A first approach to speeding-up the inter mode selection in MPEG-2/H.264 transcoders using machine learning

Gerardo Fernández-Escribano · Hari Kalva · Pedro Cuenca · Luis Orozco-Barbosa

Published online: 2 May 2007
© Springer Science + Business Media, LLC 2007

Abstract The H.264 standard achieves much higher coding efficiency than the MPEG-2 standard, due to its improved inter and intra prediction modes which come with a cost of higher computation complexity. Transcoding MPEG-2 video to H.264 is important to enable gradual migration to H.264. However, given the significant differences between the MPEG-2 and the H.264 coding algorithms, transcoding is much more complex and new approaches to transcoding are necessary. In this paper, we introduce and evaluate a low complexity macroblock partition mode decision algorithm, to be used as part of a high-efficient inter-frame prediction in MPEG-2 to H.264 transcoder. The proposed tools are used to compute an optimal MB coding mode decision with significantly reduced computational complexity. Specifically, we achieve the computational savings by using the following MB information coming from MPEG-2: the MB coding modes, the coded block pattern (CBPC) in MPEG-2, and the mean and variance of the 16 $4 \times 4$ sub blocks of the MPEG-2 residual MBs. We use data mining algorithms to develop a decision tree for H.264 coding mode decisions. The decision trees are built using RD optimized mode decisions and result in highly efficient mode decisions, with significantly reduced computational complexity. The proposed transcoder is 35% faster than the RD optimized H.264 reference transcoder without a significant PSNR degradation (0.05 dB on average).
The proposed transcoder performs over 0.4 dB better on average than the SAE cost based H.264 transcoding.

**Keywords** Video transcoding · MPEG-2 · H.264 · Inter prediction · Data mining

### 1 Introduction

Though H.264/AVC [12] is highly efficient compared to MPEG-2 [11], the wide and deep penetration of MPEG-2 creates a need for co-existence of these technologies and hence creates a need for transcoding. Transcoding MPEG-2 video to H.264 is important to enable gradual migration to H.264. However, given the significant differences between the MPEG-2 and the H.264 coding algorithms [10], transcoding is a much more complex task, compared to the task involved in other heterogeneous video transcoding architectures; and new approaches to transcoding are necessary. The main problems that need to be addressed in the design of an efficient heterogeneous MPEG-2/H.264 transcoder are: the inter-frame prediction, the transform coding and the intra-frame prediction. These problems are being examined in various research efforts that are underway [1, 2, 4–7, 9, 15–17, 19, 21–23]. In this paper, we focus our attention on a part of the inter-frame prediction: the macroblock partition mode decision, one of the most computationally intensive tasks involved in the encoding process.

The H.264 video coding standard uses the notion of macroblock partition to refer to the group of pixels in a macroblock that share a common prediction. The encoder selects the coding-modes for the macroblock, including the best macroblock partition and mode of prediction for each macroblock partition, such that the video coding performance is optimized. In this paper, we introduce and evaluate a novel macroblock partition mode decision algorithm, with emphasis on Rate Distortion Optimized transcoding and complexity reduction in the inter-frame prediction in an MPEG-2 to H.264 transcoder. Complexity is reduced by reducing the complexity of the H.264 encoding stage of the transcoder using the information gathered in the MPEG-2 decoding stage. We achieve the computational savings by reusing the following MB information coming from the MPEG-2 decoding stage: the MB coding modes, the coded block pattern (CBPC) in MPEG-2, and the mean and variance of the 16 4×4 sub blocks of the MPEG-2 residual MBs. From an exhaustive analysis of this information, we derive a decision tree suitable for its integration into our algorithm. Our results show that the proposed transcoder is 35% faster than the RD optimized H.264 reference transcoder with a negligible PSNR loss (0.05 dB on average) and a slight increment in bit rate (1.65% on average). The proposed transcoder performs over 0.4 dB better on average than the SAE cost based H.264 transcoding.

The rest of the paper is organized as follows. Section 2 reviews the principles of operation of the prediction of inter-coded macroblocks in P-slices in the H.264 encoding standard. Section 3 describes the data mining tools and the process of building a decision tree for MB mode estimation. Section 4 introduces our macroblock partition mode decision algorithm specifically designed for MPEG-2 to H.264 transcoders. In Section 5, we carry out a performance evaluation of the proposed algorithm in terms of its computational complexity and Rate-Distortion results. We compare the performance of our proposal to the SAE-cost and the RD-optimized methods proposed by the JM10.2 H.264 implementation. Finally, Section 6 draws our conclusions and outlines our future research plans.