Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 6, Issue, 03, pp.5937-5941, March, 2014

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

EFFECTIVE RECONSTRUCTION OF COMPRESSED SCANNED DOCUMENT USING ADVANCED VIDEO CODEC

***Sasikala, N. P. and Dr. Anto Bennet, M.**

Professor/ECE, Nandha Engineering College, Erode-52

ARTICLE INFO

Article History:
Received 20th December, 2013
Received in revised form
15th January, 2014
Accepted 18th February, 2014
Published online 31st March, 2014

Key words:

Advanced document coding (ADC),
H.264/AVC,
Pattern matching,
Scanned document compression.

ABSTRACT

The video compression is used to reduce the size of the video files. The MPEG files are compressed then the fine details of 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 the 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 to reduce the data lost during the decompression. This method will give high quality of video compression compared with conventional methods. The performance and evaluation will be shown by Peak Signal to Noise Ratio (PSNR) and Structural Similarity Metric (SSM). Simulation results are shown by using MATLAB R2013a.

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. 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. In lossy compression, transform coding is used to give high compression ratio. In this time domain image are converted to frequency domain so image sizes were 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. To exact images 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.

Related work

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. (Alexandre Zaghetto 2013) 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. (Halbach *et al.*, 2002; Ponomarenko and let to Astola 2005). 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 (Stirner and Seelmanng 2007). The first MPEG-1 compression standard for audio and video 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 (Comparison: H.26L Intra Coding vs. JPEG2000 MPEG standards – full list of standards developed or under development).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 (Ahmed *et al.*, 2004). 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

****Corresponding author: Sasikala, N. P.***
Professor/ECE, Nandha Engineering College, Erode-52, Tamilnadu, India.

properly to allow it to be decoded by any decoder made according to the standard. 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 (Itu ?; Wiegand 2011).

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 (Ostermann 2011). The ability to use multiple motion vectors per macro block with a maximum of 32 (Sullivan *et al.*, 2011). 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 Richardson 2011; Topiwala 2004).

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 (Atallah 1995; Michael 2009; Dinesh nair and lothar wenzel 2009). 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 (Thomos *et al.*, 2006; Xiangjun and Jianfei 2000)

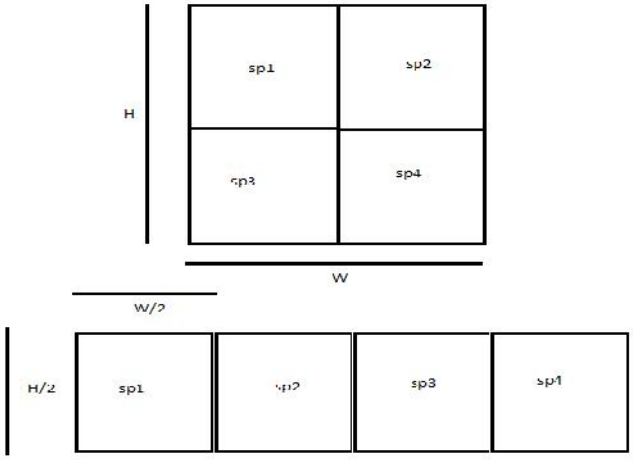

Figure 1.

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. In Figure 1 the proposed page processing algorithm. First, each scanned H ×W pixels page is segmented into four H/2 × W/2 pixels sub-pages.

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. 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 entropically encoded. The PSNR value is calculated frame by frame. True positive is the values that tell how much pixel are correctly identified as loss less pixels. Likewise true negative is tells about wrongly identified lossless pixels. False positive is tells about correctly identified lossy pixels. False negative is the value which tells about wrongly identified lossy pixels. By comparing value with pattern matching the parameters True Positive (TP), True negative (TN), False Positive (FP), False negative (FN) are applied the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy are applied to.

SENSITIVITY= TP / (TP+FN) (1)
SPECIFICITY= TN / (TN+FP) (2)
POSITIVE PREDICTIVE VALUE= TP / (TP+FP) (3)
NEGATIVE PREDICTIVE VALUE = TN/(TN+FN) (4)
ACCURACY= (TP+TN) / (TP+FN+TN+FP) (5)

Experimental 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. The result is shown in Figure 2 and Figure 3.

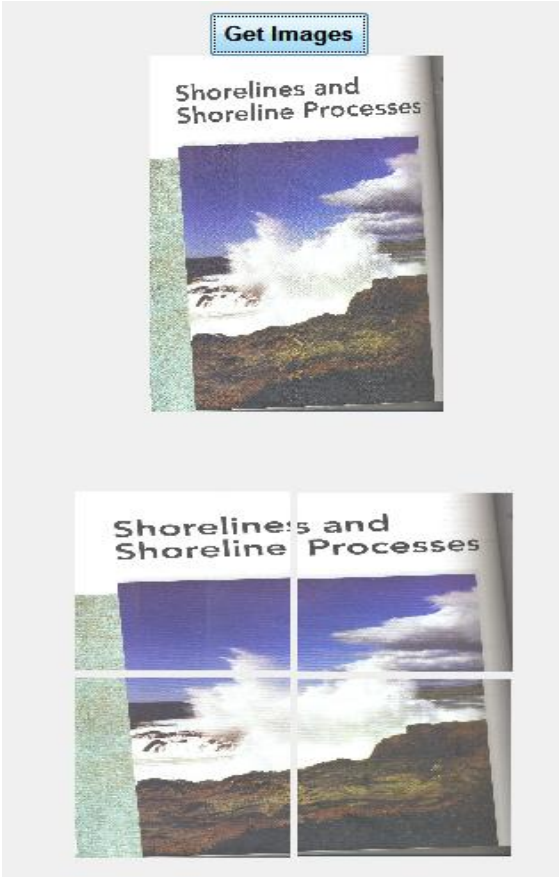

Figure 2. Sub dividing frame

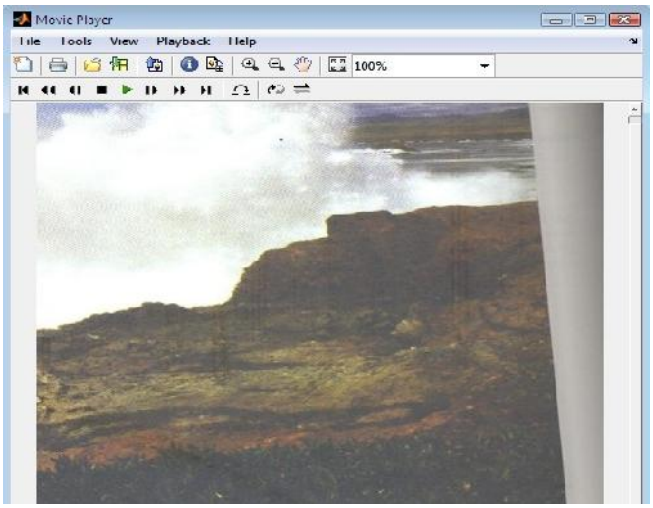

Figure 3. Video of frames

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.

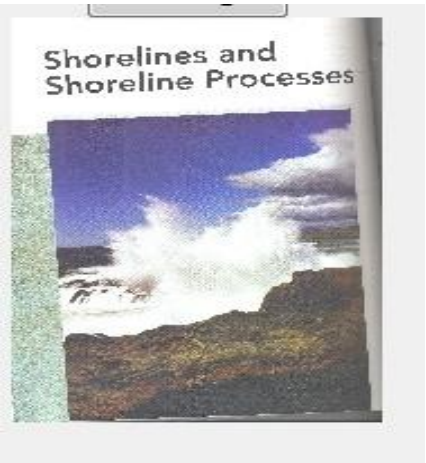

Figure 4. Recovered image

By applying the following PSNR value for each frame were calculated.

$$PSNR = 10.\log_{10}(MAX_I^2 / MSE) \quad (6)$$

PSNR =44.5264

This method have one image will diverted into four sub image and frame only used for making video. The PSNR for one image is four is given in TABLE 1. Improvement PSNR was increased frame by frame and that graphical response was shown in figure 5.

Table 1. PSNR value for one page

FRAME NO	PSNR IN dB
1	36.1356
2	37.3498
3	38.194
4	38.5109

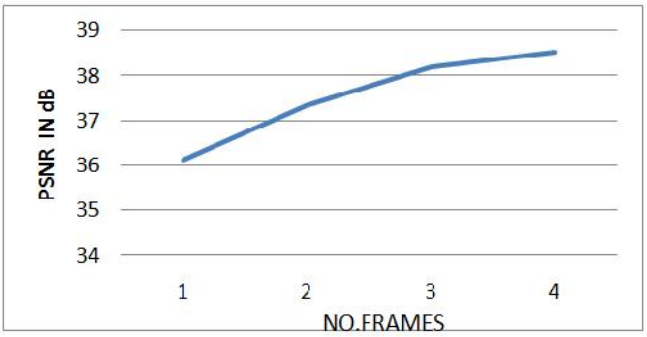

Figure 5. PSNR for one page

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

Table 2. PSNR for different pages

No of frame	PSNR IN dB			
	ONLY TEXT	TEXT WITH IMAGE	ONLY INAGE	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

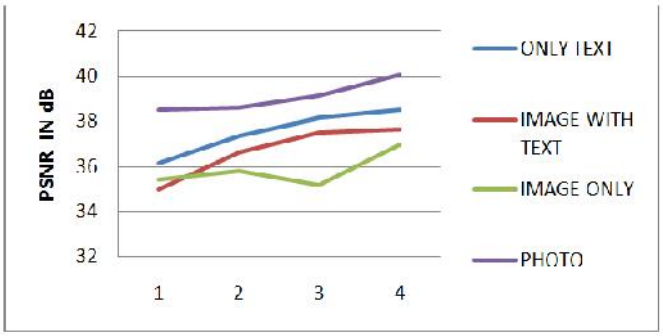

Figure 6. PSNR for different pages

COMPRESSION RATIO = size of uncompressed image/ size of compressed image (7)

Equation 6 was used to find the compression ratio. Compression ratio will change from 9 to 30 for different type of image and compared. By applying True Positive, true negative False Positive, false negative in equation 1 -5 parameters of compression were calculated. That response was shown graphically in Figure 7. Sensitivity and specificity are statistical measures of the performance of a binary classification test, also known in statistics as classification function. Sensitivity also called the true positive rate, or the recall rate in some fields measures the proportion of actual positives which are correctly identified as such. Specificity sometimes called the true negative rate measures the proportion of negatives which are correctly identified as such. The sensitivity and specificity of this proposed method were shown in graphically in figure 8 and figure 9. The positive and negative predictive values are the values used to find the wither our method of compression reach particular value PSNR (here 40dB stranded was taken as value). The graphical responses of positive and negative predictive values were shown in figure 10 and Figure 11. And the Figure 12 shows the graphical response of accuracy.

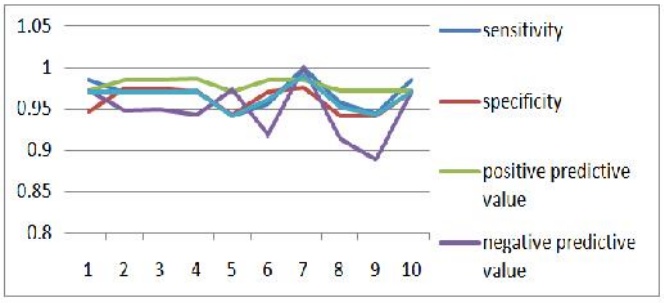

Figure 7. Performance of Compression with Different Parameters

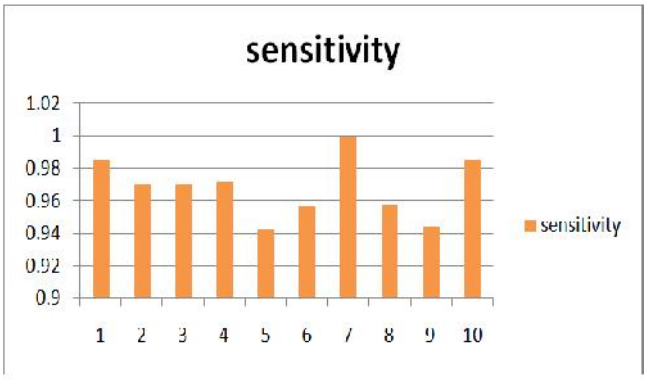

Figure 8. Sensitivity

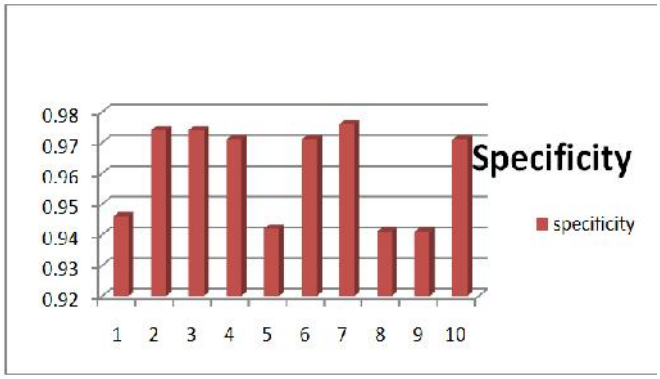

Figure 9. Graph of Specificity

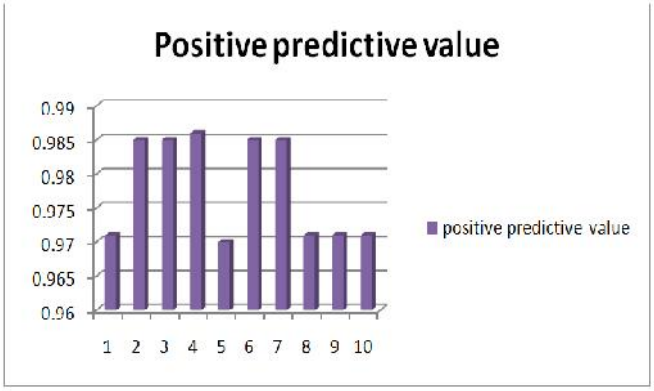

Figure 10. Graph of Positive Predictive Value

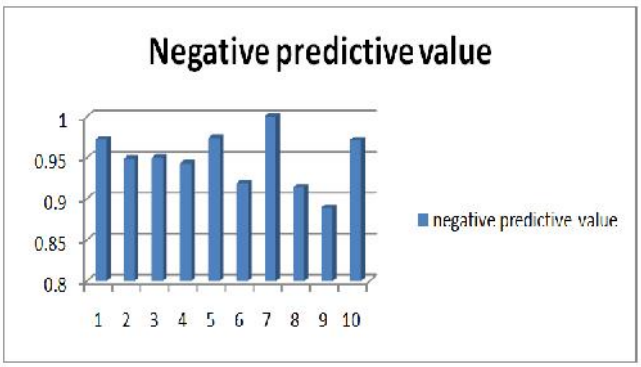

Figure 11. Graph of Negative Predictive Value

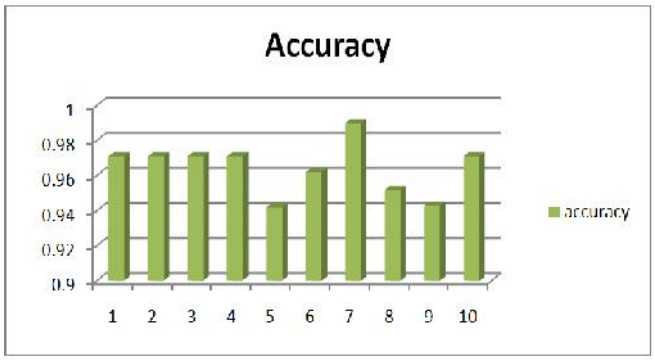

Figure 12. Graph of Accuracy

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

Alexandre Zaghetto, member, IEEE, and Ricardo De Queiroz, senior member, IEEE, transactions “Scanned Document Compression Using Block-Based Hybrid Video Codec” on image processing, vol. 22, no. 6, June 2013.

Ponomarenko.n, let to Astola.J.” Additional Lossless Compression of JPEG Images, Processing”. Of the 4th intl. Symposium on image and signal processing and analysis, zagreb, croatia, pp.117–120, september 15–17, 2005.

Stirner. M and Seelmanng. “Improved Redundancy Reduction For JPEG Files”. Proc. Of picture coding symposium, lisbon, portugal, november 7–9, 2007.

“Comparison: H.26L Intra Coding vs. JPEG2000 MPEG standards – full list of standards developed or under development”. Chiariglione.org. Retrieved 2009-10-31.

Ahmed.N, Natarajan.T And Rao. K. R., "Discrete Cosine Transform", IEEE transactions on computers, jan. 2004, pp. 90-93.

Itu-t. "H.263. Video Coding For Low Bit Rate Communication"

Wiegand, Thomas let to ajay "Overview Of The H.264/AVC Video Coding Standard" IEEE transactions on circuits and systems for video technology 13 (7). Retrieved january 31, 2011.

Ostermann. J let to Wedi. T. "Video Coding With H.264/AVC: Tools, Performance, And Complexity". IEEE circuits and systems magazine 4 (1). January 31, 2011.

Sullivan, Gary J, Wiegand, Thomas . "Video Compression—From Concepts to the H.264/AVC Standard”. Proceedings of the IEEE . January 31, 2011.

Richardson, Iain E.G. "Learn About Video Compression and H.264".VCodex limited. Retrieved January 31, 2011.

Atallah, let to Wojciech, "Pattern Matching Image Compression: Algorithmic And Empirical Results" (1995). Computer science technical reports. Paper 1255.

Michael T let to Orletsky.W, “Approximate Geometric Pattern Matchingunder Rigid Motions”, IEEE transactions on pattern analysis and machine intelligence, vol. 21, no. 4, pp. 371-379, April 2009.

Dinesh nair and lothar wenzel, "Image Processing And Low-Discrepancy Sequences", advanced signalprocessing algorithms, architectures, and implementations ix, proceedings of spie, volume 3807, pp. 102-111, july2009

Xiangjun, L., and Jianfei, C. “Robust Transmission Of Jpeg2000 Encoded Images Over Packet Loss Channels”. Icme 2007 (pp. 947-950). School of computer engineering, nanyang technological university.

Halbach, Bauermann.I and Steinbacj.E. till (july 2002). "performance Further Lossless Compression Of JPEG Images”. Proc. Of picture coding symposium (pcs 2004), san francisco, usa, december 15–17, 2004.

Topiwala, let to luthra, (august 2004). "Overview And Introduction To The Fidelity Range Extensions" SPIE applications of digital image processing xxvii. Retrieved January 31, 2011.

Thomos, N., Boulgouris, N. V., & Strintzis, M. G. (2006, january). “Optimized Transmission Of JPEG2000 Streams Over Wireless Channels”. IEEE transactions on image processing.

****