ISSN: 2320-1363

An Efficient Fully Exploiting Spatial Correlation of Compress Compound Images in Advanced Video Coding

Ali Mohsin Kaittan*1

President of the Association of scientific research and development in Iraq

Abstract

The Mixer Raster Content ITU document compression standard (T.44) specifies a multi-layer multi-resolution representation of a compound document. The compound images consists of text, graphics and natural images, which present strong anisotropic featured. It makes existing image coding standards inefficient on compressing them. It is expected that higher compression can be achieved if more efficient compression standards are used to compress each layer. This paper proposes a novel coding scheme based on the H.264 intra frame coding. Two new intra modes are proposed to better exploit spatial correlations in compound images. The first one is Residual Scalar Quantization mode, where intra-predicted residues are directly quantized entropy coded. The second is base colors and index map mode that can be viewed as an adaptive vector quantization. These two new methods as well as previous intra modes in H.264 are selected by the rate distortion optimization method in each block of image. The result is an unrivaled performance for compressing compound documents and natural images as demonstrated by our experiments.

Keywords: H.264/Advanced Video Coding, Joint Photographic Expert Group-2000, Image Coding, Compound Document.

----------***IJMTARC***----------

1. Introduction

The Mixed Raster Content, ITU document compression standard specifies a multi layer multi resolution representation of a compound document. In this paper we present a basic 3-layer Mixed Raster Coding codec that uses the H.264/Advanced Video Coding operating in INTRA mode to encode the binary Mask layer. The main objective is not to propose a new layer separation or a data-filling algorithm, but to show that Maximal Ratio Combining coding based on H.264/Advanced Video Coding and Joint Bi-level Image Expert Group2 can achieve better compression rates than schemes that use other state-of –the-art still image coders. They usually are complicated and consist of text, graphics, and natural images. When they are displayed in remote client's wireless projectors or thin clients, how to efficiently compress then has become a prevalent and critical problem in many applications.

Well recognized that the state-of the-art image coding standards (Joint Photographic Expert Group, JPEG2000 and intra frame coding of H.264) are all designed for natural images. They generate a compact description in frequency domain if the correlation of samples can be modeled by an isotropic autocorrelation function. The compound images cannot be modeled as isotropic absolutely. In addition for the compression of pure text images such as binary documents, they are already mature algorithms such as JBIG and JBIG2. But they are not good at handling natural images. Therefore some schemes have been reported to efficiently compress compound images. They can be categorized into layer based and block based approaches. The layer based approaches are mainly proposed to efficiently compress compound images. They segment an image into foreground layer, background layer and mask. The colors of text are presented in the foreground layer and the position in the mask. Both foreground and background layers are compressed by the approaches similar to those for natural images.

This paper takes H.264 intra-frame coding as a benchmark to develop a coding scheme for compound images because the mode-based design in H.264 provides an easy way to incorporate proposed techniques and efficiently do mode selection. At the same time we consider the advantages reported by previous studies on compound image compression. Our main contribution is to develop a comprehensive and systematic coding scheme by fully taking the properties of compound images into account.

2. Related Work

Compound documents are examples of raster images that contain a mix of text and pictorial contents. When

ISSN: 2320-1363

compressing text it is important to preserve the edges and shape of characters accurately to facilitate reading. An example of compound image is shown in Fig.1.

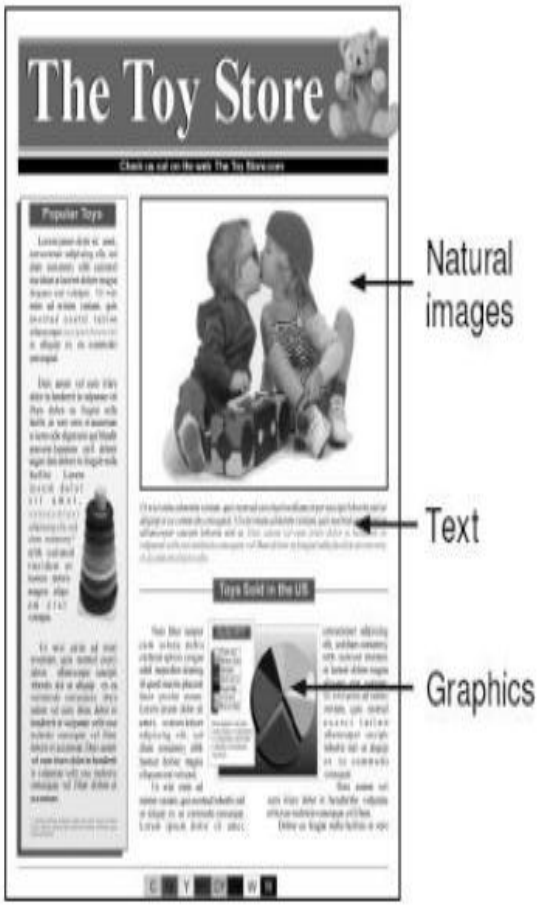

Fig.2.1. **Example of Compound Image.**

The rest of this paper is organized as follows. Section 2 gives a brief overview of the proposed coding scheme. Two new intra modes and the mode selection are discussed in detail in section 3. The experimental results presented and excellent performance of the proposed scheme in Section 4. Finally the conclusion is drawn in section 5.

2.2 Methodology of Compression Scheme

The main feature of two new intra modes is to fully exploit the correlation among samples without transformation techniques. First of all the strong anisotropic of compound images, transforms are not efficient under the isotropic assumption. Non transform inter coding has been proposed in considering the inefficiency of the transform coding on marginally correlated samples. However for compound images direct quantization and entropy coding of residues can bring big coding gain on anisotropic parts. Because of the complicated shape of text, the simple directional intra prediction form boundary samples in H.264 are not efficient some time. It not only enables us not to consider directional intra prediction but also has high coding efficiency.

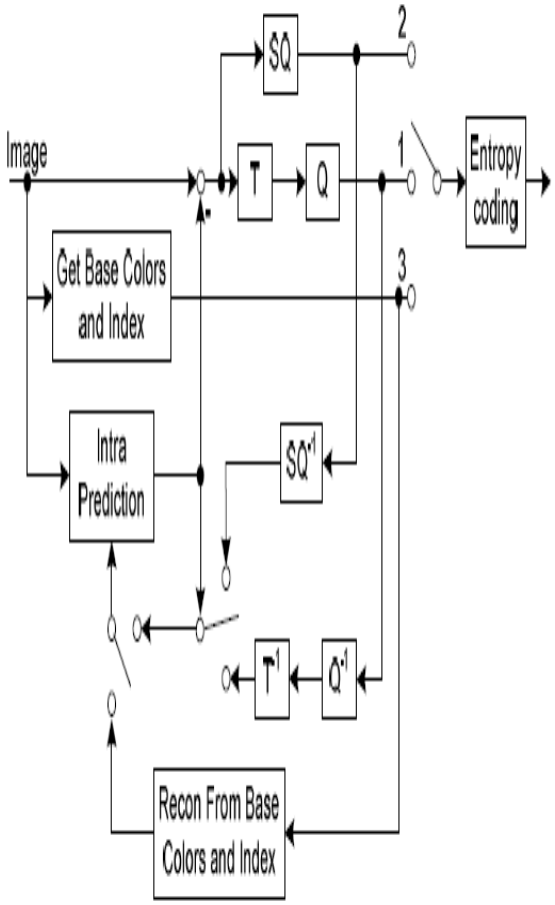

Fig.2. 2.1 the block diagram of the proposed scheme.

The block diagram of the proposed scheme for coding compound images is depicted in Figure 2. There are three possible paths for coding each block. The path 1 is the existing method in H.264 for intra block coding. The path 2 is the proposed Residual Scalar Quantization method. After the intra prediction the residues are quantized and entropy coded. The path 3 is the proposed BCIM method. It does not need intra prediction. The input block is converted to base colors and an index map to compress. The methods can be applied to different block size from 16x16, 8x8 to 4x4. The mode selection is completed by the rate distortion optimization algorithm similar to the reconstructed samples will update the reference buffer as prediction for neighboring block coding.

ISSN: 2320-1363

3. Proposed Compression Scheme

In this paper the proposed Residual Scalar Quantization mode, Binary Coded Image Map mode and mode selection are discussed in this section.

Step 1:

The directional intra prediction method is similar to that used in H.264. The newest video coding standard the H.264/Advanced Video Coding has been well explained in the literature many papers have illustrated its performance showing many comparative results against coders such as Moving Picture Expert Group-2. H.264/Advanced Video Coding is a video compression standard and it was not conceived to be applied as a still image compression tool. Nevertheless the many coding advances brought into H.264/Advanced Video Coding, not only set a new benchmark for video compression, but they also makes it a formidable compressor for still images. One of the components of these advances is the intra-frame macro block prediction method. Thus in Residual Scalar Quantization mode, the boundary samples are directly used for prediction without any low-pass filtering. After prediction the residues are directly quantized. It is a core part in the Residual Scalar Quantization mode. It is a sort of dead-zone plus uniform threshold quantization and can be described by the function C(r) for the residual signal r.

$$C(r) = \text{sign}(r) \times \max(0, \text{floor}(|r|/S + 1 - Z + f)) \quad (1)$$

S is the quantization step and f is the adaptive rounding offset. Parameter Z control the dead zone that is equal to 2S(Z-f). in order to maintain equal expected values before quantization and after inverse quantization, the rounding offset f is adaptively updated within a block as

$$F_i = f_{i-1} + c \times (|r| - |R(C(r)|/s, \quad i=1,2,\ldots, N-1). \quad (2)$$

C is a positive weighting factor and f_o is set as zero at the beginning of each block. N is the pixel number in that block R(x) is the inverse quantization function of x and gives as

$$R(x) = \text{sign}(x) \times s \times (|x| - 1 + Z). \quad (3)$$

Step 2:

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 block depends on the content. Instead of quantizing each block to a fixed number if colors, the allowed maximum distortion is set to a constant value e.g. $q^2/4$. Where q is the quantization step used in H.264/Advanced Video Coding intra coding. Thus the number of base colors of a 16x16 macro block may vary for 1 to 8. A Tree Structure Vector Quantization method is used. Where each pixel is treated as a vector. The maximum distortion is the criterion to split a tree in Tree Structure Vector Quantization. Both base color and the index map are compressed with context adaptive arithmetic coder. The YUV component of base color is first quantized. The context and the remapping are used to exploit the similar patterns to enhance the compression.

Compound Image Compression:

The above explained text block coding procedure is incorporated into H.264 intra coding as a new mode to compress compound images with both picture content and text content efficiently . H.264/Advanced Video Coding intra coding is adopted in the existing scheme for its high coding efficiently on natural images. Further the mode selection algorithm is adopted based on natural images. Further the mode selection algorithm is adopted based on Rate Distortion Optimization in H.264/Advanced Video Coding reference software to distinguish the text blocks. The Rate Distortion Optimization -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 target bit rate. All modes in proposed scheme can be categorized into two categories those are Spatial Domain and Frequency Domain. There is a flag in bit stream to distinguish them. The whole structure of all modes in the proposed scheme is depicted in the below figure.

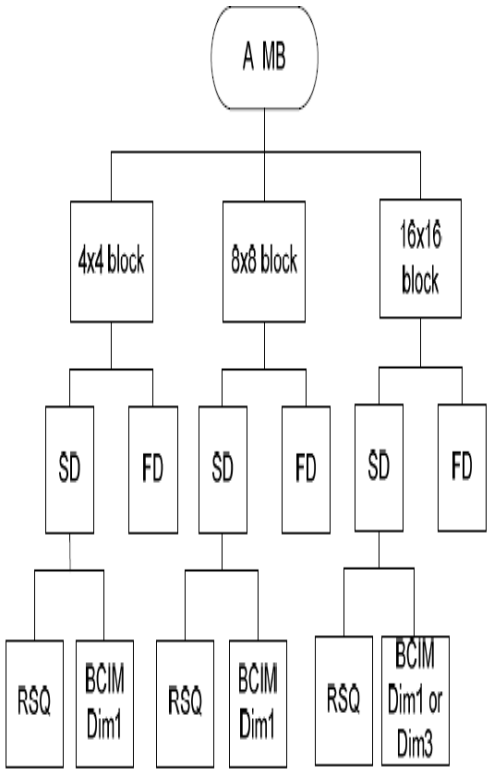

Fig.3.1. The structure of all modes in the proposed scheme.

ISSN: 2320-1363

FD indicated all original intra modes in H.264. SD indicates the proposed RSQ and BCIM modes. For 4x4 and 8x8 blocks, eight directions are used for prediction in the RSQ mode. The Binary Coded Image Map mode is only applied on the luminance component for such size of blocks. The 16x16 blocks the BCIM mode is applied to the luminance component or three color component jointly clustered, there is only an index map. It would save more bits. For the Residual Scalar Quantization mode, three prediction directions are available to get the whole residual block on luminance. But the mode selection between the direct quantization and the transform coding on residues is performed on 4x4 sub blocks which is consistent with the transform block size. In addition the flag that indicated FD or SD will be coded using the flags of neighboring blocks as contexts.

4. Simulation Results

The proposed coding scheme for compound images is implemented into the H.264/AVC reference software JM14.0. In that only intra frame coding is used. Qp in H.264 is set from 47 to 7 with an decreasing step of %. The de-blocking filter is disabled because of the anisotropic feature of compound images

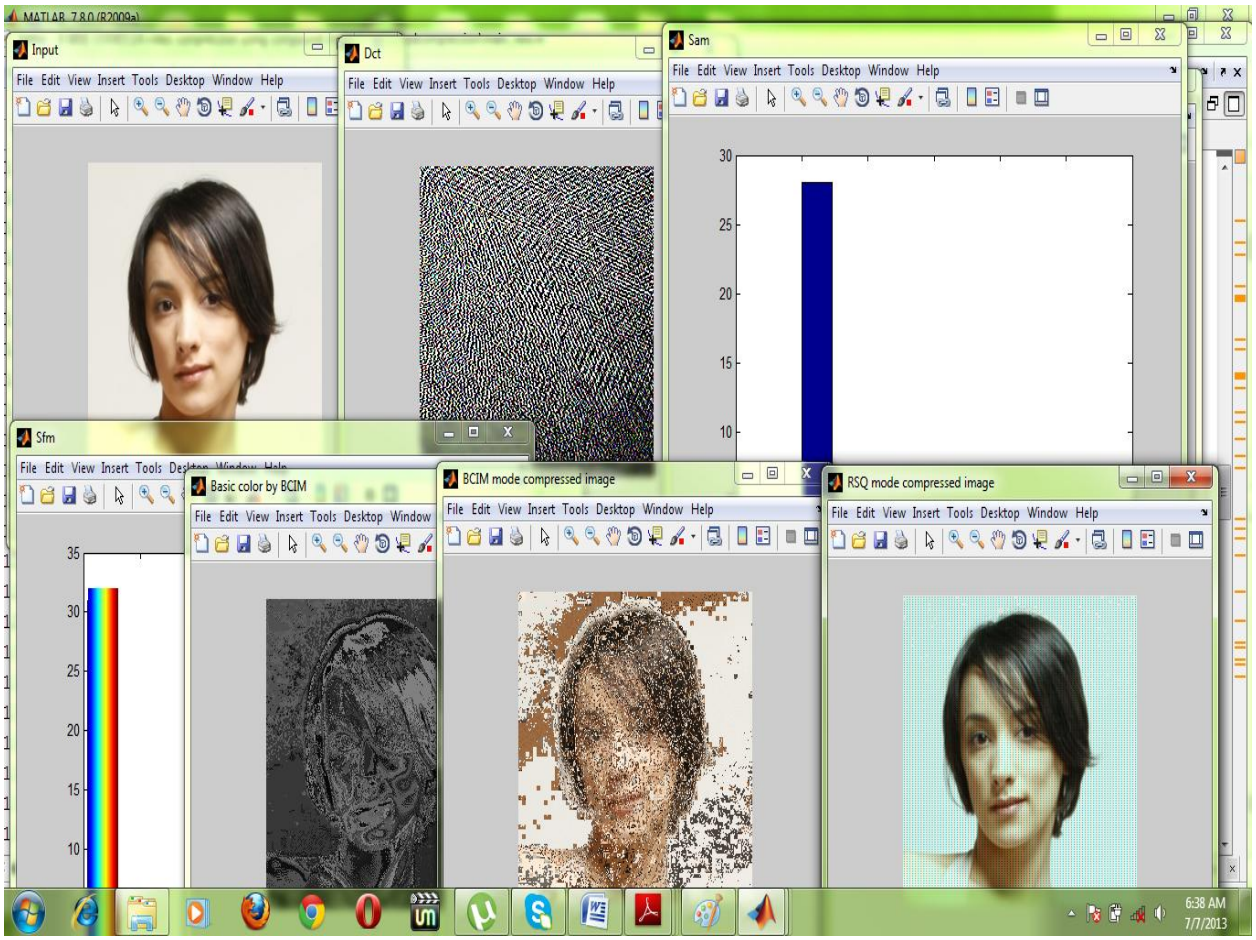

Fig.4. 1 Different Types of testing images we get these results in our experiment.

All results of PSNR for the luminance vs overall bit rate curves are depicted in figure 4. The curves of H.264 with the Residual Scalar Quantization mode are marked by Residual Scalar Quantization. The results of H.264 cooperated with the proposed modes Residual Scalar Quantization and Binary Coded Image Map are marked by proposed. Three schemes are selected for compression Joint Photographic Expert Group2000, H.264 intra-frame coding and the scheme proposed in marked by Ding. We only compress the luminance plane by all bits. The result of Joint Photographic Expert Group2000 should be better than that all three components are compressed. We can observe that the proposed scheme outperforms Residual Scalar Quantization and Binary Coded Image Map modes. The similar results are observed in different images.

ISSN: 2320-1363

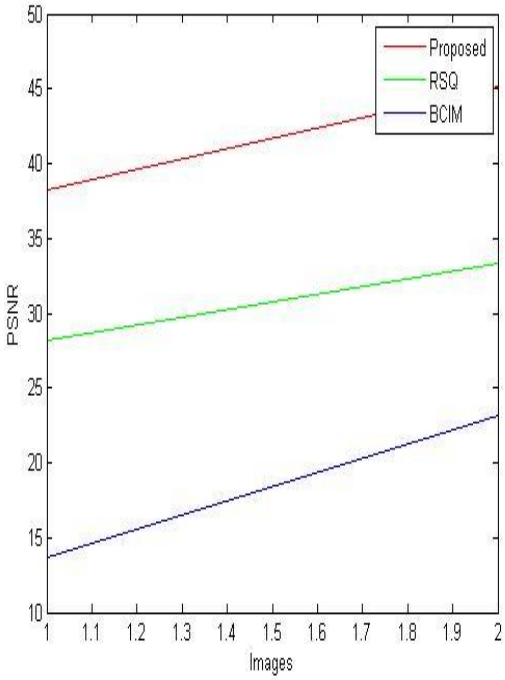

Fig.4.2 Experimental results and comparisons

5. Conclusion

This paper proposes how to extend the H.264 intra-frame coding for compound images to meet the increasing requirements in various applications. The two new intra modes (RSQ and BCIM) are proposed to fully exploit spatial correlations of samples in compound images. Simulation results demonstrate the performance of H.264 intra frame coding is improved even more than 10db on compressing compound images by integrating our proposed two modes. Meantime it keeps the original performance in compressing natural images.

References

[1] De Queiroz, R.L. (2011) Compressing Compound Documents, Document and Image Compression Handbook, edited by M. Barni, Marcel-Dekker.

[2] Ding, W., Lu, Y. and Wu,F. (2007) Enable Efficient Compound Image Compression in H.264/AVC Intra Coding, IEEE International Conference on Image Processing, Vol. 2, Pp. II - 337 - II – 340.

[3] Fan, J. (2008) Text extraction via an edge-bounded averaging and a parametric character model", Proc. Of the SPIE Document Recognition and Retrieval X, Santa Clara, CA, USA , Vol. 5010, Pp. 8-19.

[4] Zaghetto, A. and de Queiroz, (2007) Using H.264/AVC-Intra for Segmentation-Driven Compound Document Coding, IEEE ICIP 2007, Pp. II-333–II-336.

[5] R. L. de Queiroz, R. Buckley, M. Xu, "Mixed raster content (MRC) model for compound image compression", *Proc. SPIE, Visual Communications and Image Processing*, Vol. 3653, pp. 1106-1117, Jan 2009.

[6] R. L. de Queiroz, "Compressing Compound Documents", in *The Document and Image Compression Handbook*, edited by M. Barni, Marcel- Dekker, 2005.

[7] D. Mukherjee, N. Memon, A. Said, , "JPEG-matched MRC Compression of Compound Documents", *Proc. IEEE Intl. Conf. on Image Processing*, ICIP, Vol. 3, Thessaloniki, Greece, pp. 434-437, Oct. 2007.

Ali Mohsin Kaittan Received technical computer engineering from Baghdad-Iraq Al- Rafidain University. He was a President of the Association of scientific research and development in Iraq. Currently, he is studying his M.E on Computer Science & Engineering at University of engineering Osmania University Hyderabad in AP