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

A NON-RECTANGULAR PYRAMID CODING SYSTEM

S. Sallent-Ribes L. Torres-Urgell

E.T.S.I. Telecomunicación
Apdo 30.002
08080 Barcelona, Spain

Rectangular sampling lattices have been extensively applied to image coding. Sampling is a fundamental operation in image coding systems as a proper selection of the sampling lattice minimizes the physical support of the information affecting the overall performance of the system and increasing the compression ratio. Non-rectangular sampling lattices have been occasionaly used in specific image systems, but their application to coding schemes has not been reported in the literature.

In this paper we present a characterization of multidimensional systems whose input and output are defined on arbitrary sampling structures. We also introduce the system response that eases the analysis of the sampling structure conversions. In the second part of the paper this theory is applied to Laplacian Pyramid Coding. The result is the modelization of the algorithms used in Pyramid Coding. In the third part of the paper this model is used to describe an efficient adaptive Pyramid image structure defined on arbitrary non-rectangular sampling lattices.

1 INTRODUCTION

Image coding is usually a two or a three dimensional process that uses arrays of numbers which have been obtained by rectangular sampling of images. Although this sampling strategy is the most common used in this type of process, it is neither the only one that can be used nor the most efficient and compact.

For every coding scheme, one or several arbitrary sampling lattices may exist that optimize the coding process. The basic idea is to find an adequate sampling lattice that minimizes the bit rate increasing the coding performance. The optimum sampling lattice (or structure) minimizes the physical information support of the coding process. It becomes necessary, to characterize multidimensional signals and systems with proper tools.

In this paper we introduce the system function and the transmission function which characterize, in the spatial and frequency domain, multidimensional signals and linear systems defined on arbitrary sampling lattices.

The use of the system function in one dimension has been presented by Crochiere and Rabiner in [1], with the analysis of unidimensional systems with associated signals defined at different sampling rates being the main application. Our approach here, is to extend these previous works to multidimensional periodically varying and invarying linear systems by means of the system function using a compact and powerful notation. This representation will allow us to describe properly all the linear periodically varying processes used in image coding such as upsampling and downsampling conversion, modulation, etc.

Multichannel image coding, such as multidimensional sub-band coding, and other schemes based on different kinds of filter banks can be analyzed and generalized analytically to non-rectangular sampling lattices [2,3]. In the second part of this paper, the transmission function will be used to obtain the frequency domain representation of the Laplacian Pyramid coding[7]. This analysis will allow a generalization of the Pyramid coding using arbitrary sampling lattices. To that end, efficient and fast algorithms based on the Hierarchical Discrete Correlation (H.D.C.) [4] defined on arbitrary sampling lattices, will be introduced to build the non-rectangular Laplacian Pyramid.

In the third part of this paper, an adaptive Pyramid image coding scheme is introduced. In the frequency domain, an algorithm assigns the optimal sampling geometry (sampling lattice) to each image depending on the information content. The spectrum of the image is matched to the unit cell of the reciprocal sampling lattice thus resulting in each image having its proper sampling structure.

Excellent performance is shown in both the signal to noise ratio and visual quality for a bit rate of 0.4 bits/pixel.

* This work has been partially supported by the CAICYT grant 2196/84 of the Spanish Government.

2 MULTIDIMENSIONAL LINEAR SYSTEMS

In this section we characterise in terms of the system response and the transmission function, a multidimensional linear digital system for which its input and output are defined on arbitrary sampling structures. This will be particularly useful, from a practical point of view, to develop and characterize linear periodically varying and invarying systems used in image coding. Interpolation and decimation are also introduced for these kind of systems. Fig. 1 shows a multidimensional linear system.

Let $\mathcal{L}$ be a linear system with input $x(\overline{n})$ and output $y(\overline{m})$ defined on sampling structures Ψ_1, and Ψ_2 of dimensions N_1 and N_2 respectively.

Ψ_i is a sampling structure defined as a discrete set of points in R^{N_i} such that Ψ_i is the union of selected cosets of a sublattice Ω in a lattice φ [5]

$$\Psi_i = \bigcup_{i=1}^{P}(\overline{Z_i} + \Omega) \tag{1}$$

where $\overline{Z_i} - \overline{Z_j} \notin \Omega$ if $i \neq j$ and $\overline{Z_i} \in \varphi$.

We define the system response or the Green's function $K(\overline{m}, \overline{n})$ as the response of the system at the output $\overline{m}$ to a unit sample applied in $\overline{n}$, where $\overline{n} \in \Psi_1$ and $\overline{m} \in \Psi_2$. Then the response of the system relates the output with the input through a superposition sum as follows:

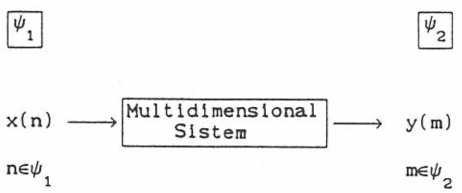


Figure 1: Multidimensional linear system

$$\begin{aligned} y(\overline{m}) &= \sum_{\overline{n}\in\Psi_1} k(\overline{m}, \overline{n})x(\overline{n}) = \\ &= \sum_{i=1}^{p}\sum_{\overline{n}\in\Omega_1} k(\overline{m}, \overline{n} + \overline{Z}_i)x(\overline{n} + \overline{Z_i}) \end{aligned} \tag{2}$$

If $\overline{\omega}$ and $\overline{\omega}'$ are the frequency domain associated to the reciprocal sampling structures Ψ_1^* and Ψ_2^* respectively, the frequency domain relationship between $X(\overline{\omega})$ and $Y(\overline{\omega}')$ is:

$$Y(\overline{\omega}') = \int_{B\in\Psi_1^*} K(\overline{\omega}', \overline{\omega})X(\overline{\omega})d\overline{\omega} \tag{3}$$

Were $K(\overline{\omega}', \overline{\omega})$ is the transmission function defined as

$$K(\overline{\omega}') = | \det\varphi_1 | \sum_{\overline{m}\in\Psi_2}\sum_{\overline{n}\in\Psi_1} K(\overline{m}, \overline{n})e^{-j\overline{\omega}'^T\overline{m}}e^{-j\overline{\omega}^T\overline{n}} \tag{4}$$

The relationship between the frequency and the spatial domain on sampling structures through Fourier transforms has been developed in [6].

In the particular case that the sampling structures become sampling lattices and if the system is space-invariant, the response of the system reduces to the conventional definition of the impulse response $h(\overline{n})$ and frecuency response $H(\overline{\omega})$. $H(\overline{\omega})$ defines then, the one to one mapping between $X(\overline{\omega})$ and $Y(\overline{\omega}')$ where $\overline{\omega} = \overline{\omega}'$ and $\overline{n} = \overline{m}$. With the aid of signal flow graphs, transmission function and response system, the concepts of duality and transposition can be generalizated to multidimensional linear systems or their associated networks.

It can be easily shown that in this case the transposition theorem states that two transpose networks are dual networks. In some kind of image coding applications it is particularly useful to define the upsampling or interpolation process as a linear operator in the following way

$$E : \mathcal{L}(\Psi_1) \rightarrow \Psi_2 \tag{5}$$

where

$$E\{X(\overline{n})\} = \begin{cases} X(L^{-1}\overline{m} + \overline{Z}_i) & \text{if } (L^{-1}\overline{m} + \overline{Z}_i) \in \Psi_1 \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

In this case the system response is

$$K(\overline{m}, \overline{n}) = \delta(\overline{n} - L^{-1}\overline{m} + \overline{Z_i}) \tag{7}$$

where the lattices $\Omega_1, \Omega_2, \varphi_1$ and φ_2 are characterized by the matrices T_1, J_1, T_2 and J_2 related to the rational matrices N, D, L as it is shown in Figure 2.

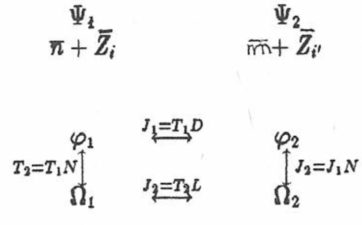


Figure 2: Relationships between the sampling structures Ψ_2 and Ψ_1 in the upsampling process.

In the same manner we define the downsampling or decimation process as a linear operator

$$\Re : \mathcal{L}(\Psi_1) \rightarrow \Psi_2 \tag{8}$$

where $\Re\{X(\overline{n})\} = X(\overline{n})$ if $\overline{n} \in \Psi_2$, being the system response

$$K(\overline{m},\overline{n}) = \delta(\overline{n} - M\overline{m} - \overline{Z}_i) \quad (9)$$

If the sampling structures Ψ_1 and Ψ_2 reduce to sampling lattices $\Psi_1 \equiv \Omega_1$ and $\Psi_2 \equiv \Omega_2$ then in the upsampling process

$$K(\overline{m},\overline{n}) = \delta(\overline{n} - L^{-1}\overline{m}) \quad (10)$$

and in the downsampling process

$$K(\overline{m},\overline{n}) = \delta(\overline{n} - M\overline{m}) \quad (11)$$

3 NON RECTANGULAR PYRAMID CODING

The Laplacian Pyramid Coding is an efficient method that uses the Laplacian Pyramid structure to produce an approximate frequency decomposition based on algorithms defined on rectangular sampling lattices. This method is not adaptive and the pyramidal structures are always built from data sampled on rectangular lattices.

The pyramidal structures can be analized in the frequency domain applying the transmission function on the H.D.C. fast algorithms [8] which are used to build the Gaussian and the Laplacian Pyramids that can be modeled as bidimensional upsampling and downsampling processes. Thus, each level of the Laplacian Pyramid can be modeled as a modulation process as it is shown in Fig. 3, with $M = L$.

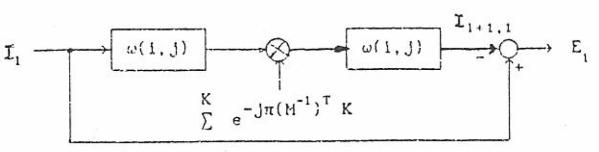


Figure 3: Laplacian Pyramid modeled as a modulation process.

$\{I_i\}$ are a set, $0 \leq i < N$ of N low-pass versions of the original image defined on sampling lattices $\{M^i\}$.$\{E_i\}$ are a set of band-pass images which build the Laplacian Pyramid, $\omega(i,j)$ is the generating nucleus and

$$M = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad (12)$$

Fig. 4. shows an example of the first level I_0 and the second level I_1 of the Gaussian structure with

$$M = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \quad (13)$$

The coding efficiency can be improved if we choose an adequate sampling structure with associated sampling matrix M thus compacting the $\{E_i\}$ set and minimizing the bit rate. Thus for each original image an optimum geometry M exists that minimizes the physical support of the information.

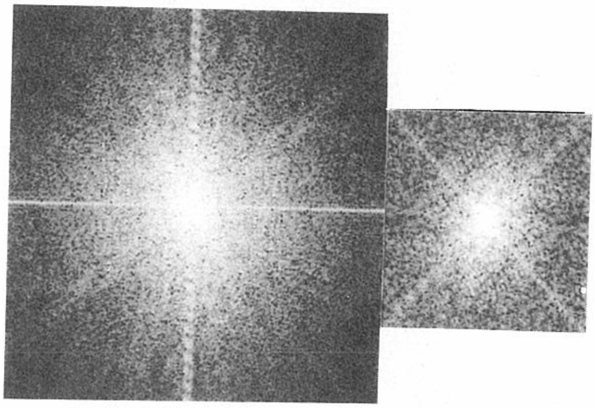


Figure 4: Spectrum of the first and second levels of the Gaussian structure.

4 APPLICATIONS AND RESULTS

As it has been stated above the theory developed here may be applied to efficient coding images. As a particular case we have applied it to the Laplacian Pyramid. Finding the optimum sampling structure of the image to be encoded, the bit rate will be improved. To that end, an adaptive algorithm in the frequency domain has been implemented. The algorithm is as follows. Details can be found in [4].

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 $\{B_i\}$, and optimum sampling structure is then selected.

To check the feasibility of the proposed algorithms, they have been tested on the very well known images LENA and MIT. Fig. 5 shows the original images and Fig. 6 the same images coded with their optimum sampling geometry as obtained with the adaptive algorithm. Bit rate is 0.38 bits/pixel and 0.53

bits/pixel respectively. Signal to noise ratios are 22 db and 20 db.

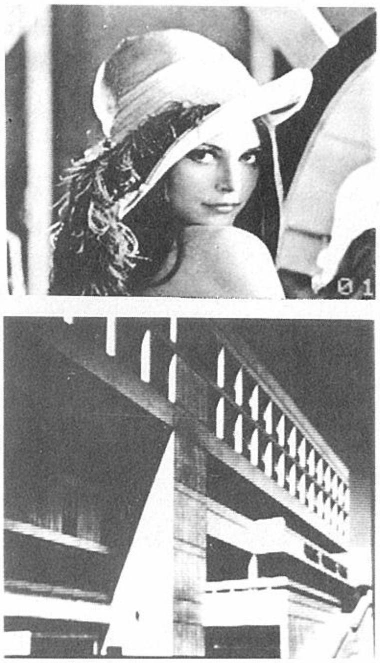
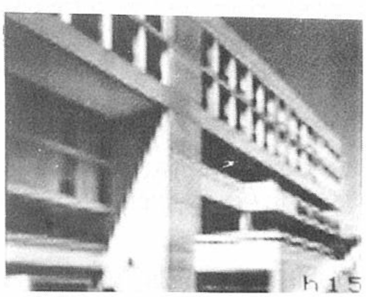


Figure 5: Original images Lena and MIT.

a)



b)

Figure 6: a)Lena at 0.38 bits/pixel. b)Mit at 0.53 bits/pixel.

5 CONCLUSIONS

Multidimensional linear systems whose inputs and outputs are defined on arbitrary sampling structures have been characterized. This analysis is based on the theory of numbers and linear systems and provides a useful and compact notation . It has been also shown that multichannel image processes can be modeled using this methodology. Image coding methods may be then generalized using arbitrary geometries. Examples of Pyramid coding with optimum geometries have been presented.

References

[1] R.E. Crochiere and L.R. Rabiner. *Multirate Digital Signal Processing.* Englewood Cliffs, N.J.: Prentice-Hall, 1983

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

[3] M. Vetterli, *"Multi-Dimensional Sub-band Coding: Some Theory and Algorithms,"* Signal Procesing, vol. 6, pp.97-112, April 1984.

[4] S. Sallent-Ribes and L. Torres-Urgell, *"An Adaptive Pyramid Image Coding System,"* Proc. ICASSP-88, April 1988, New York.

[5] C. G. Lekkerkerker, *Geometry of Numbers.* North-Holland, 1969.

[6] E. Dubois, *"The Sampling and Reconstruction of Time-Varying Imagery with Applications in Video Systems,"* Proc. IEEE, vol.73, pp. 502-22, April 1985.

[7] P. J. Burt and E.H. Adelson, *"The Laplacian Pyramid as a Compact Image Code,"* IEEE Trans. on Com., vol. 31, pp. 532-540, April 1983.

[8] P.J. Burt, *"Fast Filter Transforms for Image Processing,"* Comput. Graphics, Image Processing, vol. 16, pp. 20-51, 1981