TREIZIÈME COLLOQUE GRETSI - JUAN-LES-PINS DU 16 AU 20 SEPTEMBRE 1991

PYRAMID DECOMPOSITION AND CODING OF SELF-AFFINE IMAGES

François J. Malassenet and Russell M. Mersereau

School of Electrical Engineering, Georgia Institute of Technology
Atlanta, GA 30332, U.S.A.

RÉSUMÉ

Cet article introduit un nouveau modèle que l'on appelle *weighted multiresolution process* (WiMP) pour l'analyse bidimensionnelle des signaux. Le modèle combine les propriétés spatio-temporelles de la décomposition en ondelettes avec celles des signaux auto-affines à modéliser. Un WiMP peut être vu comme une chaine de Markov constituée de translations entre niveaux et voies d'une décomposition pyramidale. Les cas périodiques et apériodiques sont étudiés ainsi que les algorithmes de décomposition et de reconstruction.

ABSTRACT

A new model for two-dimensional signal analysis, called the *weighted multiresolution process* (WiMP), is introduced. The model combines the scale and time-frequency localization properties of the wavelet representation with the self-affine characteristics of signals to be modeled. A WiMP can be viewed as a Markov chain of weighted translations between levels and channels of a pyramid decomposition. Both aperiodic and periodic signals are investigated and the corresponding decomposition/reconstruction algorithms are presented.

1 INTRODUCTION

This paper presents a two-dimensional extension of the model introduced in [1]. The modeling technique is restricted to *self-affine* signals. A signal is called self-affine if any part of the signal after being transformed by an affine transformation is almost identical to the entire signal. This qualitative definition was formalized by Barnsley when he introduced the iterated function systems with probability, (IFS) [2]. The mathematical definition of a self-affine signal, f, in two dimensions is then:

$$f(\mathbf{x}) = \sum_{i=1}^{N} |A_i|^{-1} p_i f\left(A_i^{-1}(\mathbf{x} - \mathbf{b}_i)\right), \qquad (1)$$

where A_i is a 2×2 matrix, $|A_i|$ is the Jacobian of A_i, $\mathbf{b}_i$ is a translation vector, and p_i is a probability, i.e., $p_i \geq 0$ and $\sum_{i=1}^{N} p_i = 1$. f is called the attractor of the IFS defined by $\{A_i\mathbf{x} + \mathbf{b}_i, p_i\}_{i\in\{1,\cdots N\}}$. When one of the matrices A_i is singular, the function f can still be defined using the measure notation. This model has been proven to generate texture-like images [3, 2]. It is particularly suited for images since the functions fulfilling Equation (1) are positive and have $\mathbf{L}^1$ norm equal to one. A direct computation of the IFS code is ill-posed. This paper presents a model that would make the computation easy to perform while not compromising the diversity of images that the IFS model can generate.

In the following sections, a subset of self-affine signals, the *scale invariant* signals are considered. A signal is scale invariant when the linear parts of the affine transformations, A_i, are constant for all i's. The WiMP model introduced applies to scale invariant signals but also performs quite well for the case of self-affine signals.

2 WiMP

2.1 Wavelet Representation

This section presents a quick review of the work introduced by Mallat [4] on the wavelet representation in two dimensions. When an approximation representation of $\mathbf{L}^2(\mathbf{R}^2)$ is known, i.e., the approximation function, $\Phi(\mathbf{x})$ is known, one can construct a set of three functions Ψ_1, Ψ_2, Ψ_3, such that

$$\left\{2^j \Psi_k\left(2^j\mathbf{x} - \mathbf{n}\right)\right\}_{\substack{k\in\{1,2,3\}\\ \mathbf{n}\in\mathbf{Z}^2}} . \qquad (2)$$

forms a complete orthonormal basis of $\mathbf{L}^2(\mathbf{R}^2)$. Any function of $\mathbf{L}^2(\mathbf{R}^2)$ can thus be defined uniquely by the three three-dimensional sequences of $\ell^2(\mathbf{Z}^3)$ $\left\{d_1^j[\mathbf{n}], d_2^j[\mathbf{n}], d_3^j[\mathbf{n}]\right\}_{j\in\mathbf{Z},\mathbf{n}\in\mathbf{Z}^2}$.

$$f(\mathbf{x}) = \sum_{k=1}^{3} \sum_{j=-\infty}^{\infty} \sum_{\mathbf{n}\in\mathbf{Z}^2} d_k^j[\mathbf{n}]\, 2^j \Psi_k\left(2^j\mathbf{x} - \mathbf{n}\right). \qquad (3)$$

This work was supported by the Joint Services Electronics Program contract number DAAL-03-90-C-0004

The parameter j corresponds to the level of decomposition while k is the channel number of the decomposition.

The advantage of this decomposition compared to the regular Fourier transform is that it can be constructed to have very good localization in time and frequency for a well chosen approximation function. A local change in time or frequency will not cause a major global change in the decomposition sequences. Moreover, the decomposition can be performed using pyramid filter banks.

The signal does not have to be decomposed over all scales 2^j, its decomposition can be truncated leaving the low-frequency approximation unchanged. The truncated representation of the signal is defined by the equation:

$$\begin{aligned} f(\mathbf{x}) = \sum_{k=1}^{3} \sum_{j=M}^{\infty} \sum_{\mathbf{n}\in\mathbf{Z}^2} d_k^j[\mathbf{n}]\, 2^j \Psi_k\left(2^j\mathbf{x}-\mathbf{n}\right) \\ + \sum_{\mathbf{n}\in\mathbf{Z}^2} a^M[\mathbf{n}]\, 2^j \Phi\left(2^j\mathbf{x}-\mathbf{n}\right), \end{aligned} \quad (4)$$

where the parameter M is called the *truncation level.*

If the decomposition used is identical to the one used by Mallat, the basis functions are all separable and can be constructed from one-dimensional wavelets [4]. In many cases f does not have infinite resolution, the range of scales is usually finite and not semi-infinite as in (4), i.e., the $d_k^j[\mathbf{n}]$'s are zero, for $j > P > M$.

2.2 Self-affine signals

The space of self-affine signals generated by IFS's is not part of $\mathbf{L}^2(\mathbf{R}^2)$ but is the set of Borel measures on $[0,1]\times[0,1]$. However, if the signals are considered at finite resolution, they can be viewed as elements of $\mathbf{L}^2(\mathbf{R}^2)$, thus having a wavelet decomposition. If the signal is scale invariant, i.e., fulfills Equation (1), such that the scaling matrix is of the form $\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$, the wavelet decomposition has the following property:

$$d_k^j[\mathbf{n}] = \sum_{i=1}^{N} 2\, p_i\, d_k^{j-1}[\mathbf{n} - 2^j \mathbf{b}_i], \quad (5)$$

if both components of $\mathbf{b}_i 2^j$ are integers, otherwise the relationship involves the characteristics of the wavelet used in the decomposition. There exists a relationship between levels and channels of the wavelet decomposition. Level j can be viewed as the output of a FIR filter excited by level $j-1$. When the IFS is homogeneous and the common affine transformation is of the form 0.5 S, where S is an isometry that keeps the unit square, $[0,1]^2$ invariant, relationships between levels and channels of the decomposition are also simple. In this case, the filters are all related to each other. The positions of the non-zero filter taps are scaled by a factor of 2 while their values remain unchanged. Since the frequencies corresponding to $d_k^j[\mathbf{n}]$ increase when j increases, Equation (5) shows that the high frequency content of the signal can be reconstructed from the low frequency content.

2.3 WiMP definition

The 2-D WiMP model uses the relationships between the levels and introduces flexibility in the choice of the filters. Before presenting the 2-D WiMP model, some definitions need to be introduced. $I = \{1,2,3\}$, and $\mathbf{X} = I \times \mathbf{Z}^3$.

Definition 1 *A* channel-k level-j translation, τ_k^j, *with time translation factor* $\mathbf{t}_k^j$, *channel translation factor* χ_k^j, *and level translation* λ_k^j *is defined as follows:*

$$\tau_k^j : \ell^2(\mathbf{X}) \longrightarrow \ell^2(\mathbf{X})$$

$$\underline{\mathbf{y}} = \tau_k^j(\underline{\mathbf{x}}) \Longleftrightarrow \begin{cases} y_l^m[\mathbf{n}] = x_{l-\chi_k^j}^{m-\lambda_k^j}[\mathbf{n}-\mathbf{t}_k^j], & \text{if } m=j+\lambda_k^j, \text{ and } l=k+\chi_k^j. \\ y_l^m[\mathbf{n}] = 0, & \text{otherwise.} \end{cases}$$

$$\underline{\mathbf{y}} = \left(y_k^j[\mathbf{n}]\right)_{\substack{k\in I \\ (j,\mathbf{n})\in\mathbf{Z}^3}},\ \mathbf{t}_k^j \in \mathbf{Z}^2,\ \chi_k^j \in \{-2,-1,0,1,2\},\ \textit{and}\ \lambda_k^j \in \mathbf{Z}.$$

This transformation is the combination of two operations. First, all channels and levels except the channel-k level-j sequence are set to zero while the channel-k level-j is kept unchanged, then the new series is translated with a four-dimensional translation vector, See Figure 1. Now the formal definition of a 2-D WiMP can be given.

Definition 2 *A* two-dimensional weighted multiresolution process *(2-D WiMP) is defined by*

$$\left\{ \mathbf{X}, a^M, \left\{d_k^M\right\}_{k\in I}, \{\tau_{k,i}^j, p_{k,i}^j\}_{\substack{k\in I \\ j\in\mathbf{Z}, j>M \\ i\in\{0,\cdots,N_k^j-1\}}} \right\},$$

where $\tau_{k,i}^j$ *is a channel-k, level-j translation with translation factor* $t_{k,i}^j$, *channel translation factor* $\chi_{k,i}^j$, *and level translation factor* $\lambda_{k,i}^j$, $p_{k,i}^j$ *is a real number corresponding to the weight associated with the transformation* $\tau_{k,i}^j$, *and* a^M, *and* $\left\{d_k^M\right\}_{k\in I}$ *are sequences of* $\ell^2(\mathbf{Z}^2)$. a^M *is called the low-pass sequence and* $\left\{d_k^M\right\}_{k\in I}$ *are the seed sequences.*

To each WiMP corresponds an operator that generates a truncated wavelet representation from the four two-dimensional sequences a^M and $\left\{d_k^M\right\}_{k\in I}$. The operator is called the WiMP-operator.

Definition 3 *The* WiMP-operator, T, *associated with the above WiMP operates on a truncated wavelet representation and is defined by:*

$$T : \mathbf{X} \longrightarrow \mathbf{X}$$

$$T(\underline{\mathbf{x}}) = \sum_{j=M}^{\infty} \sum_{k=1}^{3} \sum_{i=0}^{N_k^j-1} p_{k,i}^j\, \tau_{k,i}^j(\underline{\mathbf{x}}).$$

This transformation defines all the links between different levels and channels of the truncated wavelet representation. It allows total freedom on how to relate levels and channels to each other.

The reconstruction of the function from the WiMP is performed at the wavelet representation level. The truncated wavelet representation of the function defined by the WiMP is $\left\{a^M, \lim_{n\to\infty} T^{\circ n}(\underline{\delta})\right\}$ where $\underline{\delta}$ is a semi-finite four-dimensional sequence where $\underline{\delta}_k^m[\mathbf{n}]$ is zero everywhere except at level M where $\underline{\delta}_k^M[\mathbf{n}] = d_k^M[\mathbf{n}]$. Some restrictions have to be put on the level translations τ_k^j's so that the above limit exists. In our representation all level translation factors λ_k^j's are usually set to 1 since in most cases adjacent levels are the most similar.

3 PERIODIC DECOMPOSITION

When a function of $\mathbf{L}^2(\mathbf{R}^2)$ has finite support, the support may be considered as included in $[0,1]^2$ without loss of generality. Any function with finite support can thus be periodically extended by the vectors $((0,1),(1,0))$. Since nonzero periodic functions have infinite power, they do not belong to $\mathbf{L}^2(\mathbf{R}^2)$. Perrier and Basdevant studied the wavelet representation of periodic signal on the circle [5]. The theory can be easily extended to the unit torus, $T = [0,1]^2$. The basis functions ensuring the decomposition are aliased versions of the ones used in the previous section.

$$\left\{\sum_{\mathbf{m}\in\mathbf{Z}^2} 2^j\Psi_k\left(2^j(\mathbf{x}+\mathbf{m})-\mathbf{m}\right)\right\}_{k\in I,\,(j,\mathbf{n})\in\mathbf{Z}^3} \tag{6}$$

form an orthonormal basis of $\mathbf{L}^2(T)$, where Ψ_k's are the wavelets of the aperiodic decomposition seen in the previous section.

The wavelet representation for a periodic signal is thus defined by the sequences

$$\left\{\left\{\tilde{d}_k^j[\mathbf{n}]\right\}_{\substack{k\in I\\ j\in\mathbf{Z}\\ \mathbf{n}\in\mathbf{Z}^2}}, \tilde{a}^0[0]\right\}.$$

This is not a truncated decomposition, $\tilde{a}^0[0]$ is the mean value of the function f, and the sequences $\left\{\tilde{d}_k^j[\mathbf{n}]\right\}_{\mathbf{n}\in\mathbf{Z}^2}$ correspond to the band-pass part of the periodic signal and are called the channel-k level-j sequences. Since IFS attractors have finite support by nature, the periodic decomposition can be used. The choice of the periodic decomposition is motivated by the non expansiveness of the decomposition, and the efficient computation using FFT's.

If the attractor f is generated from an homogeneous IFS whose common linear part is half identity, the decomposition has the same characteristics as in the aperiodic case, See Equation (5):

$$\tilde{d}_k^j[\mathbf{n}] = \sum_{i=1}^{N} 2p_i\,\tilde{d}_k^{j-1}[\mathbf{n}-2^j\mathbf{b}_i], \tag{7}$$

when both components of $2^j\mathbf{b}_i$ are integers. Since the WiMP model defined earlier operates on four-dimensional sequences, any periodic wavelet representation can be used in the WiMP operator. The periodic or aperiodic nature of the signal becomes apparent only in the decomposition/reconstruction phase of the wavelet representation.

4 COMPUTATION OF THE SOLUTION

Now that the model has been presented, we need to find a way to determine the parameters of the model given any signal. The problem of determining the WiMP transformation, T, is equivalent to the one presented by Cheng and Etter [6]. They considered the case of two one-dimensional signals u and v, where v is assumed to be the sum of weighted translations of u, or v is the output of a very sparse FIR filter with input u. Our method uses a straightforward extension of this technique to two dimensions.

In this approach a mean square error, (MSE), is iteratively minimized. First, sequences from contiguous levels and identical channels are considered. If $u[\mathbf{n}]$ is the channel-k level-j sequence, i.e., $u[\mathbf{n}] = \tilde{d}_k^j[\mathbf{n}]$, and $v[\mathbf{n}]$ is the residual signal initially set to $\tilde{d}_k^{j+1}[\mathbf{n}]$, the optimum translation minimizing the error $\sum_{\mathbf{n}}(v[\mathbf{n}]-u[\mathbf{n}])$ is defined by the location of the absolute maximum of the cross-correlation between $u[\mathbf{n}]$ and $v[\mathbf{n}]$. The associated optimum is the ratio of this maximum to the power of the original signal $u[\mathbf{n}]$. The residual signal, $v[\mathbf{n}]$, is updated by subtracting the weighted and shifted original signal, $u[\mathbf{n}]$.

The algorithm is optimal when the input signal $u[\mathbf{n}]$ is white. When this assumption is not true, the algorithm becomes suboptimal and behaves in an oscillatory manner. The estimated weight associated with the first translation is usually overestimated. After several iterations, the initial translation is found again, this time the associated weight reduces the magnitude of the combined weight. In this manner, erroneously estimated weights are updated after a few iterations. The oscillations can be removed by estimating simultaneously the weights each time a new translation is estimated.

The method does not have to be restricted to contiguous levels and identical channels. The above algorithm can be improved by comparing the resulting MSE after considering all possible levels and channels as long as the level considered is lower that the one to be decoded. Another freed parameter can be added by introducing isometries leaving the unit square invariant

When the above inverse solution is carried along increasing level ranks, the WiMP model and its truncated representation can be obtained simultaneously if the input of the filter between levels j and $j+1$ is the approximated sequence of level j, i.e., the output of the FIR filter between level $j-1$ and j. Moreover, since the levels and channels are orthogonal to each other, the total MSE between the original and reconstructed signal can be found by adding the MSE of each channel weighted by a rank factor.

5 CONCLUSIONS

A new algorithm for coding self-affine images has been presented. The main characteristic of the WiMP model is its estimation of the relationships between channels and levels of a pyramid decomposition. With an iterative estimation of the translations, the wavelet representation gave better results than other pyramids decomposition representations, due to time-frequency localization.

REFERENCES

[1] F. J. Malassenet and R. M. Mersereau, "Wavelet representations and coding of self-affine signals," in *Int. Conf. Acoust. Speech Signal Proc., ICASSP'91*, pp. 677–680, 1991.

[2] M. F. Barnsley, *Fractals Everywhere*. Academic Press, 1988.

[3] M. F. Barnsley, A. E. Jacquin, F. J. Malassenet, L. Reuter, and A. D. Sloan, "Harnessing chaos for image synthesis," in *SIGGRAPH '88 Conference Proceedings*, pp. 131–140, 1988.

[4] S. G. Mallat, "Multiresolution approximations and wavelet orthonormal bases of $\mathbf{L}^2(\mathbf{R})$," *Trans. Am. Math. Soc.*, vol. 315, no. 1, pp. 69–87, 1989.

[5] V. Perrier and C. Basdevant, "Periodical wavelet analysis, a tool for inhomogeneous field investigation. Theory and algorithm," *Recherche Aérospatiale (English Edition)*, vol. 3, pp. 53–67, May-June 1989.

[6] Y.-F. Cheng and D. M. Etter, "Analysis of an adaptive technique for modeling sparse systems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-37, pp. 254–264, Feb. 1989.

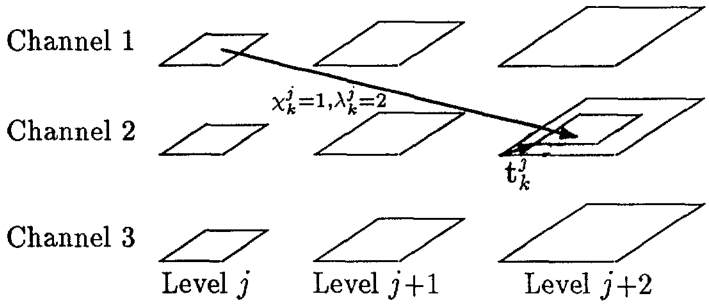

Figure 1. Example of a channel-k level-j translation, where $k = 1$, the level translation $\lambda_k^j = 2$, and the channel translation $\chi_k^j = 1$, defined in Definition 1.