Algorithm for Improving the Performance of JPEG Image Coding

Demin Wang
Communications Research Centre of Canada
Ottawa, Ontario, Canada

Abstract

The standard JPEG algorithm is improved by increasing its compression efficiency and reducing block artifacts. The improved algorithm first uses spatial block prediction to reduce spatial redundancy, then uses DCT, quantization, and entropy coding to compress the prediction errors. The spatial block prediction has two modes: DC prediction and plane prediction, each of them followed by a block smoother. Simulation results show that the improved algorithm achieves a bit rate saving of up to 13.7%, compared with the standard JPEG algorithm. It also improves subjective image quality by reducing blocking artifacts.

1. Introduction

JPEG algorithm is a standard algorithm that is widely used for compression of continuous-tone (multilevel), gray scale and color, still images.[1] It has been shown that the JPEG algorithm offers good image quality at bit rates above 0.5 bit/pixel, but results in unacceptable blocking artifacts at bit rates under 0.25 bit/pixel.[2] Many applications require lower bit rates with higher image quality. Hence, numerous researchers are working to improve the JPEG algorithm, or to propose new and more efficient algorithms for image coding.

The JPEG algorithm is based on the discrete cosine transform (DCT), scalar quantization and entropy coding. The image to be compressed is first divided into blocks of 8×8 pixels. The pixels of each block then undergo the two-dimensional DCT. The resulting DCT coefficients are scalar quantized and compressed using entropy coding.

This paper proposes a simple algorithm for increasing the compression efficiency and for reducing blocking artifacts of the JPEG algorithm. This algorithm first uses spatial block prediction to reduce spatial redundancies, then uses DCT, quantization, and entropy coding to compress the prediction errors. Hence, the novelty of the algorithm resides in the spatial block prediction.

2. Spatial Block Prediction

The standard JPEG algorithm involves a straightforward spatial prediction. After DCT, the DC and AC coefficients are compressed using two different entropy encoders. The DC coefficient of the current block is first predicted using that of the previous block. The prediction error is then coded using one of the entropy encoders. This prediction has the following two shortcomings. The first shortcoming is that it does not predict AC coefficients. The second is that some pixels in the previous block are located far from the current block, and therefore cannot accurately predict the current block.

In the following, a simple algorithm is proposed in order to achieve more accurate prediction. This spatial block prediction is performed before DCT, as shown in Figure 1. Let f(x, y) with $0 \le x, y < 8$ denote the pixels within the current block, and p(x, y) denote the prediction of f(x, y). The image to be compressed is divided into blocks of 8×8 pixels. If the current block is neither on the first row nor the first column of blocks, the pixels above and to the left of the current block have been previously encoded and decoded according to the conventional coding order of blocks. Instead of all the pixels within a previously decoded block, the pixels on the two lines above and on the two columns to the left of the current block are used to predict the current block, as shown in Figure 2. The current block will be predicted using one of the two different prediction modes: DC prediction and plane prediction.

2.1. DC Prediction Mode

If both $f(x, -1)$ and $f(-1, y)$ for $0 \le x, y < 8$ are inside the image,

$$p(x, y) = \frac{1}{16}\left[\sum_{i=0}^{7} f(i,-1) + \sum_{j=0}^{7} f(-1, j)\right]$$

If only $f(x, -1)$ for $0 \le x < 8$ is inside the image,

$$p(x, y) = \frac{1}{8}\sum_{i=0}^{7} f(i,-1)$$

If only $f(-1, y)$ for $0 \le y < 8$ is inside the image,

$$p(x, y) = \frac{1}{8}\sum_{j=0}^{7} f(-1, j)$$

Otherwise, the current block is the first block of the image, $p(x, y)$=128.

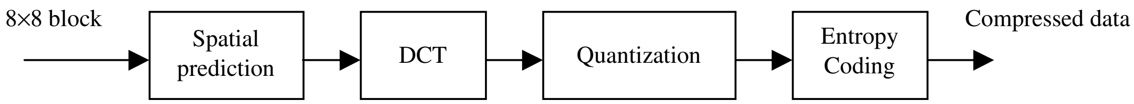

Figure 1. Block diagram of the improved JPEG encoder.

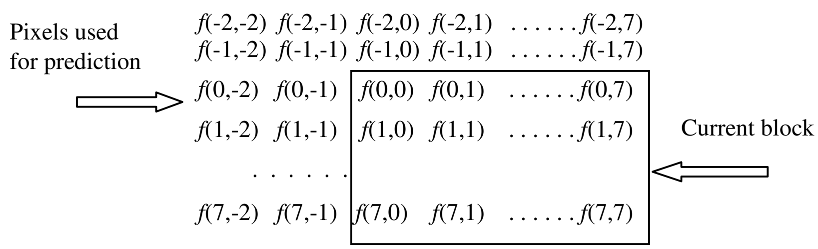

Figure 2. Current block and the pixels used for prediction.

2.2. Plane Prediction Mode

$$p(x,y) = ax + by + c$$

where a, b, and c are the parameters of the plane. The plane approximates the decoded pixels by minimizing the mean square error. If $f(x, -2)$, $f(x, -1)$, $f(-2, y)$ and $f(-1, y)$ for $0 \le x, y < 8$ are inside the image, let

$$\begin{cases} A = \sum_{x=-2}^{-1}\sum_{y=-2}^{7} x f(x,y) + \sum_{x=0}^{7}\sum_{y=-2}^{-1} x f(x,y) \\ B = \sum_{x=-2}^{-1}\sum_{y=-2}^{7} y f(x,y) + \sum_{x=0}^{7}\sum_{y=-2}^{-1} y f(x,y) \\ C = \sum_{x=-2}^{-1}\sum_{y=-2}^{7} f(x,y) + \sum_{x=0}^{7}\sum_{y=-2}^{-1} f(x,y) \end{cases}$$

The parameters are calculated by

$$a = (11204A + 6400B - 12714C) / 2349156$$
$$b = (6400A + 11204B - 12714C) / 2349156$$
$$c = (83619C - 12714A - 12714B) / 2349156$$

If only $f(x, -1)$ for $0 \le x < 8$ is inside the image,

$$\begin{cases} A = \sum_{x=0}^{7} x f(x,1) \\ C = \sum_{x=0}^{7} f(x,-1) \end{cases}$$

$$\begin{cases} a = (2A - 7C)/84 \\ b = 0 \\ c = (35C - 7A)/84 \end{cases}$$

If only $f(-1, y)$ for $0 \le y < 8$ is inside the image,

$$\begin{cases} B = \sum_{y=0}^{7} y f(-1,y) \\ C = \sum_{y=0}^{7} f(-1,y) \end{cases}$$

$$\begin{cases} a = 0 \\ b = (2B - 7C)/84 \\ c = (35C - 7B)/84 \end{cases}$$

Otherwise, the current block is the first block of the image, $a = 0$, $b = 0$, and $c = 128$.

2.3. Block Smoother

The prediction modes described above may result in abrupt changes between the predictions and the neighboring blocks that are already decoded. This abrupt changes may lead to large prediction errors and visible blocking artifacts.

In order to get a smooth transition between the predictions and the neighboring blocks, a block smoother is applied to the results of the predictions. The block smoother is performed by the following average operation:

$$f(x, y) = [f(x, y-1) + f(x-1, y) + f(x, y+1) + f(x+1, y)]/4, \text{ for } 0 \leq x, y < 8$$

The block smoother is applied to each pixel within the current block, starting from the top left corner of the block and proceeding pixel by pixel, and row by row, to the bottom right corner. For the pixels on line 0 and column 0 (shown in Figure 2), one or two pixels from decoded blocks are used in the operation. For the pixels on the last line and column, one or two of the four pixels are not available. The average is then calculated using two or three available pixels. The average result is immediately assigned to pixel (x, y) and is used for calculating the averages of the following pixels. Hence, it is a recursive procedure. The values of decoded pixels are partially propagated into the block, making a smooth transition between the predictions and the neighboring blocks. This block smoother is similar to the low pass extrapolation used in MPEG-4 for block padding.[3,4]

2.4. Prediction Mode Selection

For each block, one of the two prediction modes is used to predict the block. The criterion for prediction mode selection is the sum of absolute prediction errors. The prediction mode that yields the smaller sum of absolute prediction errors is chosen.

3. Coding Issues

For each block, the prediction mode is encoded using one bit. ‘0” indicates DC prediction, while “1” indicates plane prediction.

Run length coding is used for encoding both DC and AC coefficients of DCT. The run length coding is the same as that used for encoding AC coefficients in the standard JPEG algorithm. With the prediction of the standard JPEG algorithm, the variance of DC coefficients is usually much greater than those of AC coefficients since the prediction is not effective. Hence, two different coding methods are used for DC and AC coefficients. With the new prediction described above, the variance of DC coefficients is significantly reduced. Therefore, DC coefficients can be effectively encoded together with AC coefficients.

4. Simulation Results

The improved algorithm is more efficient than the standard JPEG algorithm in terms of compression efficiency. These algorithms have been used to encode the gray scale images of *Barbara*, *Goldhill*, and *Lena*. Images *Barbara* and *Goldhill* have 512×512 pixels. *Lena* has two formats, 256×256 pixels and 512×512 pixels. Each pixel in these images is quantized with 8 bits/pixel. Figure 3 illustrates the bit rate reduction (%) achieved using the improved algorithm. The improved algorithm reduces the bit rates by at least 2%, and up to 13.7%, compared with the standard JPEG algorithm.

Besides bit rate reduction, the improved algorithm also improves subjective image quality by reducing blocking artifacts at low bit rates. It is well known that the human perception is very sensitive to blocking artifacts in the regions with slowly varying luminance. The plane prediction can accurately predict slowly varying luminance, and thus reduce blocking artifacts in these regions. The block smoother creates a smooth transition between blocks, and also reduces blocking artifacts. Figure 4 shows two decoded images obtained using the improved and standard algorithms. These images have almost the same PSNR. The improved algorithm requires 9.6% less bits than the standard algorithm. It was found that the image obtained using the improved algorithm has much less blocking artifacts than that obtained using the standard algorithm.

5. Conclusions

The improved algorithm for image coding adds a spatial prediction module to the standard JPEG algorithm. The spatial prediction uses pixels, which are previously decoded and are close to the current block, to predict all the pixels in the current block. It has two modes: DC prediction mode and plane prediction mode followed by a block smoother. The improved algorithm is simple to implement. It achieves lower bit rates and higher subjective image quality than the standard JPEG algorithm.

Acknowledgment

The author would like to thank A. Vincent, J. Tam, P. Corriveau and P. Blanchfield for their assistance.

References

1. W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image data Compression Standard*, New York: Van Nostrand Reinhold, 1993
2. B. G. Haskell, et al., ‘Image and video coding – emerging standards and beyond,” *IEEE Trans, Circuits and Systems for Video Technology*, Vol. 8, No. 7, pp. 814-837, Nov. 1998.
3. ISO/IEC JTC1/SC29/WG11, ‘MPEG-4 video verification model version 7.0,” Bristol, April 1997.
4. A. Kaup, ‘Object-based texture coding of moving video in MPEG-4,” *IEEE Trans, Circuits and Systems for Video Technology*, Vol. 9, No. 1, pp. 5-15, Feb. 1999.

Biography

Demin Wang received the B. S. and M. S. degrees in electrical engineering from Shandong University of Technology, China, in 1982 and 1985, respectively, and the Ph. D. degree from the Institut National des Sciences Appliquees (INSA) de Rennes, France, in 1992. He joined Shandong University of Technology in 1985, where he served as a Professor of Electrical and Computer

Engineering from 1992 to 1993. He was a Visiting Researcher at the University of Sherbrooke, Canada, from 1993 to 1994, and a Visiting Professor at the IRISA Rennes, France, from 1994 to 1995. Since 1996, he has been with the Communications Research Centre, Canada, as a Research Scientist. His research interests include image processing, video processing and coding, digital TV, multimedia communication, and nonlinear filtering

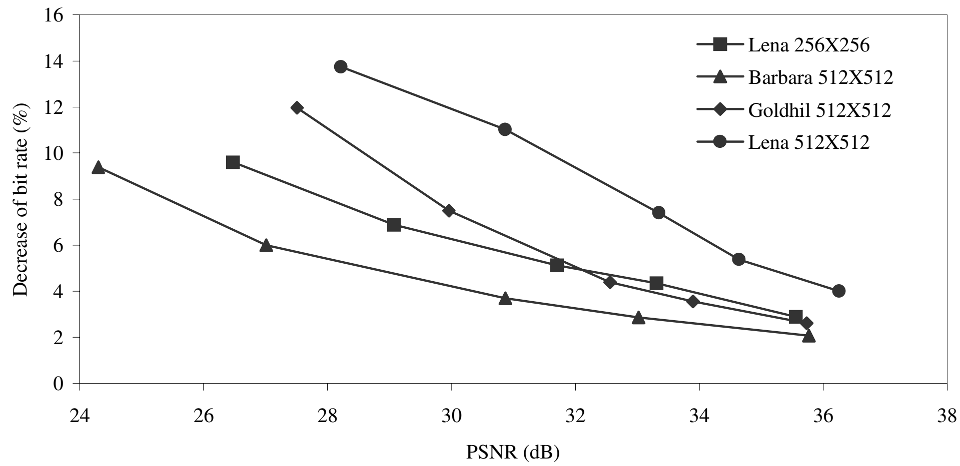

Figure 3. Bit rate reduction achieved using the improved algorithm as a function of PSNR, compared with the JPEG algorithm.

(a) *(b)*

Figure 4. Blocking artifacts resulting from different algorithms, (a) improved algorithm requiring 0.28 bit/pixel, (b) standard JPEG algorithm requiring 0.31bit/pixel.