Short Paper

Inter-Subband Correlation for Predictive Pyramid Coding

CHUNG J. KUO AND CHIA L. LEE
Signal and Media (SAM) Laboratory
Graduate Institute of Communication Engineering
National Chung Cheng University
Chiayi, 621 Taiwan

Recently, fractal and subband coding were combined to achieve good picture quality at very low bit rates, and the technique was named predictive pyramid coding (PPC) [5]. In this paper, we improve PPC performance by utilizing the inter-subband correlation between different groups of subbands. As a result, we can reduce the bit rate by 0.011—0.043 bpp without degrading the coded picture quality, or we can increase the coded picture quality by 0.29—0.63 dB without increasing the bit rate.

Keywords: fractal coding, subband coding, predictive pyramid coding, prediction, quadrature mirror filter

1. INTRODUCTION

Coding images at very low bit rates is always desirable due to the need to transmit images across a bandwidth-limited channel. Among existing coding techniques for images, subband and wavelet coding have received much attention due to their ability to code images at different resolutions and very low bit rates. On the other hand, fractal models are also effective for image coding and fractal coding (which exploits the self-similarity of images) is also suitable for low bit rate applications.

In [5], a predictive pyramid coding (PPC) is proposed to exploit similarities in the multiresolution decomposition of an image. An image is first decomposed into different subbands. The lowest subband is then PCM encoded while other frequency subbands are fractal encoded. Unlike the conventional fractal codec, no contraction is performed on the domain block because the domain and range block are of the same size. As a result, the proposed PPC decoder does not require an iteration process, and the decoding speed is significantly increased. PPC is indeed a delicate combination of the two well-known techniques—fractal and subband coding.

It is well known that similarity exists between different subbands. Therefore, some researchers have tried to utilize the inter-subband correlation in the vertical or horizontal direction to improve the efficiency of subband coding [4, 6, 7, 9]. For example, they have tried to estimate the nth vertical (or horizontal) subband based on the \((n - 1)\)th vertical (or horizontal) subband. However, the relationship between the diagonal and vertical (or
horizontal) subbands has not been used.

In this paper, we study the inter-subband correlation to find the relationship between the diagonal subband and the subbands in the vertical and horizontal directions. With this relationship, the diagonal subband can be effectively estimated (at the decoder) from the vertical and horizontal subbands with a certain degree of accuracy. In this case, the coding efficiency of PPC is improved because the fractal code for the diagonal subband is not needed.

The organization of this work is as follows. Section 2 reviews the key ideas in PPC. The proposed coding scheme and the simulation results are presented in Sections 3 and 4, respectively. Finally, a brief conclusion is given in Section 5.

2. REVIEW OF PREDICTIVE PYRAMID CODING

In PPC, multilevel pyramid decomposition of an input image \( y = y_0 \) is first achieved. At each decomposition level \( i \), the image \( y_i \) is decomposed into four subimages \( y_i^h, y_i^h, y_i^h, y_i^h \) for \( i = 1, 2, \ldots, 5 \) by using nine-tap symmetric quadrature mirror filters (QMF) described in [8] and the symmetric extension method [10]. For example, an input image \( y_0 \) is first split into four subimages \( y_1^h, y_1^h, y_1^h, y_1^h \), and \( y_2^h \), where the superscript letters (l or h) denote the row-column (low- or high-pass) filtering operation performed on input image \( y_0 \). Specifically, subband \( y_1^h \) is obtained by lowpass filtering the rows and highpass filtering the columns of image \( y \), followed by a factor of two subsampling in the horizontal and vertical directions.

After multilevel pyramid decomposition is achieved, an eight-bit uniform quantizer and seven-bit Laplacian quantizer are used to quantize subband \( y_5^h \) and subbands \( y_4^h, y_3^h \) and \( y_3^h \), respectively. The remaining subbands are then encoded using a block prediction scheme described below.

In PPC, level \( i \) subband \( y_i^h, y_i^h, y_i^h \) or \( y_i^h \) is first divided into a set of nonoverlapping range blocks \( r \) whose sizes range from \( 4 \times 4 \) to \( 2^{i/2} \times 2^{i/2} \). For example, the block size can be \( 4 \times 4, 8 \times 8 \) or \( 16 \times 16 \) for a level 2 image. For each range block \( r \) in the subband \( y_i^h \), or \( y_i^h \), we locate the domain block \( d \) in the corresponding subband \( y_i^h, y_i^h, y_i^h, y_i^h \) such that the fractal mapping minimizes the mean-squared error \( D(r, f) \) with respect to all the possible choices of isometries and contrast scaling \( \alpha \). An important feature of PPC is that the range block and domain block are of the same size. In this case, contraction of the domain block is not needed in the fractal codec.

The original fractal coding has eight different choices for isometry. In PPC, only rotations around the center of a block, through 0°, 90°, 180° and 270°, are considered. In this case, two bits are enough to represent isometry code. For each of the four rotation angles, the optimal contrast scaling factor is chosen so as to minimize the distortion, and then a six-bit Laplacian quantizer is specifically designed for the purpose of quantization. As for the position of the domain block, it is represented by means of eight-bit resolution.

It is possible that some \( 4 \times 4 \) blocks will have large quantization errors as a result of the block prediction described above. That is, the quantization error may be larger than a threshold value \( P \) that is predefined by the system designer. In this case, each coefficient in these \( 4 \times 4 \) blocks is coded by using a Laplacian quantizer. As for the number of bits needed to code each coefficient, it is decided by using an optimal bit allocation strategy described in [3].
3. PROPOSED CODING SCHEME

Several schemes [4, 6, 7] based on horizontal (or vertical) inter-subband correlation have been proposed to improve the efficiency of subband coding. In this paper, we propose another scheme to exploit inter-subband correlation. It is well known that different groups of subbands contain edge information in different directions. For example, subbands $y_{lh}^i$ and $y_{hl}^i$ (for $i = 1, 2, \ldots, 5$) contain edge information in the vertical and horizontal directions, respectively, and subband $y_{hh}^i$ has edge information in the 45° and 135° directions. Since edges in any direction exhibit edge strength in both the vertical and horizontal directions, we can find edge orientation in any direction (such as 45° and 135°) by using only two 1-D (vertical and horizontal) edge detectors [1].

Here we exploit the above idea to further reduce the bit rate required by PPC. Our key idea is to estimate the edge strength in the 45° and 135° directions based on the edge strength in the vertical and horizontal directions. Equivalently, we estimate the diagonal subband based on the horizontal and vertical subbands that are available.

In PPC, each directional (horizontal, vertical and diagonal) subband is fractal encoded independently. Therefore, our procedure is to fractal encode the horizontal and vertical subbands first, and then uses the encoded subbands to estimate the diagonal subband.

Tables 1-3 in [5] report a summary of coding results (0.26, 0.36 and 0.25 bpp) for “Lena,” “Building” and “Clown,” respectively. The number of bits required to encode each 4 × 4, 8 × 8, 16 × 16 and 32 × 32 block is given there. According to these tables, subband $y_{lh}^1$ was set to zero during encoding so no bit was allocated to it. In this case, it was not needed to estimate diagonal subband $y_{hh}^1$.

On the other hand, $y_{lh}^2$ subbands (i.e., $y_{lh}^1$, $y_{lh}^2$, and $y_{hh}^2$) are all required to encode images at low bit rates. The following estimation rule is proposed to predict the diagonal subbands $y_{hh}^2$:

$$y_{hh}^2 = a_{lh} \hat{y}_{lh}^2 + a_{hl} \hat{y}_{hl}^2 + a_{hh} \hat{y}_{hh}^2,$$

where $\hat{y}_{lh}^2$ and $\hat{y}_{hl}^2$ denote the fractal encoded subbands of $y_{lh}^2$ and $y_{hl}^2$, respectively, and $\hat{y}_{hh}^2$ denotes the estimated subband. Apparently, $a_{lh}$ and $a_{hl}$ are image and QMF dependent. Since $y_{hh}^2$ is the diagonal subband being used in PPC that has the most frequency information, its value must be the smallest one compared to other diagonal subbands, such as $y_{hh}^1$ and $y_{hh}^3$. Therefore, a moderate error in $y_{hh}^2$ should not have a significant impact on image quality.

For a given 512 × 512 image, we can generate subbands $y_{lh}^2$, $y_{hl}^2$ and $y_{hh}^2$ with size 128×128. Based on Eq. 1, we have

$$\begin{bmatrix} y_{lh}^{2,1} & y_{lh}^{2,2} & \ldots & y_{lh}^{2,16384} \\ y_{hl}^{2,1} & y_{hl}^{2,2} & \ldots & y_{hl}^{2,16384} \\ y_{hh}^{2,1} & y_{hh}^{2,2} & \ldots & y_{hh}^{2,16384} \end{bmatrix} = \begin{bmatrix} a_{lh} & a_{hl} \end{bmatrix} \begin{bmatrix} \hat{y}_{lh}^{2,1} & \hat{y}_{lh}^{2,2} & \ldots & \hat{y}_{lh}^{2,16384} \\ \hat{y}_{hl}^{2,1} & \hat{y}_{hl}^{2,2} & \ldots & \hat{y}_{hl}^{2,16384} \\ \hat{y}_{hh}^{2,1} & \hat{y}_{hh}^{2,2} & \ldots & \hat{y}_{hh}^{2,16384} \end{bmatrix}.$$

where $y_{2,j}^m$ denotes the $j$th element in subband $y_{m}^2$. Since there are 16,384 equalities with two unknown parameters ($a_{lh}$ and $a_{hl}$) in Eq. 2, the least-squared solution of these two parameters can be found such that subband $y_{hh}^2$ can be estimated from the fractal encoded subbands $\hat{y}_{lh}^2$ and $\hat{y}_{hl}^2$. 
The above estimation scheme can be directly applied to other diagonal subbands, such as \( y_{3h}^{hh} \) and \( y_{4h}^{hh} \). However, the estimation error turns out to be large, and the image quality is visually degraded. Therefore, the estimation process is only applied to the diagonal subband \( y_{2h}^{hh} \).

**4. SIMULATION RESULTS**

Simulation results are reported in this section. Values of \( a_{lh} \) and \( a_{hl} \) for four different 512 \( \times \) 512 images (Lena, F-16, Building and Elaine) are shown in Table 1. Apparently, \( a_{lh} \) and \( a_{hl} \) are image dependent and must be found for each image to be encoded. To make these two parameters available at the decoder, 64 extra bits must be used. As a result, the bit rate of the coded image is increased by 0.000244 bpp which is very low and can be ignored. The advantage is that no fractal code is needed to represent subband \( y_{2h}^{hh} \) while a significant number of bits are saved (or, equivalently, the bit rate of the encoded image is reduced).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>( a_{lh} )</th>
<th>( a_{hl} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lena</td>
<td>(-1.1954 \times 10^{-2})</td>
<td>(-2.0172 \times 10^{-2})</td>
</tr>
<tr>
<td>F-16</td>
<td>(-1.6655 \times 10^{-3})</td>
<td>(-6.4192 \times 10^{-3})</td>
</tr>
<tr>
<td>Building</td>
<td>(+5.3426 \times 10^{-3})</td>
<td>(+1.6580 \times 10^{-3})</td>
</tr>
<tr>
<td>Elaine</td>
<td>(-8.1079 \times 10^{-3})</td>
<td>(-8.0170 \times 10^{-3})</td>
</tr>
</tbody>
</table>

Although \( y_{2h}^{hh} \) can be estimated using our scheme, difference exists between the original and estimated subbands (\( \tilde{y}_{2h}^{hh} \) and \( \hat{y}_{2h}^{hh} \)). According to our simulation, the mean-squared error between the original image and the image reconstructed using the estimated subband \( \tilde{y}_{2h}^{hh} \) and other original subbands respectively, is only 0.445. (This error is the average result from the images shown in Fig. 1.) Therefore, the estimation scheme proposed here provides very good prediction results. At the same time, we find that the error increases and becomes visible when subband \( y_{3h}^{hh} \) is estimated using our proposed scheme. Therefore, only subband \( y_{2h}^{hh} \) is estimated here.

A visual comparison of the coded Lena images (obtained using PPC and the proposed modified version) is shown in Fig. 2. Here, the two coded images have almost the same PSNR and visual quality; however, our proposed scheme can reduce the bit rate by 11%. Apparently, estimating the diagonal subband according to the horizontal and vertical subbands is a feasible approach. Fig. 3 shows the PSNR and bit rate of the PPC and the proposed scheme for tested images. Our estimation scheme can reduce the bit rate of the coded image by 0.011 – 0.043 bpp with the same image quality, or it can increase the PSNR of the coded image by 0.29–0.63 dB with the same bit rate. Although our scheme assumes that the image has negligible energy content at high frequencies, which may not be true, Fig. 3 shows that our scheme works pretty well for an image (such as Building) with many details. On average, the proposed scheme can reduce the bit rate by 0.025, 0.029, 0.029, and 0.024 bpp for the images, Lena, F-16, Building and Elaine, respectively.
Fig. 1. Images used for simulation. (a) Lena. (b) F-16. (c) Building. (d) Elaine.

Fig. 2. Reconstructed Lena images. (a) Proposed method at 0.235 bpp (PSNR = 32.77 dB). (b) PPC at 0.263 bpp (PSNR = 32.78 dB).
In our scheme, the least-squared solution of $a_{lh}$ and $a_{hl}$ must be found and transmitted. Since 16,384 equations are needed to determine $a_{lh}$ and $a_{hl}$ for an image, the computational process requires a large amount of computation. On the other hand, $h_{hy}^2$ can be easily found at the decoder once $h_{ly}^2$ and $l_{ly}^2$ are available. As for PPC, fractal encoding of subband $h_{hy}^2$ requires more computations than does decoding because of the need to search for the best fractal parameters. On average, we find that the encoding and decoding time of the proposed technique is about the same as that of PPC.

5. CONCLUSIONS

Inter-subband correlation has been exploited in this paper to improve the coding performance of the predictive pyramid coding (PPC) scheme proposed in [5]. Here, the diagonal subband is estimated based on the subbands in the vertical and horizontal directions. Although the parameters for estimation are image dependent, the proposed scheme reduces the bit rate of the coded image by 0.011–0.042 bpp with the same coded image quality or improves the coded image quality by 0.29–0.63 dB without increasing the bit rate.
ACKNOWLEDGEMENT

This research was partly supported by the National Science Council, Taiwan, under contract NSC-87-2213-E-194-027. The authors also thank the anonymous reviewers for their valuable comments.

REFERENCES


Chung J. Kuo (郭鍾榮) received B.S. and M.S. degree in Power Mechanical Engineering from National Tsing Hua University, Taiwan, in 1982 and 1984, respectively, and PhD degree in Electrical Engineering (EE) from Michigan State University (MSU) in 1990. He joined EE Department of National Chung Cheng University (NCCU) in 1990 as an associate professor and then became a full professor in 1996. He is now the chairman of Graduate Institute of Communications Engineering of NCCU. Dr. Kuo was a visiting scientist at Opto-Electronics & System Lab, Industrial Technology Research Institute in 1991 and IBM T.J. Watson Research Center from 1997 to 1998 and a consultant to several international/local companies.

Dr. Kuo interests in image/video signal processing, VLSI signal processing, and optical information processing/computing. He is the co-director of the Signal and Media (SAM) Labs. At NCCU. Dr. Kuo received the Distinguished Research Award from Na-
tional Chung Cheng University in 1998, Overseas Research Fellowship from National Science Council (NSC) in 1997, Outstanding Research Award from College of Engineering, NCCU in 1997, Medal of Honor from NCCU in 1995, Research Award from NSC every year since 1991, Best Engineering Paper Award from Taiwan’s Computer Society in 1991, EE Fellowship from MSU in 1989, and Outstanding Academia Achievement Award from MSU in 1987. He was a guest editor for two special sections of Optical Engineering and 3D Holographic Imaging (to be published by John, Wiley and Sons) and an invited speaker and program committee chairman and member for several international/local conferences. He also serves as an Associate Editor of IEEE Signal Processing Magazine and President of SPIE Taiwan Chapter (1998-2000). Dr. Kuo is a senior member of IEEE and a member of Phi Kappa Phi, Phi Beta Delta, OSA, and SPIE and listed in Who’s Who in the World.

Chia-Ju Lee (李佳儒) was born in Chiayi, Taiwan on August 4, 1973. He received B.S. degree in Electrical Engineering from the Feng-Chia University in 1995 and M.S. in Electrical Engineering from the National Chung Cheng University in 1997. Currently, he is working on server design at the ACER Computer corporation. His research interests include digital image and video processing and the applications of fractals.