J. Vis. Commun. Image R. 22 (2011) 749–759

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

journal homepage: www.elsevier.com/locate/jvci

Performance analysis, parameter selection and extensions to H.264/AVC FRExt for high resolution video coding

Chenwei Deng [a], Weisi Lin [a,*], Bu-sung Lee [a], Chiew Tong Lau [a], Ming-Ting Sun [b]

[a] Nanyang Technological University, 639798 Singapore, Singapore
[b] University of Washington, Seattle, WA 98195, USA

ARTICLE INFO

Article history:
Available online 1 February 2011

Keywords:
H.264/AVC FRExt
Motion JPEG 2000
Video coding
Inter-frame coding
Intra-frame coding
High resolution
Digital cinema
Residual entropy

ABSTRACT

H.264/AVC FRExt (Fidelity Range Extensions) and Motion JPEG 2000 are the current respective inter-frame and intra-frame coding standards for high resolution (HR) (e.g., 4096 × 2160) visual signals. It is commonly believed that an inter-frame method could achieve higher coding efficiency compared with an intra-frame one, due to the exploitation of video temporal redundancy. However, Motion JPEG 2000 has been selected as the digital cinema compression standard, and some existing work has demonstrated that JPEG 2000 is more suitable at HR situations. In this paper, we compare the rate–distortion (R–D) performance of these two different schemes and give more insight from both theoretical and experimental point of view. We derive an entropy-based R–D model to analyze the test results and the impact of residual entropy and quantization for inter-frame coding. Several extensions are introduced into H.264/AVC FRExt for HR video content for better performance. Experimental results show that these extensions lead to significantly higher coding efficiency and make our extended version more suitable for HR video coding

1. Introduction

With the advance of digital signal processing technology, high definition TV (HDTV) and digital cinema are revolutionizing entertainment industry due to the superior video quality they can offer. HDTV refers to the TV having resolution substantially higher than the standard definition TV (SDTV) systems and the typical frame size is 1280 × 720 (i.e., 720p) and 1920 × 1080 (i.e., 1080p) [1,2]. In the digital cinema field, a number of significant technology improvements (e.g., digital cinematography, post-production, mastering and projection, etc.) have occurred in the past few years. The most common digital cinema system is 2K (2048 × 1080) and 4K (4096 × 2160) resolutions, and either 24 fps or 48 fps [5]. Due to the high resolution and high frame rate, HDTV and digital cinema signals have extremely large amounts of data, often hundreds of megabytes per second.

To manage such a huge data flow, it is necessary to perform effective compression for the raw data captured by HDTV or digital-cinema cameras. The mechanisms used for video compression can be roughly divided into two categories: inter-frame methods and intra-frame methods. The H.264/AVC is the latest international video coding standard developed by the ITU-T Video Coding Experts Group (VCEG) together with the ISO/IEC Moving Picture Experts Group (MPEG). It exploits both the spatial and temporal redundancy (i.e., inter-frame coding). In 2004, the so-called Fidelity Range Extensions (FRExt) of H.264/AVC was established for the applications like HDTV coding [3]. This is in contrast with another video coding standard: Motion JPEG 2000, which is derived from Part 3 of the ISO/IEC 15444-1 standard [4] and only exploits spatial redundancy to achieve data compression (i.e., intra-frame coding).

Due to the temporal redundancy reduction capability, inter-frame coding is usually more efficient than intra-frame coding in terms of rate–distortion (R–D) performance. However, the intra-frame Motion JPEG 2000 has been selected as the digital cinema compression standard by Digital Cinema Initiatives (DCI) [5]. Apart from the random access property (i.e., access the frames in a random order), it is interesting to compare the R–D performance of Motion JPEG 2000 and H.264/AVC FRExt in the case of HR video coding (e.g., 4096 × 2160 digital cinema sequences).

The coding efficiency of H.264/AVC intra-mode has been compared with that of JPEG 2000 in literature [6–8]. However, only a few tests have been conducted to compare the R–D performance between the full H.264/AVC FRExt and the intra-frame Motion JPEG 2000 for HR video content. Smith and Villasenor [9] tested the differences of coding efficiency between inter-frame and intra-frame architectures for high quality, high-resolution sequences. Zeng and Fan [10] reported the state-of-the-art compression techniques for digital cinema systems. Shi and Xu [11] compared the objective and subjective performance of JPEG 2000, H.264/AVC intra coding and the full H.264/AVC. Baruffa et al. [12]

* Corresponding author.
E-mail address: wslin@ntu.edu.sg (W. Lin).

1047-3203/$ - see front matter
doi:10.1016/j.jvcir.2011.01.004

presented an assessment of JPEG 2000-based motion compensated temporal filtering (MCTF) and the H.264/AVC FRExt. Bojkovic and Samcovic [13] showed some simulation results of the Motion JPEG 2000 and H.264/AVC for 4K video content. These experiments give a similar conclusion that Motion JPEG 2000 is more suitable for the compression of HR digital cinema sequences; however, such claim was mainly based on experimental data and without any further theoretical analysis.

In this paper, we attempt to understand the related issues and give more insights from both theoretical and experimental viewpoints. The R–D performance between H.264/AVC FRExt and Motion JPEG 2000 is first compared. The comparison results are then analyzed using an entropy-based R–D model, and the impact of residual entropy and quantization on the bit-rate is discussed for inter-frame coding. Based upon the results of the analysis, we discuss how the coding efficiency of the H.264/AVC FRExt for HR visual signals coding can be improved. Specifically, some extensions are proposed for H.264/AVC FRExt to deal with HR video: different from the DCT-like transform in standard H.264/AVC FRExt, here directional DCT (DDCT) is applied along the one-dimensional structure in each block and 1-D DCT is further performed if the neighboring blocks are with the same size and direction mode. In addition, in our extended version, larger values for motion search range, macroblock size, and skipped block are also used to further reduce the entropy of the motion compensated residuals and allocating more bits for the quantized transform coefficients.

The remainder of the paper is organized as follows: Section 2 compares the R–D performance between the H.264/AVC FRExt and the Motion JPEG 2000 using commonly-used HR sequences. Section 3 derives an entropy-based R–D model to analyze the comparison results and discusses the impact of residual entropy and quantization on the bit-rate. Section 4 discusses the parameter selection and also extends standard H.264/AVC FRExt for HR video coding, according to the R–D analysis in Section 3. Section 5 demonstrates the experimental results of our extension, while Section 6 concludes this paper.

2. Comparison on R–D performance

In this section, we test the coding efficiency of the full H.264/AVC FRExt and Motion JPEG 2000 for HR visual signals.

2.1. Overview

H.264/AVC FRExt is a hybrid video coding scheme, where each video frame is encoded in a block-by-block manner. The coding principles include: 32×32 macroblock partition, variable block size (4×4, 8×8 and 16×16) motion compensated inter-frame prediction and intra-frame prediction, block-based (4×4 and 8×8) discrete cosine transform (DCT) and Hadamard transform (HT), quantization, entropy encoding, and in-loop deblocking.

Unlike H.264/AVC FRExt, only spatial statistical dependencies are exploited in Motion JPEG 2000 and the discrete wavelet transform (DWT) is utilized as the de-correlation engine. This multi-level DWT decomposes each input video frame into different spatial-frequency components called DWT subbands, and the resultant subbands are further splitted into several code-blocks, which are independently encoded using EBCOT (embedded bitplane coding with optimal truncation) and adaptive context-based binary arithmetic coding.

2.2. Testing materials

To evaluate the R–D performance of H.264/AVC FRExt and Motion JPEG 2000, we select three 4096×2160, 50 Hz digital cinema clips "*DucksTakeOff*", "*Crown Run*" and "*Park Joy*" (as shown in

Fig. 1. Three 4096×2160, 50 Hz high resolution (HR) test sequences: (a) *DucksTakeOff*, (b) *Crowd Run* and (c) *Park Joy*.

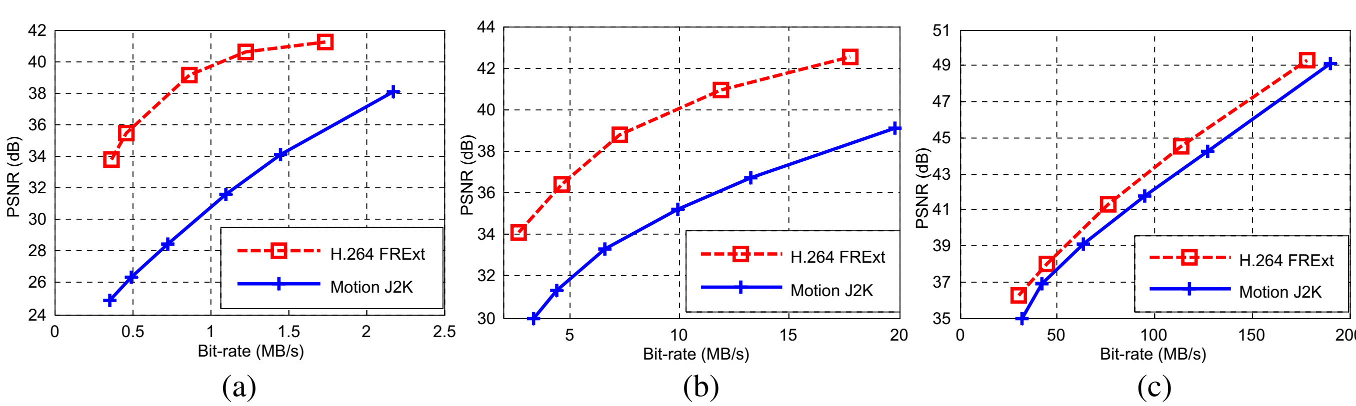

Fig. 2. The R–D performance (luma component) comparison for "*DucksTakeOff*" with different resolutions: (a) 352×288, (b) 1280×720 and (c) 4096×2160.