EFFICIENT DOCUMENT COMPRESSION USING INTRA FRAME PREDICTION TECTHNIQUE

Chitradevi S[1], Vijayalakshmi K[2]

[1]*M.E. Final year, Information Technology, Mepco Schlenk Engineering College, Tamil Nadu, India*
[2]*Professor, Information Technology, Mepco Schlenk Engineering College, Tamil Nadu, India*

Abstract

Compression is the reduction of the irrelevant or redundant data, in order to save space or transmission time. Scanned document is compressed to reduce the requirement of the storage space. This paper proposes, an Advanced Document Coding (ADC), which was a H.264/AVC (Advance Video Coding) video coding standard for compressing the scanned document. During encoding phase, intra-frame and inter-frame prediction technique is used to reduce the redundancy of the image data. Intra-frame prediction technique is performed relative to information that is contained only within the current frame and not relative to any other frame. The Inter-frame prediction technique means only changes in the image from the previous subpages. This prediction generates residual data that can be transformed and encoded using Context Adaptive Binary Arithmetic Coding (CABAC). In the decoding process, Arithmetic decoding, inverse transform are used to assemble the final document page. The proposed work is analyzed by Peak Signal to Noise Ratio (PSNR), which is calculated by using global Mean Square Error (MSE). The result reveals that, ADC improves the encoding efficiency than the JPEG2000 in terms of PSNR value.

Keywords: *Advanced Document Coding (ADC), H.264/AVC, Intra- frame prediction, Scanned Document Compression*

----------***----------

1. INTRODUCTION

Compression is one of the active areas of image processing. Compression of scanned document can be very difficult. Many compression methods have been proposed in literature such as JBIG [1], JBIG2 [2], JPEG [3], and JPEG2000 [4]. JBIG is a lossless image compression, designed for compression of binary images and uses arithmetic coding known as QM-coder. This is particularly used for faxes. However, the problems result from the fact that the arithmetic coding method is patented by IBM. JBIG2 is suitable for both lossless and lossy bi-level compression. The encoder will segment the image and compressed using a context dependent arithmetic coding algorithm. JPEG uses a lossy compression; the encoder is based on the Discrete Cosine Transform (DCT), quantization and encoded using Huffman encoding method. It may not be as well suited for line drawings and fails to compress bi-level images. JPEG2000‘s spatial redundancy prediction can contribute more to the compression process and encoded using wavelet transform. A document typically consists of presents repetitive symbols. So that dictionary- based compression methods become very efficient. We propose an encoder that explores such a recurrence through the use of pattern matching prediction and the efficient transforms encoding of the residual data.

In the proposed method three postulates are assumed. Firstly, we must be avoided complex multi-coder schemes such as mixed raster content (MRC). Secondly, the compression method is standard as possible. Because we are dealing with standard image encoders such as JPEG and JPEG2000are the most adequate. Thirdly, we required high quality reconstructed version of scanned documents. Third postulate is very important for the historical value of rare books need to be stored digitally. A standard single coder approach that operates on natural images and delivers high-quality reconstructed compound documents. The proposed coder makes heavy use of the H.264/AVC. H.264/AVC is well explained in the literature [5]-[9]. H.264/AVC leads to outperformance than the existing standards.

2. H.264/AVC CODING STANDARD

In this section we explained the H.264/AVC video coding scheme and we also discuss the various prediction techniques such as intra and inter frame prediction, transform coding and entropy coding as well as decoding process. Fig.1. Shows the overview of the proposed model.

2.1 Page processing Algorithm

For the compression, first the book will be scanned. The size of the each scanned page is H×W pixel. The proposed encoding method, the scanned page is subdivided into nine H/3 × W/3 pixels sub-pages or frames. Then, these frames are used to create a video sequence. The page segmentation is used for similar text patterns are more likely to be found on the same frames rather than on different frames.

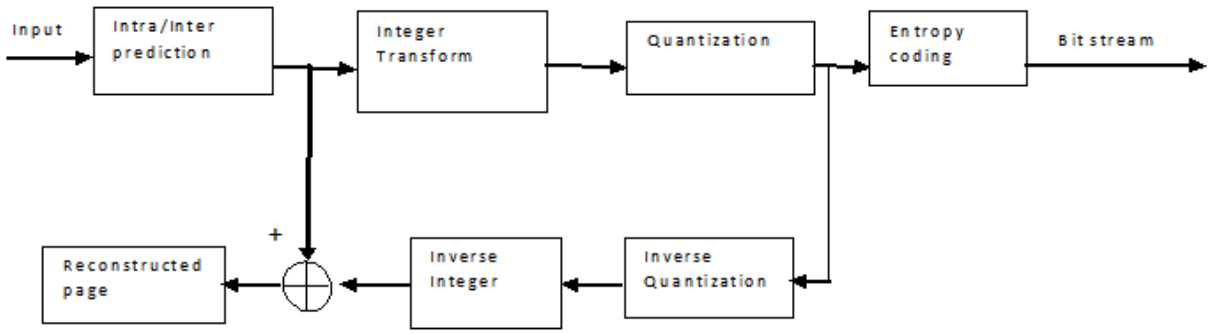


Fig -1: Overview of the Advanced Document Coding

2.2 Intra Frame Predictions

Intra frame prediction technique is used for predict the neighboring pixel values. In INTRA 4x4, a macroblock of size 16x16 pixels are divided into 16 4x4 sub blocks. Intra prediction scheme is applied individually to these 4x4 sub blocks. Nine different prediction modes supported in intra frame prediction and named as vertical, horizontal, DC, diagonal down left, diagonal down right, vertical right, horizontal down, vertical left and vertical right. It is shown in Fig.2. The predicted samples are determined from a weighted average of the previously predicted samples A to M.

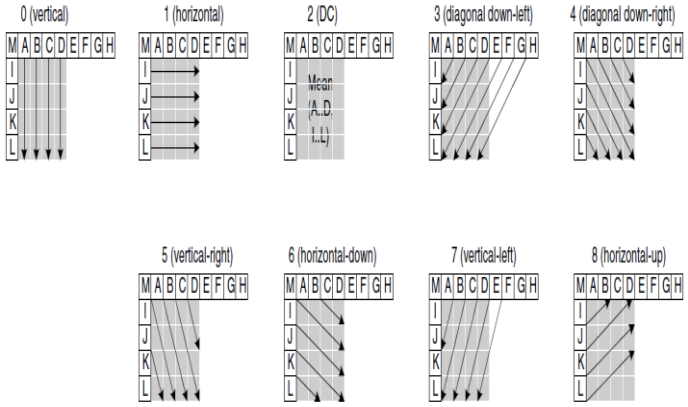


Fig-2: 4x4 intra-prediction modes

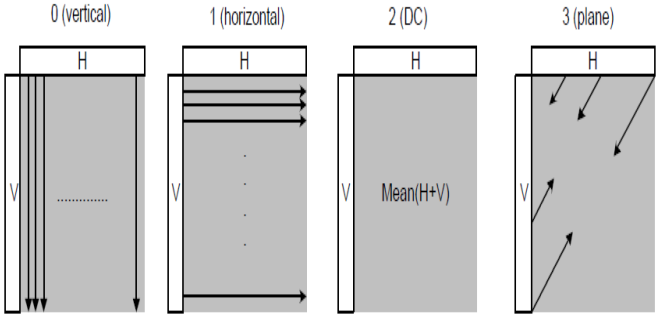


Fig-3: 16x16 intra prediction modes

In the 16x16 intra prediction four modes are used and named as vertical, horizontal, DC and plane. The fig. 3. Shows the 4 prediction modes of intra 16x16. The vertical, horizontal and DC modes are similar to the prediction modes for intra 4x4. In the fourth mode, Plane prediction use weighted combination of horizontal and vertical adjacent pixels.

2.3 Inter frame prediction

The inter frame prediction is used to exploit similarities between subpages or frames in order to reduce the amount of information to be encoded. Based on previously encoded blocks, first we built a prediction of the current frame and then create a prediction error by subtracting the prediction from the current frame. In the inter frame prediction, each current macroblock is predicted as one 16 × 16 partition, two 16 × 8, two 8 × 16 or four 8 × 8 macroblock partitions. In the case of 8×8 partition, the sub-macro blocks may be further partitioned in one 8×8 partition, two 8×4, two 4×8 or four 4×4 sub macroblock partitions. The fig.4.shows the inter frame prediction.

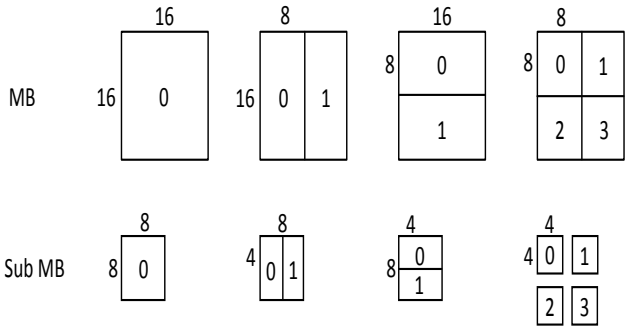


Fig-4: Inter prediction (block size: 16 × 16, 16 × 8, 8 × 16 8× 8, 4× 8, 8× 4 and 4 × 4 pixels respectively).

2.4 Residual Coding

The residual macro blocks are transformed by integer transform. The forward 4x4 integer transform [10] of a sample block can be expressed as,

$$\mathbf{Y= (CXC^T) \oplus E_f} \quad \textbf{(1)}$$

Where, C is core value of the integer transform and Ef is the Scaling factor that are defined by,

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & -2 & -1 \end{bmatrix}$$

$$E_f = \begin{bmatrix} a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \\ a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \end{bmatrix}$$ Here, a=$1/2$; b= $\sqrt{2/5}$

The resulting integer transformed coefficients are quantized and entropy encoded using CABAC.

2.5 CABAC

CABAC provides much better compression than most other entropy encoding method. It is one of the primary advantages of the proposed encoding scheme. This process has following three steps: binarization, Context modeling and binary arithmetic coding.

- **Binarization:** In this step, a non-binary valued symbol value is converted into a binary symbol value. This binary code is further encoded prior to transmission.
- **Context model selection:** This stage a probability model for one or more bins of the binarized symbol. This model chooses a selection of available models depending on recently-coded data symbols and stores the probability of each bin being "1" or "0".
- **Arithmetic encoding:** According to the selected probability model, each bin being encodes based on the Arithmetic encoder.
- **Probability update:** The context model is updated based on the actual coded value.
- Stages are repeated for each bin of the binarized symbol.

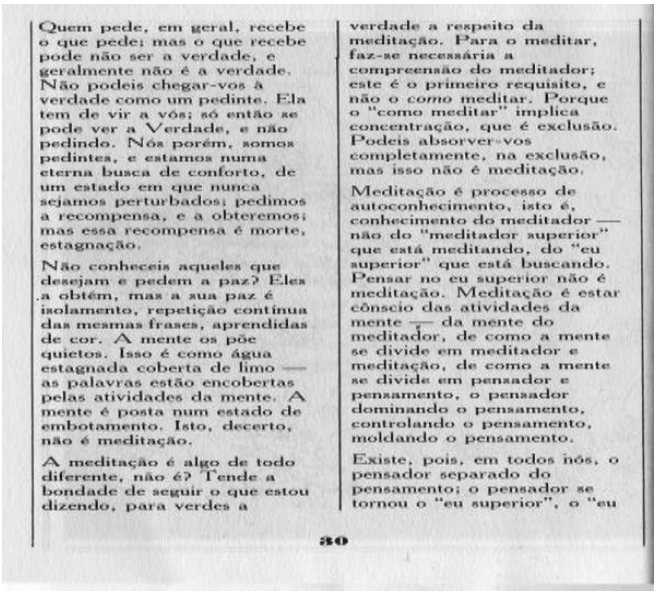
2.6 Decoder Summary

In the decoder process, inverse quantization and inverse integer transformation are done. By adding the predicted frames to the values of an inverse transform and produce the reconstructed frames and assemble into the final document book pages.

3. RESULTS

In our test, the different scanned pages are compressed using JPEG2000 and the proposed ADC. The reason why we select the JPEG2000 for comparison is that, it meets the three postulates mentioned in Section I. The test set is composed by 18 documents divided into the following 4 classes. The entire test set is available in [11].

1) Class 0: multi- page text only documents.
2) Class 1: single- page text only documents.
3) Class 2: multi- page compound documents.
4) Class 3: single –page compound documents.



Quem pede, em geral, recebe o que pede; mas o que recebe pode não ser a verdade, e geralmente não é a verdade. Não podeis chegar-vos à verdade como um pedinte. Ela tem de vir a vós; só então se pode ver a Verdade, e não pedindo. Nós porém, somos pedintes, e estamos numa eterna busca de conforto, de um estado em que nunca sejamos perturbados; pedimos a recompensa, e a obteremos; mas essa recompensa é morte, estagnação.

Não conheceis aqueles que desejam e pedem a paz? Eles a obtêm, mas a sua paz é isolamento, repetição contínua das mesmas frases, aprendidas de cor. A mente os põe quietos. Isso é como água estagnada coberta de limo — as palavras estão encobertas pelas atividades da mente. A mente é posta num estado de embotamento. Isto, decerto, não é meditação.

A meditação é algo de todo diferente, não é? Tende a bondade de seguir o que estou dizendo, para verdes a verdade a respeito da meditação. Para o meditar, faz-se necessária a compreensão do meditador; este é o primeiro requisito, e não o *como* meditar. Porque o "como meditar" implica concentração, que é exclusão. Podeis absorver-vos completamente, na exclusão, mas isso não é meditação.

Meditação é processo de autoconhecimento, isto é, conhecimento do meditador — não do "meditador superior" que está meditando, do "eu superior" que está buscando. Pensar no eu superior não é meditação. Meditação é estar cônscio das atividades da mente — da mente do meditador, de como a mente se divide em meditador e meditação, de como a mente se divide em pensador e pensamento, o pensador dominando o pensamento, controlando o pensamento, moldando o pensamento.

Existe, pois, em todos nós, o pensador separado do pensamento; o pensador se tornou o "eu superior", o "eu

30

Fig-4 (a): Input scanned page

Fig-4 (b): splitted frames

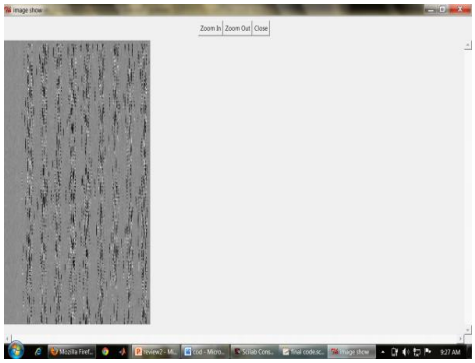


Fig-4(C): Predicted Error

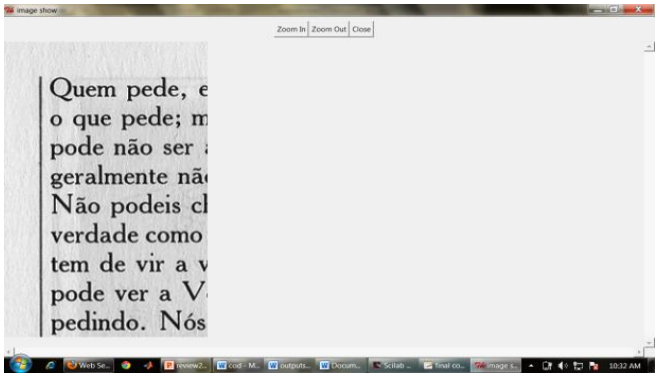
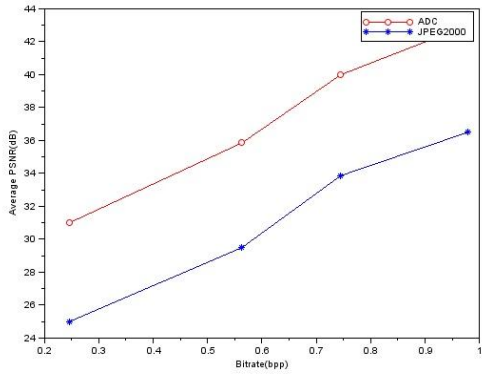
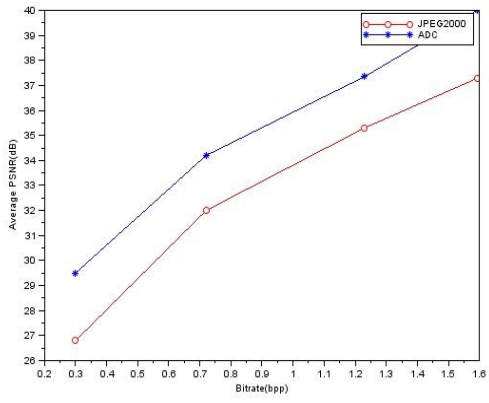


Fig-4 (d): Reconstructed Frame

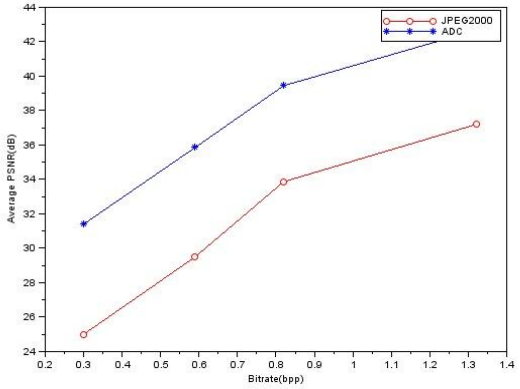
The Fig. 4 (a) shows the input scanned page of single page text only document classes and the Fig. 4 (b) shows pages are splitted into subpages. Here page is splitted into 9 frames. The Fig. 4(c) shows the residual frame which was calculated by subtracting the original frame and predicted frame. Then apply the integer transform, quantization and CABAC. After that decoding was done and then reconstruct the frames. The Fig.4 (d) shows the reconstructed frames.



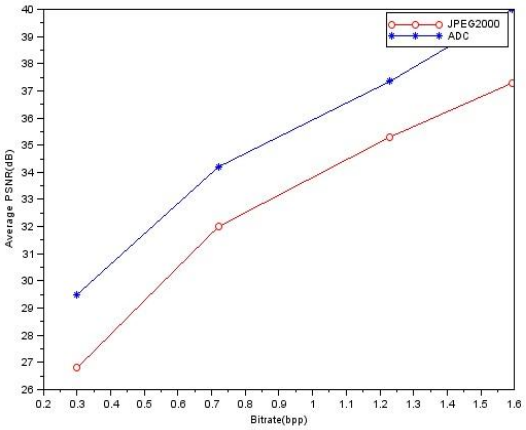
(a) Average Class Over 0



(b) Average Class Over 1



(c) Average Class over 2



(d) Average Class Over 3

Fig-5: Comparison of ADC against JPEG 2000 in terms of PSNR averaged for documents. (a) Class 0. (b) Class 1. (c) Class 2. (d) Class 3.

Distortion metrics based on visual models such as Structural Similarity (SSIM) and Video Quality Metric (VQM) has been used for test the pictorial content. Readability is very important in the compression process. Hence, the performance is analyzed by PSNR which is calculated using MSE. The quality between original and compressed image is measured by PSNR. The MSE represents the cumulative squared error between the compressed and the original image, whereas PSNR represents a measure of the peak error. The lower value of MSE represents the lower the error.

In JPEG2000 and ADC compression, the pages are separately encoded. Bitrates Vs. Average PSNR was plotted for evaluating the performance. Fig. 5 shows average PSNR improvement of ADC over JPEG2000 for each of the document class test sets. This explains that when bit rate increased the value of average PSNR also increased and the

ADC achieves high PSNR rate over JPEG2000 Standard. In all cases, ADC outperforms JPEG2000 by up to 2.7dB.

4. CONCLUSIONS

The proposed method ADC, which uses an H.264/AVC video coding standard for compress the scanned document. The reason why we chose the H.264/AVC tools to implement the proposed method is because intra and inter frame prediction scheme to reduce the storage space of the image data. In addition the integer transform and the CABAC contribute to improve the encoding efficiency. The proposed method obtained more significant gains than the existing system. Our work can be summarized as split the document into many pages, form the frames, and feed the frames to H.264/AVC. In that encoding process, intra and inter frame prediction, integer transformation, quantization and CABAC are performed. Then, PSNR was calculated using MSE. Then, in the decoding phase, inverse quantization and inverse transformation are done and the frames are reconstructed.

This project can be further enhanced for improving encoding efficiency and maximizing the PSNR rate by replacing the H.264/AVC by High Efficiency Video Codec standard (HEVC) which would yield even larger gains. The performance of the approach will also be improved further and the size of the compressed data will be reduced to a greater extend.

REFERENCES

[1] Information Technology—Coded Representation of Picture and Audio Information—Progressive Bi-Level Image Compression, ITU-T Standard T.82, Mar. 1993.

[2] Information Technology—Coded Representation of Picture and Audio —Lossy/Lossless Coding of Bi-level Images, ITU-T StandardT.88, Mar. 2000.

[3] Ricardo L. de Queiroz, Dec1998,'processing JPEG compressed images and Documents,' IEEE Trans. Image processing, Vol.7, no.12 pp.1661-1672.

[4] D. S. Taubman and M. W. Marcellin, 2002,'JPEG 2000: Image Compression Fundamentals, Standards and Practice'. Boston, MA, USA: Kluwer.

[5] E. G. Richardson, 2003, 'H.264 and MPEG-4 Video Compression'. New York, USA: Wiley.

[6] G. J. Sullivan, P. Topiwala, and A. Luthra, Aug. 2004 'The H.264/AVC advanced video coding standard: Overview and introduction to the fidelity range extensions,'Proc. SPIE, vol. 5558, pp. 53–74.

[7] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, Mar. 2004, 'Video coding with H.264/AVC: Tools, performance, and complexity,'IEEE Circuits Syst. Mag., vol. 4, no. 1, pp. 7–28.

[8] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, 'Overview of the H.264/AVC video coding standard,'IEEE Trans. Circuits Syst.

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

[10] H.S.Marver, A.Hallapuro, M.Karczewicz and L.kerofsky, july, 2003, 'Low Complexity Transform and Quantisation in H.264/AVC' IEEE Transactions on circuits and systems for video technology, vol. 13, No 7, pp. 560-576.

[11] http://image.unb.br/queiroz/test