EFFICIENT IMAGE REPRESENTATION BY ANISOTROPIC REFINEMENT IN MATCHING PURSUIT.

Pierre Vandergheynst and Pascal Frossard

Signal Processing Laboratory
Swiss Federal Institute of Technology
CH- 1015 Lausanne, Switzerland
Pierre.Vandergheynst@epfl.ch

ABSTRACT

This paper presents a new image representation method based on anisotropic refinement. It has been shown that wavelets are not optimal to code 2-D objects which need true 2-D dictionaries for efficient approximation. We propose to use rotations and anisotropic scaling to build a real bi-dimensional dictionary. Matching Pursuit then stands as a natural candidate to provide an image representation with an anisotropic refinement scheme. It basically decomposes the image as a series of basis functions weighted by their respective coefficients. Even if the basis functions can a priori take any form bi-dimensional dictionaries are almost exclusively composed of two-dimensional Gabor functions. We present here a new dictionary design by introducing orientation and anisotropic refinement of a gaussian generating function. The new dictionary permits to efficiently code 2-D objects and more particularly oriented contours. It is shown to clearly outperform common non-oriented Gabor dictionaries.

1. INTRODUCTION

Non-orthogonal transforms presents several interesting properties which position them as an interesting alternative to orthogonal transforms like DCT or wavelet based schemes. Decomposing a signal over a redundant dictionary improves the compression efficiency, especially at low bit rates where most of the signal energy is captured by only few elements. Moreover it provides a great flexibility in image representation and more particularly offers the possibility to use true 2-D dictionaries. Such dictionaries are necessary for an efficient approximation of bi-dimensional objects. The main limitation of non-orthogonal transforms is however the encoding complexity, since the number of possible decompositions becomes infinite.

Matching Pursuit algorithms [1] provide an interesting way to iteratively decompose the signal in its most important features with a limited complexity. It outputs a stream composed of atoms or basis functions along with their respective coefficients. The atoms are chosen among a redundant dictionary of basis functions. The dictionary design is completely open but it has to be carefully completed since it directly drives the coding performance. In image coding, the two-dimensional basis functions have to efficiently capture the features of natural images while offering efficient coding possibilities. This paper presents a new oriented and anisotropically refined dictionary based on gaussian functions which have shown to maximize the uncertainty principle. The new dictionary efficiently captures oriented contours as well as weakly textured regions. It moreover favorably compares to commonly used two-dimensional separable Gabor functions.

This paper is structured as follows. Sec. 2 first overviews the Matching Pursuit algorithm and the convergence properties of the coding process. Sec. 3 then emphasizes the need for true 2-D dictionaries and presents a new design method based on oriented and anisotropically refinement of gaussian basis functions. Experimental results and comparisons with common Gabor atoms are presented in Sec. 4. Finally, concluding remarks are given in Sec. 5.

This work has been partly supported by Hewlett-Packard.

2. MATCHING PURSUIT OVERVIEW

In contrast to orthogonal transforms, overcomplete expansions of signals are not unique. The number of feasible decompositions is infinite, and finding the best solution under a given criteria is a NP-complete problem. In compression, one is interested in representing the signal with the smallest number of elements, that is in finding the solution with most of the energy on only a few functions. Matching Pursuit [1] is one of the sub-optimal approaches that greedily approximates the solution to this NP-complete problem.

Matching Pursuit (MP) is an adaptive algorithm that iteratively decomposes any function f in the Hilbert space $\mathcal{H}$ in a possibly redundant dictionary of functions called *atoms* [1]. Let $\mathcal{D} = \{g_\gamma\}_{\gamma\in\Gamma}$ be such a dictionary with $\|g_\gamma\| = 1$ and Γ represents the set of possible indexes. The function f is first decomposed as follows :

$$f = \langle g_{\gamma_0}|f\rangle\, g_{\gamma_0} + \mathcal{R}f\,, \tag{1}$$

where $\langle g_{\gamma_0}|f\rangle\, g_{\gamma_0}$ represents the projection of f onto g_{γ_0} and $\mathcal{R}f$ is a residual component. Since all elements in $\mathcal{D}$ have by definition a unit norm, it is easy to see from eq. (1) that g_{γ_0} is orthogonal to $\mathcal{R}f$, and this leads to

$$\|f\|^2 = |\langle g_{\gamma_0}|f\rangle|^2 + \|\mathcal{R}f\|^2\ . \tag{2}$$

To minimize $\|\mathcal{R}f\|$, one must choose g_{γ_0} such that the projection coefficient $|\langle g_{\gamma_0}|f\rangle|$ is maximum. The pursuit is carried out by applying iteratively the same strategy to the residual component. After N iterations, one has the following decomposition for f :

$$f = \sum_{n=0}^{N-1} \langle g_{\gamma_n}|\mathcal{R}^n f\rangle g_{\gamma_n} + \mathcal{R}^N f\,, \tag{3}$$

where $\mathcal{R}^N$ is the residual of the N^{th} step with $\mathcal{R}^0 f = f$. Similarly, the energy $\|f\|^2$ is decomposed into :

$$\|f\|^2 = \sum_{n=0}^{N-1} |\langle g_{\gamma_n} | \mathcal{R}^n f \rangle|^2 + \|\mathcal{R}^N f\|^2 \ . \tag{4}$$

Although Matching Pursuit places very few restrictions on the dictionary, the latter is strongly related to convergence speed and thus to coding efficiency. Any collection of arbitrarily sized and shaped functions can be used as dictionary, as long as the completeness is respected. The completeness property ensures that the Matching Pursuit is able to perfectly recover the input signal after a possibly infinite number of iterations.

The convergence speed of the Matching Pursuit corresponds to its ability to extract the maximum signal energy in a few iterations. In other words, it corresponds to the decay rate of the residue and thus the coding efficiency of the Matching Pursuit. The approximation error decay rate in Matching Pursuit have been shown to be bounded by an exponential [1, 2]. From [3], there exists $\alpha > 0$ and $\beta > 0$ such that for all $m \geq 0$:

$$\|\mathcal{R}^{m+1} f\| \leq (1 - \alpha^2 \ \beta^2)^{\frac{1}{2}} \ \|\mathcal{R}^m f\| \ , \tag{5}$$

where $\alpha \in (0, 1]$ is an optimality factor. This factor depends on the algorithm that, at each iteration, searches for the best atom in the dictionary. The optimality factor α is set to one when the MP browses the complete dictionary at each iteration. The parameter β depends on the dictionary construction. It represents to ability of the dictionary functions to capture features of any input function f and satisfies :

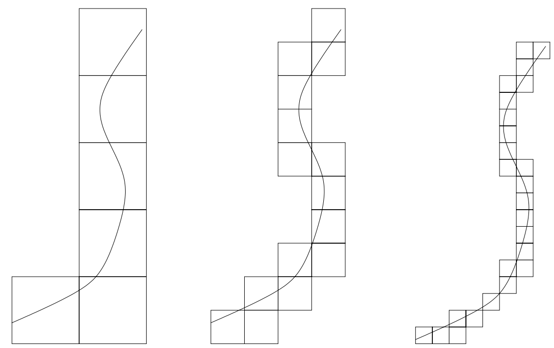
$$\sup_{\gamma} |\langle f, g_{\gamma_n} \rangle| \geq \beta \|f\|. \tag{6}$$

The redundancy factor β corresponds thus to the cosine of the maximum possible angle between a direction f and its closest direction among all dictionary vectors. A general formulation of the redundancy can be found in [4].

3. ORIENTED AND ANISOTROPICALLY REFINED DICTIONARIES

3.1. The need for anisotropic refinement schemes

In the last few years, wavelet based image compression schemes have been highly optimized. They are nowadays acknowledged to be a very efficient solution allowing for high compression ratios while keeping good quality and low complexity. The final step in the rise of wavelet techniques will be the release of the JPEG2000 compression standard which is entirely based on wavelets [5]. The efficiency of wavelets to achieve good compression ratios lies in some of their intrinsic mathematical properties. One of the most important is probably the ability of wavelets to capture the essential features of a signal with a small number of coefficients. The sparsity of the wavelet representation comes from its near optimal non-linear approximation rate when considering signals in the Besov class (piecewise smooth with any number of singularities). Intuitively, it is a simple consequence of the vanishing moment and compact support properties of wavelet basis : locally polynomial parts of the signal are filtered out by vanishing moments. Only singularities will give high wavelet coefficients. Now, since wavelet basis have short support, few wavelets will occasionally intersect singular parts of the signal, giving rise to a highly sparse representation [6, 7]. These nice properties are unfortunately not available in two dimensions and this opens the door to new image representations. Indeed, an image can still be modeled as a piecewise smooth 2-D signal with singularities, but the latter are not point like anymore. Higher dimensional singularities may be highly organized along embedded submanifolds and this is exactly what happens at image contours for example. Figure 1 shows that wavelets are inefficient at representing contours because they cannot deal with smoothness of the contours themselves. This is mainly due to the isotropic refinement implemented by wavelet basis : the dyadic scaling factor is applied in all directions, where clearly it should be fine along the direction of the local gradient and coarse in the orthogonal direction.



(a) 6 coefficients (b) 15 coefficients (c) 25 coefficients

Fig. 1. Inadequacy of isotropic refinement for representing contours in images. The number of wavelets intersecting the singularity is roughly doubled when the resolution increases.

Candes and Donoho [8] have recently proposed a construction called *the curvelet transform* which aims at solving this problem. Basically curvelets can be first oriented along a given direction and then arbitrarily refined in the orthogonal direction. Curvelet frames have been shown to achieve a much better non-linear approximation rate than wavelets thanks to this anisotropic scaling scheme. Unfortunately, even though the curvelet representation is very sparse, it yields a tremendous expansion factor in the number of data [9]. Matching Pursuit, as already stressed before, iteratively chooses the best matching terms in a dictionary. Since there is almost no constraint on the dictionary itself, MP stands as a natural candidate to implement an efficient anisotropic refinement scheme and such a construction is detailed in the next section.

3.2. Anisotropic refinement using Matching Pursuit

Our dictionary is built by acting on a generating function of unit L^2 norm by means of a family of unitary operators U_γ :

$$\mathcal{D} = \{U_\gamma, \ \gamma \in \Gamma\} \ , \tag{7}$$

for a given set of indexes Γ. Basically this set must contain three types of operations :

- Translations $\vec{b}$, to move the atom all over the image.
- Rotations θ, to locally orient the atom along contours.
- Anistropic scaling (a_1, a_2), to adapt to contour smoothness.

A possible action of U_γ on the generating atom g is thus given by :

$$U_\gamma g = \mathcal{U}(\vec{b}, \theta) D(a_1, a_2) g \tag{8}$$

where $\mathcal{U}$ is a representation of the Euclidean group,

$$\mathcal{U}(\vec{b}, \theta) g(\vec{x}) = g\big(r_{-\theta}(\vec{x} - \vec{b})\big) \,, \tag{9}$$

r_θ is a rotation matrix, and D acts as an anisotropic dilation operator :

$$D(a_1, a_2) g(x, y) = \frac{1}{\sqrt{a_1 a_2}} g\big(\frac{x}{a_1}, \frac{y}{a_2}\big) \,. \tag{10}$$

It is easy to prove that such a dictionary is overcomplete using the fact that, when $a_1 = a_2$ one gets 2-D continuous wavelets as defined in [10]. It is also worth stressing that, avoiding rotations, the parameter space is a group studied by Bernier and Taylor [11]. The advantage of such a parametrization is that the full dictionary is invariant under translations and rotations. Moreover, it is also invariant under isotropic scaling, e.g. $a_1 = a_2$.

The choice of the generating atom g is driven by the idea of efficiently approximating contour like singularities in 2-D. To achieve this, the atom must be a smooth low resolution function in the direction of the contour and must behave like a wavelet in the orthogonal (singular) direction. In our experiments, we chose a combination of a gaussian and its second derivative, that is :

$$g(x, y) = (4\, x^2 - 2)\ \exp(-(x^2 + y^2)) \,. \tag{11}$$

This choice is motivated by the optimal joint spatial and frequency localization of the gaussian kernel. We also noticed that degradations caused by truncating the MP expansion are visually less disturbing with this choice.

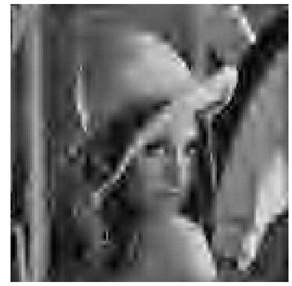
For practical implementations the range of all parameters in the dictionary must be discretized. We have chosen to discretize the scaling parameters using a dyadic grid while we kept a simple uniform grid for the position and rotation parameters. The following section describes numerical results obtained by comparing this technique with wavelets and other MP dictionaries.

4. EXPERIMENTAL RESULTS

4.1. Comparison with wavelets

We first compare our scheme with wavelet approximation. It should be stressed that this comparison is based on measuring the quality obtained by using the biggest N terms in a wavelet or MP expansion of images. We are thus only interested in the non-linear approximation power of these techniques. In order to convert these results into compression ratios, one would need to implement a fair coding of both wavelet and MP coefficients. Even though wavelet coders have been well studied [12], efficient MP encoding is still an essentially open subject. Preliminary work nevertheless show very promising results in this direction [13]. To provide readers with a fair comparison, Figure 2 compares the quality of 500 MP iterations with 500, 1000 and up to 1500 wavelet coefficients. It can be seen that the Matching Pursuit representation outperform all the wavelet expansions. Even with three times more terms wavelets are not able to provide a quality equivalent to Matching Pursuit decomposition. This clearly shows the power of a truly 2-D scheme with adaptive refinement.

(a) PSNR = 29.69 dB



(b) PSNR = 23.47 dB

(c) PSNR = 26.36 dB

(d) PSNR = 28.45 dB

Fig. 2. Reconstructed version of *Lena* encoded with (a) 500 anisotropically refined atoms and respectively (b) 500, (c) 1000 and (d) 1500 wavelet coefficients.

4.2. Comparison with other dictionaries

A natural question at this point is whether the quality of these results is due to the choice of the dictionary or to the use of matching pursuit solely. Indeed, MP is being considered as a valid alternative to wavelets or DCT in low bit rate coding schemes [14]. We have implemented two different dictionaries. The first one uses oriented Gabor atoms generated by translation, rotation and isotropic scaling of a modulated gaussian :

$$U\big(a, \theta, \vec{b}\big) g(\vec{x}) = \frac{1}{a}\, g\big(a^{-1} r_{-\theta}(\vec{x} - \vec{b})\big) \,, \tag{12}$$

whith

$$g(\vec{x}) = e^{i\vec{\omega_0}\cdot\vec{x}} e^{-\|\vec{x}\|^2/2} \,. \tag{13}$$

The other dictionary used in this comparison is an affine Weyl-Heisenberg dictionary built by translating, dilating and modulating the Gabor generating atom of Eq. 13 :

$$U\big(a, \vec{\omega}, \vec{b}\big) g(\vec{x}) = \frac{1}{a}\, e^{i\vec{\omega}\cdot(\vec{x}-\vec{b})} g\big(a^{-1}(\vec{x} - \vec{b})\big) \,. \tag{14}$$

Figure 3 shows the reconstructed PSNR as a function of the number of iterations in the MP expansion. Clearly, anisotropic scaling outperforms the other dictionaries. This comparison shows that the use of rotations is also of interest since the oriented Gabor dictionary gives better results than the modulated one. It is worth noticing that rotations and anisotropic scaling are really 2-D transformations and this shows that, in order to efficiently approximate 2-D objects, one has to use 2-D dictionaries. Separable transforms, although they provide faster implementations, are unable to cope with the geometry of edges.

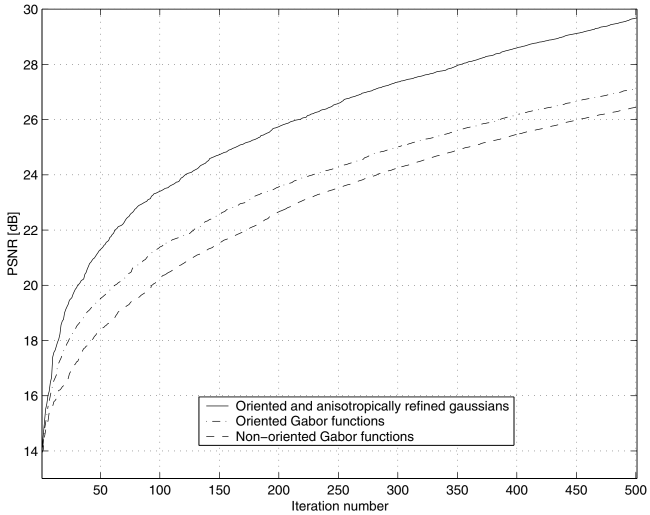


Fig. 3. Comparison of the quality of MP reconstruction when using three different dictionaries : anisotropic scaling, Gabor wavelets and Weyl-Heisenberg dictionary.

5. CONCLUSIONS

This paper has shown that true bi-dimensional transforms are necessary to provide an efficient image representation. Anisotropic refinement and orientation are true 2-D transformations that offer the possibility to advantageously approximate the bi-dimensional geometrical pattern. Matching Pursuit has made possible a simple implementation of anisotropic refinement which has been shown to clearly outperform common Gabor dictionaries. Finally, complexity has voluntarily not been addressed in this paper. Modifications of the basic Matching Pursuit used in this paper will lead to significant complexity reduction. In the same time, an efficient coding scheme for the Matching Pursuit coefficients and atoms indexes will also be investigated.

6. REFERENCES

[1] Mallat S.G. and Zhang Z., "Matching Pursuits With Time-Frequency Dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397–3415, December 1993.

[2] Davis G., Mallat S. and Avellaneda M., "Adaptive Greedy Approximations," *Journal of Constructive Approximations*, vol. 13, pp. 57–98, 1997.

[3] Mallat S., *A Wavelet Tour of Signal Processing*, Academic Press, 2 edition, 1999.

[4] Frossard P. and Vandergheynst P., "Redundancy in Non-Orthogonal Transforms," in *IEEE International Symposium on Information Theory*, Washington D.C., June 2001, submitted paper.

[5] ISO/IEC JTC1/SC29/WG1, "JPEG 2000 - lossless and lossy compression of continous-tone and bi-level still images: Part I minimum decoder," Final Committee Draft Version 1.0, March 2000.

[6] DeVore R.A., *Acta Numerica*, vol. 7, chapter Nonlinear Approximations, pp. 51–150, 1998.

[7] Donoho D.L., Vetterli M., DeVore R.A. and Daubechies I., "Data Compression and Harmonic Analysis," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2435–2476, October 1998.

[8] Candès E.J. and Donoho D.L., "Ridgelets : A key to higher-dimensional intermitency ?," *Phil. Trans. R. Soc. Lond. A*, 1999, to appear.

[9] Donoho D.L. and Duncan M.R., "Digital curvelet transform: strategy, implementation and experiments," Tech. Rep., Department of Statistics, Stanford University, 1999.

[10] Antoine J.-P., Murenzi R. and Vandergheynst P., "Directional wavelets revisited: Cauchy wavelets and symmetry detection in patterns," *Applied and Computational Harmonic Analysis*, vol. 6, pp. 314–345, 1999.

[11] Bernier D. and Taylor K., "Wavelets from square-integrable representations," *SIAM Journal on Mathematical Analysis*, vol. 27, no. 2, pp. 594–608, 1996.

[12] Said A. and Pearlman W.A., "A New, Fast, and Efficient Image Codec Based on Set Partitioning in Hierarchical Trees," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 3, pp. 243–250, June 1996.

[13] Rabiee H.R., Kashyap R.L. and Safavian S.R., "Multiresolution Segmentation-Based Image Coding with Hierarchical Data Structures," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1996, vol. 4, pp. 1870–1873.

[14] Neff R. and Zakhor A., "Very Low Bit-Rate Video Coding Based on Matching Pursuits," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 1, pp. 158–171, February 1997.