M2.18

AN ADAPTIVE PYRAMID IMAGE CODING SYSTEM

Sebastià Sallent-Ribes and Lluis Torres-Urgell

E.T.S.I. Telecomunicacion
08034 Barcelona, Spain

ABSTRACT

The Laplacian Pyramid is a new and efficient method for image encoding. The method is of increasing interest as band-pass pyramids and multiresolution images are being used in other image processing applications. The Pyramid Coding system is a hierarchical structure well suited for progressive image transmission over low-speed channels and hierarchical image retrieving in computerized image storage.

The goal if this paper is to present a new pyramidal image structure that uses algorithms defined on arbitrary non-rectangular sampling lattices. An adaptive technique is also developed that assigns to the pyramidal structure an optimum sampling lattice as a function of the frequency content of the original image.

1. INTRODUCTION

Burt and Adelson have proposed an efficient Pyramidal Coding method |1|. Their approach uses the Laplacian Pyramid to produce an approximate frequency decomposition using algorithms defined on rectangular sampling lattices.

The Pyramidal Coding is a predictive and non-causal method. The original image I_0 is predicted by a low-pass version I_1 of it. I_n is formed using local weighting averaging with unimodal Gaussian-like bidimensional operators. The prediction error is given by

$$E_0 = I_0 - I_1 \tag{1}$$

Coding E_0 and I_1 is equivalent to directly code the picture itself. Data compression is achieved because I_1 is a low-pass image what implies that the sampling density may be reduced. E_0 is a decorrelated error high-pass image, therefore, it may be represented with less bits, because the low sensitivity of the eye at these frequencies. The Gaussian and Laplacian pyramidal structures, are constructed based on both the low-pass versions $\{I_1\}$ and error images respectively. These images are defined on rectangular lattices.

The coding efficiency can be improved since the method is independent of the image content, i.e., is not adaptive and the pyramidal structure is always built from data sampled on rectangular lattices.

More accurate prediction and, hence, greater data compression is achieved if for each original image, the low-pass images I_1 are defined on appropiate bidimensional sampling lattices thus compacting the content information and minimizing the redundancy of the error images.

In this paper, non-rectangular sampling lattices are applied to construct the pyramidal coding. The basic idea is to find for each image a suitable sampling lattice associated to the pyramidal structure to provide the best information compaction. That is accomplished by using an adecuate sampling lattice in such a way that the Gaussian and Laplacian pyramids can be represented with minumum entropy and minimum physical support.

The new pyramidal structure is built using the fast algorithms described in |2| which have been generalized for arbitrary sampling lattices. These algorithms use unimodal non-separable generating nucleus similar to normal distributions.

A frequency-domain approach based upon both the frequency content of the original image, and the shape of the reciprocal unit cell associated to a certain sampling lattice, assigns the

CH2561-9/88/0000-0808 $1.00 © 1988 IEEE

optimal sampling lattice to pyramidal coding.

We combine both the Laplacian Pyramid coding and the bidimensional sampling techniques |3| resulting in a new pyramidal structure having better performance than other coding schemes |4|, |5|, |6| for lower bit rates.

In section II we briefly discuss a summary of the generalization of the Laplacian Pyramid based on the Generalized Hierarchical Discrete Correlation, GHDC. In section III the adaptive scheme is presented and results are discussed. Finally in section IV conclusions are drawn.

2. NON-RECTANGULAR LAPLACIAN CODING

The generalized pyramid coding with an associated sampling lattice defined by a matrix M, is based on the transmission of a set of band-pass images $\{E_l\}$ defined on arbitrary sampling lattices M^l. These images are stacked as a secuence of decreasing dimensions thus constituting the pyramid data structure. The set of band-pass images can be obtained convolving the original image with D.O.G. functions |7|.

D.O.G. functions are defined as the difference between bidimensional normal distributions of different variances $\bar{\sigma}_1$, $\bar{\sigma}_2$ and parameters r_1 r_2. The bidimensional normal distribution is given by

$$G(x,y) = \frac{1}{\sqrt{2\pi}\sigma_{ix}\sigma_{iy}\sqrt{1-r_i^2}}$$

$$\exp\{-\frac{1}{2(1-r_i^2)}\left|\frac{x^2}{\sigma_{ix}^2} - \frac{r_i xy}{\sigma_{ix}\sigma_{iy}} - \frac{y^2}{\sigma_{iy}^2}\right|\} \qquad (2)$$

where $\bar{\sigma}_i = (\sigma_{ix}, \sigma_{iy}) \quad 1 \leqslant i \leqslant 2$

The normal distribution can be closely approximated by an equivalent nucleus $N_l(\bar{x})$ built with a generalized bidimensional version of the Hierarchical Discrete Correlation associated to a sampling lattice characterized by a 2x2 matrix $M=(M^1,M^2)$. The G.H.D.C. applied over an arbitrary F_0 function is defined as |8|

$$F_l(\bar{x}) = \sum_{\bar{i}=-\bar{L}_1}^{\bar{L}_2} N_l(\bar{i})F_0(\bar{x}+\bar{i}) = \sum_{\bar{i}=-\bar{K}_1}^{\bar{K}_2} \omega(\bar{i})F_{l-1}(\bar{x}+M^{l}\bar{i}) \qquad (3)$$

where l is the level of the G.H.D.C., $N_l(\bar{x})$ is the equivalent l-level nucleus and $\omega(\bar{x})$ is the generating nucleus. This algorithm shows that the correlation of the function $F_0(\bar{x})$ with the equivalent nucleus can be recursively computed as l correlations with the generating nucleus. The generating nucleus has a support region defined by $\bar{L}_1$ and $\bar{L}_2$.

The generalized pyramidal coding is obtained by applying the reduced version of G.H.D.C.. A key difference between our coding scheme and the Burt-Adelson's method is that we introduce the pyramidal coding defined on an arbitrary sampling lattice rather than on a rectangular lattice. The new scheme is implemented in four stages.

1. A set $\{I_l\}$, $0 \leqslant l < N$, of N low-pass versions of the original image defined on sampling lattices $\{M^l\}$ is obtained by recursively applying a reduced form of the GHDC algorithm.

$$I_{l+1}(n_1,n_2) = \sum_{i=K_1}^{K_2} \sum_{j=K_3}^{K_4} \omega(i,j) I_l\left(M\binom{i}{j}+\binom{n_1}{n_2}\right) \qquad (4)$$

$$0 \leqslant l < N-2$$

where I_0 is the original image. M is the associated matrix to the pyramidal coding which defines the sampling lattice or geometry lattice. $\omega(i,j)$ is the generating nucleus. The low-pass images defined on $\{M^l\}$ can be stacked on a pyramidal structure of N levels, where the dimension of the l-level image is $(L_l+1)x(K_l+1)$. The implementation of this scheme is realizable if for each level l and geometry M, some integers (K_1,K_2) can be found such that the expression

$$\binom{L+1}{K+1} = M^l\binom{K_1}{K_2} + \binom{1}{1} \qquad 0 \leqslant l < N \qquad (5)$$

is fulfilled. $(L+1)x(K+1)$ are the dimensions of the original image.

2. A set $\{E_l\}$ of band-pass images are constructed by applying an inverse form of the G.H.D.C. as a suitable difference between two levels of the low-pass structure. The band-pass or error images are recursively obtained through the expression

$$E_l(n_1,n_2) = I_l(n_1 n_2) - |\det M| \sum_{i=K_1}^{K_2} \sum_{j=K_3}^{K_4} \omega(i,j) I_{l+1}\left(M^{-1}\binom{i-n_1}{j-n_2}\right) \qquad (6)$$

$$0 \leqslant l < N-1$$

with $E_{N-1}=I_{N-1}$
and $0\leq n_1<L_1+1$, $0\leq n_2<K_1+1$.

This algorithm is only evaluated for integer values of I_{l+1}.

The similarity between the equivalent nucleus and the bidimensional normal distribution, subject to the constraint of the convergence conditions of both pyramidal algorithms, allows to design the generating nucleus $\omega(n_1,n_2)$. Since the properties of the generating nucleus are translated to the equivalent nucleus $\omega(n_1,n_2)$ must verify the same properties of the normal distribution, e.g., normalization, unimodality, and symmetry. If $M=(M^1,M^2)$, the symmetry is defined with respect to M^1 and M^2.

3. The set of error images $\{E_l\}$ are properly quantized depending on the statistics of each level. The histograms of the first levels of the error images fit approximately the Laplacian probability density function, therefore, non-uniform quantizers are implemented. The parameters of the quantizer are designed depending on both the compression and the quality of the desired reconstructed image. The set of quantized images $\{E'_l\}$ are coded and transmited using variable length codewords.

4. At the receiver, the set $\{E'_l\}$ is decoded and the original image is reconstructed by successively applying the expression

$$I'_l(n_1,n_2)=E'_l(n_1,n_2)+ |\text{Det } M| \sum_{i=K_1}^{K_2} \sum_{j=K_3}^{K_4} \omega(i,j) I_{l+1}\left(M^{-1}\binom{i-n_1}{j-n_2}\right) \quad (7)$$

$$0\leq l\leq N-2$$

where $I'_{n-1}=E'_{n-1}$, and $I'_0(n_1n_2)$ is the reconstructed image.

Figure 1. shows the block diagram of the generalized Pyramidal Coding with an associated geometry M. The algorithms (4) and (6) are modeled as bidimensional decimation and interpolation techniques. The bidimensional decimation techniques are obtained through low-pass filtering and down-sampling. The bidimensional interpolation techniques are obtained through up-sampling and low-pass filtering.

The pyramidal coding proposed by Burt and Adelson becomes a particular case if $M=\begin{pmatrix}2 & 0\\ 0 & 2\end{pmatrix}$.

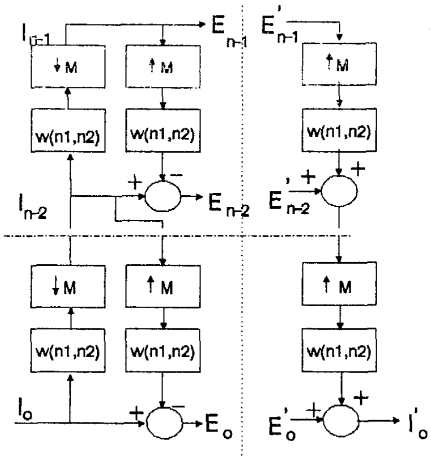

Figure 1. Block Diagram of the generalized pyramidal coding.

3. ADAPTIVE PYRAMIDAL CONDING

The use of non-rectangular geometries in the construction of the band-pass images allows to split the spectrum in regions of similar statistics, increasing the efficiency of the coding and adapting the coding process to the image characteristics.

The support region, characterized by its shape and dimensions, of the band-pass images, is selected through the geometry associated to the Pyramidal Coding. The optimum sampling lattice is assigned to each image using frequency criteria.

In the frequency domain the optimum geometry is obtained in three steps.

1. A binarized image of the spectrum magnitude is constructed, where one level corresponds to 98% of the signal energy.

2. A set of binary images $\{B_i\}$ representing the reciprocal sampling lattice shape for each geometry M_i, is built.

3. The spectrum of the image is compared against a reference set $\{B_i\}$. This comparison gives the optimum sampling lattice.

It has also been developed an algorithm on the spatial domain that finds the optimum geometry for which the bit rate is minimized for a given original image. This algorithm is based on the shape of the autocorrelation function of the image.

The results obtained are comparable to those obtained using the adaptive frequency domain approach.

The bit rate of the pyramidal coding with associated geometry M is a function of the number of levels and also of the levels dimension and entropy

$$B=\left|(L+1)\cdot(K+1)\right| / \sum_{l=0}^{N-1} B_l \times J_l \qquad (8)$$

where B_l is the entropy and J_l is the number of pixels of the level l.

4. CODING RESULTS AND CONCLUSIONS

Figure 2. shows the four levels of the pyramidal structures $\{I_l\}$ and $\{E_l\}$ for "LENA" for an associated geometry $M=\begin{pmatrix}2 & -2\\ 2 & 2\end{pmatrix}$.

Figure 3. shows "LENA" coded at 0.47 and 0.36 bits per pixel using a $\begin{pmatrix}2 & -2\\ 2 & 2\end{pmatrix}$ geometry. The signal-to-Noise ratios are 22.8 and 22 dB respectively. The results have been obtained using the two adaptive algorithms explained above. Notice that the frequency domain approach and the spatial domain approach give the same results.

The work explained above has shown that the efficiency of the pyramidal coding scheme can be improved if appropriate sampling lattice are chosen. Two algorithms, one in the frequency domain and the other in the spatial domain are presented which allow to select the optimum sampling lattice.

a) b)

Figure 2. Four levels of the pyramidal structures.
a) $\{E_l\}$ b) $\{I_l\}$

5. REFERENCES

|1| P.J. Burt and E.H. Adelson. "The Laplacian Pyramid as a compact image code", IEEE Trans. Commun, Vol. COM-31, pp. 532-540, April 1983.

|2| P.J. Burt, "Fast filter transforms for image processing", Comput. Graphics, Image Processing, vol. 16, pp. 20-50, 1981.

|3| E. Dubois, "The sampling and Reconstruction of Time-Varying Imagery with Applications in Video Systems", Proc. IEEE, vol. 73, pp. 502-522, April 1985.

|4| C.H. Chen, "Laplacian Pyramid image data compression", Proc. ICASSP-87, pp. 737-739, April 87, Dallas, Texas.

|5| P.H. Westerink, J. Biemond, D.E. Boekee, "Sub-band coding of images using predictive vector quantization", Proc. ICASSP-87, pp. 1378-1381, April 87, Dallas, Texas.

|6| J.W. Woods and S.D. O'Neil, "Subband Coding of Images", IEEE ASSP, vol. ASSP-34, pp. 1278-1288, October 1986.

|7| E.C. Hildreth, "Implementation of Theory of edge detection", A.I. LAB. MIT, April 1980, Massachusetts.

|8| Sallent S., "Optimización de los métodos multicanal en codificación de imagen",Ph. D. Thesis, Barcelona, January 1988.

a)

b)

Figure 3. a) LENA at 0.47 bits/pixel
b) LENA at 0.36 bits/pixel