A TWO LAYER HYBRID IMAGE CODER

LEVICKÝ D., MATÚŠ E. and KRÁL P.
Department of Radioelectronics
Technical University in Košice
Park Komenského 13
040 01 Košice
Slovak Republic
E-Mail: matus@ccsun.tuke.sk

Abstract

This paper presents a new idea of a two layer hybrid image coder based on visual pattern image coding of original image and residual image by means of the wavelet transform. At first contours are extracted from the original image. As a contour extractor a VPIC coder is used. In the second, the residual image is computed and coded by the wavelet transform. At the decoder side the sum of contours and residual (texture) image parts is made to obtain the reconstructed image. Some results of coding simulation on picture Lena (256x256x8) are presented. The proposed coding technique is well suitable for image coding and progressive image transmission.

Keywords:

Image coding. Progressive image transmission. Wavelet transform. Vector quantization.

1. Introduction

The canonical digital representation of an image is very ineffective and requires a very large number of bits. The main goal of image coding is reduce, as much as possible, the number of bits necessary to represent and reconstruct an image with defined quality. Kunt in [1] has divided image coding techniques into two generations. The first generation methods are based on the principles of classical information and coding theory. Image coding by using the above mentioned methods is based on exploiting of the statistical redundancy and subjective irrelevancy in the statistical sense. Major techniques of this generation are DM, PCM, DPCM, bit-plane coding, interpolative coding, transform coding and hybrid coding. The most effective techniques decompose the image into an orthogonal basis. The basis is chosen in order to minimise the entropy of decomposition coefficients. Several orthogonal bases were used for this purpose: the Karhunen-Loĉve [2], the discrete cosine transform [2] and different orthogonal subband decomposition [3],[4]. Due to nonstationarity of image signals and specific properties of the human visual system (HVS) the first group of these methods are limited in achieving high compression ratio. The second generation methods are based on eliminating of redundancy that results from the structural properties of the images, and their perceiving by HVS. The well-known are contour-texture oriented techniques, directional decomposition based coding and pyramidal coding [1].

In this paper was proposed a two layer hybrid image coder based on visual pattern image coding of original image and the wavelet transform of residual one. In the section 2 the basic principle of this coder is described. In the parts 3 and 4 the main blocks of this coder are described. Part 5 discuses about decoding problems and in the next section the experimental results are summarised.

2. Hybrid VPIC-WT Coder

The basic idea in the contour-texture coding is to separate image parts with different structural properties in the perceptual sense. It divides image information in two categories: contours (edges) and texture. Contours are very important for image understanding. Textures are less structured then contours and texture distortion is less visible then contour distortion. So, one can use different coding techniques for each part. The scheme of the contour-texture coder is in Fig.1. At first, contours X_C are

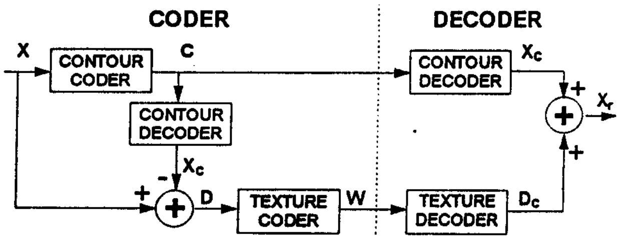

Fig.1
A block scheme of the contour-texture coder-decoder

extracted from the original image X. It is generally a difficult (usually iterative) non-linear problem and requires very much number of operations [6],[7]. This reason limits the practical implementation of such coder. At second, residual image D, witch contains texture information is computed and coded. In the decoder, the sum of the contour and texture image parts (X_C+D_C) is made to obtain reconstructed image X_r. Due to computational complexity of proposed technique, we have modified this

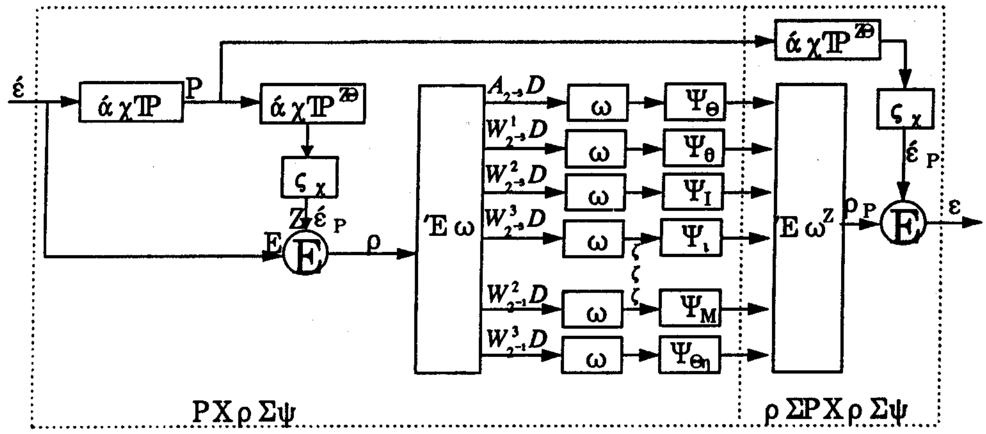

Fig.2
A block scheme of the two layer hybrid VPIC-WT coder.

method by using of suboptimal operations. As contour extractor we have used visual pattern image coder (VPIC) [10] and texture information is coded with orthogonal wavelet transform (WT). The block scheme of the proposed (VPIC-WT) coder is in Fig.2.

In the proposed technique the VPIC coder is used as a contour extractor. Coded information C is transmitted to the receiver and then the inverse VPIC operation is made (look up table operations). A smoothing postfilter H_P is used to eliminate blocking effect caused by the VPIC coder. Residual image D is computed as difference between original image X and VPIC coded image X_C and then the three-level dyadic wavelet transform of residual image D is computed. Block thresholding operation on the wavelet coefficients is performed to remove no texture information in the residual image. Finally, quantization of the wavelet coefficients is made via optimal scalar quantization. On the receiver end inverse WT and VPIC is computed and sum of the parts $X_C + D_C$ is made.

3. Visual pattern image coder

The VPIC coder is used as a main image feature extractor. We have used modified algorithm of the VPIC coder that can be classified as a conditional vector quantization. The main advantage over contour extractors used in [4],[5] is in the very fast decoding. Due to this fact, the complexity of decoder will decrease, but the VPIC coder is not ideal contour extractor and for very low bit rates the blocking effect appears.

The modified VPIC coder uses the set of visual patterns. Visual patterns were chosen based on a simply viewing geometry model [5]. Some example of visual patterns is in Fig.3 and the block scheme of the modified VPIC coder is in Fig.4.

The vectors of the codebook are ordered into S classes so that the vectors of the j-th class corresponds to the j-th visual pattern. For the image block **x** to be coded the signed block **r** is calculated by the following way:

$$r_i = \begin{cases} -1 & \text{if } x_i < m, i = 1,2,..M \\ +1 & \text{otherwise} \end{cases} \quad (1)$$

where m is the mean block intensity and M is the dimension of the code vector.

The signed block **r** is mapped onto the set of visual patterns by the neural network Hamming classifier with the following measure

$$d_j = \sum_{i=1}^{M} r_i . p_{ij} \quad \text{for } j = 1,2,...,S \qquad (2)$$

The maximum value d_L identifies the pattern L with the highest correlation to the signed block **r**.

Very important task in the VPIC coding is the codebook design. In the proposed method the codebook was designed by the modified Kohonen's LVQ algorithm [9],[10]. The codebook vectors are initialised by random values at first and then the vectors of the j-th class

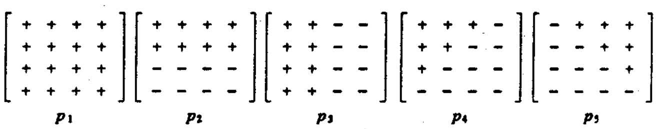

Fig. 3
Basic visual patterns, the signs + and - denote the intensity of pixels considering the mean block intensity

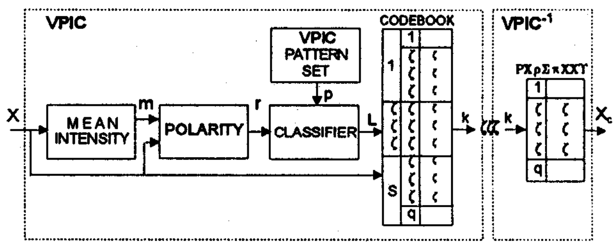

Fig.4
A block scheme of the modified VPIC coder- decoder

correspond to the j-th visual pattern. For the image block **x** the code vector $\mathbf{v}_k$ with minimal Euclidean distance in the

L-th visual class is searched. In the training phase the code vector $\mathbf{v}_k$ is adapted by the following relation

$$v_{ik}(t+1) = v_{ik}(t) + \alpha(t).[x_i(t) - v_{ik}(t)] \tag{3}$$

where $i = 1, 2, ...M$ and $\alpha(t) \in <0, 1>$ is the non-increasing function of time. During decoding the indices L and k were used for simple look up in the codebook.

4. Wavelet Transform of the Residual Image

Wavelet transform is powerful tool in many signal processing applications. Due to its decorrelation and energy compaction property it is very important in the signal compression and in the field of multiresolution signal decomposition. Discrete approximation $A_{2^j}D(m,n)$ of an image $D(u,v)(m,n \in Z; u,v \in R)$ at resolution 2^j ($j \in Z$) can be computed by dilating of so called scaling function $\phi(u,v)$ by factor 2^j and its translating over grid 2^{-j}:

$$A_{2^j}D(m,n) = \langle D(u,v), \phi(2^j u - m, 2^j v - n) \rangle \tag{4}$$

where $\langle . \rangle$ is the scalar product. Set of the functions $(\phi(2^j u - m, 2^j v - n))_{m,n \in Z}$ is an orthogonal basis. Difference between the approximation of the image $D(u,v)$ at resolutions 2^{j+1} and 2^j (high frequencies) is represented by discrete detail signals $W^k_{2^j}D(m,n)$ where $k = 1, 2, 3, ...$ is number of detail signals. It can be extracted by the orthogonal wavelet transform

$$W^k_{2^j}D(m,n) = \langle D(u,v), \psi^k(2^j u - m, 2^j v - n) \rangle . \tag{5}$$

The orthogonal basis functions $\psi^k(2^j u - m, 2^j v - n)$ are derived from the mother wavelet function $\psi^k(u,v)$. In our case, we have used the separable model in two dimensions, so we can build the mother wavelets by combination one dimensional scaling and wavelet functions $\phi(u), \psi(u)$, $u \in R$:

$$\psi^1(u,v) = \phi(u)\psi(v) \tag{6a}$$

$$\psi^2(u,v) = \phi(v)\psi(u) \tag{6b}$$

$$\psi^3(u,v) = \psi(u)\psi(v) \tag{6c}$$

Discrete approximation $A_{2^j}D$ and the detail signals $W^k_{2^j}D$, $k = 1, 2, 3$ of the image D can by computed from the finer approximation $A_{2^{j+1}}D$ by use of the filter bank with one dimensional analysing quadrature mirror filters (QMF) $H_1(z)$ (lowpass) and $H_2(z)$ (highpass) and decimating by factor 2. The principle of this decomposition is in Fig.5. By means of this rule we have computed recursively three-level (j=-1,-2,-3) discrete orthogonal wavelet representation of image D

$$\left(A_{2^{-3}}D, \left(W^k_{2^j}D \right)_{\substack{-3 \le j \le -1 \\ k=1,2,3}} \right) \tag{7}$$

On the opposite side, the synthesis rule is made via $G_1(z)$ and $G_2(z)$ and by putting the zeros between samples. $G_1(z)$ and $G_2(z)$ are conjugate pairs to $H_1(z)$ and $H_2(z)$, respectively. Block scheme of the proposed method is in Fig.6 and the WT coder-decoder is in Fig.7.

The wavelet orthogonal basis decomposes an image into details appearing at different resolutions, and within different spatial orientations [8]. As the prototyping filter was chosen 4 tap FIR QMF filter designed in [4] with maximal energy compaction property and three level WT was performed. At each scale 2^j three components $W^k_{2^j}$, $k = 1, 2, 3$, of size $2^j N$ (image size in pixels) are computed, where N is size of the original image and $j = -1, -2, -3$ is the level number. The total number of samples of this orthogonal representation is equal to the number of samples of the original error image D.

In order to reduce the number of the wavelet coefficients to be coded, we set to zero as many coefficients as possible. After performing of the WT, the block thresholding of the wavelet coefficients is done (blocks T in Fig.2). We divide the wavelet components $W^k_{2^j}D$, $k = 1, 2, 3$, $j = -1, -2, -3$ at each resolution into 8x8 non-overlapped blocks and then energy of each block is computed. The values of pixels in the block are set to zero if energy of the block exceeds of the predefined threshold.

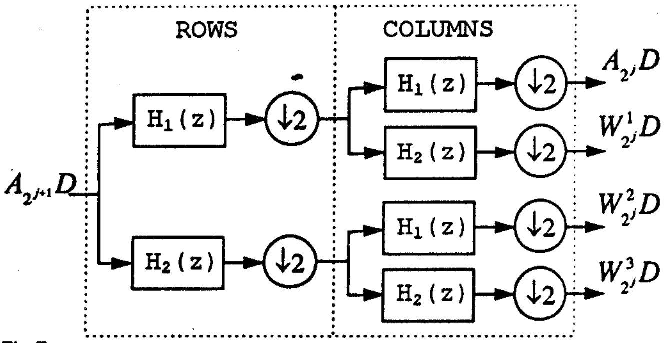

Fig.5
Four band analysis filter bank for fast wavelet transform computing. Filters H1 and H2 are quadrature mirror filter pairs. Circle with down arrow denotes a decimator and the decimation factor.

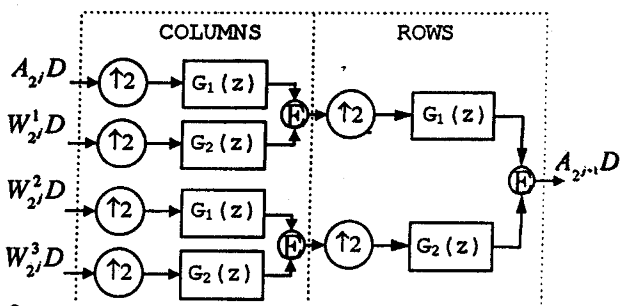

Fig.6
Four band synthesis filter bank for fast inverse wavelet transform computing. Filters G1 , G2 are conjugate to filters H1 and H2 (fig.5), respectively. Circle with up arrow denotes putting zeros between samples and the interpolation factor.

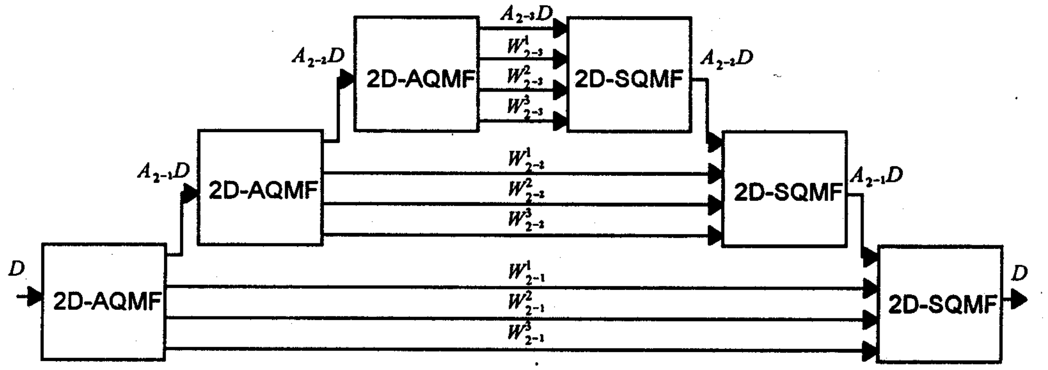

Fig.7
A block scheme of three-level fast direct and inverse wavelet transform computing based on quadrature mirror filter bank. The analysis blocks 2D-AQMF (2 dimensional analysis QMF) and synthesis blocks 2D-SQMF (2 dimensional syntesis QMF) are in Fig.5 and Fig.6, respectively.

This operation removes low visible texture information of the image.

At last, the quantization of the wavelet coefficients is performed. In this quantization the optimal scalar quantizers was used. An exponential quantizer is well adapted to the statistical distribution of the wavelet coefficients, in order to produce a small mean-square distortion. The bit allocation for wavelet coefficients was made on dependence of distortion contribution of each coefficient to overall distortion. More bits are allocated for low spatial frequency components and fewer bits for high spatial frequency components. Bit allocation strategy is based on the principles of marginal analysis theory. In our experiments the full optimal operational rate distortion function was computed, also all optimal bit allocation sets was available.

5. Decoding

In comparison with the algorithms proposed in [6],[7] our method offers very fast image decoding. First, the VPIC coded image X_C is computed via inverse VPIC^{-1} decoding. This represents look up table operations likewise vector quantizer decoding and due to blocking effect the postfiltering by filter H_p is made. This part of the decoder is identical with the coder part. Second, the coded residual image D_C via inverse wavelet transform WT^{-1} [8] is computed and added to the decoded image X_C, so that the reconstructed image X_r is obtained.

6. Simulation

For the simulation the grayscale image Lena of 256x256 pixels was used. Each pixel in this image was coded by 8 bits. The VPIC coder was designed with respect to good contour extraction property and the vector dimension 4x4 and 8x8 was used. The set of visual patterns contained nine visual patterns: mean, horizontal, vertical and diagonal direction extractors and ten code vectors for each visual pattern were assigned. Together 90 code vectors are computed for the VPIC coder. The postfilter H_p was chosen as a smoothing filter in order to eliminate blocking effect. It was shown, that 5 tap Gaussian like FIR filter is well suitable for this application. The QMF filters for fast WT computing were chosen in order to minimise aliasing effect and maximise the energy compaction. We have used 4 tap QMF filter proposed in [4]. The threshold values were chosen from the interval 0-100 and optimal scalar quantizer was computed for scalar quantization of the wavelet coefficients.

Simulation results in the form of the signal to noise ratio (SNR) versus bit rate for various parameters are shown in Fig.8. SNR is computed by the following way

$$SNR = 10\log_{10}\left(\frac{\sum_{i,j} X^2(i,j)}{\sum_{i,j}[X(i,j)-X_r(i,j)]^2}\right). \qquad (8)$$

The bit rate is computed as the sum of achievable bit rate of the VPIC coder and WT coder. As is shown the best results was obtained with the VPIC code vector dimension 8x8 and followed by postfiltering.

The proposed image coding technique is well

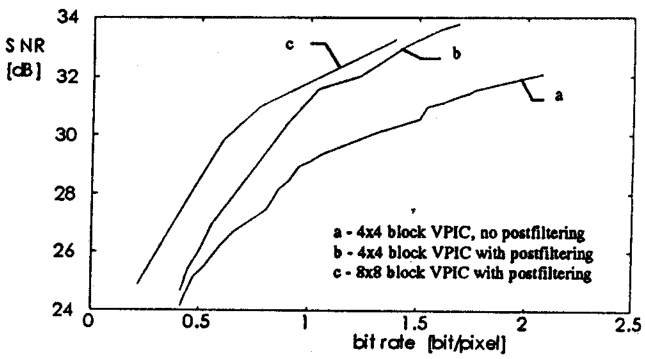

Fig.8
SNR versus bit rate for hybrid VPIC-WT coder

suitable for progressive image transmission. In this method the VPIC coded image C is transmitted at first followed by progressive transmission of elementary levels

$W_{2j}^k D$, $k=1,2,3$, $j=-1,-2,-3$ of wavelet transformed residual image.

7. Conclusion

This paper presents some results of a hybrid image coding technique based on dividing of image information into contours (edges) and texture. A new idea of a two layer image coder is presented. Contours are extracted from original image by means of VPIC coder and the wavelet transform for coding of the residual image is used. The best results in simulation was obtained with the VPIC code vector 8x8 followed by postfiltering.

8. References

[1] Kunt M., Ikonomopulos A., Kocher M. : Second-Generation Image-Coding Techniques. Proceedings of the IEEE, Vol. 73, No. 4, pp. 549-572, April 1985

[2] Jayant N.S., Noll P. : Digital Coding of Waveforms. Prentice-Hall, Inc., Englewoods Cliffs, New Jersey 1984

[3] Woods J.W., O'Neil S.D.: Subband Coding of Images. IEEE TR-ASSP, Vol.34, pp. 1278-1288, Oct. 1986

[4] Akansu A.N., Haddad R.A.: Multiresolution Signal Decomposition. Academic Press, Inc., New York, 1993

[5] Chen D., Bovik A. : Visual Pattern Image Coding. IEEE TR-COM, Vol. 38, No. 12, pp 2137-2146, Dec. 1990,

[6] Carlsson S. : Sketch Based Coding of Grey Level Images. Signal Processing NorthHoland, Vol. 15, No. 1, pp. 57-83, 1980

[7] Froment J., Mallat S.: Second Generation Compact Image Coding with Wavelets. C.K.Chui ed.: Wavelets - A tutorial in theory and aplications. Academic Press, Inc., pp. 655-678, 1992

[8] Mallat S.G. : A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. IEEE TR-PAMI, Vol.11, No.7, pp. 674-692, July 1989

[9] Kohonen T.: Self-Organization and associative memory. Third edition, Berlin: Springer-Verlag, 1989

[10] Levický D., Král P. :Visual Pattern Image Coding Using Neural Network. Proceedings of the Scientific Conference with International Participation Košice, sept. 1994, pp.28-32

About authors,...

Dušan Levický was born in Slanec (Slovak Republic) in 1948. He received the M.S. degree and PhD. degree at Technical University in Košice and now he is proffesor at the Department of Radioelectronics Faculty of Electrical Engineering and Informatics in Košice and his research interests include digital image processing, image transmission and neural networks.

Emil Matúš was born in Gelnica, Slovak Republic in 1965. He received the M.S. degree in electrical engineering from TU Košice in 1988. In 1988 he joined Technical University in Košice. He was visit researcher at TU Erlangen, Germany in 1992-1993. Currently he is the research assistent in TU Košice. His research interests are image coding, progressive image transmission, wavelet transform.

Peter Král was born in Partizánske (Slovak Republic) in 1969. He received the M.S. degree at the Technical University in Košice. Currently he is the PhD. student at the Department of Radioelectronics the Faculty of Electrical Engineering and Informatics in Košice. His research interests are image coding and neural networks.