SIGNAL PROCESSING IV: Theories and Applications
J.L. Lacoume, A. Chehikian, N. Martin, and J. Malbos (Editors)
Elsevier Science Publishers B.V. (North-Holland)

COMPRESSION OF DITHERED IMAGE BLOCKS

A. Israel and D. Malah

Electrical Engineering Department
Technion - Israel Institute of Technology
Technion City, Haifa 32000, Israel

ABSTRACT In this paper, two methods for block-coding of dithered images are examined: Block-quantization and lossless block-coding. The first block compression method uses a dictionary of blocks which is iteratively designed by the LBG algorithm, modified to deal with the binary case. In addition, an appropriate distance measure between binary blocks is applied. With this method a compression factor of 3.2 is achieved without any visible degradation in the reconstructed dithered image. For lossless block-coding we present a simple new variable-length coding technique for dithered images. The compression factor obtained by this technique is 3.7. By combining both methods high quality reconstructed dithered images are obtained with a compression factor of about 5 (i.e. 0.2 bit/pixel).

1. INTRODUCTION

Dithering techniques are well known to provide effective means for displaying continuous-tone images on bi-level devices [1]. Some existing facsimile apparatuses employ these techniques to make it possible to deal with continuous-tone pictures such as photographs. The ordered dither technique is the most often utilized one, because it is simple and results in good quality of the displayed pictures [2].

The dithering technique consists of comparing the multilevel image with position dependent thresholds and setting only those pels (picture elements) in the reproduced picture to white (or '1') for which the input gray level exceeds the threshold value. Other pels are set to black (or '0'). The thresholds are organized in a $n \times m$ matrix denoted as the 'dither matrix'. The original image is divided into blocks of the same size as the dither matrix, and each element in a block is compared to the corresponding threshold. In this work a 4×4 dither matrix is used with threshold values as proposed by Bayer [3].

Fig. 1(a) shows an example of a dithered image obtained by using the Bayer dither matrix.

The binary dithered image still has some redundancy which can be reduced by appropriate coding. Due to the special texture of dithered images, standard run length coding methods, such as MH (Modified Huffman) or MR (Modified READ) [10], are not effective for this type of images.

Several dither coding schemes that were useful are: Bit interleaving, threshold reordering, and predictive (Markov model) coding. All these schemes are pel by pel coding techniques and are reviewed in [2].

This paper presents lossy and lossless *block-coding* techniques for dithered images. Section 2 of the paper describes previous methods for block compression which use a dictionary of blocks. Section 3 describes a design algorithm for the dictionary, and a distance measure which is particularly appropriate for binary images. In section 4 we propose and test a new variable-length block-coding technique for dither images. In section 5 we combine the two previous techniques : lossy coding using a dictionary followed by variable-length coding.

2. PREVIOUS APPROACHES TO BLOCK CODING OF DITHERED IMAGES

Block coding of dithered images was tested in [4] using lossless and lossy coding approaches (and is also mentioned briefly in [9]).

In the *lossless coding* approach, all the unique blocks that occur in the dithered image are found and used to form a dictionary, which is transmitted previous to the blocks coding.

In [4] this method provided an average information rate of 0.38 bit/pel (i.e.,a compression ratio of 2.6:1), for an ensemble of pictures.

According to the *lossy-coding* method examined in [4], only a subset of the different image blocks are used to form the dictionary. This coding method can be considered to be a block-quantization technique. Basically, block quantization consists of an encoder mapping of an input block B to an R bit binary code word. The decoder then assigns a reproduction block $\hat{B}$ to the received code word. The set of reproduction blocks is called 'dictionary' or code book. The information rate, in bits per block, for a dictionary with $M = 2^R$ reproduction blocks is R, and the information rate, in bits per pel, is given in equation (1) for blocks of $n \times m$ pels. This method is particularly attractive for fixed rate transmission.

$$r = \frac{\log_2 M}{n \cdot m} \text{ bit/pel} \qquad (1)$$

The compression ratio that is achieved in [4], using this method, is 4:1.

In [4], The dictionary content is set to be the most frequent blocks in a group of dithered images. A block in the image that is not found in the dictionary is replaced by that block in the dictionary which has the minimum Hammming distance from the original block.

3. PROPOSED APPROACH TO LOSSY BLOCK-CODING

In an attempt to obtain a lower average distortion than with the method suggested in [4], we apply in this paper an iterative algorithm for designing the dictionary. The iterative algorithm is based on the LBG algorithm [6], originally suggested for vector quantization (VQ). This algorithm converges to a local minima of the average distortion.

In addition, a different distance measure is applied. This distance measure, known as 'match-distance' [8], better reflects the "subjective closeness" between binary blocks than the Hamming distance, and hence is more suitable for coding dithered images.

3.1 Iterative Dictionary Design (LBG Algorithm)

For a given dictionary size - M, the goal is to find an optimal dictionary and mapping, i.e. to minimize the average distortion $E[d(B,\hat{B})]$, where $d(\cdot,\cdot)$ is a non negative distance measure.

The common technique for designing a good dictionary is to use a large enough training sequence, and to apply the iterative LBG algorithm [6]. The basic steps of the LBG algorithm are as follows:

(0) Based on the empirical statistics obtained from the training sequence assign an initial block-dictionary of size M.

(1) For a given dictionary, map every block of the training sequence to the nearest reproduction block in the dictionary.

This mapping defines a partition of clusters. Where all the blocks that are mapped to the same reproduction block form a cluster.

(2) Given the partition of clusters, find for every cluster the optimal reproduction block.

The new set of reproduction blocks forms the new dictionary.

(3) Repeat steps (1) and (2) until no significant decrease in the average distortion is obtained.

Further details of the algorithm can be found in [5].

3.2 Distance Measures

The simplest distance measure for the binary case is the Hamming distance. The Hamming distance between two blocks B and $\hat{B}$ with binary pels b_{ij} and $\hat{b}_{ij}$, respectively ($1 \le i \le n$, $1 \le j \le m$), is

$$d(B,\hat{B}) = \sum_{i=1}^{n} \sum_{j=1}^{m} \delta(b_{ij},\hat{b}_{ij}) \tag{2.1}$$

where

$$\delta(x,y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{otherwise} \end{cases} \tag{2.2}$$

However, this distance measure dose not reflect well the "subjective closeness" or difference between two binary blocks.

As mentioned earlier, a more suitable distance measure for this purpose is the 'match-distance' [8].

A matching of B and $\hat{B}$ is defined as a 1-1 pairing of the '1'-valued elements of B and $\hat{B}$. The 'match-distance' $d_m(B,\hat{B})$ between B and $\hat{B}$ is the minimum pairwise distances for all possible matchings of B and $\hat{B}$.

In formal notation - let π be the permutation giving a minimal matching between B and $\hat{B}$. The matching-distance between B and $\hat{B}$ is

$$d_m(B,\hat{B}) = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ b_{ij}=1}}^{m} \delta_m((i,j), \pi(i,j)) \tag{3}$$

where δ_m can be taken, e.g., as any of the city-block, chessboard,or Euclidean distances. In this paper we use the city-block distance, for which the distance between block indeces (i,j) and (k,l) is defined as $\delta_m((i,j),(k,l)) = |i-k| + |j-l|$.

Several problems arise by utilizing this distance measure. One of these is the high computational complexity needed to find the minimum matching between two blocks having a large number of elements to be paired. The second problem occurs when we need to compute the match-distance between blocks having a different number of '1'-valued elements. This problem is solved by normalization. Several approaches for normalization are described in [8].

An additional increase in the complexity occurs when the LBG algorithm is applied with the match-distance measure. In this case, in order to find the optimal reproduction block (centroid) of a given cluster, an exhaustive search over all possible blocks is needed. This is not so for the Hamming distance measure [7].

3.3 Simulation Results

In this paper five 512×512 dither images were used as the training sequence (i.e. 81920 4×4 binary blocks), and five other images were used as test images.

Four values of M, the dictionary size, were examined: 64,32,16 and 8 - blocks. Table 1 gives the average distortion per pel for the initial dictionaries and for the final dictionaries that were obtained by using the binary LBG algorithm and with the 'match-distance' as described above.

In those instances where the initial dictionaries were chosen to be the M most frequent blocks in the training images, only one iteration of the LBG algorithm was needed, and the final dictionary contents was identical to the initial one. Hence, the average distortion for these cases was not changed by the application of the iterative technique.

Another approach for selecting the initial dictionaries was random selection of M blocks from the training images (without repetition). The final dictionaries obtained with this approach resulted in a better average distortion for $M = 8$ and $M = 32$. An examination of the content of these dictionaries showed a great similarity between them and the previous ones which contain the M most frequent blocks in the training images. This fact explains why only a small improvement is obtained by the final dictionaries mentioned above.

For comparison, simulation were also carried out with the Hamming distance measure. In this case the final dictionaries were very similar to the previously mentioned final dictionaries.

Subjective judgement of the training and test reproduced dither images, using the best dictionaries in our simulation, results in the following observations (see also Fig. 1) :

$M = 64$ The quality of the reproduced images is judged to be same as the original dither image (for both distance measures).

$M = 32$ Negligible differences between the reproduced dither images and the original one (for both distance measures).

$M = 16$ Small but noticeable differences between the reproduced images and the original. Examination of zoomed regions showed a small advantage of the 'match-distance' over the Hamming distance.

$M = 8$ Bad quality of the reproduced dither images (for both distance measures).

From the above results it is concluded that this class of dithered images can be coded at a rate of R=5 bit/block (corresponding to $M = 32$), i.e. $r = 0.313$ bit/pel, without any visible

TABLE I

Average distortion (per pel) of different block dictionaries

Dictionary size M		64	32	16	8
Initial dictionary contents	average distortion $\times 10^{-2}$				
Frequent blocks in the training images	Initial	0.47	1.35	2.61	10.36
	Final	0.47*	1.35	2.61*	10.36
Random selection from the training images	Initial	0.59	1.34	3.36	10.01
	Final	0.57	1.34*	3.36	9.12*

Note: The best dictionaries that were found are marked by '*'.

degradation; whereas at the rate of $R=4$ bit/block ($M=16$), i.e. $r=0.25$ bit/pel, a slight visible degradation in the reproduced dithered image is expected.

4. LOSSLESS VARIABLE-LENGTH BLOCK-CODING

In this section we propose and examine a lossless (i.e. without distorsion) variable-length block-coding method that is different from the lossless block-coding method reported in [4]. In this method we exploit two important statistical properties of dithered images:

(1) The probability that two successive blocks are identical is about 0.5 .

(2) A small subset of all dither blocks occurs more frequently than all the others.

Let $B(k)$ denote the current block for coding, and $B(k-1)$ the previous block, which its value is known to the decoder. Let $\mathbf{B}$ denote a subset of 2^q blocks. We define three events that can occur during the coding of the dithered image blocks. Accordingly, the proposed bit assignment for these events is as follows:

event	prefix bits	data bits	code word length
$E_1: B(k)=B(k-1)$	0	None	1
$E_2: B(k)\neq B(k-1)$, and $B(k)\in \mathbf{B}$	10	$xx\cdots x$ q bits	$2+q$
$E_3: B(k)\neq B(k-1)$, and $B(k)\notin \mathbf{B}$	11	$xx\cdots x$ $n\cdot m$ bits	$2+n\cdot m$

The prefix bits identifies the event, while the data bits that are added in events E_2 and E_3 identify a particular block. In the case of event E_3 the $n\cdot m$ bits are the values of the $n\cdot m$ pels of the block. This code is UD (Uniquely Decodable) code since it satisfy the prefix condition.

The bit rate per pel, using $n\times m$ pel blocks is

$$r=\frac{1}{n\cdot m}\,[p(E_1)+(2+q)\cdot p(E_2)+(2+n\cdot m)\cdot p(E_3)] \text{ bit/pel} \quad (4)$$

In this coding technique the subset $\mathbf{B}$ of 2^q blocks is taken as the most frequent blocks from the training images.

The average bit rate for the five training images is 0.281 bit/pel for q=4, and 0.270 bit/pel for q=5 (compression ratios of 3.6:1 and 3.7:1 respectively). The compression ratio obtained for the five test images is 3.1:1 and 3.2:1 for q=4 and q=5 respectively. For some images a higher compression ratio could possibly be achieved by increasing q, but the coding scheme becomes more complicated for implementation.

The above mentioned compression ratios are better than those reported in [4] and are the same as those quoted in [2], but here with a very simple coding procedure.

5. COMBINING LOSSY AND VARIABLE-LENGTH CODING

A further increase in compression ratio can be obtained by combining the techniques described in sections 3 and 4. Namely, representing the original dither image using a dictionary of 2^q dither blocks and then coding the new dither image using the proposed variable-length block- coding technique.

In this method only events E_1 and E_2 can occur. Accordingly, only one prefix bit (of value '1') is assigned to identify event E_2.

The average bit rate that is obtained for the training images is 0.169 bit/pel for q=4 (16 blocks dictionary), and 0.201 bit/pel for q=5 (32 blocks dictionary), i.e. compression ratios of 5.9:1 and 5:1, respectively. The compression ratios that are obtained for the test images are 5.3:1 and 4.4:1 for q=4 and q=5, respectively.

The dictionary contents in these cases is the most 2^q frequent blocks in the training images, and the match-distance is used as the distance measure. The subjective observations of the reproduced training and test dither images, for the mentioned above dictionary sizes, are as discused in section 3.3.

6. SUMMARY AND CONCLUSIONS

Lossy and lossless block-coding methods of dithered image were presented.

In the lossy coding (block-quantiziation) method we applied the LBG algorithm for the dictionary design, and used the 'match-distance' as a distance measure. For several cases the final dictionaries which resulted from the LBG algorithm achieved a lower distortion than those suggested by a previous approach. However, the results of this paper justifies the earlier approach. Namely, using the most frequent blocks as a dictionary. A careful examination of the reconstructed dithered images shows that better results are obtained with the new distance measure. The compression ratio that was achieved was 3.2:1 without noticeable degradation in the reproduced images.

In the lossless coding method, the proposed variable-length block-coding technique resulted in an average compression ratio of 3.7:1 with a simple coding procedure.

A further increase in the compression ratio was obtained by combining the lossy and the variable-length lossless coding techniques. A compression ratio of about 5:1 was achieved by this method without any visible degradation in the quality of the reconstructed images.

REFERENCES

[1] J.C.Stofel, J.F.Moreland, "A Survey of Electronic Techniques for Pictorial Image Reproduction," IEEE Trans. Commun., vol. COM-29, pp. 1898-1925, Dec 1981.

[2] Yasuhiko Yasuda et al. ,"Advances in FAX," Proc. IEEE , pp. 706-730, April 1985.

[3] B.E.Bayer, "An Optimum Method for Two-Level Rendition of Continuous Tone Pictures," IEEE ICC 1973, pp. 26.11-26.15 .

[4] C.N.Judice, A.B.White, R.L.Johnson, . "Transmission and Storage of Dither Coded Images Using 2-D Pattern Matching," Proc. of the SID ,Vol. 17/2, pp. 85-91, Second Quarter 1976.

[5] R.M.Gray, "Vector Quantization," IEEE ASSP Magazine pp. 4-29, April 1984.

[6] Y.Linde, A.Buzo, R.M.Gray, "An Algorithm for Vector Quantizer Design," IEEE Trans. Commun. , vol. COM-28, pp. 84-95, January 1980.

[7] V.R.Udpikar, J.P.Raina, "BTC Image Coding Using Vector Quantization," IEEE Trans. Commun., vol. COM-35, pp. 352-356, Mar 1987.

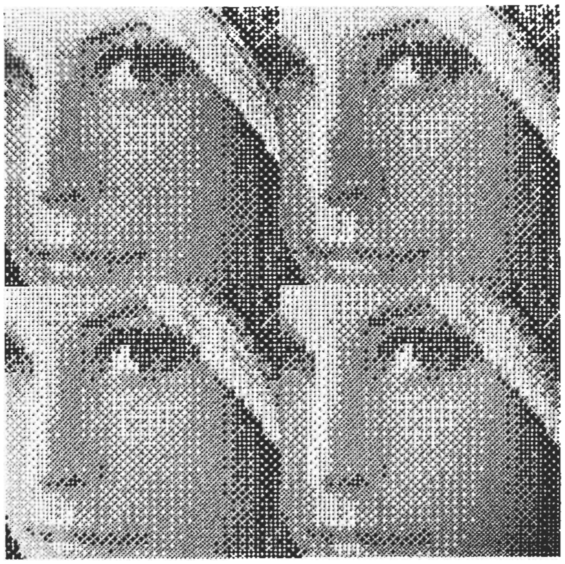
[8] M.Werman, S.Peleg, A.Rosenfeld, "A Distance Metric for Multidimensional Histograms," Computer Vision, Graphics, and Image Processing, Vol 32, pp. 328-336, 1985.

[9] C.N.Judice, "Dithervision - A Collection of Techniques for Displaying Continuous Tone Still and Animated Pictures on Bilevel Displays," SID Symposium Digest, 1975.

[10] R.Hunter, A.H.Robinson, "International Digital Facsimile Standards," Proc. IEEE , vol 68, no 7, pp. 854-867, July 1980.

FIGURE 1: Original and reproduced dither images:

(a) Original dithered image.



(b) Zoomed regions, from top left in clockwise direction: (1) original image. (2) 32 block dictionary, match distance. (3) 16 block dictionary, match distance. (4) 16 block dictionary, Hamming distance.

(c) 8 Most frequent blocks dictionary, Hamming distance measure.

(d) 8 Blocks dictionary obtained by the LBG algorithm, match-distance measure.