration of a system where such transcoders can be employed. The H.264 coded video can be delivered over IP networks and can be transcoded at the head end of a Cable TV network for delivery as

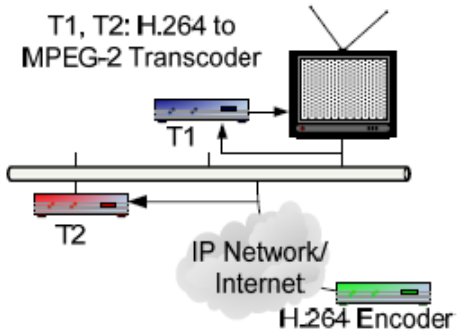

Figure 1. H.264 to MPEG-2 Transcoder

MPEG-2 video. Alternatively, the H.264 video is downloaded to the set-top-box or digital video recorder (DVR) where it is transcoded for playback in MPEG-2 on the existing digital TV or DVR.

While transcoding video is trivial – decode H.264 and re-encode as MPEG-2 – reducing the complexity of transcoders with negligible drop in quality is a challenging problem. A number of techniques have been developed to reduce the complexity of transcoders [1,2]. However, the work on transcoding H.264 to MPEG-2 is very limited; our prior work on transcoding P-frames in H.264 to MPEG-2 was reported in [3]. In this paper we present out work on B-frame transcoding. B-frames in MPEG-2 contribute the most to the compression efficiency but also take the most computational resources. An efficient B-frame transcoder will reduce the complexity of transcoding significantly.

B-FRAMES

B pictures present a special problem for the H.264 to MPEG-2 transcoding case as these standards' definition of B frames is different. MPEG-2 B pictures require exactly one forward and one backward reference frame that has been previously decoded. MPEG-2 B pictures themselves are not used as references for the prediction of other pictures. MPEG-2 B macroblocks may contain either one forward motion vector (MV), one backward MV, or both. H.264 describes a generalized B picture which may refer to any two reference pictures; not necessarily forward and backward frames. In difference to MPEG-2, the reference frames chosen can have a differing display-order relative to the target picture.

Certain differences in H.264 and MPEG-2 B picture specification can be overcome through buffering and application of ME information gained during the decoding stage where it is appropriate to do so. However, because the reference frames used for H.264 B pictures may not have a one-to-one relationship with MPEG-2 implementing a

H.264/MPEG-2 transcoder for B pictures is not a straightforward process. To make best use of information gained during the H.264 decoding stage, a technique must be applied to enhance MPEG-2 B picture prediction where ME information for the desired reference frame may not exist. For this reason we also examine how MPEG-2 B picture prediction may be enhanced were H.264 motion vectors may not be applied.

The coding mode decisions in MPEG-2 are made at a macro block level and mode mapping is used to select the target coding mode for the MPEG-2 B-frame MBs. The B-frame transcoding is based on the mode mapping as shown in Table 1.

H.264 MB Mode	MPEG-2 MB-Mode
Intra	Intra
P-MB	P-MB
B-MB, both forward	B with one forward
B-MB, both backward	P-MB
B-MB, one forward and backward	B-MB

Table 1. MB Mode Mapping for B-frame Transcoding

Frames are mapped such that I-frames are coded as I, P-frames as P and B-frames as B in MPEG-2. The MB mode mapping reduces the complexity as the search space for motion estimation is reduced. Based on the MB mode mapping, the H.264 motion vectors are used to find a candidate motion vector using an approach described in [3]. The performance is compared against a reference transcoder comprised of a full H.264 decoding stage followed by a full MPEG-2 encoding stage. The MPEG-2 and H.264 videos have the same GOP structure. Flexible frame mapping will be explored in the future.

EXPERIMENTAL SETUP

Simulations were conducted using the following experimental setup:

- H.264 QP = 22
- MPEG-2 target bitrates = 6 Mbps ~ 12 Mbps
- IPBBP GOP structure
- H.264 reference frame distance = 1

RESULTS

Experimental results for an application of dynamic range techniques (reported in [3]) for a b-frame GOP structure are listed in Table 2. Simulations show improvements in time between 63% and 73% when compared to a baseline cascaded transcoder. PSNR changes for several sequences show a negligible change, below a 0.5 dB difference.

Dynamic Range				
Seq.	Mean d PSNR Y	Mean d PSNR U	Mean d PSNR V	Mean d Time
ayersroc	0.23	1.00	1.11	67.16%
flower	-0.18	0.11	0.16	73.61%
hook	0.38	1.24	1.30	69.08%
martin	-0.31	0.51	0.77	63.40%
mobile	-0.19	0.07	0.08	62.20%
tennis	-0.26	0.38	0.74	68.68%

Table 2. Dyn. Range Mean change in PSNR and Time

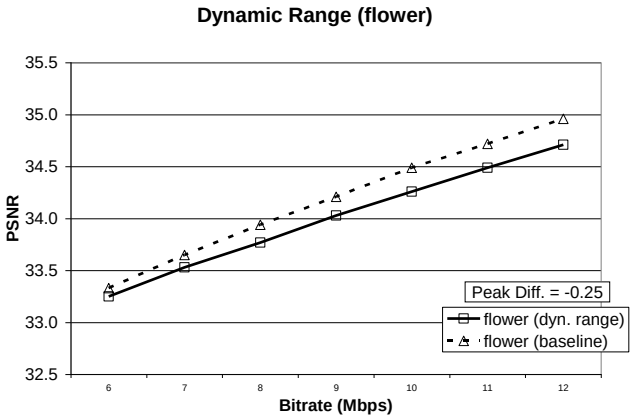

Figure 2. RD Curve (flower) for Dynamic Range

CONCLUSIONS

Reduced complexity transcoding for b-frames requires the resolution of several issues. An H.264 B-MB typically contains several different prediction directions at the 8x8 level. H.264 ME information for a particular MPEG-2 MB may either be unavailable or not useful for the desired MB mode. MB mode mapping, illustrated by Table 1, provides a useful means for translation of H.264 ME information to MPEG-2 in these instances. Simulations show that MB mode mapping for B-frames and an application of reduced complexity transcoding techniques provide a

significant gain in transcoding time at a negligible cost in picture quality.

REFERENCES

[1] G. Fernandez-Escribano, H. Kalva, P. Cuenca, and L. Orozco-Barbosa, “Speeding-up the Macro block Partition Mode Decision for MPEG-2 to H.264 transcoding,” Proceeding of ICIP 2006, Atlanta, September 2006.

[2] J. Chu et. al., “H.264/MPEG-2 Transcoding based on Personal Video Recorder Platform,” Proceedings of the Ninth International Symosium on Consumer Electronics pp. 438-440, June 2005.

[3] P. Kunzelmann and H. Kalva, "Reduced Complexity H.264 to MPEG-2 Transcoder," Proceedings of the IEEE International Conference on Consumer Electronics, Las Vegas, USA, January 10-14, 2007.

[4] J. Bialkowski, A. Kaup, K. Illgner, “Fast transcoding of intra frames between H.263 and H.264,” International Coference on Image Processing, 2004. ICIP ’04. vol. 4, pp. 2785-2788.

[5] Y.-K. Lee, S-S. Lee, and Y.-L. Lee, “MPEG-4 to H.264 Transcoding using Macroblock Statistics.” Proceedings of ICME 2006, Toronto, Canada. uly 2006.