THREE DIMENSIONAL ADAPTIVE LAPLACIAN PYRAMID IMAGE CODING

S.Sallent*

L.Torres**

L.Gils*

* Department of Applied Mathematics and Telematics, ** Department of Signal Theory and Communications

ETSETB-UPC, Apartado 30002, Barcelona 08034, Spain.

In this paper we propose a three dimensional Laplacian Pyramid coding scheme. The input image sequence is subsampled both spatially and temporally into different channels using proper three dimensional sampling structures. This results in the image sequence being represented by a series of bandpass sequences through three dimensional Gaussian and Laplacian pyramid data structures. In order to build the spatio-temporal pyramid structure several filters are discussed.

1 INTRODUCTION

The Laplacian Pyramid is a new and efficient method for image encoding (1). The method is of increasing interest as bandpass pyramids and multiresolution images are being used in other image processing applications. The pyramid image structure can be naturally adapted for progressive image transmission over low-speed channels and hierarchical image retrieving in computerized image storage.

The method obtains good compression rates and excellent visual quality for static images. It was then logical to extend the pyramid image structure using arbitrary non-rectangular sampling lattices and characterize the sampling lattices by matrices, thus providing a compact and powerful notation (2).

On the other hand increasing interest is focused on image sequence coding. Applications such as videoconference, videotelephone and low bit rate image coding in general, are a key issue in current video communication systems. The paper we present proposes a new three dimensional coding scheme as an extension of the previously reported work on static images. The input image sequence is subsampled both spatially and temporally into different channels using proper three dimensional sampling structures. This results in the image sequence being represented by a series of bandpass sequences through three dimensional Gaussian and Laplacian pyramid data structures.

In order to build the temporal pyramid structure several spatial-temporal filters

are discussed along with different sampling strategies. It is shown, as in the static image case, that the performance of the pyramid encoding system can be improved by proper selection of the sampling structures thus resulting in an adaptive and efficient encoding method. Motion compensation algorithms are also introduced in the scheme to further decrease the bit rate as is the case for hybrid methods.

2 GAUSSIAN AND LAPLACIAN DATA STRUCTURES

The temporal and spatial pyramid data structure represents the original image sequence into a set of code elements which are localized in spatial and temporal frequencies as well as in space and time. Each element in the new data structure is obtained by applying an appropriate three dimensional weighting function defined on an arbitrary sampling structure.

The image sequence, and particularly the video-conference sequences, are characterized by the high correlations of the neighboring pixels in the spatial and temporal dimensions. The three dimensional pyramid coding reduces the correlation by subtracting the original sequence $S_0^o(l,m,n)$ from the low-pass version sequence of itself $S_1^o(l,m,n)$, where l,m,n are the temporally and spatially coordinates respectively.

The code elements are obtained from those sequences as a suitable difference which represents the prediction error

$$D_o(l,m,n) = S_0^o(l,m,n) - S_1^o(l,m,n) \quad (1)$$

being $D_o(l, m, n)$ more decorrelated than the original sequence.

Then rather, than encode $S_o^o(l, m, n)$ it encodes the set of band-pass sequences obtaining data compression. The compression is achieved because the low-pass sequences are built through a decimation process associated with M_i matrix. Consequently the low-pass sequences are encoded at a reduced sample rate, and the high-pass sequences can be described with fewer bits.

Iterating this process over the low-pass sequence a set of low-pass $\{S_i^o(l, m, n)\}$ and band-pass $\{D_i(l, m, n)\}$ image sequences is obtained whose support regions are defined on sampling lattices characterized by a 3×3 M_i matrix.

Both data sets can be modeled as a pyramid data structures where each level is a sequence of decreasing dimension and resolution, the original sequence being the bottom of the pyramid.

The low-pass sequence set is constructed applying recursively the decimation algorithm

$$S_{i+1}^o(l, m, n) = \text{dec}\{S_i^o(l, m, n)\} = \sum_o \sum_p \sum_q W(o, p, q) \cdot S_i^o(M_i[l, m, n]^T + [o, p, q]^T)^T \quad (2)$$

where o, p, q belong to the support region of the three dimensional weighting function W , i is the pyramid level, with $i=0$ the bottom of the pyramid structure and $0 \leq i < L-1$.

The band-pass sequence set is built by applying recursively the interpolation algorithm

$$\begin{aligned} D_i(l, m, n) &= S_i^o(l, m, n) - S_{i+1}^o(l, m, n) = \\ &= S_i^o(l, m, n) - \text{inter}\{S_{i+1}^o(l, m, n)\} = \\ &= S_i^o(l, m, n) - |\det(M_i)| \cdot \\ &\sum_o \sum_p \sum_q W(o, p, q) \cdot S_{i+1}^o(M_i^{-1}(l - o \ m - p \ n - q))^T \end{aligned} \quad (3)$$

This algorithm is only evaluated for integer values of $S_{i+1}^o(l, m, n)$.

When the decimation and interpolation process uses like-Gaussian filters, the data

sets are named Gaussian and Laplacian data structures.

3 THREE DIMENSIONAL LAPLACIAN PYRAMID CODING

The three dimensional Laplacian Pyramid Image coding with associated sampling lattices defined by a three by three matrices set $\{M_i\}$ is based on the transmission of the quantized set of band-pass sequences $\{D_i(l, m, n)\}$, L the number of total levels of the pyramid structure and $D_{L-1}(l, m, n) = S_{L-1}^o(l, m, n)$ is the top sequence of the pyramid.

The new coding scheme is implemented in four stages.

a- A set of L low-pass versions of the original sequence defined on sampling $\{M_i\}$ lattices is obtained by applying recursively the decimation algorithm. For the construction of each level $i+1$ an appropriate sampling lattice is chosen in order to provide the best information compaction. Thus resulting in each level having its proper sampling structure. In the frequency-domain the shape of the reciprocal unit cell associated to the sampling lattice M_i is compared to the frequency content of the low-pass sequence $S_i^o(l, m, n)$. The algorithm can be modeled as a low-pass filtering and down-sampling process.

b- The set of band-pass sequences are constructed by applying recursively the interpolation algorithm over the low-pass sequence. This algorithm can be modeled as an up-sampling process, low-pass filtering and a suitable difference. The weighting function and sampling lattice used are the same that have been used in the construction of the equivalent level in the low-pass sequence.

c- The set of band-pass sequences are quantized by laplacian quantizers. The parameters of the quantizers are selected according to the statistics, the total chosen compression and the quality of the desired reconstructed sequence. The set of quantized

sequences $\{D_i(l,m,n)\}$ are transmitted using variable length codewords.

d- At the receiver, the original sequence is reconstructed applying recursively

$$S_i(l,m,n) = D_i(l,m,n) + |\det(M_i)| \cdot$$

$$\sum_o \sum_p \sum_q W(o,p,q) \cdot S_{i+1}(M^{-1}(l-o \ m-p \ n-q))^T \quad (4)$$

where $S_{L-1}(l,m,n) = D_{L-1}(l,m,n)$ and $S_0(l,m,n)$ is the reconstructed sequence. The use of arbitrary sampling lattices in the construction of low and band pass sequences allows to split the spectrum in regions of similar statistics adapting the coding process to the spatial and temporal characteristics of the sequence.

Figure 1 shows the block diagram of the Three Dimensional Laplacian Pyramid Coding.

The prediction error can be improved applying motion compensation to the weighting function. A regular decomposition

quadtree method (3) is applied to segment the interframe differential signal into homogeneous regions of different block sizes. Each region is characterized by a motion vector and used to correct the local weighted average of each pixel.

4 SIMULATIONS AND RESULTS

The results presented here were derived from two video sequences known as "Miss America" and "Walter" which are 256x256 pixels per frame with eight pixels per bit and twenty five frames per second.

In this coding method, the type of the filter has been chosen according to the appropriate matrix M_i associated to each level. For instance, figure 2 shows the first level of the Three Dimensional Laplacian Pyramid for "Miss America" sequence. In this case we use 3D spatio-temporal filter with 125 taps and 1D temporal filter 5 taps associated to

$$M_i = \begin{pmatrix} 200 \\ 020 \\ 002 \end{pmatrix}$$

and

$$M_i = \begin{pmatrix} 200 \\ 010 \\ 001 \end{pmatrix}$$

video-conference sequences
Miss America" and "Walter"
a b

Figure 2 The first level of the Laplacian Pyramid $D_0(l,m,n)$ for "Miss America" and "Walter" using a) 125 tap spatio-temporal filter, and b) 5 tap temporal filter.

Figure 3 shows the histograms of the first level using the spatio-temporal and temporal filters described above of the walter and Miss America sequences.

Figure 4 shows results of consecutive coded frames 6, 7, 8 and 9 of the "Miss

Figure 4 Reconstructed frames 6,7,8,9 of "Miss America" and 1, 2, 3 of "Walter" with a signal to noise ratio of 24 and 27 dB respectively.

America" with a compression ratio of 40, and frames 1, 2, and 3 of "Walter" sequence with a compression ratio of 20. Computer results are presented at 64 x 4 kbits/sec with excellent visual quality and a signal to noise ratio of 24 and 27db respectively. This rate can be lowered by applying motion compensation to the three dimensional weighting function.

The simulation results indicated that the proposed method is capable of operating very efficiently for a wide class of video applications such as 192 Kbits/second high definition video conferencing and in the range of B-ISDN hierarchies. It is shown that the scheme does not present any block effect, offers low computational complexity, and can be implemented in parallel.

5 CONCLUSION

We presented the three dimensional Laplacian Pyramid Coding used to obtain a high compression rate for wide variety of video applications. This method is the result of extending the Pyramid Coding to the three dimensions, using arbitrary sampling lattices. Also the suitability to apply motion compensation on the weighting functions was emphasized, achieving a significant improvement in the compression ratio and visual quality.

6 REFERENCES

- [1] "The Laplacian Pyramid as a Compact Image Code", P.J. Burt, E.H. Adelson, IEEE Transactions on Communications, Vol. COM-31, n°4, April 1983, pp.532-540.
- [2] "An Adaptive Pyramid Image Coding System", S.Sallent, L.Torres, Proceedings ICASSP 1988, New York, April 11-14, 1988.
- [3] "Simulation Of A Teleconference Codec For ISDN", S.Sallent, A.Artero, J.Haro, EUSIPCO-90, Barcelona, September, 1990.
- [4] "The Sampling And Reconstruction Of Time-Varying Imagery With Applications in Video Systems", E. Dubois, Proceedings IEEE 1988, Vol. 73, 502-522, April 1985.