ENABLE EFFICIENT COMPOUND IMAGE COMPRESSION IN H.264/AVC INTRA CODING

Wenpeng Ding[1,*] *Yan Lu*[2] *Feng Wu*[2]

[1]Dept. of Computer Science, University of Science and Technology of China, Hefei 230027, China
Email: wpding3@mail.ustc.edu.cn
[2]Internet Media Group, Microsoft Research Asia, Beijing 100080, China
Email: yanlu@microsoft.com, fengwu@microsoft.com

ABSTRACT

This paper presents an efficient compound image compression approach based on H.264/AVC intra coding. The text blocks are distinguished from the picture blocks and compressed with a new coding mode. In particular, the text blocks are represented by base colors and index map in spatial domain. A color quantization algorithm optimized for compression is designed to generate this representation. As for the entropy coding of text blocks, a structure-aware context-based arithmetic coder is developed. The mode selection algorithm based on rate distortion optimization is used to select the text blocks along with H.264/AVC intra modes, which can adapt to the targeted bit-rate. Experimental results show that the proposed scheme can achieve 2.8dB gain on average for compound images compared with H.264/AVC intra coding.

***Index Terms* —** Compound image compression, color quantization, entropy coding, mode decision.

1. INTRODUCTION

With the fast development of Internet and widespread rich media applications, we produce not only natural images but also compound images such as web pages, slides and posters. For natural pictures, the existing image and video coding standards (e.g., JPEG2000 [1] and H.264/AVC [2]) have shown good coding performance. However, they are not good at compressing the compound images. Thus, it is necessary to have more research work on compound image compression. Attempts in this field can be grouped into two categories: layer-based and block-based approaches.

Most of layer-based coding algorithms use the standard three-layer mixed raster content (MRC) representation [3], which separates images into two image layers with a binary mask. DjVu [4] is a typical example, which uses a wavelet-based codec (IW44) for image layers, and JBIG2 for mask layer. Hierarchical color clustering on filtered images is used to segment images into two layers. Wu et al. propose another segmentation method to distinguish text from compound images [5]. While these layer-based approaches show satisfactory coding performance on some compound images, their coding performances on purely natural image decrease, because the segmentation algorithms are not robust enough to adapt to both natural and compound images.

Block-based approaches classify images into blocks of several types and use different coding methods for different types of blocks. Said et al. propose a simple block-based scheme, which compresses text blocks using JPEG-LS and picture blocks using JPEG [6]. In [7], Lin et al. propose a classification algorithm, based on the thresholding of the number of colors in each block. These block-based approaches own very flexible coding structure. However, the coding performances are not good enough for both natural and compound images.

In this paper, we propose an efficient compound image compression scheme based on H.264/AVC intra coding, which introduces a new text mode into H.264/AVC intra coding. Blocks in text mode are represented by base colors and index map in spatial domain. This representation is generated using a color quantization algorithm optimized for compression. Since the distribution of color-quantized text blocks usually vastly vary in a compound image, the context model should be able to quickly adapt to the local distributions. Towards this goal, a structure-aware context-adaptive arithmetic coder is proposed for entropy coding.

The rest of this paper is organized as follows. Section 2 analyzes the characteristics of text blocks. Section 3 presents the coding algorithms for text blocks. In Section 4, the text block coding is incorporated into H.264/AVC intra coding. Experimental results are presented in Section 5. Finally, Section 6 concludes this paper.

2. TEXT BLOCKS ANALYSIS

During the past years, various representation methods have been proposed to represent images compactly so as to improve coding efficiency. Transforms such as DCT and wavelet are usually used to represent natural images, due to their features of energy convergence. However, the transform-domain representations are not very suitable for

*This work was done when W. Ding was with Microsoft Research Asia as an intern.

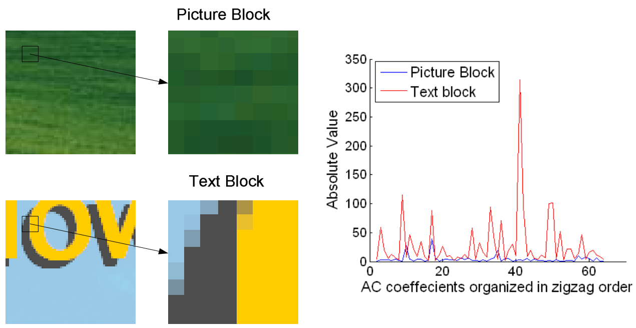

Fig. 1. Comparison of AC coefficients from the picture block and the text block.

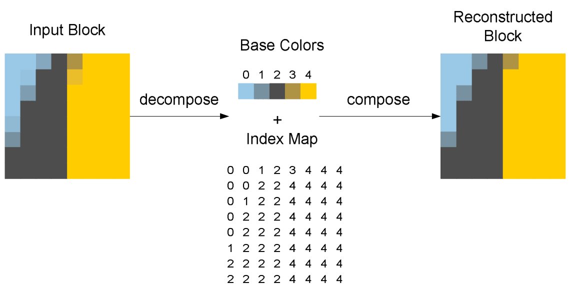

Fig. 2. Text block composed of base colors and index map.

the content with complex structures such as the text content. Fig. 1 shows the exemplified picture block and text block as well as the absolute values of their AC coefficients after DCT. It can be observed that the picture block is compactly represented in DCT domain, because only a few DCT coefficients at low frequencies are large. However, for the text block, its energy is scattered in many DCT coefficients. In spatial domain, text blocks exhibit many features different from picture blocks. For example, the number of colors is limited, and the edges are usually sharp and regular. Moreover, similar structures can be observed among text blocks, which make it more suitable to represent text blocks in spatial domain.

3. TEXT BLOCK CODING

Based on the above observations, we propose a spatial domain representation for text blocks, which describes text blocks using several base colors and an index map. An exemplified text block is shown in Fig. 2. Typically, the index map indicates the structure information of a text block. To decompose a text block, a color quantization algorithm is designed to generate this representation, followed by the structure-aware context-adaptive arithmetic coding.

3.1. Color Quantization

Traditional color quantization methods such as vector quantization and K-means [8] are designed to minimize the distortion for the display of limited number of colors. However, the entropy of the color-quantized image may be high, which makes it difficult to compress. In the proposed approach, the number of bits for compression is also considered in color quantization. Since base colors cost much less bits than the index map, the entropy of a text block usually depends on the index map. To reduce the entropy, we need to constrain the diversity and disorder of the index map. We propose a two-step quantization method to achieve the rate and distortion trade-off.

First, a local quantization is performed to cluster the neighboring similar pixels into a group, which guarantees that neighboring pixels tend to be quantized to the same base color. As a result, the disorder of the index map will be reduced. In particular, given the allowed maximum distortion among neighboring pixels, all neighboring pixels with distance under a threshold are clustered to the same group, which is represented by their average value.

Second, the block after local quantization is further quantized to several base colors. Note that the pixels in the same group are quantized to the same color. The number of base colors of a block depends on the content. Instead of quantizing each block to a fixed number of colors, we set the allowed maximum distortion to be a constant, e.g. $q^2/4$, where q is the quantization step used in H.264/AVC intra coding. Thus, the number of base colors of a 16x16 macroblock may vary from 1 to 8. In practice, we employ the tree structure vector quantization (TSVQ) method, where each pixel is treated as a vector. The maximum distortion is the criterion to split a tree in TSVQ.

3.2. Entropy Coding for Text Block

Both base colors and the index map are compressed with context-adaptive arithmetic coder. The YUV components of a base color are first quantized. The index map shows similar patterns among text blocks. We use contexts and the remapping to exploit the similar patterns to enhance the compression. When coding the index map, the indices are arranged in scan line order. The context for an index to be coded is deduced from its neighboring index values. The current index is then mapped to a symbol according to its context and neighboring indices, and the symbol is coded with an arithmetic coder using the context.

In particular, we define 15 contexts according to their four neighboring indices (i.e. left, left-top, top, right-top). As shown in Fig. 3, each shape represents a neighboring index, while the different shapes indicate the different index values. The 15 contexts are classified into 5 categories. The contexts of the same category have the same number of different color indices. The contexts in Figs. 3(a) and (b) indicate that the number of identical neighboring color indices is four and three, respectively. The contexts in Fig. 3(c) indicate that there are two pairs of identical neighboring

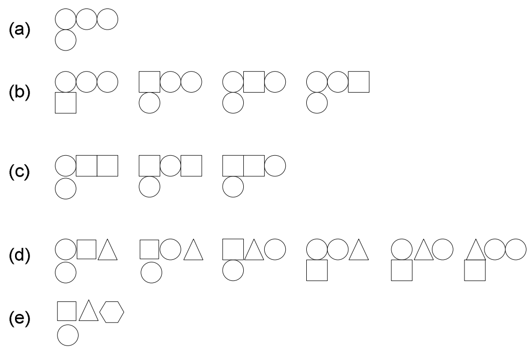

Fig. 3. Contexts for the coding of index map in text block.

Table I: Mapping methods for all the contexts

Context Category	Mapping
(a)	○ → 0
(b)	○ → 0 □ → 1
(c)	○ → 0 □ → 1
(d)	○ → 0 □ → 1 △ → 2
(e)	○ → 0 □ → 1 △ → 2 ⬡ → 3

indices. The contexts in Fig. 3(d) indicate that only two identical neighboring indices exist. When all neighboring indices are different, the context in Fig. 3(e) is used.

The index values represented by the symbols (e.g., "○") will be remapped. The reason is that, in the same context, the probability distributions of indices still depend on the actual values of neighboring indices. Generally, neighboring index values have higher probabilities than other values to occur in the current index. We remap indices with high probability to small values. Taking the context of type (b) as an example, the index value of "○" has the highest probability in occurrence. We then remap this index value to 0, which makes the compression of the current index more efficient. The detailed mapping methods for all cases are shown in Table I.

Moreover, other twenty-two contexts are developed for boundary indices, whose neighboring indices are partially unavailable. These boundary indices are compressed with the similar method.

4. COMPOUND IMAGE COMPRESSION

We incorporate the text block coding into H.264 intra coding as a new mode to compress compound images with both picture content and text content efficiently. H.264/AVC

Fig. 4. Exemplified test images.

Fig. 5. Block classification results.

intra coding is adopted in the proposed scheme for its high coding efficiency on natural images. More details about H.264/AVC intra coding can be found in [2].

Further, we adopt the mode selection algorithm based on rate distortion optimization (RDO) in H.264/AVC reference software to distinguish the text blocks. The RDO-based mode selection can balance the quality of picture content and text content in terms of mean square errors. It can also select the mode adaptively according to the targeted bit-rate.

5. EXPERIMENTAL RESUTLS

The proposed compound image coding approach is implemented into H.264/AVC reference software JM11 [9]. We test the proposed scheme on some compound images (1024X768). Some example images are shown in Fig. 4. The text blocks in the exemplified images are depicted in Fig.5. Fig.6 draws the PSNR-bitrate curves for luma components of the exemplar images. Our scheme achieves 2.8dB gain on average and up to 7dB gain on luma component of the test compound images compared with H.264/AVC intra coding. We observe more improvement for chrome components, because the text mode codes the YUV components jointly.

Fig.7 compares the visual quality of a test image compressed at 0.2 bpp using JPEG2000, H.264/AVC intra coding, DjVu [4] and the proposed scheme. It can be easily observed that the overall quality of our scheme is better than that of the H.264/AVC intra coding. In particular, the text part coded with our scheme is free from ringing artifact, which can always be observed in traditional image coding methods. The proposed scheme also outperforms DjVu by better preserving the chrome components.

We also test the scheme on several standard test images such as Lena, Barbara and the Kodak test images. On these natural images, our scheme achieves almost the same coding performance as the H.264/AVC intra coding.

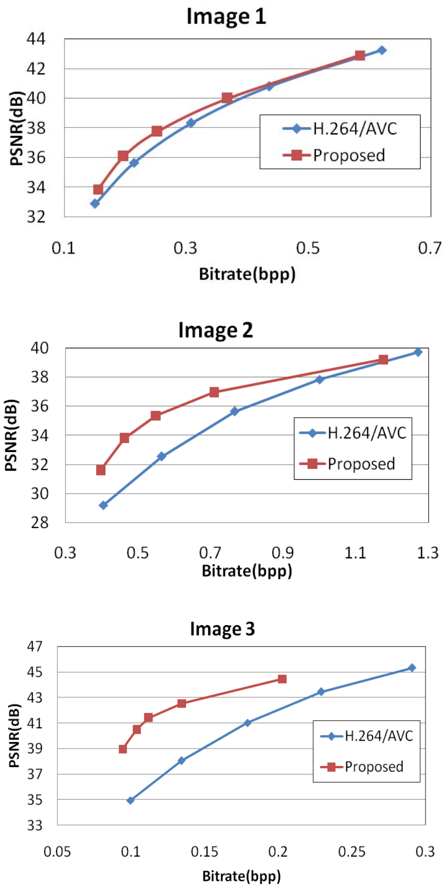

Fig. 6. Rate-distortion curves for the test images.

Original

JPEG2000

H.264/AVC

DjVu

Proposed

Fig. 7. Visual quality comparison.

6. CONCLUSION

This paper has presented a compound image coding approach, which enables H.264/AVC to efficiently compress images with text regions. The text mode is introduced, based on a spatial-domain representation of a text block with base colors and color index map. A color quantization algorithm optimized for compression is designed to generate this representation. Further, a structure-aware context-based arithmetic coder is developed to compress the color index map. Experimental results show that the proposed approach yields significant improvement on compound images and still achieves similar coding performance on natural images compared with H.264/AVC intra coding.

REFERENCES

[1] D. Taubman and M. Marcellin, JPEG2000: image compression fundamentals, standards, and practice, Kluwer Academic Publishers, 2001.

[2] “Draft ITU-T recommendation and final draft international standard of joint video specification (ITU-T Rec. H.264/ISO/IEC 14 496-10 AVC,” JVT of ISO/IEC MPEG and ITU-T, JVTG-050, 2003.

[3] “Mixed raster content (MRC)”, ITU-T Recommendation T.44, Study Group-8 Contribution, 1998.

[4] P. Haffner, L. Bottou, P.G. Howard, P. Simard, Y. Bengio, Y. LeCun, “High Quality document image compression with DjVu”, Journal of Electronic Imaging, pp. 410-425, July 1998.

[5] B. Wu, C. Chiu and Y. Chen “Algorithms for compressing compound document images with large text/background overlap”, IEE Proc. Vis. Image signal Process., vol. 151 no. 6, pp. 453-459 , Dec. 2004.

[6] A. Said and A. Drukarev, “Simplified segmentation for compound image compression”, in Proceeding of ICIP 1999, pp. 229-233, 1999.

[7] T. Lin and P. Hao, “Compound image compression for real-time computer screen image transmission”, IEEE Transactions on Image Processing, vol. 14, no. 8, pp. 993-1005, Aug. 2005.

[8] J. P. Braquelaire and L. Brun, “Comparison and optimization of methods of color image quantization”, IEEE Transactions Image Processing, vol.6, no. 7, pp. 1048-1052, Jul. 1997.

[9] http://iphome.hhi.de/suehring/tml/download/jm11.0.zip.