Grenze International Journal of Engineering and Technology, Jan 2018

Scanned Compound Image Compression using Block based Pattern Matching

Shubhada S.Thakare[1], Dr.Sujata N.Kale[2] and Swarup S. Mathurkar[3]
[1, 3]Assistant Professor, Department of Electronics Engineering, Govt. College of Engg., Amravati (M.S.)
Email: shubh_diwan@rediffmail.com, swarupmathrkar@gmail.com
[2]Assistant Professor, Applied Electronics Department, Sant Gadgebaba Amravati University
Email: sujatankale@rediffmail.com

***Abstract*—Scanned document compression is needed to store the conventional paper formats in to digital formats. It can be implemented by digitizing scanned document and then optimized memory through compression process. Scanned document may contain different parts such as text, graphics, image and handwritten text. This method compresses the scanned document based on blocked based pattern matching. Here scanned document is first segmented into the blocks and then each block is processed separately. These blocks are matched with dictionary block, if matched found then only the index of that block is used to encode the original block otherwise parse the original block. In this way whole process is carried out for each block and the whole document is compressed. The flags and dictionary indexes generated are compressed with adaptive arithmetic coding. The quality of reconstructed image is measured on the basis of certain performance parameter such as Peak Signal to Noise Ratio, Mean Square Error, Compression Ratio etc.**

***Index Terms*— component; formatting; style; styling; insert.**

I. INTRODUCTION

Digital documents can be classified into either computer-generated or scanned documents. Though both of them can be modeled by compound image source, their noise characteristics vary and lead to different compression goals[1].Unlike computer-generated documents, scanned documents typically contain visually noticeable noise. Therefore visually lossless compression becomes a more reasonable goal than lossless compression for scanned documents. Scanned document is the compound type of document, may contain different parts such as image, text, graphics etc.

Many document imaging applications such as document storage and archiving, scan to print, scan to mail, image coding, internet fax, wireless data transmission and teleconferencing require the use of scanned representations of documents. Document images have been easier and cheaper to use after converting them into electronic form from the conventional paper format. This digital form is then stored in small space and also can send on to the internet. Document's digital form is used for on-line libraries and some publishing sites which make available electronic copies of documents that were originally created or for which the original versions are no longer available. The document pages get degraded after period of times. If these documents are converted into digital format then important document can be saved as it is.

Grenze ID: 01.GIJET.4.1.507_1

The scanned document is either compressed as continuous tone picture, or it is binarized before compression. The binary document can then be compressed using any available two level lossless compression algorithm such as JBIG and JBIG2, or it may undergo character recognition. Binarization may cause strong degradation to object contours and textures, such that, whenever possible, continuous- tone compression is preferred. [1]
Size of scanned image depends on number of factors like scanning resolution, color space and physical dimension of document. A document scan at 300dpi requires less memory space than a document scan at 600dpi.A binary image requires less memory space than color or grayscale image.
This paper is organized as follows: Section I give the introductory knowledge about the compression processes whereas section II is related with compression techniques that are available. In section III, proposed methodology is discussed. Experimental details and results are discussed in section IV. Finally conclusions are given in section V and references are available in the last section.

II. Related Compression Techniques

First step to compress a compound image is to segment that image into different regions. Compound image compression normally based on the following segmentation.
1. Object Based
2. Layer Based
3. Block Based

A page may consist of different objects such as a photograph, a graphic, a letter etc. In object based method, a page is decomposed into object regions. These regions are represented using three layers mask, foreground, and background. Background layer contains picture, mask layer would contain text characters, line art, regions of graphics and text whereas foreground layer contains the colors of the shapes in the mask layer.[2] This method is more complex than other methods .
Layer-based method divides a page into layers. Most of these methods follow the 3-layer mixed raster content (MRC) model as shown in fig. 1. This model represents a color image as two color layers, foreground (FG) and background (BG), third layer is a binary layer (mask). Mask layer is used to recreate original image from FG/BG layers.[3] There are some limitations of this method such as imperfect segmentation, coding of mask layer. Coding of this layer should be accurate because it contains information about foreground and background.

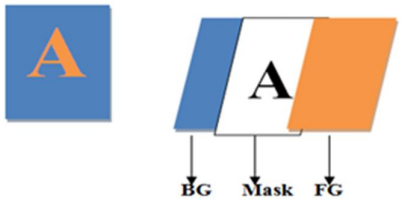


Fig. 1 Basic 3-layer model

Block-based method divides a page into rectangular blocks. Block based segmentation is based on 1) histogram of block and 2) AC-coefficient based. The histogram based segmentation uses different thresholds for separation of the blocks.[4] The AC coefficient based segmentation uses AC coefficient of DCT to decide block type. Background block is identified by adding AC energy to user defined threshold. Other blocks are then classified as text or image block based on k-means clustering by using feature vectors.[4] The advantages of this approach are easy segmentation and use of existing compression algorithms.
In traditional facsimile page compression run-length coding is applied on the document image, but it does not provide high compression ratio. There are several standards developed for binary image compression such as JPEG, JBIG1, JBIG 2, TIFF, CCITT group 3 and 4, GIF, Photo CD, PNG *etc*. CCITT is used for facsimile transmission of document which uses Huffman coding technique.[5] JBIG is used as lossy or lossless compression technique for bi-level images.
The Djvu coder is the most popular algorithm for lossy compression of scanned document. It is based on Mixed Raster Content (MRC) model. The mask layer is compressed by the JBIG2 coder and the background or foreground layers are compressed by a wavelet-based image coder (IW44). For entropy coding the both techniques use a new adaptive binary arithmetic coder i.e. Z-coder. A typical magazine page scanned at 300 dpi can be compressed with Djvu coder to approximately 5 to 10 times better than JPEG for a similar level of

subjective quality.[6] Though Djvu has achieved high compression ratio, the quality of decoded images is merely satisfactory for the web-browsing purpose. With this method document images can be transmitted at faster rate over low speed connections. The ancient document should be pre-processed first to remove noise. For compression document image is first segmented into the text and background images. Binarized text image is compressed by using the pattern matching and substitution (PMS) compression technique.[7] In this technique, dictionary of patterns is dynamically created from the image. Each new component in the image is compared with the elements in the dictionary and coded with a reference element. The patterns and the dictionary are then entropy coded using standard Ziv-Lampel dictionary based technique. After text removal from image, the background image is obtained by filling the holes. The background image is then compressed by using the lossy compression technique.

Document image is compressed using structural analysis and pattern matching[8]. Image is segmented into symbols according to the relation of foreground to background and symbol connectivity and a symbol set is formed. This initial symbol set is then converted to final symbol set using multistage structure clustering and removing repeated and identical patterns. The position of symbols and parameters of representative patterns are compressed using the adaptive arithmetic coder whereas library image is compressed using Q-coder. This technique is independent of symbol and size and can be used for various symbolic images. This method is used for compression of Chinese document images[9].

Compound image is compressed using block based segmentation and adaptive coding algorithm [10]. The method segments image into blocks, classifies into four groups text block, image block, graphic block and smooth block based on histogram analysis. Different compression techniques are used for different blocks. The smooth block is compressed by adaptive arithmetic coder. The text bock is binarized first and then compressed using adaptive binary coder. The graphic block is compressed by a context-based adaptive arithmetic coder. The image block is compressed using the wavelet based coding algorithm.

As our document is scanned textual document, on the basis of review, here by taking combination of methods, new method is introduced which is based on block based pattern matching which provide high PSNR with good subjective quality.

III. PROPOSED METHODOLOGY

This method uses approximate block matching and adaptive dictionary. After pre-processing, there is a partitioning of image into number of blocks. Block size is fixed for an image. For text block, each block X^l of the image, it first searches the dictionary for the element S^l_i that minimizes the Lagrangian cost function of the approximation. The distortion between these two blocks is then comparing with threshold. Distortion between two blocks is calculated by mean square error of two blocks.

$$D = d(X^l, S^l_i) \quad D_M = 1/M*N \sum_{y=1} \sum_{x=1} [X^l(x,y) - S^l_i(x,y)]^2$$

If distortion is less than threshold S^l is consider as a match for X^l otherwise X^l is segmented into two blocks with half pixels of original block and the dictionary is searched for the elements that minimizes the distortion functions for each of the sub-blocks. If a block is again segmented, the same procedure is recursively applied to each half. The bisections are repeated successively in the vertical and horizontal directions.

The flag and dictionary indexes are then encoded by adaptive arithmetic encoder. Save the compressed information in the header file and got the compressed file of the original image. For decompression, exactly reverse operation is done whereas for image block, Prediction is used with compression [11]. A block is taken and predicted its value by using its neighbor pixels which are at top and left side of that block. By evaluating all modes, one mode is chosen giving minimum SAD (sum of absolute difference) value [12]. Comparing this SAD with its threshold, if it is greater, then portioned the original block into two halves with half of the pixels of original block. If the prediction is successful then flag1 is sent with prediction mode and is encoded residue with text block algorithm else same procedure is recursively applied. Flags, mode and dictionary indexes are encoded by adaptive arithmetic encoder. The compressed information is saved in the header file and the compressed file of the original image is obtained. The decoding process is reverse process of encoding.

In the compound image compression, Image is partitioned into number of blocks; each block undergoes the process of identification of block type and form text and image block with a mask which represent each block type. After segmentation, text and image blocks were applied to text algorithm and image algorithm respectively. Arithmetic coder was used to compress binary mask. Then saved the compressed information in

the header file and got the compressed file of the original, applied exactly opposite steps to reconstruct the image.

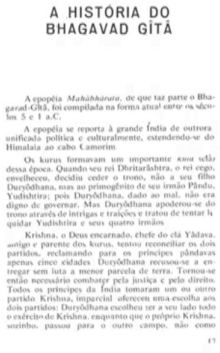
IV. Experimental Details and Results

Performance Analysis was done on the basis of the amount of distortion, which was calculated using mean square error (MSE) and peak signal-to-noise ratio (PSNR) measured in decibels(dB) and compression ratio(CR) measured were used as performance indicators. A good compression algorithm would reconstruct the image with low MSE and high PSNR.

This image compression algorithm is applied on the gray scale images. If image is not gray scale image then it is converted into gray scale format. Here images are first partitioned into number of blocks, block size may vary and then apply designed image compression and decompression algorithm on the blocks. The designed technique is implemented and run using MATLAB 2013. Some images are taken from http://image.unb.br.queiroz/testset and some images are scanned at 600 dpi.

Different original images and their reconstructed images using designed techniques with block size of 16x16

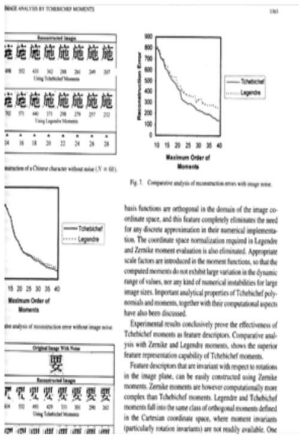
(a) Original Test image



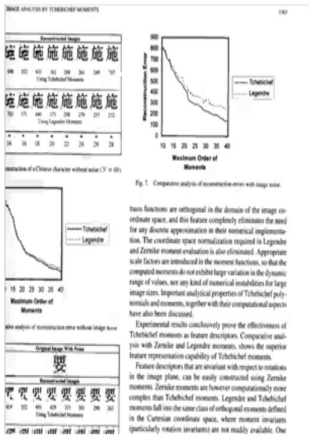
(b) Reconstructed Test image

(c) Original IMG1 image

(d) Reconstructed IMG1 image

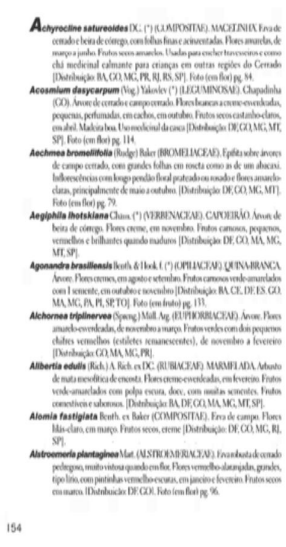


(e) Original Scan0004 image



(f) Reconstructed Scan0004 image

(g) Original Cerrado Image



(h) Reconstructed Cerrado Image

The following tables I and II shows analysis of image compression algorithm with different block sizes.

TABLE I: PERFORMANCE MEASUREMENT FOR BLOCK SIZE 8X8

Name of Image	Original Size in KB	Compressed size in KB	Compression ratio (%)	PSNR (dB)	Compression Time (min)	De- Com-pression Time (sec)	BPP
Test	257	21	8.09	31.79	3.38	1.63	0.65
IMG1	257	34	12.94	32.82	8.71	3.01	1.04
Scan 0004	257	35	13.29	32.15	9.34	3.74	1.07
Cerrado	257	35	13.45	31.31	9.534	4.75	1.08

Comparison with Existing Technology Here result of designed algorithm is compared with existing algorithm. DjVu is one of the well known techniques for scanned document compression. Result of designed algorithm is compared with DjVu-VO method. Comparison is done on the basis of PSNR. Here the graph is plot PSNR verses bit rate.

TABLE II: PERFORMANCE MEASUREMENT FOR BLOCK SIZE 16X16

Name of Image	Original Size in KB	Compressed size in KB	Compression ratio (%)	PSNR (dB)	Compression time (min)	De-Compression Time (sec)	BPP
Test	257	22	8.25	31.38	3.66	2.01	0.66
IMG1	257	34	13.22	31.41	9.3	3.36	1.06
Scan 0004	257	36	13.78	31.58	10.16	4.26	1.11
Cerrado	257	37	14.08	31.50	10.44	5.59	1.13

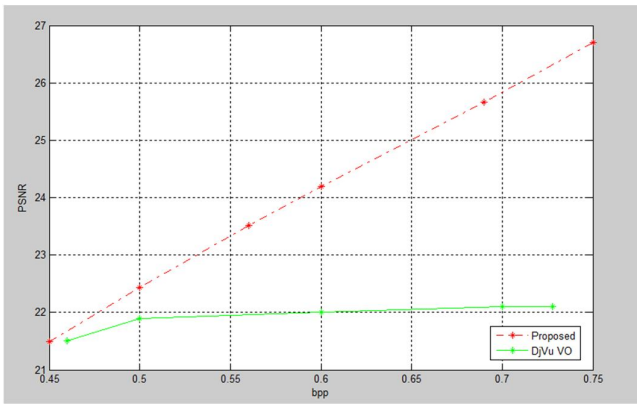


Fig 2: Comparison of Designed Technique with DjVu-VO for Scan0004 image

Here results of designed technique for image scan0004 with dimension (512 x 512) are compared against result of DjVu method. It is observed that designed algorithm provides high PSNR than Djvu method. Below 0.5 bpp PSNR is nearly equal while after 0.5 bpp PSNR is constantly increasing for designed method while DjVu method gives constant PSNR. From the graph designed technique provides better PSNR than DjVu method. Also subjective quality of reconstructed images with designed technique is also good.
Following graph shows the result for cerrado image with dimension (1056 x 1568).

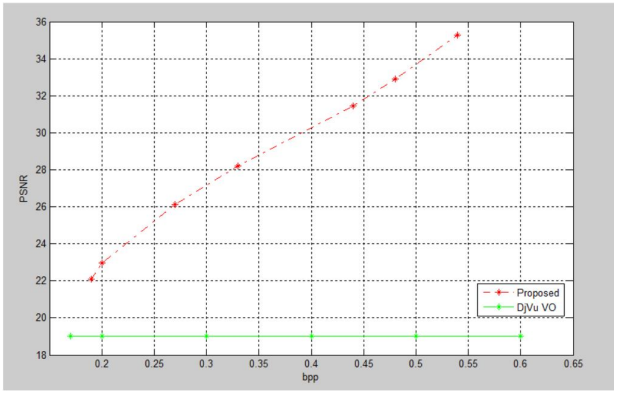


Fig 3: Comparison of Designed Technique with DjVu-VO for Cerrado Image

DjVu technique gives constant PSNR at each bitrate while PSNR is increased for designed technique. In designed technique at high bitrate PSNR increases and subjective quality also increases. From graph it is observed that designed technique gives better PSNR value than DjVu technique for cerrado image.

V. CONCLUSIONS

From the result it is found that PSNR with 8x8 block size is nearly 1db greater than PSNR with 16x16 block size. Time required for compression decreases as block size increases. Here PSNR are high for 8x8 block sizes and subjective quality is also better than 16x16 block size. Compressed size is increased by 1or 2 KB

with 16x16 block size than with 8x8 block size. Here all the results are shown for image dimension 512x512. Designed technique has comparable result with existing method thus this method is different than traditional methods based on transform and quantization approach, can effectively used to compress scanned compound document. In future the study will concentrate on improvement in algorithm such as minimizing compression and decompression timing and maximizing PSNR for low bit rate.

REFERENCES

[1] S. Jayaraman, S. Esakkirajan, T. Veerakumar, "Image Compression," in *Digital image processing*, Ed. New Delhi McGraw-Hill, 2013, pp 493-494.

[2] R. L. de Queiroz, "Compressing Compound Documents," in the Document and Image Compression Handbook, M. Barni, Ed. New York, USA: Marcel Dekker, 2005.

[3] D. Maheswari, Dr. V. Radha, "Comparison of Layer and Block Based Classification in Compound Image Compression," International Journal of Computer Science and Information Technologies, Vol. 2(2), pp. 888-890, 2011.

[4] D. Maheshwari, Dr. V. Radha, "Improved Block Based Segmentation and Compression Techniques for Compound Images," International Journal of Computer Science and Engineering Technology, Vol. 2 No. 3, pp. 49-53,

[5] H. G. Musmann, D. Preuss, "Comparison of Redundancy Reducing Codes for Facsimile Transmission of Documents," IEEE Transactions on Communications, Vol. Com-25, No. 11, November 1977.

[6] L. Bottou et al., "High Quality Document Image Compression Using DjVu," Journal of Electronic Imaging, Vo1.7, pp.410-425, July 1998.

[7] G. Calvagno, G, A. Mian, R. Rinaldo, L. Zimolo, "Ancient Document Compression and Archiving via Page Decomposition", proceeding of X European signal processing conference, 2000.

[8] Y. Yang, H. Yan, D. Yu, "Content Lossless Document Image Compression Based on Structural Analysis and Pattern Matching," Pattern recognition 33, 2000, pp 1277-1293.

[9] Maggie M. K. Tsui, Alan W. C. Liew, H. Yan, "Pattern Based Content Lossless Compression of Chinese Document Images," International symposium on Intelligent Multimedia, Video and Speech processing, October 2004, 20-22.

[10] X. Li, S. Lei, "Block-based Segmentation and Adaptive Coding for Visually Lossless Compression of Scanned Documents," International Conference on image processing, vol. 3, pp 450-453, 2001

[11] A. Zaghetto, R. L. de Queiroz, June 2013, "Scanned Document Compression Using Block- Based Hybrid Video Codec," vol. 22, No.6, pp.2420-2428, June 2013.

[12] N. Francisco, Nuno M. M. Rodrigues, A. B. da Silva, Murilo Bresciani de Carvalho, Sérgio M. M. de Faria, Vitor M. M. Silva, "Scanned Compound Document Encoding using Multiscale Recurrent Patterns," IEEE Transaction on image processing, Vol. 19, No. 10, October 2010.

[13] N. M. M. Rodrigues, E. A. B. de Silva, M. B. de Carvalho, S. M. M. de Faria,and V. M. M. Silva, On "Dictionary Adaptation for Recurrent Pattern Image Coding," IEEE Trans. Image Process., vol. 17, no. 9, Sep. 2008, pp. 1640-1653.

[14] L. Lucas, N. Rodrigues, and S. Faria, "Compound Image Segmentation for Multiscale Recurrent Pattern Coding" in Proc.Int.Conf.Telecommunications,May2009,pp.267-270.