16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

LOSSLESS RE-ENCODING OF JPEG IMAGES USING BLOCK-ADAPTIVE INTRA PREDICTION

Ichiro Matsuda, Yukio Nomoto, Kei Wakabayashi and Susumu Itoh

Department of Electrical Engineering, Faculty of Science and Technology,
Science University of Tokyo

2641 Yamazaki, Noda-shi, Chiba 278-8510, JAPAN
phone: +81 4 7124 1501 (ext.3740), fax: +81 4 7124 9367
email: matsuda@ee.noda.tus.ac.jp

ABSTRACT

This paper proposes a kind of transcoding scheme which compresses existing JPEG files without any loss of quality. In this scheme, H.264-like block-adaptive intra prediction is employed to exploit inter-block correlations of quantized DCT coefficients stored in the JPEG file. This prediction is performed in spatial domain of each block composed of 8×8 pels, but the corresponding prediction residuals are calculated in DCT domain to ensure lossless reconstruction of the original coefficients. Moreover, block-based classification is carried out to allow accurate modeling of probability density functions (PDFs) of the prediction residuals. A multisymbol arithmetic coder along with the PDF model is used for entropy coding of the prediction residual of each DCT coefficient. Simulation results indicate that the additional reductions of coding rates obtained by the proposed scheme are 18–28 % for monochrome JPEG images.

1. INTRODUCTION

The JPEG [1], which was adopted as the first international standard for still image compression, has been the most popular format for storing and transmitting digital images. Though several superior formats such as JPEG 2000 [2] and HD-Photo [3] have been developed, the old JPEG is even now exclusively used as a common format on imaging devices like digital cameras. One reason for preventing generation shift of the image formats seems to be in existence of numerous images which have been captured and stored in the JPEG format over past decades. Of course one can transcode already compressed images, for example, from JPEG to JPEG 2000. However, additional loss of image quality is introduced in this case and, as a result, the advantage of the JPEG 2000 in coding efficiency is usually spoiled [4].

From this point of view, new transcoding methods which compress existing JPEG files without any loss of quality have been proposed in recent years [5–7]. These methods first decode the JPEG file halfway to extract quantized DCT coefficients, and then re-encode the coefficients by using modern entropy coding techniques based on arithmetic coding. In fact, this approach itself is not absolutely new and the JPEG standard already specified an optional entropy coding mode using the binary arithmetic coder called QM-coder [1]. Unfortunately, such simple replacement of entropy coding techniques from the ordinary Huffman coding to arithmetic one has a certain limit in improvement of coding efficiency. Therefore, in order to gain further improvement, the authors of the above methods [5–7] focused on inter-block correlations of DCT coefficients which are not exploited in the original JPEG standard with the exception of DC coefficients. However, because they tried to make use of the correlations always in DCT domain, only a few coefficients at the same frequency in neighboring blocks were referred to change coding order [5] or statistical models [6, 7] in adaptive arithmetic coding.

In this paper, we propose a novel lossless re-encoding scheme using block-adaptive intra prediction. The intra prediction is performed in spatial domain of each block in a similar way to intra prediction modes used in the H.264 standard [8]. This approach enables us to exploit the inter-block correlations which can be spread over multiple frequencies in the DCT domain. On the other hand, the prediction residuals are calculated in the DCT domain to ensure lossless reconstruction of quantized DCT coefficients stored in the original JPEG file. It makes the reconstructed image also lossless if the same reconstruction process as the standard JPEG decoder is applied.

2. BLOCK-ADAPTIVE INTRA PREDICTION FOR QUANTIZED DCT COEFFICIENTS

Figure 1 shows an overview of the proposed re-encoding scheme. First, Huffman decoding of a given JPEG file is carried out to extract quantized DCT coefficients $y_{k,l}(i, j)$, where k and l are indices of blocks composed of 8×8 pels and $i, j = 0, 1, \ldots, 7$ indicate frequencies of two-dimensional DCT coefficients in the block. Next, these coefficients are dequantized using a quantization matrix $Q(i, j)$ included in header information of the JPEG file, and inverse DCT (IDCT) is performed to reconstruct image values $x_{k,l}(i, j)$ of

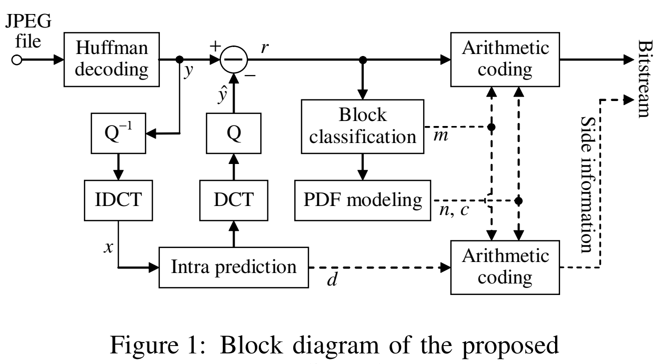


Figure 1: Block diagram of the proposed re-encoding scheme.

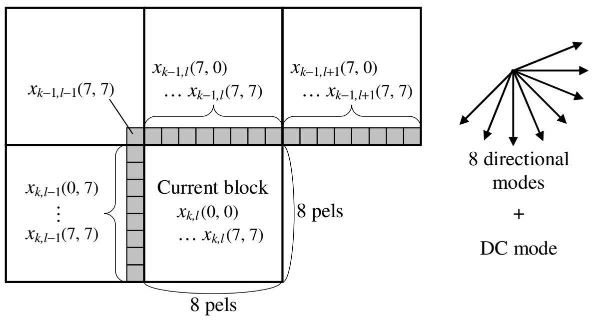


Figure 2: Block-adaptive intra prediction.

each block. By using the reconstructed values at boundaries of neighboring blocks, whose positions are shown as shaded boxes in Figure 2, adaptive intra prediction is performed in spatial domain of the current block. For the adaptive intra prediction, eight directional modes and a DC mode, which are almost same as the nine modes of 4×4 intra prediction specified in the H.264 standard [8] but size of the block is extended to 8×8 pels in conformity with that of DCT, are tested and the most suitable one is actually performed. Furthermore, predicted values in the block are transformed into DCT domain and quantized using the quantization matrix $Q(i,j)$ to obtain DCT coefficients $\hat{y}_{k,l}(i,j)$. Since the same quantization matrix $Q(i,j)$ is used for both the dequantization and the quantization processes (Q^{-1} and Q), DCT coefficients $y_{k,l}(i,j)$ and $\hat{y}_{k,l}(i,j)$ are integers with the same scale at the respective frequencies. Therefore, the prediction residuals of the DCT coefficients can be calculated as differences of both integers:

$$r_{k,l}(i,j) = y_{k,l}(i,j) - \hat{y}_{k,l}(i,j) \quad \text{for } i,j = 0,1,\ldots,7. \quad (1)$$

In general, it is more efficient to encode the residuals $r_{k,l}(i,j)$ than encode the original DCT coefficients $y_{k,l}(i,j)$ directly, because inter-block correlations can be effectively removed through the above block-adaptive intra prediction. In this case, a prediction mode $d(k,l) \in \{1,2,\ldots,9\}$ used in each block must be encoded as side information to perform the same prediction at a decoder, or an inverse transcoder.

3. PDF MODELING OF PREDICTION RESIDUALS

For accurate modeling of probability density functions (PDFs) of the prediction residuals, we prepare the following 32 kinds of the generalized Gaussian functions [9]:

$$P_n(t) = \frac{c_n \cdot \eta(c_n,\sigma_n)}{2\Gamma(1/c_n)} \cdot \exp\left\{-\left|\,\eta(c_n,\sigma_n)\cdot t\,\right|^{c_n}\right\},$$
$$\eta(c_n,\sigma_n) = \frac{1}{\sigma_n}\sqrt{\frac{\Gamma(3/c_n)}{\Gamma(1/c_n)}} \qquad (n = 1,2,\ldots,32), \quad (2)$$

where $\Gamma(\cdot)$ is the gamma function, σ_n is standard deviation of the prediction residuals and c_n is a shape parameter which controls sharpness of the function $P_n(t)$. In this paper, values of σ_n are chosen in advance to cover a wide range of activity in the prediction residuals. On the other hand, values of the shape parameter c_n can be changed image-by-image so as to fit the actual PDFs. Moreover, all the blocks composed of 8×8 pels are classified into M classes. Each class has a look-up table $n_m(i,j)$ which describes assignment of the PDF models $\{P_n(t)\,|\,n = 1,2,\ldots,32\}$ to the respective frequencies $i,j = 0,1,\ldots,7$. It should be noted that the continuous function $P_n(t)$ gives a PDF model for the prediction residuals of DCT coefficients before the quantization. In practice, a probability of each quantization level r is given by:

$$\Pr(r\,|\,n,Q) = F_n((r+0.5)\cdot Q) - F_n((r-0.5)\cdot Q),$$
$$F_n(x) = \int_{-\infty}^{x} P_n(t)\,\mathrm{d}t, \quad (3)$$

where $Q = Q(i,j)$ is a quantization step-size and $n = n_m(i,j)$ indicates the PDF model assigned to the $r_{k,l}(i,j)$ in the m-th class. $F_n(x)$ is a cumulative distribution function of $P_n(t)$ and its values are pre-calculated and stored in a table at a sampling rate of $Q_{min}/2$ for fast calculation of Eq.(3). Probabilities of all possible quantization levels calculated in this way are used for entropy coding of the actual prediction residuals $r = r_{k,l}(i,j)$. In our implementation, the randecoder [10] which is known as a fast multisymbol arithmetic coder is employed for the entropy coding.

4. OPTIMIZATION OF PREDICTION MODES AND SOME CODING PARAMETERS

In the proposed scheme, parameters listed below must be encoded as side information.

- Quantization matrix $Q(i,j)$ extracted from the original JPEG file.
- Prediction mode $d(k,l) \in \{1,2,\ldots,9\}$ used in each block.
- Class label $m(k,l) \in \{1,2,\ldots,M\}$ assigned to each block.
- Look-up table $n_m(i,j) \in \{1,2,\ldots,32\}$ used for PDF modeling of the DCT coefficients belonging to the m-th class.
- Shape parameter $c_n \in \{0.2,0.4,\ldots,3.2\}$ of each PDF model.

With exception of the quantization matrix $Q(i,j)$, values of these parameters are iteratively optimized before the actual re-encoding process so that the following cost function can be a minimum.

$$J = -\sum_{k,l}\sum_{i,j}\log_2 \Pr\left(r_{k,l}(i,j)\,\middle|\,n_{m(k,l)}(i,j), Q(i,j)\right) + B_{side}. \quad (4)$$

The first term of the cost function represents the number of coding bits required for the entropy coding of the prediction residuals. The second term (B_{side}) indicates the amount of side information on the above parameters. Concrete procedures for the optimization are as follows:

(1) An initial prediction mode $d(k,l)$ is determined for each block. At this stage, sum of squared prediction residuals $S_{k,l}(d)$ is minimized instead of the cost function J:

$$d(k,l) = \operatorname*{argmin}_{d\in\{1,2,\ldots,9\}} S_{k,l}(d),$$
$$S_{k,l}(d) = \sum_{i,j=0}^{7}\left\{r_{k,l}(i,j)|_{\text{prediction mode}=d}\right\}^2. \quad (5)$$

In addition, an initial class label $m(k,l)$ is assigned to every block by quantizing the obtained value of $S_{k,l}(d(k,l))$ into one of $M = 32$ levels.

(2) For each class (m), the look-up table $n_m(i,j)$ is optimized by selecting one of the 32 PDF models $P_n(t)$ ($n = 1,2,\ldots,32$) at every frequency of DCT coefficients.

(3) The optimum value of the shape parameter c_n is chosen from $\{0.2, 0.4, \ldots, 3.2\}$ for each PDF model.

(4) Re-classification of the blocks is performed by selecting the optimum look-up table $n_m(i,j)$ ($m = 1,2,\ldots,M$) at every block.

(5) The optimum prediction mode $d(k,l)$ which minimizes the cost function J is determined for each block.

(6) By assuming one of M classes is being removed, importance of that class is evaluated by the cost function J. In this case, the first term of Eq.(4) would be increased by substituting the second best class (look-up table) at some blocks. However, side information on the removed class can be excluded from the second term. Taking this tradeoff into account, the least important class is actually removed and the value of M is decreased by one.

(7) The above one-by-one removal of the useless class is repeated until the resulting cost function J cannot be reduced further. This procedure enables automatic setting of the number of classes ($M \leq 32$) according to characteristics of each JPEG image.

(8) Procedures (2)–(7) are iteratively carried out until all of the parameters converge. Typically, the number of iterations needed for this step is 20–40 and the overall re-encoding process takes 7–25 seconds on a computer with a 3 GHz Intel processor.

When values of the parameters optimized in this way are encoded, the rangecoder is also used together with some ingenuity for reducing the amount of the side information. For example, we employ the move-to-front (MTF) method [11] to remove inter-block correlations of the parameters $d(k,l)$ and $m(k,l)$ both of which are needed for every block and account for large parts of the side information. The MTF method expresses a symbol by its position in a table whose elements are constantly rearranged so that the most probable symbol can be placed in the front. In our scheme, three values of top, left and top-right blocks are moved to the front positions in this order (if two of them are the same, it is placed at the first position according to a majority rule). Figure 3 illustrates this rearrangement process by an example. In this figure, values of $5,2,5,5,7\cdots$ assigned for the second row of blocks are converted into a

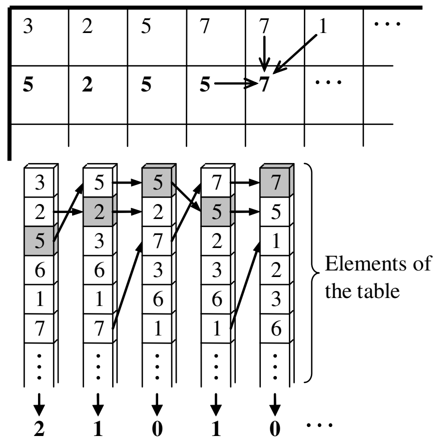


Figure 3: The move-to-front method for efficient coding of prediction modes and class labels.

sequence of $2,1,0,1,0\cdots$. This process is reversible and the obtained values tend to be small when the original values are correlated. The method is also used in calculation of the B_{side} in Eq.(4) to make the cost function J exact.

5. SIMULATION RESULTS

In order to evaluate coding performance, monochrome images are once encoded by the baseline JPEG algorithm using Huffman coding [1] and then further compressed by the proposed scheme and other lossless re-encoding schemes. StuffIt [12] is data compression software produced by Smith Micro (formerly Allume Systems) and its recent version (we tested StuffIt Deluxe 9.0) offers capability to compress JPEG files. Unfortunately, the detailed algorithm used in StuffIt is unknown because it is proprietary software. PackJPG [13] is a freely available program (V2.2) which is an implementation of the JPEG re-encoding algorithm presented in [7]. JPEG (QM-coder) means an optional mode of the JPEG standard using QM-coder [1] and is also used as lossless re-encoding scheme in this experiment. Table 1 reports coding rates obtained by the above mentioned lossless re-encoding schemes when a quality factor of 75 is used in the baseline JPEG. The best result for each image is shown in **boldface** in the table. Moreover, Figure 4 plots rate-SNR curves of the lossless re-encoding schemes together with those of the representative direct coding schemes: JPEG 2000 [2], HD-Photo [3] and the baseline JPEG (the quality factor is ranging from 30 to

Table 1: Coding rates obtained by lossless re-encoding schemes (bits/pel).

Image	SNR (dB)	Proposed scheme	StuffIt	PackJPG	JPEG (QM-coder)	JPEG (baseline)*
Camera	34.7	**1.033 (−21.0 %)**	1.063 (−18.7 %)	1.111 (−15.1 %)	1.217 (−7.0 %)	1.308
Lena	36.4	0.871 (−23.0 %)	**0.862 (−23.8 %)**	0.931 (−17.7 %)	1.062 (−6.1 %)	1.131
Baboon	31.3	**1.716 (−18.4 %)**	1.744 (−17.1 %)	1.771 (−15.8 %)	1.898 (−9.7 %)	2.103
Barb	36.6	1.008 (−23.7 %)	**1.003 (−24.1 %)**	1.097 (−17.0 %)	1.235 (−6.5 %)	1.321
Average	—	***1.157 (−21.1 %)***	*1.168 (−20.3 %)*	*1.228 (−16.3 %)*	*1.353 (−7.7 %)*	*1.466*

*quality factor = 75

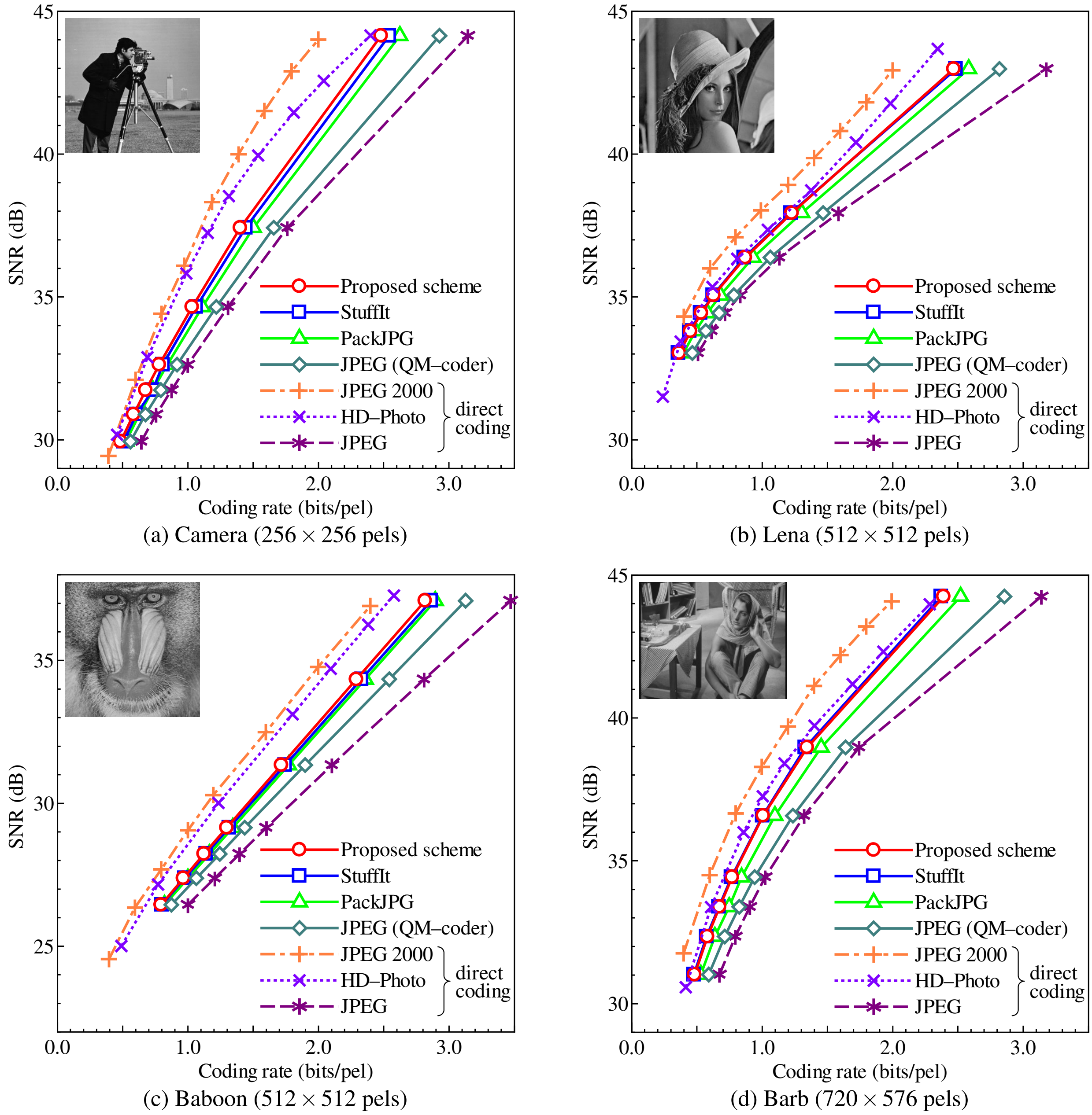


Figure 4: Coding performance.

95). It is demonstrated that the proposed scheme and StuffIt provide almost the same coding performance and have distinct advantages over the other lossless re-encoding schemes. In addition, for ‘Lena’ and ‘Barb’ images, coding efficiency of both the re-encoding schemes approaches to that of HD-Photo which has been recently developed by Microsoft as a direct coding scheme for still images and is currently under consideration for a new standard whose tentative name is JPEG XR.

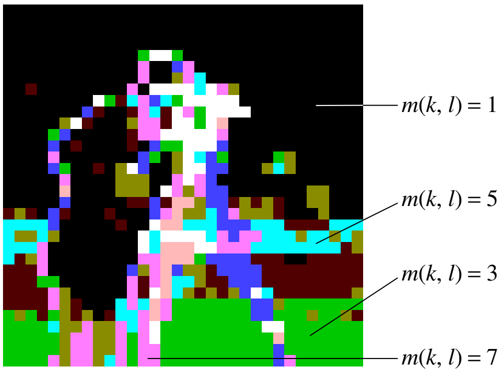
Figure 5 (b) depicts an example of the block-based classification for PDF modeling. In this example, the number of classes automatically determined through the optimization procedures is $M = 9$. Values of σ_n, which are associated with the look-up tables $n_m(i, j)$ $(m = 1, 2, \ldots, M)$ and used for the PDF modeling of DCT coefficients, are partly shown in Figure 5 (c). We can see that appropriate classification of blocks according to local characteristics of the image is carried out and, as a result, accurate PDF modeling for the prediction residuals in DCT domain is realized.

6. CONCLUSIONS

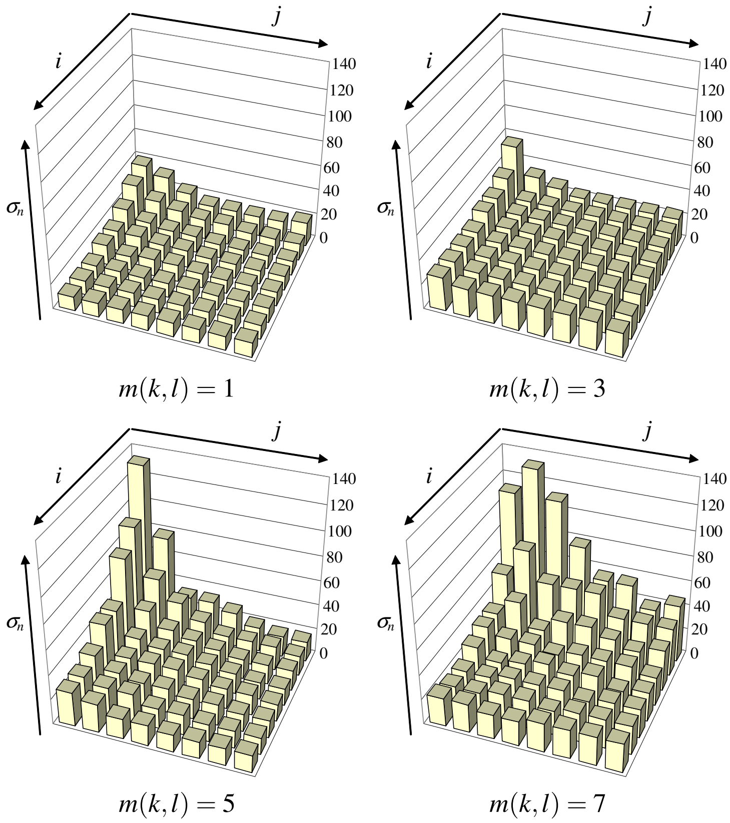
In this paper, we have proposed a lossless re-encoding scheme for existing images already encoded by the baseline JPEG algorithm. The scheme employs H.264-like block-adaptive intra prediction to exploit inter-block correlations of quantized DCT coefficients stored in the original JPEG file. Prediction residuals of the DCT coefficients are losslessly encoded by using an adaptive arithmetic coder and several sets of PDF models whose assignments are iteratively

16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, August 25-29, 2008, copyright by EURASIP

(a) Original image of 'Camera'



(b) Assignment of class labels



(c) Values of σ_n used for PDF modeling of DCT coefficients

Figure 5: An example of block-based classification for 'Camera'.

optimized for each JPEG image. Simulation results indicate that coding efficiency of the proposed scheme is comparable with that of the StuffIt scheme which was reported to give the best compression performance for JPEG files in some benchmark tests [14, 15]. Further improvement of the coding efficiency by enhancing the intra prediction technique will be our future task.

REFERENCES

[1] ITU-T Rec. T.81 | ISO/IEC 10918-1, "Information Technology — Digital Compression and Coding of Continuous-Tone Still Images: Requirements and Guidelines," 1994.

[2] ITU-T Rec. T.800 | ISO/IEC 15444-1, "Information Technology — JPEG 2000 Image Coding System — Part 1: Core Coding System," 2001.

[3] http://www.microsoft.com/windows/windowsmedia/forpros/wmphoto/default.aspx

[4] K. Matsuo, M. Hashimoto and A. Koike, "A JPEG2000 Re-encoding Scheme with High Quality Performance for JPEG Coded Pictures," *Proc. of Intl. Symposium on Signal Processing and Information Technology (ISSPIT 2003)*, Darmstadt, Germany, pp.660–663, Dec. 14–17, 2003.

[5] I. Bauermann and E. Steinbacj, "Further Lossless Compression of JPEG Images," *Proc. of Picture Coding Symposium (PCS 2004)*, San Francisco, USA, Dec. 15–17, 2004.

[6] N. Ponomarenko, K. Egiazarian, V. Lukin and J. Astola, "Additional Lossless Compression of JPEG Images," *Proc. of the 4th Intl. Symposium on Image and Signal Processing and Analysis (ISPA 2005)*, Zagreb, Croatia, pp.117–120, Sep. 15–17, 2005.

[7] M. Stirner and G. Seelmann, "Improved Redundancy Reduction for JPEG Files," *Proc. of Picture Coding Symposium (PCS 2007)*, Lisbon, Portugal, Nov. 7–9, 2007.

[8] ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC, "Advanced Video Coding for Generic Audiovisual Services," 2003.

[9] I. Matsuda, N. Ozaki, Y. Umezu and S. Itoh, "Lossless Coding Using Variable Block-Size Adaptive Prediction Optimized for Each Image," *Proc. of 13th European Signal Processing Conference (EUSIPCO 2005)*, Antalya, Turkey, WedAmPO3, Sep. 4–8, 2005.

[10] M. Schindler, "A Fast Renormalization for Arithmetic Coding," *Proc. of Data Compression Conference (DCC '98)*, Snowbird, Utah, p.572, Mar. 30–Apr. 1, 1998.

[11] J. L. Bentley, D. D. Sleator, R. E. Tarjan and V. K. Wei, "A Locally Adaptive Data Compression Scheme," *Communications of ACM*, Vol.29, No.4, pp.320–330, Apr. 1986.

[12] http://my.smithmicro.com/win/index.html

[13] http://www.elektronik.htw-aalen.de/packjpg/

[14] http://compression.ca/act/act-jpeg.html

[15] http://www.maximumcompression.com/data/jpg.php