PACKET LOSS CONCEALMENT IN BASELINE JPEG CODED IMAGES

Shahram Shirani, Faouzi Kossentini, and Rabab Ward

Department of Electrical and Computer Engineering,
University of British Columbia,
Vancouver BC V6T 1Z4, Canada

ABSTRACT

Loss of coded data can affect a JPEG decoded image to a large extent, making concealment of errors caused by data loss an important issue. Reconstruction of JPEG coded images in an error-prune channel environment is investigated in this paper. First a method for estimating the missing DC coefficients of a JPEG coded image which is required for decoding the compressed image, is suggested and evaluated. As effects of errors in estimating the missing DC value will appear as stripes across the image, a post-processing technique for removing such stripes is then developed. Finally the missing data is reconstructed by exploiting the correlation between adjacent blocks. A novel reconstruction technique which has a good performance in reconstruction of edges is proposed. A key contribution of our work is that, unlike in previously published reconstruction algorithms, differential encoding of the DC coefficients is assumed. Simulation results indicate that the performance of our algorithm is very good, even when many packets are lost during transmission of the JPEG coded image.

1. INTRODUCTION

The JPEG (Joint Photographic Expert Group) standard is intended to compress still, continuous tone, monochrome and color images. In the sequential mode of JPEG, the input source image is partitioned into 8×8 blocks along the conventional raster scan order (i.e. left to right / top to bottom) for the 2-D DCT operation. Each of the 64 resulting DCT coefficients is quantized using a different uniform quantizer. The DC coefficient is treated separately from the 63 AC coefficients. It is differentially coded, that is, the difference coefficient DIFF given by

$$DIFF = d_i - d_{i-1},$$

where d_i and d_{i-1} are the current block and the previous block DC coefficient values (respectively), is coded. The quantized 63 AC coefficients are formatted in a zigzag scan in preparation for variable-length coding. Compression is achieved by utilization of both quantization and entropy coding [1].

In this work, we will address the problem of packet based transmission of JPEG coded images over a lossy network. It is assumed that the compressed image data is packetized spatially. That is, each packet contains an integer number of entire 8×8 blocks of DCT coefficients. It is also assumed that the decoder knows which blocks belong to each packet, specifying uniquely the location of missing blocks. Because of differential encoding of DC coefficients, loss of a block affects all the blocks after it. To the best of our knowledge, in previously published work in this area, it has always been assumed that the DC coefficients of 8×8 blocks are coded non-differentially, thus constraining the impact of a block loss to only the subject block [2, 3]. Such an assumption leads to a coder that is not compliant with JPEG [4]. As a key contribution of this work, we here extend the reconstruction method to baseline JPEG coded images. The structure of this paper is as follows. In Section 2, packetization of JPEG coded images is discussed. In Section 3, estimation of the DC values of lost blocks is presented. Section 4 describes our post-processing technique. Section 5 discusses the details of our final reconstruction method. Sections 6 and 7 present the experimental results and the conclusions respectively.

This work was supported by MDSI, BC ASI and NSERC.

2. PACKETIZATION OF JPEG CODED IMAGES

Suppose that a JPEG coded image is packetized spatially as mentioned in the previous section. While transmitting the JPEG coded images over lossy packet networks, it is possible that one or more of the packets will not reach the destination. This type of loss of coded data can affect the decoded image to a large extent, making correction of errors caused by data loss an important issue. Fortunately, redundancies in JPEG coded images can be exploited to improve the quality of an image corrupted by packet loss. Loss of blocks (due to a packet loss) causes two problems at the decoder. First, because the DC coefficient of the 8×8 DCT transform is differentially coded, the decoder is unable to decode the DC coefficient of the blocks which follow the missing one. Thus it is necessary to estimate the missing DC coefficient by a proper method. Second, the data of missing blocks should be reconstructed. The DC coefficient of the missing block is estimated using the DC coefficients of adjacent blocks. The data of adjacent blocks are usually highly correlated, with correlation decreasing as the distance between the blocks increases. Thus, the best blocks for estimating the DC value of the missing block are those adjacent to it. To avoid large contiguous missing blocks in the image, and to facilitate the estimation of lost DC coefficients, it is better to assign adjacent blocks to different packets (i.e. the blocks should be interleaved such that blocks in the same packet are not spatially close to each other). In this work it is assumed that the blocks are interleaved according to the algorithm suggested in [5]. This algorithm maximizes the minimum distance between the blocks in the same packet.

3. ESTIMATION OF DC VALUES

As mentioned previously, the decoder knows the interleaving pattern and therefore the position of missing data blocks. It starts de-

0-7803-4957-1/98/$10.00©1998 IEEE

coding the data, and when it reaches an area representing a missing block, it estimates the DC values of the missing block using the DC value of adjacent blocks. The DC value of the missing block is estimated by a linear, causal estimator $\hat{d}$ given by

$$\hat{d} = w_1 d_1 + w_2 d_2 + w_3 d_3 + w_4 d_4, \quad (1)$$

where d_1, d_2, d_3, d_4 are the DC values of the previously coded blocks shown in Figure 1. The coefficients of the estimator are obtained by using well known MMSE (minimum mean square error) techniques [6, 7]. Based on a large training set of images, such values should approximately be $w_1 = 0.1$, $w_2 = 0.4$, $w_3 = 0.1$, $w_4 = 0.4$.

4. POST-PROCESSING OF STRIPES

Error in estimating the DC value of a missing block causes a level shift in all the pixels between that block and the next missing block. A we will see in Section 6 this level shift appears as a stripe. To remove the stripe, the amount of the level shift should be determined. Adjacent pixels in an image are highly correlated and their values are close to each other. Therefore the difference between the average of pixel values in the stripe and their corresponding pixels outside the stripe is an estimate of the level shift. To use the pixels that are as close as possible to each other, the average of pixel values along the upper line segment connecting two consecutive missing blocks is compared to the corresponding line segment in the previous row, as shown in Figure 2. Therefore, using the a priori information about the position of missing blocks, we compute

$$\Delta_i = l_{i,i+1} - m_{i,i+1}, \quad (2)$$

where $l_{i,i+1}$ is the average value of pixels along the upper line segment connecting missing blocks i and $i+1$ and $m_{i,i+1}$ is the average value of pixels in the corresponding line segment in the previous row (See Figure 2). If $|\Delta_i| > T$, where T is a threshold, a stripe is detected and removed by subtracting the Δ_i from all the pixels between missing blocks i and i+1. The procedure is repeated for all the missing blocks, $1 < i < M$, where M is the number of missing blocks. To maintain the probability of false alarm [8](i.e. there is no stripe but the algorithm detects one) to less than 1%, a threshold value of 4 was selected. The procedure of selecting the threshold value is explained in detail in [5].

5. RECONSTRUCTION OF MISSING DATA

As a final stage, the missing data should be reconstructed. Our method for estimating the missing data is based on the technique suggested in [2]. The method proposed in [2] estimates the data of a missing block as a linear combination of its top, bottom, right and left blocks, i.e., the estimate matrix $\hat{\mathbf{p}}_Z$ is given by

$$\hat{\mathbf{p}}_Z = w_T \mathbf{p}_T + w_B \mathbf{p}_B + w_L \mathbf{p}_L + w_R \mathbf{p}_R \quad (3)$$

where $\mathbf{p}_T$,$\mathbf{p}_B$,$\mathbf{p}_L$, and $\mathbf{p}_R$ are 8×8 matrices of pixel values on top, bottom, left and right of the missing block. The weights w_T,w_B,w_L,w_R are selected so that the total squared edge error

$$\epsilon^2 = \epsilon_T^2 + \epsilon_B^2 + \epsilon_L^2 + \epsilon_R^2 \quad (4)$$

is minimized. Each edge error is defined in terms of pixels by

$$\begin{aligned} \epsilon_T^2 &= \|(\hat{\mathbf{p}}_{Zt} - \mathbf{p}_{Tb})\|^2 , \\ \epsilon_B^2 &= \|(\hat{\mathbf{p}}_{Zb} - \mathbf{p}_{Bt})\|^2 , \\ \epsilon_L^2 &= \|(\hat{\mathbf{p}}_{Zl} - \mathbf{p}_{Lr})\|^2 , \\ \epsilon_R^2 &= \|(\hat{\mathbf{p}}_{Zr} - \mathbf{p}_{Rl})\|^2 , \end{aligned} \quad (5)$$

where the matrix $\hat{\mathbf{p}}_Z$ is generated using (5). The parameter vectors $\hat{\mathbf{p}}_{Zt}$, $\hat{\mathbf{p}}_{Zb}$, $\hat{\mathbf{p}}_{Zl}$ and $\hat{\mathbf{p}}_{Zr}$ are those elements of $\hat{\mathbf{p}}_Z$ that correspond to the top, bottom, left and right pixels in the missing block respectively. Similarly, the vector $\mathbf{p}_{Tb}$ consists of the bottom line pixels of the top block, $\mathbf{p}_{Bt}$ consists of the top line pixels of the bottom block, $\mathbf{p}_{Lr}$ consists of the right line pixels of the left block and $\mathbf{p}_{Rl}$ consists of the left pixels of the right block. This technique cannot reconstruct strong diagonal edges. The reason is that a diagonal edge cannot be constructed as a linear combination of shifted diagonal edges [2]. What yields a poor reproduction quality around the edges is the fact that only one weight for each of the top, below, right, and left blocks is used. In this work, these blocks are partitioned into two parts, and a different weight is assigned to each part. Thus the estimation relation can be modified to

$$\begin{aligned} \hat{\mathbf{p}}_Z = & \begin{bmatrix} w_{T1}\mathbf{I}_{4\times4} & \mathbf{0} \\ \mathbf{0} & w_{T2}\mathbf{I}_{4\times4} \end{bmatrix} \mathbf{p}_T + \\ & \begin{bmatrix} w_{B1}\mathbf{I}_{4\times4} & \mathbf{0} \\ \mathbf{0} & w_{B2}\mathbf{I}_{4\times4} \end{bmatrix} \mathbf{p}_B \\ & + \mathbf{p}_L \begin{bmatrix} w_{L1}\mathbf{I}_{4\times4} & \mathbf{0} \\ \mathbf{0} & w_{L2}\mathbf{I}_{4\times4} \end{bmatrix} \\ & + \mathbf{p}_R \begin{bmatrix} w_{R1}\mathbf{I}_{4\times4} & \mathbf{0} \\ \mathbf{0} & w_{R2}\mathbf{I}_{4\times4} \end{bmatrix}, \end{aligned} \quad (6)$$

where $\mathbf{I}_{4\times4}$ is the 4×4 identity matrix, $\mathbf{0}$ is the 4×4 zero matrix and $w_{T1}, w_{T2}, w_{B1}, w_{B2}, w_{L1}, w_{L2}, w_{R1}$, and w_{R2} are the weighting coefficients. Values for the latter coefficients are selected by minimizing the cost function described in (4) and (5).

6. EXPERIMENTAL RESULTS

Figure 3 shows the 512×512, 8 bits/pixel, image LENA coded and decoded using baseline JPEG. The size of compressed image is about 21 k Byte. It is assumed that the coded image data is packetized into 64 packets each containing 64 blocks [5]. Each packet is about $\frac{1}{3}$ k Byte. Figure 4 shows the same image decoded when two randomly selected packets are missing. The distance between blocks in the same packet is maximized using the algorithm proposed in [5], and the missing DC value required for decoding succeeding blocks is estimated using equation (1). The stripes appearing in Figure 4 are due to error in estimating the DC value of missing blocks. Figure 5 shows the image of Figure 4 after applying the proposed post-processing algorithm for removing the stripes. Clearly most of the error effects in estimating the DC values are removed. Figure 6 shows the reconstructed image using the method proposed in [2]. In Figure 7, where a magnified picture of the shoulder area of the image LENA is shown, the poor performance of the above algorithm is apparent. Figure 8 shows the image LENA after reconstruction using our proposed algorithm. Figure 9 shows a magnified picture of the same area shown in Figure 7. By comparing Figures 7 and 9, it is clear that our algorithm improves the reconstruction of edges. Comparing Figure 4, which shows the image LENA missing about 3% of compressed image data packets before applying the algorithm, to Figure 8, which shows the same image after removing the stripes and reconstructing the missing data, one can see that all visible artifacts have been removed. To quantify (objectively) the amount of improvement, the PSNR measure is also used. The PSNRs for the image LENA are 25.66 before and 34.85 after overall post-processing. The results shown here, as well as results obtained for other test images, demonstrate that the proposed algorithm perform very well, both subjectively and objectively.

7. CONCLUSIONS

Reconstruction of missing data in packetized transmission of sequential JPEG coded images was investigated. The DC values of the missing blocks which are necessary for decoding of the image are estimated using the DC value of their adjacent blocks. An algorithm for post-processing of the decoded image to remove the effects of the error in estimation of the DC values of the missing blocks on the rest of the image was proposed. Finally, to estimate the missing data, a reconstruction algorithm that has a good performance in reconstruction of edges was suggested. A key contribution of our reconstruction algorithm is that, unlike previously published algorithms, differential encoding of the DC values of the DCT coefficients, which is part of the JPEG standard, is assumed. The performance of the proposed reconstruction algorithm was tested using several images. The quality of the reconstructed images is relatively good both subjectively and objectively.

8. REFERENCES

[1] K. R. Rao and J. J. Hwang, *Techniques and Standards for Image, Video and Audio Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[2] S. S. Hemami and T. H. Y. Meng, "Transform coded image reconstruction exploiting interblock correlations," *IEEE Transactions on Image processing*, vol. 4, pp. 1023–1027, July. 1995.

[3] Y. Wang and Q. F. Zhu, "Signal loss recovery in DCT-based image and video codecs," *SPIE Conf. Visual Commun. Image Processing*, pp. 667–678, Nov. 1991.

[4] ISO/IEC, *Digital compression and coding of continuous tone still images: requirements and guidlines*. International Standard Organization, 1994.

[5] S. Shirani, F. Kossentini, S. Kallel, and R. Ward, "Reconstruction of JPEG coded images in lossy packet networks," *submitted to IEEE Transactions on Communications*, 1997.

[6] A. K. Jain, *Fundamentals of Digital Image Processing*. Englewood Cliffs, New Jersey: Prentice-Hall, 1989.

[7] S. M. Kay, *Fundamentals of Statistical Signal Processing*. Englewood Cliffs, New Jersey: Prentice-Hall, 1993.

[8] H. L. Van-Trees, *Detection, Estimation, and Modulation Theory*. USA: John Wiley and Sons, Inc., 1968.

d_1	d_2	d_3
d_4	$\hat{d}$	

Figure 1: Estimation of DC value.

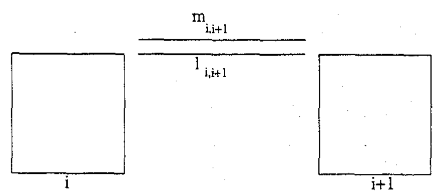

Figure 2: Post-processing for removing the stripes.

Figure 3: The original image LENA.

Figure 4: Decoded image with two packets missing.

Figure 5: The image LENA after removing the stripes.

Figure 6: Reconstructed image using the method proposed in [2].

Figure 7: Magnified picture of part of the image LENA.

Figure 8: Reconstructed image using the proposed method.

Figure 9: Magnified picture of the same area shown in Figure7.