ISSN 2348-375X

Unique Journal of Engineering and Advanced Sciences

Available online: www.ujconline.net

Research Article

HIGH QUALITY COMPRESSION OF SCANNED DOCUMENT USING ADVANCED VIDEO CODEC

Sasikala NP[1*], Anto Bennet M[2]

[1]Nandha Engineering College, PG scholar (VLSI Design), Erode-52, India
[2]Professor, Nandha Engineering College, Erode-52, India

Received: 20-02-2014; Revised: 18-03-2014; Accepted: 17-04-2014

*Corresponding Author: **N.P. Sasikala,**
Nandha Engineering College, PG scholar (VLSI Design), Erode-52, E-mail id: sasikalaperiasamy@gmail.com

ABSTRACT

The video compression is used to reduce the size of the video files. The MPEG files are compressed then the fine details of that video are lost their originality due to quantization. Lost details of video not appear after decompression. The better video processing algorithms and many techniques like pre-processing and low pass filters are used to reduce this compression error. The character recognition is used to find the data in written or printed document or any data source. In this character recognition data is scanned by a special scanner. Pattern matching algorithms is used in character recognition to find character that may be in any form. In the proposed coder scanned document are compressed using video codec method. The scanned page will decomposed in to N number of blocks. Then that will arrange in sequence to build video. Video sequence is compressed using advanced video codec (H.264/AVC). The pattern matching algorithm is used for reduce the data lost during the decompression. This method will give high quality of video compression compared with conversional methods. The performance evaluation will be shown by peak signal to noise ratio and similarity factor metric. Simulation results are shown by using MATLAB R2013a.

Keywords: Video Compression, Decompression, Video Processing, Algorithms, Scanner.

INTRODUCTION

Image scanner often abbreviated to just scanner is a device that optically scans images, printed text, handwriting, or an object, and converts it to a digital image. That digital data in a way that reduces the size of a computer file needed to store it, or the bandwidth needed to stream it. Compression may be lossy or lossless compression. Lossy compression transform coding is used. It will give high compression ratio. In this time domain image are converted to frequency domain so size are reduces but high frequency co-efficient are cut off. This high frequency co-effect has fine details of that digital information like corner of figure of letter. So this method cannot give exact image when recovering original image. In lossless compression different algorithm is used. It will give exact image when recovering original image but it gives low compression ratio. Proposed method is used to reduce the cut off error of co-efficient also improve that compression ratio.

EXISTING METHODS

In this proposed method scanned documents are arranged as sequence to built video and compressed. The JPEG compression algorithm is at its best on photographs and paintings of realistic scenes with smooth variations of tone and colour. For web usage, where the amount of data used for an image is important, JPEG is very popular[1]. On the other hand, JPEG may not be as well suited for line drawings and other textual or iconic graphics, where the sharp contrasts between adjacent pixels can cause noticeable artefacts[2,3]. JPEG 2000 is not only improving compression performance over JPEG but also adding (or improving) features. Higher-resolution images tend to benefit more, where JPEG-2000's spatial-redundancy prediction can contribute more to the compression process[4]. The first MPEG-1 compression standard for audio and video. It was basically designed to allow moving pictures and sound to be encoded into the bit rate . MPEG-2 is considered important because it has been chosen as the compression scheme for over-the-air. MPEG-3 dealt with standardizing scalable and multi-resolution compression[5].

Other MPEG standards are mainly used for multimedia application. In H.261 design the basic processing unit of the design is called a macro block, and H.261 was the first standard in which the macro block concept appeared. Each macro block consists of a 16x16 array of luma samples and two corresponding 8x8 arrays of chroma samples. The coding algorithm uses a hybrid of motion compensated inter-picture prediction and spatial transform coding with scalar

quantization[6].The H.261 standard actually only specifies how to decode the video. Encoder designers were left free to design their own encoding algorithms, as long as their output was constrained properly to allow it to be decoded by any decoder made according to the standard[7]. As H.264 provides a significant improvement in capability beyond H.263, the H.263 standard is now considered a legacy design. Most new videoconferencing products now include H.264 as well as H.263, MPEG-2 and H.261 capabilities[8-11].

The H.264 video format has a very broad application range that covers all forms of digital compressed video from low bit-rate Internet streaming applications to HDTV broadcast and Digital Cinema applications with nearly lossless coding. With the use of H.264, bit rate savings of 50% or more are reported[9]. The ability to use multiple motion vectors per macro block[10] with a maximum of 32. Its main purpose is to give examples of H.264/AVC features, rather than being a useful application. Some reference hardware design work is also under way in the Moving Picture Experts Group. The above mentioned are complete features of H.264/AVC covering all profiles of H.264. A profile for a codec is a set of features of that codec identified to meet a certain set of specifications of intended applications. This means that many of the features listed are not supported in some profiles[12].

Pattern matching is the act of checking a perceived sequence of block for the presence of the constituents of some pattern. In contrast to pattern recognition, the match usually has to be exact. The patterns generally have the form of either sequences or tree structures. Uses of pattern matching include outputting the locations of a pattern within a block sequence, to output some component of the matched pattern, and to substitute the matching pattern with some other token sequence[13-15]. PSNR is most commonly used to measure the quality of reconstruction of lossy compression codec's. The signal in this case is the original data, and the noise is the error introduced by compression. When comparing compression codec's, PSNR is an approximation to human perception of reconstruction quality. Although a higher PSNR generally indicates that the reconstruction is of higher quality, in some cases it may not. One has to be extremely careful with the range of validity of this metric; it is only conclusively valid when it is used to compare results from the same codec and same content[17].

PROPOSED METHOD

The scanned page will be compressed using H.264/AVC, the proposed encoding method organizes the scanned pages in such a way the inter frame prediction may find on previously encoded macro blocks (16 × 16 pixels blocks) text patterns that are similar to those on the macro block currently being encoded. Figure 1 illustrates the proposed page processing algorithm. First, each scanned H ×W pixels page is segmented into four H/2 × W/2 pixels sub-pages. As shown in Figure 1.

Then, these sub-pages are used to build a video sequence. The only reason page segmentation should be used is that in some cases similar text patterns are more likely to be found on the same page rather than on different pages. If the text style is constant throughout the whole book, each page may be converted into one single frame and segmentation may be skipped. The final step is to compress the resulting video using H.264/AVC

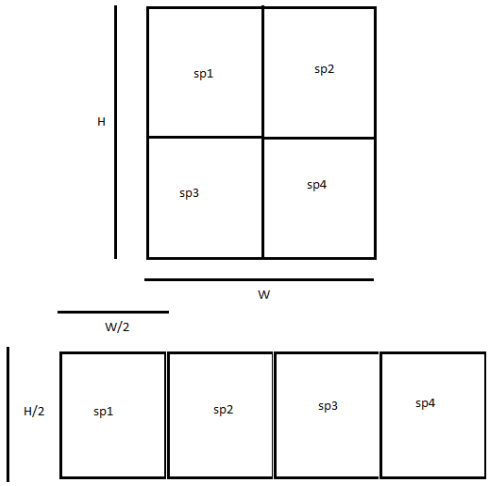

Figure: 1

The basic idea of the inter frame prediction is to exploit similarities between video frames in order to reduce the amount of information to be encoded. Based on previously encoded blocks, it first constructs a prediction of the current frame and then creates a residual frame by subtracting the prediction from the current frame. In H.264/AVC, the luma component of each current macro block is predicted as one 16 × 16 partition, two 16 × 8, two 8 × 16 or four 8 × 8 macro block partitions. In case partitions with 8 × 8 pixels are chosen, the 8 × 8 sub-macro blocks may be further partitioned in one 8 × 8 partition, two 8 × 4, two 4 × 8 or four 4×4 sub-macro block partitions.

The prediction of each luma block is constructed by displacing an area of the reference frame, determined by a motion vector and a reference frame index. That previously encoded text areas (reference frames) can be seen as a dictionary used by the pattern matching (inter frame prediction) algorithm. The dictionary is updated in parallel with the encoding process, since new reference frames become constantly available.

Furthermore, a rate-distortion optimization algorithm is used to estimate which intra/inter modes combination should be applied. Once the residual data is available, H.264/AVC uses an integer transform with similar properties as the DCT (Discrete Cosine Transform) and the resulting transformed coefficients are quantized and entropic ally encoded. Then PSNR value is calculated frame by frame.

RESULTS

First step in this method is dividing one page into four equal size frames and arrange in sequence to a video. And applying pattern matching algorithm within the frames. Taking one frame as reference frame. This result is shown in above Figure 2,Figure 3. Then frame by AVC is applied for compression to

transmit or store. When recovering image the reverse presses is down as shown in figure 4.

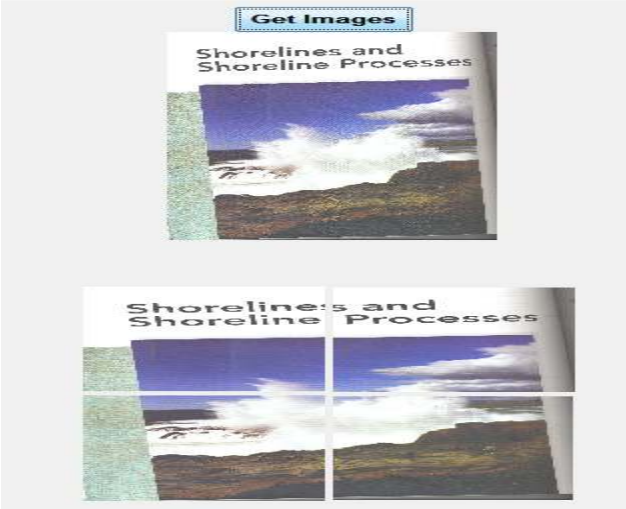

Figure 2: Sub Dividing Frame

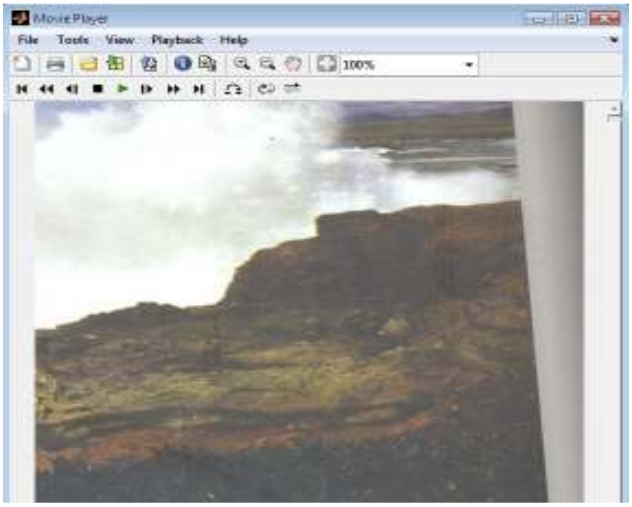

Figure 3: Video of Frames

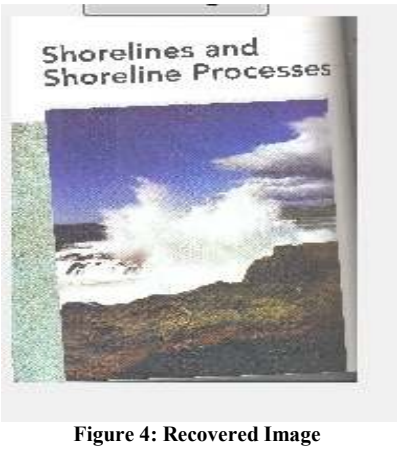

Figure 4: Recovered Image

Then applying following PSNR formula, PSNR value for each frame was calculated. In this method we get the PSNR value more than 35db. This value will vary based on the scanned pages. Because, that page may have only text, only image, or both text and image.

$$PSNR = 10.\log_{10} (MAX_I^2 / MSE)$$

In this method one image will diverted into four sub image then that frame only used for making video. The PSNR for one image that is four is given in TABLE 1 and graph FIGURE 5. That shows that PSNR was increased frame to frame.

Table 1: PSNR value for one page

Frame No	PSNR in DB
1	36.1356
2	37.3498
3	38.194
4	38.5109

PSNR is use to find the compression efficiency. This may vary based on type of page that was scanned. PANR value is high for image that have only image. If the pages have text and image then PSNR value is reduced then pure text. Pages have maximum image than text then PSNR again reduced. But these methods have high PSNR if pure image used. This compression is shown in Table 2 and Figure 6

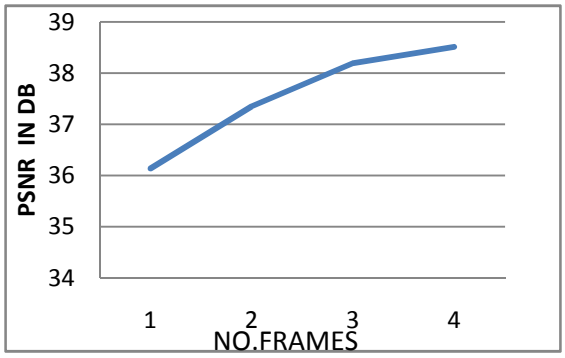

Figure 5: Graph of PSNR for one page

Table 2: PSNR for different pages

No of frame	PSNR in DB			
	Only Text	Text With Image	Only Image	Photo
1	36.1356	34.996	35.4217	38.5296
2	37.3498	36.6599	35.7963	38.6063
3	38.194	37.4989	35.1866	39.1491
4	38.5109	37.6537	36.9768	40.0573

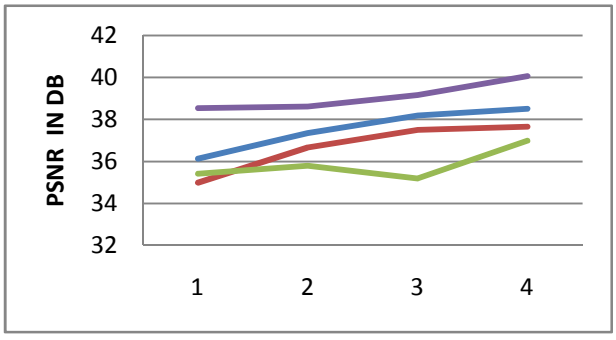

Figure 6: Graph of PSNR for different pages

Then compression ratio was calculated by applying following formula

COMPRESSION RATIO = size of uncompressed image/ size of compressed image

This will change from 9 to 30 for different type of image.

CONCLUSION

The H.264/AVC may be lossy compression if use only this compression fine information may be lost while recovering the original image. The pattern matching is lossless compression. It was used to reduce that type of loss by reducing coefficient length. So this method complains lossy and lossless compression. This gives high compression ratio like lossy compression with minimum loss in data. Finally this method was implemented by splitting scanned page and arranged one by one like video and applying video compression. Results show that the proposed method objectively outperforms of other compression by up to 4 dB to 5 dB. And PSNR valve was calculated for different type of frames, those results shows that PSNR value is depends on the type of scanned page, quality of the scanner and color of that page.

REFERENCES

1. Alexandre Z, Ricardo DQ. Scanned Document Compression Using Block-Based Hybrid Video Codec on image processing. 2013; 22(6).
2. Halbach, Bauermann I, Steinbacj E. Performance Further Lossless Compression of JPEG Images. Proc. Of picture coding symposium, san francisco, usa, December. 2004; 15.
3. Ponomarenko, Astola J. Additional Lossless Compression of JPEG Images, Processing Of the 4th intl. Symposium on image and signal processing and analysis, zagreb, Croatia. 2005;117.
4. Stirner M, Seelmanng. Improved Redundancy Reduction For JPEG Files. Proc. of picture coding symposium, lisbon, Portugal. 2007; 7.
5. Comparison: H.26L Intra Coding vs. JPEG2000 MPEG standards – full list of standards developed or under development. Chiariglione.org. Retrieved. 2009; 10.
6. Ahmed N, Natarajan T, Rao KR. Discrete Cosine Transform, IEEE transactions on computers. 2004; 90.
7. Itu t. H.263: Video Coding For Low Bit Rate Communication.
8. Wiegand, Thomas. Overview of theH.264/AVC Video Coding StandardIEEE transactions on circuits and systems for video technology. 2011;13(7).
9. Ostermann J, Wedi T. Video Coding With H.264/AVC: Tools, Performance, and Complexity. IEEE circuits and systems magazine. 2011;4(1).
10. Sullivan, Gary J, Wiegand, Thomas. Video Compression-From Concepts to the H.264/AVC Standard. Proceedings of the IEEE. 2011.
11. Richardson, Iain EG. Learn About Video Compression and H.264.VCodex limited. Retrieved. 2011.
12. Topiwala. Overview and Introduction to the Fidelity Range Extensions SPIE applications of digital image processing xxvii. Retrieved. 2004.
13. Atallah. Pattern Matching Image Compression: Algorithmic and Empirical Results. Computer science technical reports. 1995.
14. Michael T, OrletskyW. Approximate Geometric Pattern Matchingunder Rigid Motions IEEE transactions on pattern analysis and machine intelligence. 2009; 21(4): 371.
15. Dinesh N, Lothar W. Image Processing And Low-Discrepancy Sequences advanced signalprocessing algorithms, architectures, and implementations ix, proceedings of spie. 2009; 3807: 102.
16. Thomos N, Boulgouris NV, Strintzis MG. Optimized Transmission Of JPEG2000 Streams Over Wireless Channels. IEEE transactions on image processing. 2006.
17. Xiangjun L, Jianfei C. Robust Transmission of Jpeg2000 Encoded Images Over Packet Loss Channels. Icme. School of computer engineering, nanyang technological university. 2007.

Source of support: Nil, Conflict of interest: None Declared