博 士 論 文 概 要

論 文 題 目

Research on Computational Complexity Optimization for Intra and Inter Prediction in Video Compression

申 請 者

Guifen　TIAN

情報生産システム工学専攻　　マルチメディアシステム研究

2013 年 5 月

Video content is growing in volume every year for systems such as digital broadcasting, video conference, and online video-on-demand. To manipulate the huge amount of data so that it can be compressed, stored or transmitted via network, video coding technology is indispensable. As video resolution has been increased to high definition (HD) since 2003, H.264/AVC is standardized as the latest international video coding standard which provides high compression ratio while introducing heavy computational load. As video resolution continues to grow from HD to 4kx2k and 8kx4k (SHD), more powerful coding tool is desirable and High Efficiency Video Coding (HEVC) has been under standardization since 2010 and becomes next generation coding standard in Jan. of 2013. HEVC provides twice compression efficiency compared to H.264/AVC, at the cost of tremendously increased computational cost and power consumption. Thus, it is valuable and strongly desirable to develop algorithms which can greatly reduce encoding computational complexity while keeping coding performance.

In video compression, an image frame is divided into blocks and compression is performed block by block. For both H.264/AVC and HEVC, intra and inter prediction contributes significantly for compression efficiency. Intra prediction reduces data redundancy inside of one frame. It predicts the luminance and chrominance components of a block by referring to neighboring reconstructed blocks at different block sizes with a wide range of prediction modes. Inter prediction estimates and compensates motion by finding a matching block in continuous several frames at sub pixel accuracy. In both predictions, the difference between original image block and its predictor will be transformed, quantized and entropy coded before transmission. Simulations show that intra and inter prediction consume about 70-80% encoding complexity of a whole encoder. Thus, it is straightforward and urgent that their computational complexity should be reduced.

In H.264/AVC, 4x4 and 16x16 prediction block with 9 and 4 prediction modes respectively, are allowed for intra prediction; for each 16x16 macroblock (MB), motion estimation and motion compensation in 1/2 and 1/4

pixel for 7 kinds of prediction blocks are performed in inter prediction. To get the best predictor, all the above prediction blocks and prediction modes will be brutally searched which requires enormous computational complexity as well as high power consumption. In HEVC, number of intra prediction blocks is increased to 5 kinds from 4x4 to 64x64 and number of prediction modes also increases about 4 times. For inter prediction, an image block is recursively partitioned into 4 sub-blocks using a quadtree structure and number of coding unit (CU) is 4 with size from 8x8 to 64x64, each with 3~4 prediction blocks. Moreover, number of transform blocks (TU) is also doubled. Thereafter, the computational problem becomes even critical for HEVC.

As to optimization works for H.264/AVC, IEEE TCSVT 2005, ICACT2007, IEEE TM 2007 and IEEE TCSVT2009 proposed to reduce unnecessary prediction blocks for intra and inter prediction, by analyzing edge dictions and detecting motion homogeneity within one MB. If there are little edges or motions, small prediction block is not beneficial; otherwise, the MB will be partitioned into small sub-blocks. In this way, computational complexity can be partially reduced. As to complexity optimization for HEVC, JCTVC-F092 2011 brings forwarded an early termination technique when SKIP mode is the best mode to disable quadtree-based CU partitioning, and computational complexity of HEVC encoder is reduced noticeably. However, for complicated texture and when quantization parameter (QP) is low, SKIP mode happens very infrequently and this method can hardly accelerate HEVC encoder. PCS 2012 proposes to train feature vectors to decide partition or not for each CU using large amount of video samples. This method reduces up to 40% encoding complexity of HEVC, while, increment of bitrate is large due to the distortion between test videos and training samples.

This dissertation avoids the above problems and targets to reduce computational complexity of video encoders largely while keeping encoding performance in terms of image quality and bitrate, so that video data can be compressed efficiently within less encoding time or under lower power consumption. In proposed methods, spatial, temporal, and transformed features

of video signals will be evaluated instead of brutal searching. Based on these common features, unnecessary prediction blocks, non-contributable transform blocks, early termination of motion estimation and motion compensation becomes possible and the encoding computational complexity can be greatly reduced without affecting video quality and bitrate, for a wide range of videos with various characteristics.

Chapter 1 **[Introduction]** gives a brief introduction to the latest H.264/AVC standard and the emerging new standard of HEVC. The common and most computationally intensive coding tool of intra and inter predictions are discussed, followed by the motivation of this research.

Chapter 2 **[Homogeneity Detection based Mode Decision Scheme for H.264/AVC]** presents a fast mode decision algorithm to reduce computational complexity in intra and inter prediction for H.264/AVC. A novel and precise homogeneity detection method based on entropy theory for block type decision is proposed. Here, entropy feature of MB is for the first time introduced, which is calculated as the entropy value of the intensities of 256 pixels located in an MB. Entropy distribution proves that, for MB with low entropy feature, small perdition blocks are of little contribution and can be skipped; for MB with high entropy feature, large prediction blocks can be omitted without sacrificing coding performance. Based on the phenomenon, proposed method can select 16x16 or 4x4 block appropriately for intra coding; Meanwhile, either the large blocks in {16x16, 16x8, 8x16} or sub-blocks in {8x8, 8x4, 4x8, 4x4} are safely chosen for inter coding. Compared with D.Wu's method (IEEE TCSVT 2005) which detects MB's homogeneity as the sum of edge amplitude, proposed method can detect MB homogeneity more accurately. Simulations demonstrate that entropy based block size decision algorithm achieves 94.88% decision accuracy. Moreover, consistent encoding gain is achieved for all videos with different motion and spatial features. Averagely, 52.5% computational complexity of original H.264/AVC encoder has been reduced with negligible loss in bitrate and PSNR. Compared with Z.Liu's method (IEEE TCSVT 2009), original encoding process is accelerated by 12% with bitrate

improved by 0.6%.

Chapter 3 **[Content Adaptive Hierarchical Decision of Variable Block Sizes in HEVC]** brings forward an adaptive scheme that exploits and integrates features in temporal, spatial and transform domains to exclude coding blocks of little contributions, to speed up HEVC encoder, targeting at videos in resolution of HD and beyond. As utilization ratio of each coding unit among several continuous frames stay the same, to use temporal features for block size decision, analysis on utilization ratio of each coding unit is performed and rarely adopted coding unit is skipped for each 64x64 CU. Since optimal coding block is highly dependent on texture complexity, to use spatial features for block size decision, texture complexity (TC) measurement is applied to filter out none-contributable coding blocks to reduce unnecessary coding complexity. Here, a 64x64 image block is firstly down-sampled to 16x16 and texture complexity is calculated as normalized variance of 256 pixels' intensities. Moreover, a dynamic threshold setting approach is proposed to make filtering adaptable to videos and coding parameters. Thirdly, sum of absolute quantized residual coefficient (SAQC) in transform domain is introduced and utilized to prune useless coding blocks efficiently. SAQC is firstly utilized in the same way as using TC for dynamic CU pruning. Also, for zero SAQC, coding blocks from size of 4x4 to 16x16 is non-contributable to coding performance and they are pruned from using for motion estimation and motion compensation. By integrating the three proposals together, the overall block size decision scheme is able to achieve optimization effects for all kinds of coding scenario, even for low QP cases and for videos with complex texture and fast motions. Proposed scheme speeds up original HEVC encoder by a factor of up to 61.89% for 4kx2k video sequences and average complexity reduction is 33.4%. Compared with JCTVC-F092 2011 which is adopted as encoder-side optimization tool in HEVC, another 16% of encoding computational complexity is successfully reduced with equivalent compression performance. For complicated scene, about 30% reduction is further achieved.

Chapter 4 **[All-Zero Block-Based Optimization for Quadtree-Structured**

Prediction and Residual Encoding] presents a novel all zero block (AZB) detection algorithm and an optimization scheme which takes full advantage of AZB to reduce encoding complexity for HEVC. AZB is a residual block whose coefficient becomes zero after DCT and quantization. Generally, if sum of absolute difference (SAD) of a block is lower than a pre-defined threshold, the block can be detected as AZB and threshold setting is critical. In proposed method, a near-sufficient condition for detecting variable sized AZB in luma and chroma components is proposed by modeling quantized residual as a linear function of residual block without performing DCT and quantization. For detected AZB, three techniques are proposed to reduce computational complexity for HEVC. (1) Early termination of motion estimation (2) Skipping of DCT and quantization and (3) Two dimensional pruning of prediction blocks. In technique (1), SAD generally decreases during motion estimation and if it is less than AZB detection threshold, remaining motion estimation process can be early terminated. One important feature is that early termination can be applied to all prediction blocks. In technique (2), all TU for DCT and quantization are skipped for residual encoding. Most importantly, in technique (3), vertically, if either block of {64x64, 32x32} is AZB, all prediction blocks in {16x16, 8x8} will be all pruned; horizontally, if either block in {64x64, 32x32, 16x16, 8x8} is AZB, remaining prediction blocks will be skipped. The overall proposal only needs shift and addition and is suitable for hardware implementation. Compared with L.A.Sousa's method (IEEE Letter 2000), AZB detection rate for luma 8x8, 16x16, 32x32 blocks have been improved by 46.95%, 31.46% and 11.63%, respectively. Experiments on videos from 416×240 to 4k×2k show that HEVC encoder is accelerated by up to 70.46% and 50.34% on average. Compared with best results achieved by Shen's algorithm (PCS 2012), proposed approach further accelerates HEVC encoder by another 3% with bit rate improvement nearly to 1.15%.

Chapter 5 **[Conclusion]** concludes the three proposals and summarizes contribution of this dissertation.