AN EFFICIENT COMPOUND VIDEO COMPRESSION BY EXPLOITING SPATIAL CORRELATION

[1]Vijay Kumar. K, [2]Mallikarjuna Lingam. K
[1]Department of ECE, Mallareddy college of Engineering& Technology, Hyderabad, India
[2]Associate Professor Department of ECE, Mallareddy college of Engineering& Technology, Hyderabad, India

Abstract

In the past, JBIG technique produced strong anisotropic features. But JBIG has inefficient on compress. To solve the problem, this paper proposes a novel coding scheme based on the H.264 intra-frame coding. The H.264 contains two steps to compress the video. The first is residual scalar quantization (RSQ) mode. It directly quantized and entropy coded. In the entropy coding part, quantized residues are compressed in a similar manner to the coding of DCT transform coefficients. The second is base colors and index map (BCIM) mode that can be viewed as an adaptive vector quantization. In the BCIM mode, we first get the base colors of a block by using a clustering algorithm. All the base colors constitute a base color table. Then, each sample in the block will be quantized to its nearest base color. The index map indicates which base color is used by each sample. Experimental results show that the proposed scheme not only improves the coding efficiency even more than 10dB for compound images but also keeps the similar performance as H.264 for natural images.

Index term: *DCT, BCIM, RSQ and compound image compression*

I. INTRODUCTION

Computer screen images are mixed with text, graphics, and natural pictures. Only in recent three years we saw the compression of computer generated images being studied. Li and Lei [3] developed a lossless compression algorithm, including intraplane coding and interplane coding. In [4], a modified JPEG-LS algorithm was proposed for lossless/near-lossless coding. VNC [1] developed a simple rectangle-based lossless coding, based on the assumption that GUI images are composed of filled rectangles. Obviously, these 3 lossless algorithms are ineffective for natural pictures.

Another category of compound images is scanned document images, and its compression has been intensively studied in the past several years. In order to apply different compression algorithms to different image types, usually a scanned image is first segmented into different classes before compression. Layer-based and block-based algorithms are two main methods frequently used in the literature. Most layer-based approaches use the standard 3-layer Mixed Raster Content (MRC) representation [5][6]. DjVu [7][8] uses a wavelet-based coder (IW44) for the background and foreground, and JB2 for the mask. The segmentation is based on hierarchical color clustering and a variety of filters. Digipaper [9] uses JPEG for the background, color tags for the foreground, and a token-based representation for the mask. Its segmentation is a complicated procedure, which involves connected components, shape cohesiveness, token comparison, etc. In [10], a layered coding method is presented for check image compression. An adaptive morphological filter is used for the segmentation.

Block-based approaches for scanned images are studied due to its low complexity and high spatial resolution. In [11], a rate-distortion optimized segmentation was proposed by using block-thresholding. Cheng and Bouman [12] investigated two segmentation algorithms (TSMAP and RDOS) to classify 8×8 blocks into four classes (Picture blocks, Two-color blocks, One-color blocks, and Other blocks). In [13], Cheng and Bouman extended this method for the application of the standard 3-layer MRC format. JPEG-matched MRC compression [14] first decomposes each block into the standard three layers by using vector 2-means method, then uses JPEG for foreground and background layers and JBIG for mask layers. Li and Lei [15] proposed a histogram analysis to classify each 16×16 block into one of the four types: smooth block (one-color), text block (two-color), graphics block (four-color), and image block (wavelet-based coding). GRAFIT [2] classifies 8×8 blocks into four modes and use different coding methods for each mode.

In this paper is organized as follows. Proposed method in section II. Measurement analysis in section III. The simulation results are presented in Section Iv. Concluding remarks are made in Section V.

II. PROPOSED METHOD

The block diagram proposed method as shown in fig.1. The proposed method consists of two methods to compress the video. They are RSQ mode and BCIM mode. The RSQ mode is more efficient than other transform coding on compressing text and graphics residual blocks. The BCIM mode provides the ability to have a high performance improvement for the efficient representation form of the text/graphics block and minimize the block distortion.

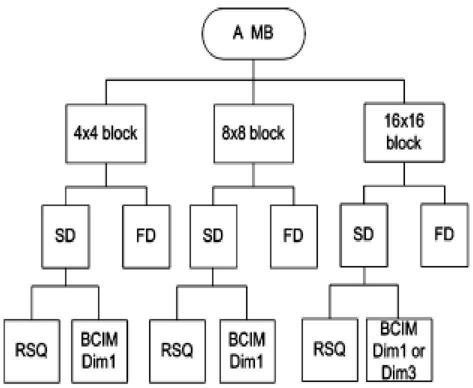

Fig.1. structure of modes in proposed scheme

In fig.1. Contain two types of domain. They are spatial domain (SD) and DCT frequency domain (FD). FD indicates the original intra modes in H.264, where the compression is performed in the DCT domain. SD indicates our proposed RSQ and BCIM modes. To adapt to the local non stationary property of compound images, the spatial domain (SD)modes are applied to 16× 16, 8 ×8, and 4 ×4 block sizes as those DCT frequency domain (FD) modes. The best mode in the spatial domain is compared with the best mode in the DCT frequency domain for the same size block in the rate-distortion sense. The better one is selected.

A. RSQ mode

In the RSQ mode, the directional intra-prediction method is similar to that used in H.264. However, considering that the RSQ mode is usually applied to the strong anisotropic parts (e.g., text and graphics), it is not a good sense that the boundary samples should be first filtered before prediction. Thus, in RSQ mode, the boundary samples are directly used for prediction without any low-pass filtering.

For text and graphic blocks edges cannot completely remove the directional correlation among samples. After intraprediction, residues still preserve strong anisotropic correlation. In this case, it is not efficient to perform a transform on them..To improve efficient performance using RSQ mode. In RSQ mode, the coding gain is defined as

$G_{PCM/TC}=D_{TC}/D_{PCM}$

Where

D_{PCM} is the distortion of PCM and is given by

$D_{PCM}=¢^{2}\sigma^{2}2^{-2R}$

D_{TC} is the distortion on transform coefficients is given by

$$D_{TC} = \left(\prod_{k=0}^{N-1} c^2_{t,k}\sigma^2_{t,k}\right)^{\frac{1}{N}} 2^{-2R}.$$

If is larger than 1, it indicates that non transform coding is more efficient than transform coding. To statistically investigate the coding gain of PCM over a transform on text and graphics residual blocks, we collect 8297 residual blocks of size 8× 8 from text and graphics parts. Taking DCT as the transform, we get the coding gain by $G_{PCM/TC}$ and then apply the residual signal is given by

C(s)=sign(s)×max(0,floor(s/q+1-z+f))

q is the quantization step and is the adaptive rounding offset value. f Parameter controls the width of the dead zone, which is equal to 2q(z-f) . floor(x) is used to get the largest integer that is not greater than x .

B. BCIM mode

In the BCIM mode, the conversion from an image block to several base colors and an index map can be granted as adaptive vector quantization. The searching of base colors is in fact a K:N clustering process.N is equal to 16 for a 4x4 block and 256 for a 16x16 block. K is the number of base colors. Assume that &x_n ,n=1,2...N, is all the samples in a block. The base colors are described as y_k ,k= 1,…,K. In this paper, an

ISSN: 2250-3676
Volume-3, Issue-3, 126-130

upper limit of base color number K_{max} is set as four to restrict the symbol set size of indexes. The selection of base colors is done by minimizing the block distortion as

$$D = \sum_k \sum_{all\ x_n \in Y_k} \|x_n - y_k\|.$$

$$y_k = \frac{\sum_{all\ x_n \in Y_k} x_n}{\sum_{all\ x_n \in Y_k} 1}.$$

Since both " and $ are not large, the clustering is not too complicated. It is done by a node splitting method called tree structure VQ (TSVQ). A suitable $ for each block is decided automatically during the clustering by a distortion restriction as in. Once base colors are determined, every sample in the block is assigned to an index pointing to the closest base color.

The BCIM mode provides the ability to have a high performance improvement for the efficient representation form of the text/graphics block. They are both able to preserve the spatial structures of the text and graphics parts, important to visual quality. A rate distortion optimal method, similar to that in H.264, simplifies the mode selection and avoids the performance loss imported by the inaccurateness of segmentation.

The procedure of BCIM mode as shown below

- In the BCIM mode, we first get the base colors of a block by using a clustering algorithm.
- Then each sample in the block will be quantized to its nearest base color. The index map indicates which base color is used by each sample.
- By mode selection the block will select their mode according to them.
- In this coding method, several lossless coding techniques such as RLE, PNG and gzip are combined into H.264 hybrid coding

III. MEASUREMENT ANALYSIS

Here we are using DCT instead of DFT, to get the Fi,j because it is the transform for compression in our scheme.
Fi,j is the corresponding signal toxi,j in the frequency domain. To analyze the properties of text and graphics blocks in quantity by introducing two features. They are

1. SAM
2. SFM

SAM is a measurement of image predictability and is given by

$$SAM = \frac{\frac{1}{M\cdot N}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}|F(i,j)|^2}{\left\{\prod_{i=0}^{M-1}\prod_{j=0}^{N-1}|F(i,j)|^2\right\}^{\frac{1}{M\cdot N}}}.$$

SFM indicates the activity level of an image and is given by

$$\begin{cases} SFM = \sqrt{R^2 + C^2}, \\ R = \sqrt{\frac{1}{M\cdot N}\sum_{j=0}^{N-1}\sum_{i=0}^{M-1}(x_{i,j} - x_{i-1,j})^2} \\ C = \sqrt{\frac{1}{M\cdot N}\sum_{i=0}^{M-1}\sum_{j=1}^{N-1}(x_{i,j} - x_{i,j-1})^2}. \end{cases}$$

R and C are defined as row and column frequency, respectively.

IV. EXPERIMENTAL RESULTS

The input video frame we are taking is compound image, that frame is shown below.

Fig 2. 12^{th} frame from input video

Discrete cosine transform for the input image is shown below, to get the Fi,j.

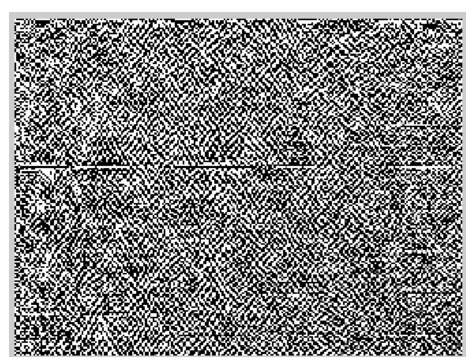

Fig 3. DCT of 12^{th} frame

Fig.4. RSQ and BCIM are combined to compress the compound image

V. CONCLUSION

This paper proposes a compound image compression scheme by exploring spatial domain properties of compound images. Two spatial domain modes, called residual scalar quantization (RSQ) and base colors and the index map (BCIM) achieve significant gains at all bit rates. Its good way to extendH.264 to compress compound images with simple technical extensions and to moderate complexity increasing because of addition mode selections.

VI. REFERENCES

[1] http://www.uk.research.att.com/vnc/index.html

[2] D. Mukherjee, C. Chrysafis, and A. Said, "Low complexity guaranteed fit compound document compression," in Proc. ICIP, vol. I, pp. 225-228, 2002.

[3] X. Li and S. Lei, "On the study of lossless compression of computer generated compound images," in Proc. ICIP 01, vol. 3, pp. 446-449, 2001.

[4] F. Ono, I. Ueno, T. Takahashi, and T. Semasa, "Efficient coding of computer generated images with acceptable picture quality," in Proc. ICIP 02, vol. 2, pp. 653-656, 2002.

[5] Draft Recommendation T.44, Mixed Raster Content (MRC), ITU-T Study Group 8, Question 5, May 1997.

[6] R. de Queiroz, R. Buckley, and M. Xu, "Mixed raster content (MRC) model for compound image compression," in Proc. SPIE Visual Communications and Image Processing, vol. 3653, pp. 1106-1117, 1999.

[7] L. Bottou, P. Haffner, P. G. Howard, P. Simard, Y. Bengio, and Y. LeCun, "High quality document image compression with DjVu," Journal of Electronic Imaging, vol. 7, no. 3, pp. 410-425, July 1998.

[8] P. Haffner, L. Bottou, P. G. Howard, and Y. LeCun, "DjVu: Analyzing and compressing scanned documents for Internet distribution," in Proc. Int. Conf. Document Analysis and Recognition, Sept. 1999.

[9] D. huttenlocher, P. Felzenszwalb, and W. Rucklidge, "DigiPaper: A versatile color document image representation," in Proc. ICIP, vol. I, pp. 219-223, Oct. 1999.

[10] J. Huang, Y. Wang, and E. K. Wong, "Check image compression using a layered coding method," Journal of Electronic Imaging, vol. 7, no. 3, pp. 426-442, July 1998.

[11] R. de Queiroz, Z. Fan, and T. D. Tran, "Optimizing block-thresholding segmentation for multilayer compression of compound images," IEEE Trans. Image Processing, vol. 9, pp. 1461-1471, Sep. 2000.

[12] H. Cheng and C. A. Bouman, "Document Compression Using Rate-Distortion Optimized Segmentation," Journal of Electronic Imaging, vol. 10, no. 2, pp. 460-474, April, 2001.

[13] H. Cheng, G. Feng, and C. A. Bouman, "Rate-distortion based segmentation for MRC compression," in Proc. SPIE Color Imaging: Device-Independent Color, Color Hardcopy, and Applications, vol. 4663, San Jose California, January 21-23, 2002.

[14] D. Mukherjee, N. Memon, and A. Said, "JPEG-matched MRC compression of compound documents," in ICIP'01, pp. 434-437, 2001.

[15] X. Li and S. Lei, "Block-based segmentation and adaptive coding for visually lossless compression of scanned documents," in Proc. ICIP, vol. III, pp. 450-453, 2001.

ISSN: 2250-3676
Volume-3, Issue-3, 126-130

Vijay Kumar. K, received Bachelor Degree in Electronics and Communications from vidya vikas institute of technology, Hyd. Presently he is Pursuing his M.Tech in systems and signal processing in Malla reddy collage of engg and technology, Hyderabad.

Mr K.Mallikarjuna Lingam is an Associate professor at Malla Reddy College of Engineering and Technology, Hyderabad. He is a Research Scholar of J.N.T.U.K, Kakinada. He received his Bachelor degree in Electronics and Tele communication engineering from IETE, New Delhi in 2004. M.Tech, VLSI Design from Satyabhama University, Chennai in 2007. He is Active Member of IETE. His current research interest includes Information Retrieval and Content based Image/Video Retrieval.