RECURSIVE INTERFEROMETRIC REPRESENTATIONS

Stéphane Mallat

CMAP, Ecole Polytechnique
91128 Palaiseau, France
June 2010
Proceedings of EUSIPCO conference 2010

ABSTRACT

Classification requires building invariant representations relatively to groups of deformations that preserve signal classes. Recursive interferometry computes invariants with a cascade of complex wavelet transforms and modulus operators. The resulting representation is stable relatively to elastic deformations and provides invariant representations of stationary processes. It maps signals to a manifold which preserves signal discriminability.

1. INTRODUCTION

Signal classes are usually invariant to certain types of deformations that may include translations, rotations, scalings or any other group of operators. Classification algorithms must then be invariant relatively to these deformations. The invariance often also applies to elastic deformations which define much larger Lie groups. However, building invariants reduces the representation dimension, which may affect its ability to discriminate different patterns. It is therefore necessary to construct representations that balance invariance, stability and discriminability requirements.

The Fourier transform modulus is translation invariant but the representation of high frequencies is highly not invariant to elastic deformations. Computer vision researchers have introduced histogram techniques to build local invariants by delocalizing high frequency information, which lead to efficient local descriptors for classification, when global invariants are not needed. Deep neural networks also also provide efficient data dependent invariant representation [2, 3] in number of applications, but are not well understood.

This paper follows a harmonic analysis approach to invariant representations. Section 2 introduces a recursive interference representation, and analyzes the properties low frequency interferences computed with cascades of wavelet transform modulus. Section 3 studies interference invariance and discriminability and Section 4 provides a fast filter bank implementation. The paper concentrates on translation invariance but generalization to any other group can be found in [4].

2. RECURSIVE INTERFEROMETRY

Recursive interferometry maps signal high frequencies to lower frequencies, with a cascade of wavelet transforms and modulus operators, which yields a progressively more invariant representation.

2.1 Wavelet Transform Modulus

A modulus operator applied on a wavelet tranform is shown to compute low frequency interferences. A wavelet transform filters a real multidimensional signal $f \in \mathbf{L^2}(\mathbf{R}^d)$ with a family of $K-1$ wavelets $\{\psi_d\}_{1\leq d<K}$ which are scaled by 2^j:

$$\forall x \in R^d \ \ , \ \ W_{j,k} f(x) = f \star \psi_{j,k}(x)$$

with

$$\psi_{j,k}(x) = 2^{-dj} \psi_k(2^{-j}x) \ .$$

It is computed up to a coarse scale 2^J where the remaining low frequencies are carried by a low-pass filtering $f \star \psi_{J,0}(x)$, where $\psi_0(x)$ is a real low frequency scaling function. Let $\hat{f}(\omega)$ be the Fourier transform of f with $\omega \in \mathbf{R}^d$. The modulus of $\omega \in \mathbf{R}^d$ is written $|\omega|$. Since $\widehat{W_{j,k} f}(\omega) = \hat{f}(\omega)\, \hat{\psi}_k(2^j\omega)$ and $\hat{f}(-\omega) = \hat{f}^*(\omega)$, if for all $\omega \in R^d$

$$(1-\delta) \leq |\hat{\psi}_0(2^J\omega)|^2 \ \ + \ \ \sum_{k=1}^{K-1}\sum_{j\geq J}\Big(|\hat{\psi}_k(2^j\omega)|^2 \ \ + \ \ |\hat{\psi}_k(-2^j\omega)|^2\Big)/2 \leq 1 \qquad (1)$$

then the wavelet transform is a complete contracting mapping

$$\|f\|^2(1-\delta) \leq \|f\star\psi_{0,J}\|^2 + \sum_{k=1}^{K-1}\sum_{j\geq J}\|W_{j,k}f\|^2 \leq \|f\|^2,$$

with $\|f\|^2 = \int |f(x)|^2\,dx$. We consider complex analytic wavelets such that $\hat\psi_k(\omega) = 0$ if $\hat\psi_k(-\omega) \neq 0$ for $k \geq 1$. At low frequencies, $\hat\psi_0(\omega)$ covers the domain $|\omega| \leq \pi$, with $\hat\psi_0(2p\pi) = 0$ for $p \in \mathbf{Z}^d$, and $\hat\psi_k(\omega)$ for $k \geq 1$ is mostly non-negligible inside a 1 octave frequency annulus $\pi \leq |\omega| \leq 2\pi$.

High frequency wavelet coefficients are mapped to low frequencies with a complex modulus which computes frequency interferences. The Fourier transform of $M_{j,k}f(x) = |W_{j,k}f(x)|^2$ is the convolution of $\widehat{W_{j,k}f}(\omega)$ with itself:

$$\widehat{M_{j,k}f}(\omega) = (2\pi)^{-d}\int \widehat{W_{j,k}f}(\xi)\,\widehat{W_{j,k}f}^*(\xi-\omega)\,d\xi. \tag{2}$$

This convolution measures the correlation between frequencies that are ω apart. In quantum physics, where probabilities are calculated as the squared modulus of complex wave functions, it is interpreted as interferences. Although $\widehat{W_{j,k}f}(\omega)$ is non-negligible inside a frequency annulus $2^{-j}\pi \leq |\omega| \leq 2^{-j+1}\pi$, (2) shows that $\widehat{M_{j,k}f}(\omega)$ is a correlation measure which is mostly non-zero at lower-frequencies $|\omega| \leq 2^{-j}\pi$.

To iterate this mapping and guarantee stability, the squared complex modulus is replaced by a modulus, which is contracting. It involves a square root operator $|W_{j,k}f(x)| = \sqrt{M_{j,k}f(x)}$, which is singular when $W_{j,k}f(x)$ vanishes. Let us write

$$|W_{j,k}f(x)|^2 = \|W_{j,k}f\|^2\,w^2(x)\,(1+\varepsilon(x)),$$

where $w(x)$ which is constant over the support of f with $\|w\| = 1$. A series expansion of $\sqrt{1+\varepsilon}$ gives

$$|W_{j,k}f(x)| = \|W_{j,k}f\|\,w(x)\left(1+\frac{1}{2}\varepsilon(x)+O(\varepsilon^2(x))\right).$$

The lower frequencies of $|W_{j,k}f(x)|$ are dominated by the squared modulus interferences term $\varepsilon(x)$ and the $O(\varepsilon^2(x))$ higher order terms produce higher frequency harmonics of low amplitude. As a result, $|W_{j,k}f(x)|$ has a Fourier tranform which is also mostly located at the lower frequencies $|\omega| \leq 2^{-j}\pi$.

2.2 Recursive Interference Tree

Recursive interferometry computes a progressively lower frequency representation by iteratively calculating complex wavelet transforms and modulus operators, which produce "interferences of interferences".

An interference tree up to a scale 2^J is a set of signals $\tilde I_j f(x,\alpha)$ located at the nodes of a tree, where $j \leq J$ gives the depth of a node and α its horizontal position in a left to right order. The wavelet transform modulus of f builds a first tree branch with $K-1$ leaves per level, which carry 1st order interferences at each scale

$$\tilde I_j f(x,k) = |f \star \psi_{j,k}(x)| \ \text{ for } \ j \leq J \text{ and } 1 \leq k < K$$

plus the low signal frequencies at the last level

$$\tilde I_J f(x,0) = f \star \psi_{J,0}(x)\ .$$

Each of the $K-1$ leaves of depths $m < J$ are subdecomposed with a second wavelet transform and modulus operator, which computes second order interferences located at the leaves of a new tree of depth J.

The interference tree is progressively constructed by decomposing the signals $\tilde I_m f(x,\alpha)$ at the leaves of a previously calculated tree, with a wavelet transform modulus up to a scale 2^J, until all the tree leaves are at the depth J, as illustrated in Figure 1. The wavelet transform modulus of $\tilde I_m f(x,\alpha)$ up to the level J defines a new tree whose leaves are

$$\tilde I_j f(x,\alpha K^{j-m}+k) = |\tilde I_m f(.,\alpha)\star\psi_{j,k}(x)| \text{ for } l < j \leq J\ ,$$

and

$$\tilde I_J f(x,\alpha K^{J-m}) = \tilde I_m f(.,\alpha)\star\psi_{0,J}(x)\ .$$

The signals $\tilde I_j f(x,\alpha)$ are recursive interferences, computed with $p(\alpha)$ wavelet transforms and modulus operators. The interference order $p(\alpha)$ at a node α is the number of non-zero digit of α written in base K.

All tree signals $\tilde I_j f(x,\alpha)$ are further filtered with the low-pass filter $\psi_{0,j}(x)$ to eliminate high frequency harmonics resulting from the last modulus computation:

$$I_j f(x,\alpha) = \tilde I_j f(.,\alpha)\star\psi_{0,j}(x)\ .$$

If $f(x) \in \mathbf{L^2}[0,1]^d$ has a period 1 along the d directions, then interference signals $\tilde I_J f(x,\alpha)$ have also a period 1. Since $\hat\psi_0(2p\pi) = 0$ for $p \in \mathbf{Z}^d$, at the maximum scale $2^J = 1$, all $I_0 f(x,\alpha)$ are constant in x. The tree leaves stores a single value $I_0 f(\alpha)$ providing a delocalized information on the whole support of f.

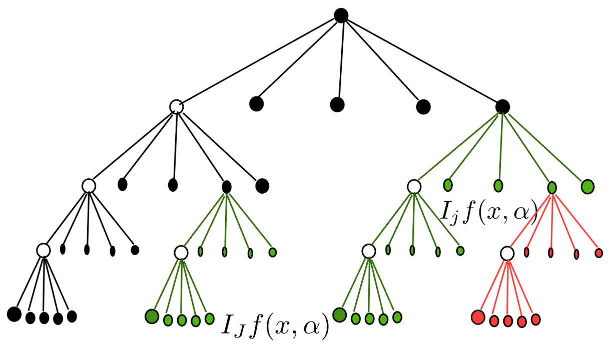

Figure 1: A recursive interference tree computes a first wavelet transform modulus (in black), and iteratively computes wavelet transform modulus of its leaves (2nd order in green and 3rd order in red), until all leaves are at the maxium depth.

3. INVARIANCE AND DISCRIMINABILITY

The classification ability of recursive interferometry relies on its invariance and discriminability properties that are reviewed. The norm of interference signals at a depth j is

$$\|I_j f\|^2 = \sum_\alpha \|I_j f(.,\alpha)\|^2$$

with $\|I_j f(.,\alpha)\|^2 = \int |I_j f(x,\alpha)|^2\,dx$. An interference tree is computed with a succession of wavelet transforms and modulus operators and a final low-pass filtering, which are all contracting operators. The resulting transform is therefore also contracting:

$$\|I_J f\| \leq \|f\|\ .$$

Let $D_\tau f(x) = f(x-\tau(x))$ be an elastic translation with $\tau(x) = (\tau_m(x))_{m\leq q} \in \mathbf{R}^d$ for $x = (x_p)_{p\leq q}$. We consider invertible deformations which satisfy:

$$|\nabla\tau(x)| = \left(\sum_{p,m=1}^{d}\left|\frac{\partial\tau_m(x)}{\partial x_p}\right|^2\right)^{1/2} < 1-a \text{ with } a>0.$$

The following theorem [4] proves that at large scales, a recursive interferometric transform is nearly invariant to such deformations. We write $\|\tau\|_\infty = \sup_x |\tau(x)|$ the maximum deformation amplitude, and $\tau\,.\,\nabla f = \sum_p \tau_p\,\partial f/\partial x_p$.

Theorem 1 *If the support of $\hat{f}$ is in $[-N\pi,N\pi]^d$ then there exists C that does not depend on f with:*

$$\|I_J D_\tau f - I_J f\| \leq C\,\|f\|\left(2^{-J}\|\tau\|_\infty + \log N\,\|\nabla\tau\|_\infty\right) \quad (3)$$

and

$$\|I_J D_\tau f - I_J f - \tau\,.\,\nabla I_J f\| \leq C\,\|f\|\Big(2^{-2J}\|\tau\|_\infty^2 + \log N\,\|\nabla\tau\|_\infty\Big)\ . \quad (4)$$

The error terms depends the maximum translation amplitude $\|\tau\|_\infty$ relatively to the scale 2^J and on the size of the elastic deformation measured by $\|\,|\nabla\tau|\,\|_\infty$. The residual error is reduced by an order of magnitude with a linearization of the deformation in (4). If we neglect the error, then at each position x the deformation $\tau(x)$ can be estimated by solving a system of linear equations

$$\forall\alpha,\ I_J D_\tau f(x,\alpha) - I_J f(x,\alpha) - \tau(x)\,.\,\nabla I_J f(x,\alpha) \approx 0.$$

This system has no solution if the error is not negligible in (4) because either the elastic deformation amplitude $|\nabla\tau|$ is too large or the scale 2^J is too small.

If f is 1 periodic then the translation error term $2^{-J}\|\tau\|_\infty$ disappears at the maximum scale $2^J = 1$, and I_0 is fully invariant to rigid translations. Computing an invariant representation relatively to a group is a form of quotient of the signal space by this group, One must ensure that the resulting dimensionality reduction is not too strong to preserve the discriminability between signals in the transformed space.

Suppose that the support of $\hat{f}(\omega)$ is included in $[-N\pi,N\pi]^d$ so that f belongs to a space $\mathbf{V}_N$ of dimension N^d. The Fourier transform modulus is a translation invariant transformation, which maps $\mathbf{V}_N$ over a half space of dimension $N^d/2$. However, the Fourier transform modulus is not stable relatively to elastic deformations. The operator I_0 maps $\mathbf{V}_N$ over a more complex non-linear manifold. Some properties of this manifold are studied in [4], in the particular case where $\hat{\psi}_k(\omega)$ are indicator functions of non-overlapping frequency bands. It proves that the manifold has a dimension larger then $N^{d/2}$, which means that the I_0 can be inverted over certain balls of dimension $N^{d/2}$ in $\mathbf{V}_N$. Continuity relatively to elastic deformations comes with a much larger dimensionality reduction then with a Fourier modulus but the manifold dimensionality remains large.

Figure 2 gives a simple classification example illustrating the translation invariance and discrimiability over deformable templates. Deformable templates [1] are obtained by applying deformation operators on deterministic signals. We consider two classes $\mathscr{C}_1$ and $\mathscr{C}_2$ whose elements are realizations of two random processes $F_1 = f_1(x - \tau_1(x))$ and $F_2(x) = f_2(x - \tau_2(x))$. The template signals f_1 and f_2 are deformed with two elastic random deformations $\tau_1(x)$ and $\tau_2(x)$ satisfying $|\tau_1'(x)| \leq a < 1$ and $|\tau_2'(x)| \leq a < 1$. Let $\tilde{F}_i = F_i + W$ be a noisy realization of F_i with an additive Gaussian white noise W. Figures 2(a,b) show two realizations of $\tilde{F}_1$ and $\tilde{F}_2$.

The probability distributions of $\|\Phi(\tilde{F}_i) - \Phi(\tilde{F}_{i'})\|^2$ is shown in Figure 2(c) for $\Phi(f) = f$ and in Figure 2(d) for a Fourier modulus $\Phi(f) = |\hat{f}|$. In these two cases, the intra class distance for $i = i'$ is of the same order as the distances across classes when $i \neq i'$. Indeed, if $\Phi(f) = f$ then the signal representation is not invariant to translation and $\Phi(f) = |\hat{f}|$ is not stable relatively to elastic deformations. Both classes can therefore not be discriminated with these distances.

Figure 2(e) gives the distribution of $\|\Phi(\tilde{F}_i) - \Phi(\tilde{F}_{i'})\|^2$ for $\Phi(f) = I_0 f$. Recursive interferences are computed with a one-dimensional Gabor wavelet $\psi(x) = \theta(x) e^{i\xi x}$, where θ is a Gaussian. The distance is larger across classes $(i \neq i')$ then within classes $(i = i')$, so both classes can be discriminated by thresholding the distance on recursive interferences.

3.1 Stationary Processes Interferences

Not all signal classes may be obtained as deformations of a deterministic template signal. In particular, realizations of a stationary texture are not elastic deformations of a single signal. Recursive interferences map the realizations of a stationary process to a small ball in the transformed space. Discriminating the realizations of two stationary processes is thus possible through the Euclidean distance of their interference representation.

If F is a zero-mean stationary process then $I_J F(x,\alpha)$ remains stationary in x. Indeed, it is computed with a cascade of wavelet transforms which are convolutions and modulus operators, which both preserve stationarity. Let $\sigma^2 = E\{|F(x) - E\{F(x)\}|^2\}$. Wavelet signals $F \star \psi_{j,k}(x)$ are stationary processes, and (1) implies that their variance $\sigma_{j,k}^2$ satisfy

$$(1-\delta)\sigma^2 \leq \sum_{j,k} \sigma_{j,k}^2 \leq \sigma^2 \,.$$

However, the modulus operator reduce these variances because of the complex phase suppression.

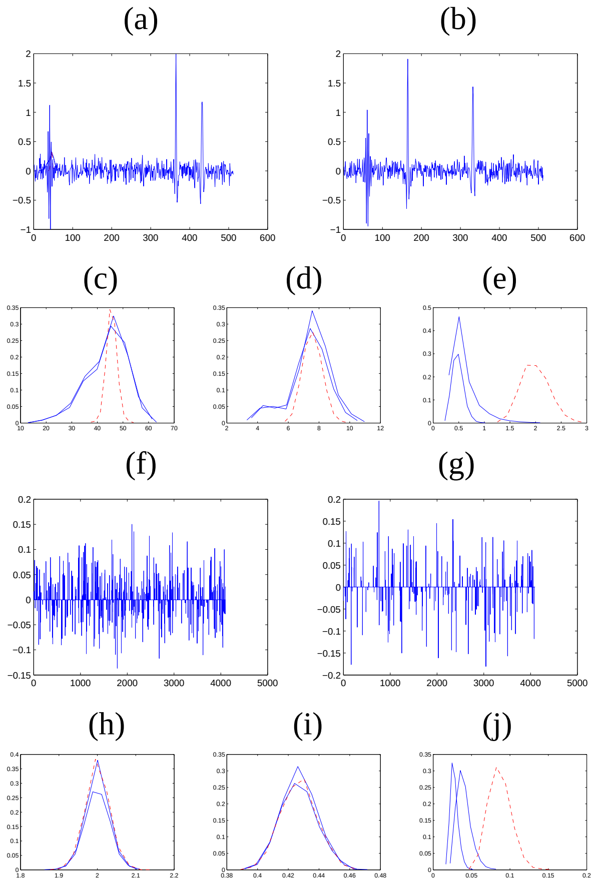

Figure 2: (a,b): noisy signals $\tilde{F}_1$ and $\tilde{F}_2$. (c,d,e): distributions of $\|\Phi(\tilde{F}_i) - \Phi(\tilde{F}_{i'})\|$ for $i = i'$ (full blue curves), and $i \neq i'$ (dashed red curves) for $\Phi(f) = f$ in (a), $\Phi(f) = |\hat{f}|$ in (b) and $\Phi(f) = I_0 f$ in (c). (f,g): realizations of two white noises F_1 and F_2. (h,i,j): distributions of $\|\Phi(F_i) - \Phi(F_{i'})\|$ as in (c,d,e).

The variance $\tilde{\sigma}_{j,k}^2$ of $|F \star \psi_{j,k}(x)|$ satisfies

$$\frac{\tilde{\sigma}_{j,k}^2}{\sigma_{j,k}^2} = 1 - \frac{m_{j,k}^2}{\sigma_{j,k}^2}$$

where $m_{j,k}$ is the mean of $|F \star \psi_{j,k}(x)|$ and hence the first order moment of $F \star \psi_{j,k}(x)$.

If F is a Gaussian process then $\tilde{\sigma}_{j,k}^2 / \sigma_{j,k}^2 = 1 - \pi/4$. If the correlation of F decreases sufficiently quickly then one can show that $I_J F(x,\alpha)$ remains nearly Gaussian for each α. If $I_J F(x,\alpha)$ are interferences of order $p(\alpha)$, obtained with a succession of p wavelet convolutions and modulus operators, then the variance σ_{α}^2. of $I_J F(x,\alpha)$ decreases like $\sigma^2 (1-\pi/4)^p$. If $\sigma_{j,k}^2 = O(\sigma^2)$, one can derive [4] that

$$\frac{E\{\|I_0 F - E\{I_0 F\}\|^2\}}{E\{\|F\|^2\}} = O(N^{-\beta}) \;\text{ with }\; \beta > 0\,,$$

where as $\|E\{I_0F\}\|^2 \sim E\{\|F\|^2\}$. It shows that I_0F remains in a ball whose spread is much smaller then its distance to 0. Realizations of two stationary processes are discriminated by measuring the distance of their interference transform.

Figures 2(f,g) show the realizations of two different white noise processes F_1 and F_2. Their supports are defined by two Bernouilli distributions $\mathrm{Prob}\{(F_i(n)) = 0\} = p_i$ and $\mathrm{Prob}\{(F_i(n)) \neq 0\} = 1 - p_i$, with $p_1 = 2\,p_2$. Over its support, each $F_i(n)$ is a Gaussian white noise. For $\Phi(f) = f$ and $\Phi(f) = |\hat{f}|$, Figures 2(h,i) show that the distribution of $\|\Phi(F_i) - \Phi(F_j)\|$ are similar within the same class $i = i'$ and and across classes $i \neq i'$, when $\Phi(f) = f$ and $\Phi(f) = |\hat{f}|$. On the opposit, intra class and across class distances are well separated by an interference representation $\Phi(f) = I_0 f$.

4. FAST ALGORITHM WITH MODULUS FILTER BANK

This section describes a fast filter bank algorithm which computes the recursive interference transform of a multidimensional discrete signal $f[n]$ of size N, with $n = (n_1, ..., n_d)$. The computational structure involves a cascade of convolutions and modulus operators as in deep neural architectures [2, 3] but involves no learning.

We consider $f[n]$ as a signal obtained by sampling a 1 period function $f(x)$ at intervals $2^L = N^{-1}$. Each $\tilde{I}_j f(x,\alpha)$ has a frequency support mostly concentrated at frequencies $|\omega| \leq 2^{-j}\pi$ but may go beyond, and it is thus uniformly sampled at intervals 2^{j-1}. We write $\tilde{I}_j f[n,\alpha] = \tilde{I}_j f(2^j n, \alpha)$.

The discrete wavelet transform of f is computed at scales $2^j < 2^L = N^{-1}$. The root of the tree is at the level L and $\tilde{I}_L f[n,0] = f[n]$. The finest scale wavelet transform of $f[n]$ is computed without subsampling, using discrete wavelet filters $\psi_{1,k}[n] = 2^{-1}\,\psi_k(2^{-1}n)$:

$$\tilde{I}_{L+1} f[n,0] = f \star \psi_{1,0}[n]$$

and

$$\tilde{I}_{L+1} f[n,k] = |\, f \star \psi_{1,k}[n]\,| \text{ for } 0 < k < K.$$

These signals are nearly oversampled by a factor 2 relatively to their frequency spread.

The filtering algorithm continues recursively by computing the K children of each $\tilde{I}_j f[n,\alpha]$, with one low-pass filter and $K-1$ complex band-pass filters, which are subsampled by a factor 2. If $\alpha \neq 0 \bmod K$ and is thus a band-pass filter output then the wavelet transform is calculated with oversampled wavelets filters $\psi_{2,k}[n] = 2^{-2}\,\psi_k(2^{-2}n)$:

$$\tilde{I}_{j+1} f[n,\alpha K] = \tilde{I}_j f[.,\alpha] \star \psi_{2,0}[2n]$$

and

$$\tilde{I}_{j+1} f[n,\alpha K + k] = |\tilde{I}_j f[.,\alpha] \star \psi_{2,k}[2n]| \text{ for } 0 < k < K.$$

If $\alpha = 0 \bmod K$ and is thus a low-pass filter output then since $\tilde{I}_j f[n,\alpha]$ was already obtained through a convolution with $\psi_{j,0}$, the next wavelet scale is calculated with the filters $g_k[n]$ whose transfer function $\hat{g}_k(\omega)$ satisfies:

$$\hat{\psi}_k(2\omega) = \hat{g}_k(\omega/2)\,\hat{\psi}_0(\omega)\ .$$

Children are then computed with

$$\tilde{I}_{j+1} f[n,\alpha K] = \tilde{I}_j f[.,\alpha] \star g_0[2n]$$

and

$$\tilde{I}_{j+1} f[n,\alpha K + k] = |\tilde{I}_j f[.,\alpha] \star g_k[2n]| \text{ for } 0 < k < K.$$

The oversampling factor 2 is finally removed by filtering $\tilde{I}_j f[n,\alpha]$ with the low-pass filter $\psi_{1,0}[n]$ and by subsampling the output

$$I_j[n,\alpha] = \tilde{I}_j f[.,\alpha] \star \psi_{1,0}[2n]\ .$$

At each level j of the tree, there are K^{j-L} indices α and each signal $I_j f[n,\alpha]$ has 2^{-dj} samples, so there is a total of $2^{-dj} K^{j-L}$ coefficients, with $2^L = N^{-1}$. If $d = 1$ and $K = 2$ there are N coefficients. The filter bank algorithm is implemented with $O(N \log_2 N)$ operations. If $K > 2$ then the $2^{-dj} K^{j-L}$ coefficients are computed with $O(2^{-dJ} K^{J-L})$ operations, which makes $O(N^{\log_2(K)})$ at the bottom of the tree.

REFERENCES

[1] Y. Amit,, U. Grenander and M. Piccioni; "Structural image restoration through deformable templates". Journal of the American Statistical Association, 86 (414), 376? 387. 1991.

[2] K. Jarrett, K. Kavukcuoglu, M. Ranzato and Y. LeCun: What is the Best Multi-Stage Architecture for Object Recognition ?, Proc. of ICCV 2009.

[3] Jake Bouvrie, Lorenzo Rosasco, Tomasso Poggio: On Invariance in Hierarchical Models. NIPS 2009

[4] S. Mallat "Invariance by Recursive Interferometry", Tech Report, CMAP, Ecole Polytechnique, 2010.