Frames for Exact Inversion of the Rank Order Coder

Masmoudi Khaled, Marc Antonini, Pierre Kornprobst

HAL Id: inria-00627075
https://hal.inria.fr/inria-00627075

Submitted on 29 Sep 2011

Frames for Exact Inversion of the Rank Order Coder

Khaled Masmoudi — Marc Antonini — Pierre Kornprobst

N° 7744

September 2011

Thème BIO

Rapport de recherche

Frames for Exact Inversion of the Rank Order Coder

Khaled Masmoudi , Marc Antonini* , Pierre Kornprobst [†]

Thème BIO — Systèmes biologiques
Équipes-Projets NeuroMathComp & UNSA–CNRS–I3S

Abstract: Our goal is to revisit rank order coding by proposing an original exact decoding procedure for it. Rank order coding was proposed by Simon Thorpe et al. who stated that the retina represents the visual stimulus by the order in which its cells are activated. A classical rank order coder/decoder was then designed on this basis [1]. Though, it appeared that the decoding procedure employed yields reconstruction errors that limit the model Rate/Quality performances when used as an image codec. The attempts made in the literature to overcome this issue are time consuming and alter the coding procedure, or are lacking mathematical support and feasibility for standard size images. Here we solve this problem in an original fashion by using the frames theory, where a frame of a vector space designates an extension for the notion of basis. First, we prove that the analyzing filter bank considered is a frame, and then we define the corresponding dual frame that is necessary for the exact image reconstruction. Second, to deal with the problem of memory overhead, we design a recursive out-of-core blockwise algorithm for the computation of this dual frame. Our work provides a mathematical formalism for the retinal model under study and defines a simple and exact reverse transform for it with up to 270 dB of PSNR gain compared to [1]. Furthermore, the framework presented here can be extended to several models of the visual cortical areas using redundant representations.

Key-words: Rank order code, bio-inspired image coding, frame theory, scalability, difference of Gaussians.

* I3S
[†] INRIA, EPI NeuroMathComp

Frames for Exact Inversion of the Rank Order Coder

Résumé :   Notre objectif est de revisiter le codage d'images statiques par rang en proposant une procédure originale de décodage exact. Le codage par rang a été proposé par Simon Thorpe et al. qui a affirmé que la rétine représente le stimulus visuel par l'ordre selon lequel ses cellules sont activées. Un codeur par ordre classique ainsi que le décodeur ont ensuite été conçus se basant sur ces résultats [1]. Cependant, il s'avère que la procédure de décodage employé engendre des erreurs de reconstruction qui limitent les performances Débit / Qualité du modèle lorsqu'il est utilisé comme un codec d'images. Les tentatives proposées dans la littérature pour surmonter ce problème prennent du temps et modifie la procédure de codage, ou manquent d'apport mathématique et de faisabilité pour des images de tailles standards. Ici nous résolvons ce problème de façon originale en utilisant la théorie des "frames", où une frame d'un espace vectoriel désigne une extension de la notion de base. Tout d'abord, nous montrons que le banc de filtres d'analyse considéré est une frame, puis nous définissons la frame duale correspondante qui est nécessaire pour la reconstruction exacte de l'image. Deuxièmement, pour faire face au problème du débordement de mémoire, nous concevons un algorithme récursif, out-of-core, et opérant par blocs pour le calcul de cette frame duale. Notre travail fournit un formalisme mathématique pour le modèle de la rétine à l'étude et définit une inversion simple et exacte de la transformée bio-inspirée définie dans [1] avec un maximum de 270 dB de gain de PSNR par rapport au modèle originel. Par ailleurs, le travail présenté ici peut être étendu à plusieurs autres modèles de zones corticales visuelles utilisant des représentations redondantes.

Mots-clés :   Codage par rang, codage d'images bio-inspiré, théorie des frames, scalabilté.

1 Introduction

Neurophysiologists made substantial progress in better understanding the early processing of the visual stimuli. Especially, several efforts proved the ability of the retina to code and transmit a huge amount of data under strong time and bandwidth constraints [2, 3, 4]. The retina does this by means of a neural code consisting in a set of electrical impulses: the spikes. Thus, our aim is to use the computational neuroscience models that mimic the retina behavior to design novel lossy coders for static images.

Although there is no clear evidence on how the spikes encode for the visual stimuli, we assume in this paper that the relevant encoding feature is the order in which the retina cells emit their first spike given the stimulus onset. This choice was motivated by Thorpe et al. neurophysiologic results on ultra-rapid stimulus categorization [2, 5]. The authors showed that still image classification can be achieved by the visual cortex within very short latencies of about 150 ms. As an explanation, it was stated that: *There is information in the order in which the cells fire*, and thus the temporal ordering can be used as a code. This code, termed as rank order code (ROC), is at the origin of a classical retina model [1].

However, one major limitation of the ROC coder defined in [1] prevents us from its use for the design of image codecs. It is the inaccuracy of the proposed decoding procedure. Indeed, the bio-inspired retina model that generates the spikes is based on a redundant filter bank image analysis, where the considered filters are not strictly orthogonal. Thus, the filter overlap yields reconstruction errors that limit the Rate/Quality performance [6, 7, 8]. Efforts to correct this issue followed two main approaches. A first one consisted in inverting directly the transform operator to obtain a reverse filter bank as in [9, 7]. The method presented is based on a pseudo-inversion. Though interesting, we will show that this method lacks mathematical support. Besides the procedure used deals with a high dimension matrix and thus is infeasible, as such, for standard size images. A second approach relies on matching pursuit (MP) algorithms as in [6, 7]. These methods are time consuming and alter the coding procedure. In addition, the MP approach depends on the order in which the "match and update" mechanism is performed, and this makes the coding procedure depend on the stimulus itself.

In this paper, we give an original solution relying on the mathematical concept of frames and dual frames, which is an extension of the notion of basis [10, 11]. First, we express the model in a matrix-based fashion. Then we prove that the analyzing filter bank as we define it is a frame. Finally, we construct the corresponding dual frame that is necessary for the exact image reconstruction. Besides we design an out-of-core algorithm for the computation of the dual frame which resolves the encountered issues of memory overhead. Our method requires the computation of a reverse operator once for all and keeps the bio-inspired coding scheme unchanged. Thanks to this approach, we show that the reconstructed image that we get is equal to the original stimulus.

This paper is organized as follows: In Section 2, we present the three stages of the rank order coding/decoding method. Then in Section 3 we define an exact decoding scheme through the construction of a dual frame. Finally in Section 4, we show the gain that we obtain in terms of Rate/Quality tradeoff.

2 The rank order codec: Three stages

This Section summarizes the three stages of the ROC coding/decoding procedure as defined in [1]. First we present in Section 2.1 the image transform as performed by a bio-inspired retina model. We then give the specification of the subsequent rank ordering and decoding procedures in Sections 2.2 and 2.3.

2.1 The image transform: A bio-inspired retina model

Neurophysiologic experiments have shown that, as for classical image coders, the retina encodes the stimulus representation in a transform domain. The retinal stimulus transform is performed in the cells of the outer layers. Quantitative studies have proven that the outer cells processing can be approximated by a linear filtering. In particular, the authors in [12] proposed the largely adopted DoG filter which is a weighted difference of spatial Gaussians that is defined as $DoG(x,y) = w^c G_{\sigma^c}(x,y) - w^s G_{\sigma^s}(x,y)$, where w^c and w^s are the respective positive weights of the center and surround components of the receptive fields, σ^c and σ^s are the standard deviations of the Gaussian kernels G_{σ^c} and G_{σ^s}, such that $\sigma^c < \sigma^s$. The DoG cells can be arranged in a dyadic grid Γ of K layers to sweep all the stimulus spectrum as shown in Figure 1(b) [1, 6, 8]. This architecture also mimics to some extent the actual retina topology as the cells density and scale are inversely proportional. Each layer $0 \leqslant k < K$ in the grid Γ, is tiled with filtering cells, denoted by DoG_k, having a scale k and generating a transform subband B_k such that:

$$DoG_k(x,y) = w^c G_{\sigma_k^c}(x,y) - w^s G_{\sigma_k^s}(x,y), \tag{1}$$

where $\sigma_{k+1}^c = \frac{1}{2}\sigma_k^c$ and $\sigma_{k+1}^s = \frac{1}{2}\sigma_k^s$. Each DoG_k filter has a size of $(2M_k+1)^2$, with $M_k = 3\,\sigma_k^s$. Authors in [1] chose the biologically plausible parameters $w^c = w^s = 1$, $\sigma_k^c = \frac{1}{3}\sigma_k^s \;\forall k$, and $\sigma_{K-1}^c = 0.5$ pixel.

In order to measure the degree of activation c_{kij} of a given retina cell, such that $(k,i,j) \in \Gamma$, we compute the convolution of the original image f by the DoG_k filter. Yet each layer k in the dyadic grid Γ is undersampled with a step of 2^{K-k-1} pixels with an original offset of $\lfloor 2^{K-k-2} \rfloor$ pixels, where $\lfloor . \rfloor$ is the floor operator. Having this, we define the function u_k, such that the c_{kij} coefficients are computed at the locations $\big(u_k(i), u_k(j)\big)$ as follows:

$$u_k(i) = \lfloor 2^{K-k-2} \rfloor + 2^{K-k-1}i, \forall k \in [\![0, K-1]\!]. \tag{2}$$

u_k can be seen as an undersampling function. We notice that $u_{K-1}(i) = i$, and that the functions $(u_k)_{k \in [\![0,K-2]\!]}$ are undersampled versions of u_{K-1}. c_{kij} is then computed as follows:

$$c_{kij} = \sum_{x=u_k(i)-M_k, y=u_k(j)-M_k}^{x=u_k(i)+M_k, y=u_k(j)+M_k} DoG_k(u_k(i) - x, u_k(j) - y) \, f(x,y). \tag{3}$$

This transform generates a vector c of $(\frac{4}{3}N^2 - 1)$ coefficients c_{kij} for an N^2-sized image (if N is a power of 2). This architecture is similar to a Laplacian pyramid [13]. An example of such a transform performed on the *cameraman* test image is shown in Figure 1.
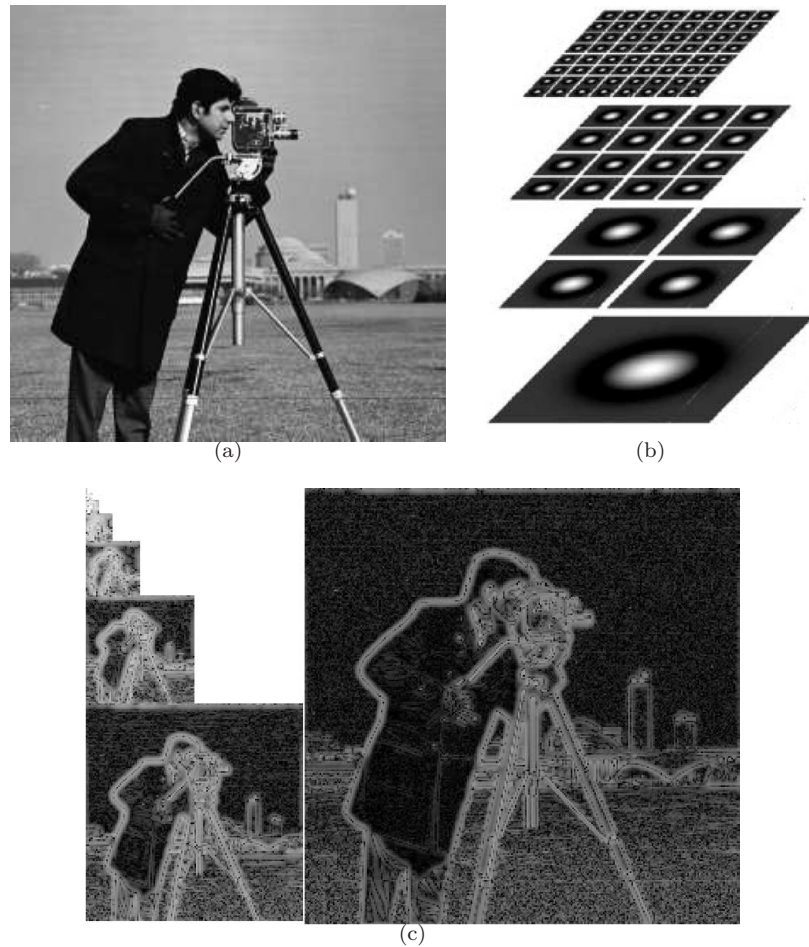
Figure 1: Illustration of the retinal image transform applied on *cameraman* 1(a). Image size is 257×257 pixels. 1(b): Example of a dyadic grid of DoG's used for the image analysis (from [1]). 1(c): The transform result showing the generated subbands in log scale.

2.2 Sorting: The generation of the rank order code

Thorpe et al [2, 5] proposed that the order in which the spikes are emitted encodes for the stimulus. This yielded the ROC which relies on the following simplifying assumptions:

(i) From stimulus onset, only the first spike emitted by each cell is considered.

(ii) The time to fire of each cell is proportional to its degree of activation.

(iii) Only the order of firing of the cells encodes for the stimulus.

Such a code gives a biologically plausible interpretation to the rapidity of the visual stimuli processing in the human visual system. Indeed, it seems that most of the processing is based on feed-forward mechanism before any feedback

occurs [5]. So, the neurons responses $(c_{kij})_{kij \in \Gamma}$ defined in Equation (3) are sorted in the decreasing order of their amplitude $|c_{kij}|$.

The final output of this stage, the ROC, is then a sorted list of N_s couples (p, c_p) such that $|c_p| \geqslant |c_{p+1}|$, with p being the index of the cell defined by $p(k, i, j) = k N_k^2 + i N_k + j$ and N_k^2 being the number of cells in the subband B_k. Here, the generated series $(p, c_p)_{0 \leqslant p < N_s}$ is the only data transmitted to the decoder.

2.3 Decoding procedure of the rank order code

We consider the set of the first N_s highest cell responses forming the ROC of a given image f. In [1], the authors defined $\tilde{f}_{N_s}$, the decoded estimation of f by:

$$\tilde{f}_{N_s}(x, y) = \sum_{p(k,i,j)=0}^{N_s - 1} c_p \, DoG_k(u_k(i) - x, u_k(j) - y). \qquad (4)$$

Equation (4) defines a progressive reconstruction depending on N_s. Indeed, one can restrict the code to the most valuable coefficients c_p, i.e the most activated cells of the retina. This feature makes the coder be scalable [8].

An example of such a reconstruction is given in Figure 2(a), with all the retina cells taken into account. Figure 2(b) also shows that the retina model decoding procedure, though giving a good approximation of the stimulus, is still inaccurate. In this example, reconstruction quality is evaluated to 27 dB of PSNR. This is due to the fact that the DoG filters are almost but not exactly orthonormal. We detail in the next section our original method to reconstruct exactly the input f.

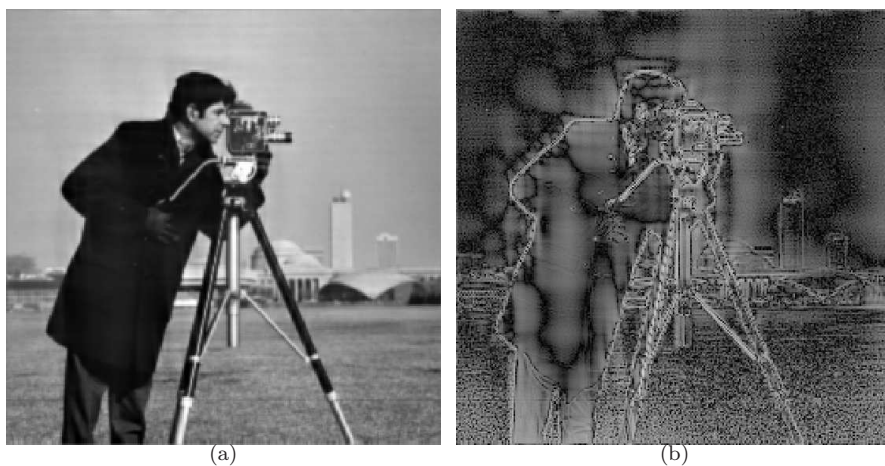(a)   (b)

Figure 2: Result of the decoding procedure with the original approach using all of the retina cells responses. 2(a): Reconstructed image. The PSNR quality measure of $\tilde{f}_{N_s}$ yields 27 dB. 2(b): Error image in a logarithmic scale: high frequencies are the ones that are the most affected by this approach.

3 Inverting the bio-inspired retina model

In this section, we define an original and exact image reconstruction algorithm starting from the ROC. First, we introduce in Section 3.1 a low-pass scaling function in the analyzing filter bank. This modification will be shown to be necessary for the transform reversibility. Then, in Section 3.2, we give a matrix-based formalism for the transform and we use it to prove that our filter bank is a frame in Section 3.3. Finally, in Section 3.4, we show the exact reconstruction results using the dual frame and introduce an out-of-core algorithm to construct it.

3.1 Introduction of a low-pass scaling function

We introduce a low-pass scaling function in the filter bank used for image analysis. This modification do not alter the ROC coder architecture and has both a mathematical and a biological justification. Indeed, the Fourier transform of a Gaussian is another Gaussian, so that $\mathscr{F}(DoG_k)$ is a difference of Gaussians. Therefore, with $w^c = w^s = 1$ (cf. Equation (1)), we have:

$$\mathscr{F}(DoG_k) = 2\pi \left(\sigma_k^c\right)^2 G_{\left(\sigma_k^c\right)^{-1}} - 2\pi \left(\sigma_k^s\right)^2 G_{\left(\sigma_k^s\right)^{-1}}. \tag{5}$$

We can easily verify that the central Fourier coefficient $\mathscr{F}(DoG_k)\big(u_0(0), u_0(0)\big) = 0 \ \forall k$, and that $\mathscr{F}(DoG_k)(i,j) > 0 \ \forall (i,j) \neq \big(u_0(0), u_0(0)\big)$.

In order to cover up the centre of the spectrum, we propose to replace the DoG_0 filter, with no change in the notation, by a Gaussian low-pass scaling function consisting in its central component, such that:

$$DoG_0(x, y) = w^c G_{\sigma_0^c}(x, y). \tag{6}$$

Figures 3(a) and 3(b) show the spectrum partitioning with the different DoG_k filters ($k \geq 1$, in blue) and the spectrum of the new scaling function DoG_0 (in red dashed line) which covers low frequencies. With no scaling function, all constant images would be mapped into the null image 0 and this would make the transform be non-invertible. Here we overcome this problem as the central Fourier coefficient $\mathscr{F}(DoG_0)\big(u_0(0), u_0(0)\big) > 0$.

The scaling function introduction is further justified by the actual retina behavior. Indeed, the surround $G_{\sigma_k^s}$ in Equation (1) appears progressively across time driving the filter passband from low frequencies to higher ones. So that, the Gaussian scaling function represents the very early state of the retina cells.

Once we introduced the DoG_0 scaling function, and in order to define an inverse for the new transform, let us demonstrate that it is a "frame". So that, we will define the adequate notations in Section 3.2 and give the proofs in Section 3.3.

3.2 Matrix notations for ROC

Unlike the implementations in [1, 6, 8], let us introduce the matrix Φ to compute the image transform through the retina model. The rows of Φ are the different DoG_k filters. This yields an "undersampled Toeplitz-block" sparse matrix
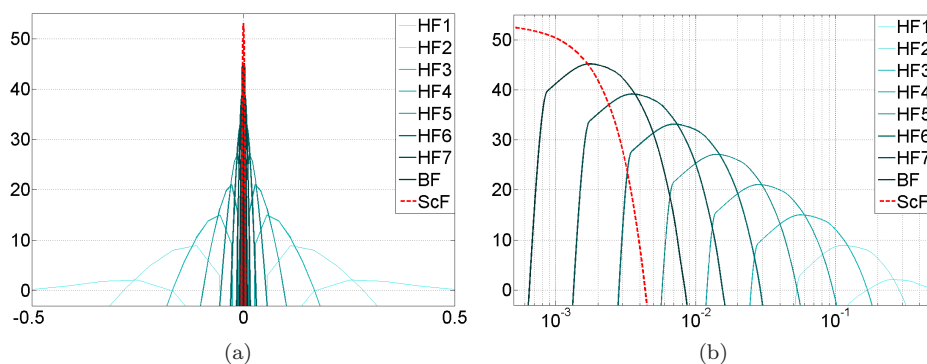
(a)

(b)

Figure 3:   3(a): Spectrum of the DoG filters. The abscissa represents the frequencies. The ordinate axis represents the different DoG_k filters gain in dB. 3(b): Half of the spectrum in 3(a) with the abscissa having a logarithmic step. The scaling function DoG_0 is plotted in red dashed line.
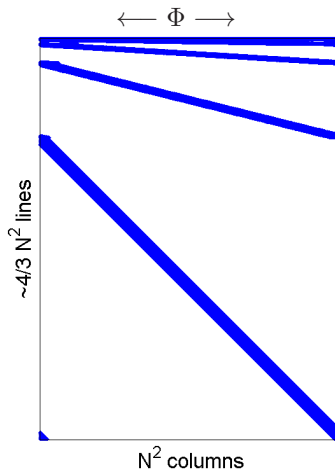

Figure 4:   Template of the DoG transform matrix: In this paper, the transform Φ is represented as a matrix where blue dots correspond to non-zero elements. Note here that Φ is a highly sparse matrix.

(see Figure 4). Such an implementation allows fast computation of the multiscale retinal transform through sparse matrix specific algorithms. This will in addition help us to construct the dual frame of Φ.

With this notation, the DoG transform is rewritten:

$$c = \Phi\, f. \tag{7}$$

Interestingly, the straightforward synthesis as defined in (4) amounts to the multiplication of the vector output c by Φ^* the Hermitian transpose of Φ. Then, the reconstruction procedure defined in (4) is rewritten:

$$\tilde{f}_{N_s} = \Phi^*\, c. \tag{8}$$

3.3 The DoG transform is a frame operator

Our aim is to prove that the bio-inspired retina image transform presented amounts to a projection of the input image f on a frame of a vector space. The frame is a generalization of the idea of a basis to sets which may be linearly dependent [10, 11]. Frames allow a redundant signal representation which, for instance, can be employed for coding with error resilience. By proving the frame nature of this transform, we will be able to achieve an exact reverse transform through the construction of a dual frame.

According to [11], to prove that our transform is a frame we need to show that $\exists\, 0 < \alpha \leqslant \beta < \infty$, such that:

$$\alpha \left\| f \right\|^2 \leqslant \sum_{kij \in \Gamma} \left(c_{kij} \right)^2 \leqslant \beta \left\| f \right\|^2, \ \forall f. \tag{9}$$

In the literature, one can find some related ideas. For example, authors in [14] proved that the Laplacian pyramid [13] is a frame and estimated experimentally α and β. The authors in [15] also designed an invertible Laplacian pyramid based on the same principles. However, though there are some similarities, the DoG transform does not use the same filters and it is based on a different algorithm. Thus, in the following, we give an original demonstration in our specific case which proves that the DoG transform is a frame.

Proposition 3.1 *Let* $\Psi(\phi) = \{\phi_{kij}, (k, i, j) \in \Gamma\}$ *be the set of vectors* ϕ_{kij}*, such that* $\phi_{kij}(x, y) = DoG_k\left(u_k(i) - x, u_k(j) - y\right)$ *as defined in Equation* (1) *for* $k > 0$*, and in Equation* (6) *for* $k = 0$*. Then* $\Psi(\phi)$ *is a frame of the* $N \times N$ *images vector space.*

Proof **Upper bounding:** We have:

$$\sum_{kij \in \Gamma} \left(c_{kij} \right)^2 = \sum_{k=0}^{K-1} \left\| B_k \right\|^2, \tag{10}$$

where B_k is the subband of scale k generated by the image transform with:

$$B_k(i, j) = \sum_{x=u_k(i)-M_k, y=u_k(j)-M_k}^{x=u_k(i)+M_k, y=u_k(j)+M_k} DoG_k(u_k(i) - x, u_k(j) - y)\, f(x, y).$$

If we denote by U_k the undersampling operator corresponding to the function u_k (cf. Equation (2)), we can write the following:

$$B_k = U_k(DoG_k * f). \tag{11}$$

Then, we have the following inequalities:

$$\left\| B_k \right\| = \left\| U_k(DoG_k * f) \right\| \leqslant \left\| U_k\left(|DoG_k| * |f| \right) \right\|$$
$$\leqslant \left\| |DoG_k| * |f| \right\|$$
$$\leqslant \left\| DoG_k \right\| \left\| f \right\|.$$

So, with (10) we infer the following bounding:

$$\sum_{kij \in \Gamma} (c_{kij})^2 = \sum_{k=0}^{K-1} \|B_k\|^2 \leqslant \left(\sum_{k=0}^{K-1} \|DoG_k\|^2\right) \|f\|^2$$
$$= \beta \|f\|^2, \tag{12}$$

which shows the first inequality.

Lower bounding: We start from the fact that:

$$\sum_{k=0}^{K-1} \|B_k\|^2 \geqslant \|B_{K-1}\|^2 + \|B_0\|^2, \tag{13}$$

which amounts to write the following inequalities:

$$\sum_{kij \in \Gamma} (c_{kij})^2 \geqslant \|DoG_{K-1} * f\|^2 + \left\|\left(DoG_0 * f\right)\left(u_0(0), u_0(0)\right)\right\|^2$$
$$= \|\mathscr{F}(DoG_{K-1})\,\mathscr{F}(f)\|^2 + \left\|\left(\mathscr{F}(DoG_0)\,\mathscr{F}(f)\right)\left(u_0(0), u_0(0)\right)\right\|^2$$
$$= \sum_{i,j=0}^{N-1} \left(\mathscr{F}(DoG_{K-1})(i,j)\,\mathscr{F}(f)(i,j)\right)^2$$
$$+ \left\|\mathscr{F}(DoG_0)\left(u_0(0), u_0(0)\right)\,\mathscr{F}(f)\left(u_0(0), u_0(0)\right)\right\|^2,$$

We know that $\mathscr{F}(DoG_{K-1})(i,j) > 0$, $\forall (i,j) \neq \left(u_0(0), u_0(0)\right)$ and that $\mathscr{F}(DoG_{K-1})\left(u_0(0), u_0(0)\right) = 0$. We also have $\mathscr{F}(DoG_0)\left(u_0(0), u_0(0)\right) > 0$ (cf. Section 3.1). So, if we define a set S_{K-1} by $S_{K-1} = [\![0, N-1]\!]^2 \setminus \left(u_0(0), u_0(0)\right)$ and α by:

$$\alpha = min\Big\{\mathscr{F}(DoG_0)^2\left(u_0(0), u_0(0)\right),$$
$$\left\{\mathscr{F}(DoG_{K-1})^2(i,j),\ (i,j) \in S_{K-1}\right\}\Big\} > 0,$$

then we get the following:

$$\sum_{i,j=0}^{N-1} \left(\mathscr{F}(DoG_{K-1})(i,j)\,\mathscr{F}(f)(i,j)\right)^2$$
$$+ \left\|\mathscr{F}(DoG_0)\left(u_0(0), u_0(0)\right)\,\mathscr{F}(f)\left(u_0(0), u_0(0)\right)\right\|^2$$
$$= \sum_{i,j \in S_{K-1}} \left(\mathscr{F}(DoG_{K-1})(i,j)\,\mathscr{F}(f)(i,j)\right)^2$$
$$+ \left\|\mathscr{F}(DoG_0)\left(u_0(0), u_0(0)\right)\,\mathscr{F}(f)\left(u_0(0), u_0(0)\right)\right\|^2$$
$$\geqslant \alpha \sum_{i,j \in [\![0,N-1]\!]^2} \left(\mathscr{F}(f)(i,j)\right)^2 = \alpha \|f\|^2,$$

so that, $\sum_{kij \in \Gamma} (c_{kij})^2 \geqslant \alpha \|f\|^2$. $\qquad\qquad\square$

3.4   Synthesis using the dual DoG frame

We introduce in this section a correction means for the reconstruction error in the retina model presented through the frame theory.

The straightforward analysis/synthesis procedure can be outlined in the relation between the input image and the reconstruction estimate: $\tilde{f}_{N_s} = \Phi^*\Phi f$. As we already demonstrated that the DoG transform is a frame, $\Phi^*\Phi$ is said to be the frame operator. To have an exact reconstruction of f, one must construct the dual DoG vectors. A preliminary step is to compute $(\Phi^*\Phi)^{-1}$, the inverse frame operator. We then get a corrected reconstruction $f^*_{N_s}$, defined by: $f^*_{N_s} = (\Phi^*\Phi)^{-1}\tilde{f}_{N_s}$. If N_s is the total number of the retina model cells, we have:

$$f^*_{N_s} = (\Phi^*\Phi)^{-1}\tilde{f}_{N_s} = (\Phi^*\Phi)^{-1}\Phi^* \, c = (\Phi^*\Phi)^{-1}\Phi^* \, \Phi f = f.$$

As made clear through the equation above, the dual vectors are the rows of $(\Phi^*\Phi)^{-1}\Phi^*$. If we reconstruct f starting from the ROC output c and using the dual frame vectors, we get the results shown in Figure 5. The reconstruction obtained is accurate and requires only a simple matrix multiplication. In this example, reconstruction quality is evaluated to 296 dB of PSNR.

An example of dual vectors constructed as the rows of $(\Phi^*\Phi)^{-1}\Phi^*$ is shown in Figure 6. Dual vectors resemble the DoG analyzing filters. This is obvious as the straightforward image reconstruction $\tilde{f}_{N_s}$ is already close to f, which means that $\Phi^*\Phi$ is close to identity. Figure 6 shows also that the exact reconstruction of f is obtained by a relaxation in the symmetry constraint of the DoG filters.

The recursive out-of-core blockwise inversion algorithm Though the mathematical fundamentals underlying this work are simple, the implementation of such a process is a hard problem. Indeed, in spite of the sparsity of Φ and Φ^*, the frame operator $\Phi^*\Phi$ is an N^4-sized dense matrix for an N^2-sized image f. For instance, if $N = 257$, $\Phi^*\Phi$ holds in 16 Gbytes, and 258 Gbytes if $N = 513$. The solution is to recourse to the so-called out-of-core algorithms [16].

The frame operator $\Phi^*\Phi$ is constructed block by block, and each block is stored separately on disk. The inversion is then performed using a recursive algorithm that relies on the block-wise matrix inversion formula that follows:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}BQ^{-1}CA^{-1} & -A^{-1}BQ^{-1} \\ -Q^{-1}CA^{-1} & Q^{-1} \end{pmatrix},$$

where Q is the Schur complement of A, such that $Q = D - C \, A^{-1} \, B$. Thus, inverting a matrix amounts to the inversion of two matrices that are 4 times smaller. The inversion consists then in subdividing the problem by a factor 4 at each recursion level until we reach a single block problem. Obviously, this algorithm requires out-of-core blockwise matrix routines for multiplication, subtraction and addition, that we implemented in a multi-threaded fashion to accelerate the computation.

Advantages of our approach $(\Phi^*\Phi)^{-1}$ is a square, definite positive, and invertible matrix [11]. Another advantage of our method is that $(\Phi^*\Phi)^{-1}$ is well conditioned, with a conditioning number estimated to around 16, so that its inversion is stable. This is a crucial issue as previous work aimed at conceiving the DoG reverse transform tried to invert the original filter bank with no scaling function DoG_0 [9, 7]. This is obviously mathematically incorrect as the filter bank thus defined is not a frame and thus its pseudo inverse $(\Phi^*\Phi)^{-1}\Phi^*$ does not exist. The solution proposed by the authors of [9, 7] gives only a least squares

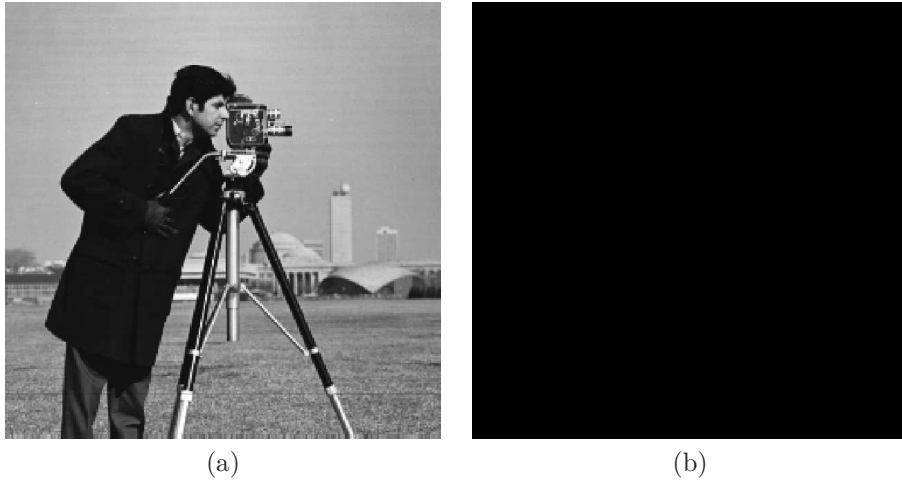(a)   (b)

Figure 5: Result of the decoding procedure with the dual DoG frame using all of the retina cells responses. 5(a): Reconstructed image. The PSNR quality measure of $f_{N_s}^*$ yields 296 dB. 5(b): Error image in logarithmic scale. This shows that the reconstruction using the dual frame is very accurate.
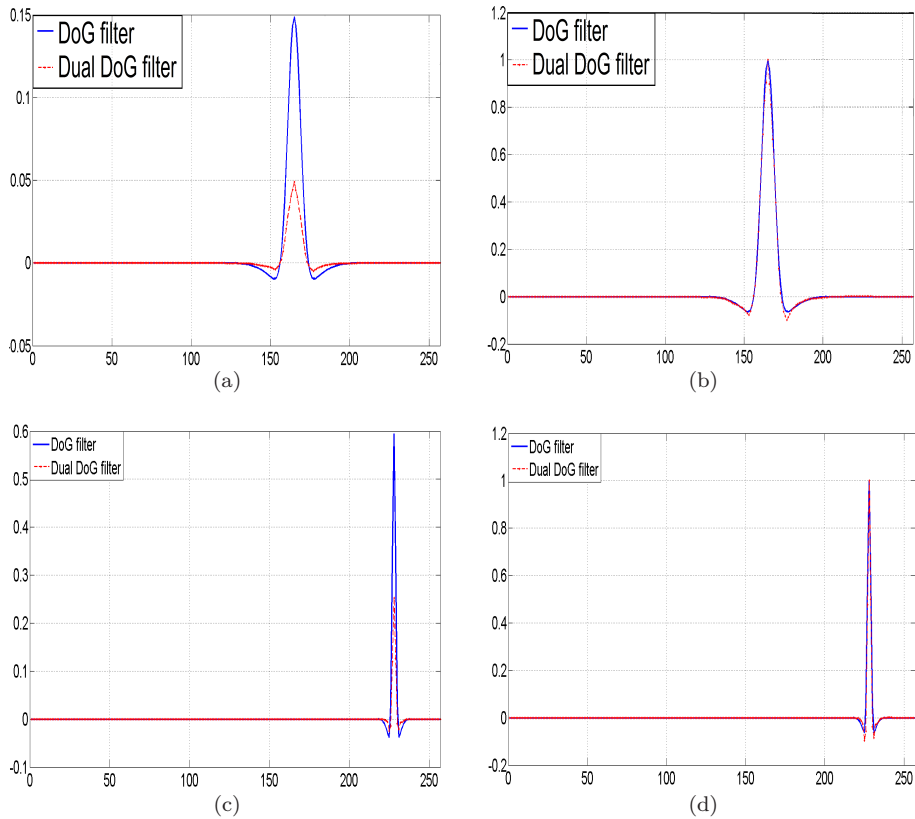

Figure 6: 6(a) : Comparison between the DoG filters and their duals. 6(b) : Same as previous with the highest value normalized to 1. Though close in shape to the DoG filters, the dual DoG filters are asymmetric.

Figure 7: Reconstruction of the *cameraman* image f using different percentages of the highest c_{kij}. Upper line shows the straightforward progressive synthesis $\tilde{f}_{N_s}$. Bottom line shows the corrected reconstruction $f^*_{N_s}$ using the dual frame. PSNR for the upper/lower image is from left to right: (19.2 dB/19.5 dB), (20.4 dB/20.8 dB), (24.08 dB/25 dB), (25.8 dB /27.5 dB), and (27.9 dB/296 dB).

solution to an ill-conditioned problem. Our method instead is stable. Besides through the out-of-core algorithm that we designed we can invert $(\Phi^*\Phi)$ even for large images whereas authors in [9, 7] are restricted to a maximum size of 32×32.

Furthermore, correcting the reconstruction errors using the adequate dual frame does not alter the coding procedure. Indeed, methods introduced in [6, 7] are based on a matching pursuit (MP). MP is time consuming and depends on the order in which the "match and update" mechanism is performed. Our method keeps the coding procedure straightforward, fast and order-independent.

4 Comparison to the original ROC codec

We experiment our new decoder in the context of scalable image decoding. We reconstruct the *cameraman* test image using an increasing number N_s of highest responses taken into account (cf. Equations (4) and (8)). We compare the results when using the original DoG filters in Φ^* and the dual DoG filters in $(\Phi^*\Phi)^{-1}\Phi^*$ for the decoding procedure. Figure 7 summarizes the results obtained, with the upper line showing the progressive straightforward reconstruction $\tilde{f}_{N_s}$ and the bottom line showing the corrected progressive reconstruction $f^*_{N_s}$. The gain in PSNR is significant for low rates (around 0.3 dB) and increases with N_s up to 270 dB. Figure 8 compares the Rate/Quality curves of the two methods and shows the high quality gain obtained with ours. Besides, our method does not alter the coding procedure and keeps it straightforward, fast and order-independent.
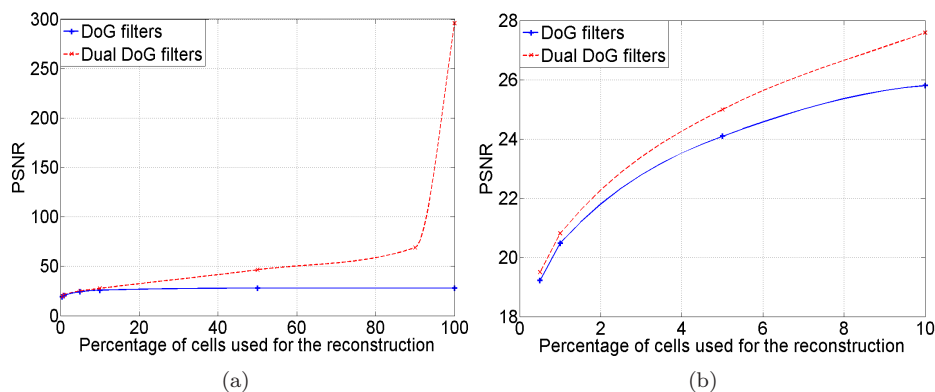
Figure 8: Quality of the reconstruction using the DoG filters [1] (in blue solid line) and the dual DoG filters (in red dashed line) as a function of the number of cells taken onto consideration. The abscissa represents the percentage of the highest responses used for the reconstruction. The ordinate represents the reconstruction quality in terms of PSNR. 8(a): Results shown for percentages between 0% and 100%. 8(b): Results shown for percentages between 0% and 10%.

References

[1] R. Van Rullen and S. Thorpe, "Rate coding versus temporal order coding: What the retinal ganglion cells tell the visual cortex," *Neural Comput.*, vol. 13, pp. 1255–1283, 2001.

[2] S. Thorpe, "Spike arrival times: A highly efficient coding scheme for neural networks," in *Parallel Processing in Neural Systems and Computers*, R. Eckmiller, G. Hartmann, and G. Hauske, Eds.   Elsevier, 1990, pp. 91–94.

[3] M. Meister and M. Berry, "The neural code of the retina," *Neuron*, pp. 435–450, 1999.

[4] T. Gollisch and M. Meister, "Rapid neural coding in the retina with relative spike latencies," *Science*, vol. 319, no. 5866, pp. 1108–1111, 2008.

[5] S. Thorpe, D. Fize, and C. Marlot, "Speed of processing in the human visual system," *Nature*, vol. 381, pp. 520–522, 1996.

[6] L. Perrinet, M. Samuelides, and S. Thorpe, "Coding static natural images using spiking event times: do neurons cooperate?" *IEEE Trans. Neural Netw.*, vol. 15, no. 5, pp. 1164–1175, 2004.

[7] B. S. Bhattacharya and S. Furber, "Biologically inspired means for rank-order encoding images: A quantitative analysis," *IEEE Trans. Neural Netw.*, vol. 21, no. 7, pp. 1087 –1099, 2010.

[8] K. Masmoudi, M. Antonini, P. Kornprobst, and L. Perrinet, "A novel bio-inspired static image compression scheme for noisy data transmission over

low-bandwidth channels," in *Proceedings of ICASSP*.   IEEE, 2010, pp. 3506–3509.

[9] B. S. Bhattacharya and S. Furber, "Maximising information recovery from rank-order codes," in *Proceedings of SPIE*, vol. 6570, no. 1, Orlando, FL, U.S.A., 2007.

[10] R. J. Duffin and A. C. Schaeffer, "A class of nonharmonic fourier series," *Trans. Amer. Math. Soc.*, vol. 72, no. 2, pp. pp. 341–366.

[11] J. Kovacevic and A. Chebira, *An introduction to frames.*   Foundations and Trends in Signal Processing, Now Pub, 2008.

[12] D. Field, "What is the goal of sensory coding?" *Neural Comput.*, vol. 6, no. 4, pp. 559–601, 1994.

[13] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, 1983.

[14] S. Rakshit and C. Anderson, "Error correction with frames: the Burt-Laplacian pyramid," *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 2091–2093, 1995.

[15] M. Do and M. Vetterli, "Framing pyramids," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2329–2342, 2003.

[16] S. Toledo, "A survey of out-of-core algorithms in numerical linear algebra," in *External Memory Algorithms and Visualization*, J. Abello and J. S. Vitter, Eds.   Amer. Math. Soc. Press, 1999, pp. 161–180.