Two-dimensional affine frames for image analysis and synthesis

Chen Sagiv [a,*], Nir A. Sochen [b], Yehoshua Y. Zeevi [c]

[a] *Eliezer Yaffe 37, Ra'anana 43451, Israel*
[b] *Department of Applied Mathematics, University of Tel Aviv, Ramat-Aviv, Tel-Aviv 69978, Israel*
[c] *Department of Electrical Engineering, Technion – Israel Institute of Technology, Technion City, Haifa 32000, Israel*

Abstract

An affine-group-based design methodology of Gabor-type filter bank is presented for the purpose of image analysis and synthesis. Various tessellations of the combined spatial-feature space are considered. We combine ideas introduced by Daugman [J.G. Daugman, Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, J. Opt. Soc. Am. 2 (7) (1985) 1160–1169], Lee [T.S. Lee, Image representation using 2D Gabor-wavelets, IEEE Trans. PAMI 18 (10) (1996) 959–971] and Manjunath and Ma [B.S. Manjunath, W.Y. Ma, Texture features browsing and retrieval of image data, IEEE Trans. PAMI 18 (8) (1996) 837–842], and extend them by applying the action of the full affine group on Gaborian-type mother wavelets. In this approach we adopt optimality criteria of minimal spatial-features combined uncertainty, as well as tightness of the frame tessellating this combined space. In this work, scalings in the x and y directions, allowing for independent dilations in these two directions, as well as rotations and translations are allowed. For each discrete set of scalings, rotations and translations the frame bounds are calculated. For frames where the frame operator is well approximated by a multiple of the identity, we use the same set of functions in the analysis and synthesis, as though the frame is equivalent to an orthogonal basis. Moreover, we show that in the case of independent scalings in the x and y directions, the number of dominant (characteristic) orientations of the filter bank may depend on scale. We further show that the orientation bandwidths thus obtained, resemble those attained under the constraint imposed by the uncertainty principle.
© 2007 Elsevier Inc. All rights reserved.

1. Introduction

Information regarding local features of images is essential for image understanding and for implementation of various image processing and computer vision algorithms. For example, quantification of the local frequency, scale and orientation in the neighborhood of a specific pixel can be useful in determining the structure and local properties of the image.

A common approach to specification of this type of information is via image analysis by means of a filter bank. A set of such bank of filters can be obtained by applying the action of a group of operations (e.g., translations,

* Corresponding author.
E-mail addresses: chensagivron@gmail.com (C. Sagiv), sochen@post.tau.ac.il (N.A. Sochen), zeevi@ee.technion.ac.il (Y.Y. Zeevi).

rotations, scaling and modulations) on some basic generating ("mother") function [11,14]. The local features are then obtained by the convolution of the generated filters with the image [20]. This can also be understood as a projection of the image on the subspace spanned by the filter.

Two interesting questions arise in this context. The first regards the selection of the optimal generating function. As there can be several optimality criteria a possible one is to select the function that provides maximal accuracy with respect to the local features of interest. This is equivalent to the selection of a generating function that minimizes the uncertainty of the local features. In 1946 Gabor has shown that representation of images using Gaussian functions modulated by complex exponentials is optimal in the sense of minimizing the joint uncertainty in the combined time–frequency space [9]. Thus, these functions provide the best trade-off between time- and frequency-resolution. These findings were later extended to two dimensions by Daugman [4], where he noted the reciprocal relations between the spatial domain resolution and the frequency or orientation resolution in the Fourier domain. He developed a relation between the orientation half-bandwidth and the spatial frequency bandwidth.

The problem of finding the minimizer for a generalized uncertainty principle was also discussed in the past in the context of harmonic analysis and more recently in the context of group theory. Researchers [1,6] have considered the affine group in one dimension and the similitude group in two dimensions. In these studies, it was shown that there does not exist a non-trivial canonical function which minimizes the uncertainty equation associated with the similitude group of $\mathbb{R}^2$. In a recent study, we have extended these results to the case of the full affine group and also considered the affine Weyl–Heisenberg group which accounts for spatial and frequency translations as well as for spatial dilations [18,19].

Once an optimal mother function is selected, the second question regards the optimal tessellation of the spatial and/or the features' space. The present study is devoted to the discrete sampling of the Lie group parameter-space such that an optimal filter bank is provided. For the sake of this analysis we choose to work with the Gabor filter bank. This is the set of modulations of a sine wave by functions that result from the application of some group action (e.g., translations, dilations and rotations) to a specific mother wavelet, a Gaussian in this case. The Gabor filter bank is chosen because of its simplicity and its frequency-scale sensitivity. It is not the only possible choice though. Other approaches that employ translation, rotation and dilation to generate a filter bank and extract the local properties of images were suggested in recent studies. The curvelet transform [3] accounts for directional parabolic scaling of the mother function. The shearlet transform [10,12] is a new approach, that employs affine systems, and therefore possesses mathematical properties similar to those of wavelets. More specifically, it accounts for a single generating mother shearlet function parameterized by scaling, shear, and translation parameters, where the shear parameter captures the direction of singularities [12]. Finally the contourlet transform [7] provides both multi-resolution and multi-direction expansion using non-separable filter banks that rely on contour segments.

Next, we would like to have a filter bank with good coverage of the combined spatial-features domain. Thus, the measure of optimality is adopted here as the synthesis ability, in the sense that the filter bank has a tight frame property. This property guaranties a good reconstruction of the image from the filter bank by linear combination of these filters. This turns the filter bank to become similar to an orthonormal system from the viewpoint of reconstruction (i.e., synthesis).

In this study, we combine ideas offered by Manjunath and Ma [15], Lee [13] and Daugman [4] to provide a systematic approach to tessellation of the combined space so that the related frame bounds can be evaluated.

Manjunath and Ma [15] have considered the issue of designing a Gabor-wavelets filter bank in the context of browsing and retrieval of images in a large database. Their task was to define a signature which characterizes the contents of an image. They have selected this signature to be the vector of the absolute values of the responses obtained when calculating the inner product between the image and a set of Gabor filters. They have used Gabor wavelets that were generated by the action of the SIM(2) group operations; i.e., dilations (with the same scaling factor for the x and y axes), rotations and translations. The similitude group of $\mathbb{R}^2$ as a topology space is isomorphic to: $SIM(2) = (\mathbb{S}^1 \times \mathbb{R}^+) \ltimes \mathbb{R}^2$. Choosing the parameterization of this space by $(\theta, a, \vec{b})$, where $a > 0$, $\theta \in [0, 2\pi]$, $\vec{b} \in \mathbb{R}^2$, the group law is then given by

$$(\theta, a, \vec{b})(\theta', a', \vec{b}') = (\theta + \theta', aa', \vec{b} + aR_\theta \vec{b}'), \tag{1}$$

where R_θ implies a rotation matrix by an angle of θ. The design strategy of Manjunath and Ma was based on the constraint that the curves of the filters' half-peak magnitude, in the frequency domain, kiss each other, e.g., they have one point in which they share the same tangent line.

Lee [13] has considered image representation in the context of a frame approach. Thus, the bank of filters selected in his study constitutes a frame. The frame criterion for one-dimensional wavelets was first developed by Daubechies [2]. It was further developed into a matrix algebraic approach for the analysis of multi-window Gabor-type schemes [21], and extended to two dimensions [13]. In the latter study, the spatial frequency space was paved using a single dilation parameter for the x and y directions. Moreover, the calculations of frame bounds were only carried out for the octave or sub-octave cases.

We follow the guidelines of the studies by Daugman, by Manjunath and Ma and the study by Lee, and offer various tessellations of Gabor filters. There are several novelties in our study. First we treat the scaling along the x and y directions separately. The consequence is an increase in the number of possible filter banks available. Next we choose the biologically motivated Gabor mother wavelet and calculate the frame bounds for different scalings in the x and y directions and present the results of reconstruction of images with the linear summation formula (we do not use the dual frame for the synthesis procedure. The interested reader is referred to [2,8,21]). We show that the higher is the overlap between the Gabor wavelets in the frequency domain, the tighter is the frame they form. Moreover, we show that when the scalings in the x and y directions are independent of each other, the number of orientations required to guarantee a certain degree of overlap between the frequency responses of the filters depends on the scales. We find that the calculated orientation bandwidth has an interesting resemblance to the orientation bandwidth derived by Daugman in the context of uncertainty.

The rest of this paper is organized as follows: First, we address the issue of tessellation of the position–frequency combined space with Gabor functions. We refer to the work of Manjunath and Ma, and extend it for parameters of the full affine group. Next, we generalize the work of Lee and calculate the frame bounds for Gabor filters for affine group parameters. Then, we present several filter banks along with the tightness of the frames they form, as well as reconstruction results of images. We provide a detailed analysis of the calculation of the frame bounds in Appendix C.

2. Tessellation of the spatial-frequency space with Gabor functions

Recall that a Gabor function centered at the 2D frequency coordinates (ω_x, ω_y) has the general form of:

$$\psi(x, y) = g(x', y') \exp\big(2\pi i (\omega_x x + \omega_y y)\big), \tag{2}$$

where

$$(x', y')^t = R(x, y)^t = \big(x \cos(\phi) + y \sin(\phi), -x \sin(\phi) + y \cos(\phi)\big)^t, \tag{3}$$

where R is a 2×2 rotation matrix, and $g(x, y)$ is a non-symmetric two-dimensional Gaussian

$$g(x, y) = \frac{1}{2\pi \sigma_x \sigma_y} \exp\left(-\frac{x^2}{2\sigma_x^2} - \frac{y^2}{2\sigma_y^2}\right). \tag{4}$$

The parameters σ_x, σ_y denote the effective widths in the x and y directions respectively, whereas the major axis of the Gaussian is oriented at angle ϕ relative to the x-axis. Accordingly, the Fourier transform of the Gabor function is

$$\hat{\psi}(\xi, \nu) = \exp\big(-2\pi^2 \big(\sigma_x^2 (\xi' - \omega_x')^2 + \sigma_y^2 (\nu' - \omega_y')^2\big)\big), \tag{5}$$

where (ξ', ν') and (ω_x', ω_y') are rotated frequency coordinates. Thus, $\hat{\psi}(\xi', \nu')$ is a bandpass Gaussian with its minor axis oriented at angle ϕ from the ξ-axis, and the radial center frequency ω is defined by $\omega(x, y) = (\omega_x^2 + \omega_y^2)^{1/2}$, with orientation $\theta(x, y) = \arctan(\omega_y/\omega_x)$.

Studies of the Human Visual System (HVS) pointed to the fact that simple cells in the HVS can be modeled by Gabor functions [5,16]. Therefore, researchers [4,13,17] have suggested the usage of physiological constraints to reduce the number of degrees of freedom for the Gabor function:

- Constraint 1: The aspect ratio of the elliptical Gaussian envelope is 2:1 [4].
- Constraint 2: The plane wave with frequency (ω_x, ω_y) tends to have its "propagating direction" along the short axis of the elliptical Gaussian, i.e., $\omega_x = \omega \cos(\phi)$, $\omega_y = \omega \sin(\phi)$, where ϕ has already been defined as the orientation of the filter. Thus the orientation of the sine/cosine is aligned with the major axis of the elliptical Gaussian.

Applying these constraints, the following mother wavelet is obtained:

$$\psi(x, y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{8\sigma^2}(4x^2 + y^2)} e^{i\omega x}. \tag{6}$$

In our study we use this mother wavelet [13], and employ it in the context of the tessellation of the frequency plane. The frequency space is tessellated according to translations of the frequency parameters in the x and y directions. However, it is possible to generate Gabor wavelets from a single mother-Gabor-wavelet by the action of a group, e.g., the SIM(2) group, upon it, e.g., translations, rotations and dilations.

One question in this context concerns the choice of the group that acts on the mother wavelet. Other questions arise after we choose a group and want to apply a numerical algorithm. In this case one should discretize *the group's parameters*. The choice of the discretization affects the tessellation of the frequency domain and the ability to sinthesize the original function from its projections on the filters.

Choosing the SIM(2) group, the following discretization may be used: We use a set of filters for a finite number of scales, S, and orientations K. The filters that are generated from these choices are

$$\psi_{mn}(x, y) = a^{-m} \psi\left(\frac{x'}{a^m}, \frac{y'}{a^m}\right), \tag{7}$$

where (x', y') are the spatial coordinates rotated by $\frac{\pi n}{K}$ and scaled by powers $m = 0, \ldots, S - 1$.

Manjunath and Ma [15] have considered the generation of a Gabor filter bank for applications such as browsing and retrieval of images in large databases. Their design strategy was to ensure that the curves of half-peak magnitude of the filters in the frequency domain are kissing, i.e., they share a tangent line in one point of the curve. This design strategy leads to a set of constraints on the filter bank parameters: the scaling parameter a, the variances in the x and y directions, the lowest and highest frequencies allowed for the filter bank, and the number of scales. Out of theses seven parameters four can be defined, and the other three are determined by the constraints of the design strategy.

We extend the design methods of Manjunath and Ma in two directions: the first is to use different scaling parameters for the x and y directions. The second direction is to investigate the frame properties of this design methodology. The basic assumption of Manjunath and Ma is that the half-peaks of the filters in the frequency domain should touch each other. However, does this constraint necessarily lead to a design of a tight frame? Or should the filters touch each other not at half-peak, but at 0.7-peak or perhaps 0.3-peak. In order to answer these questions, we refer to the work of Lee [13]. Lee has extended the frame criterion developed by Daubechies for one-dimensional wavelets to two dimensions. He has also computed the frame bounds for the case of Gabor wavelets based on the SIM(2) group. We extend the two-dimensional frame criterion to account for different scalings of the x and y dimensions. We elaborate on this issue in the next section.

Next, we calculate the filters' parameters, under the constraint that their λ-peak curves "kiss" each other. In the analysis of Manjunath and Ma, λ is fixed to $\frac{1}{2}$.

Let $\psi(x, y)$ be the (non-admissible) mother Gabor wavelet:

$$\psi(x, y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{8}(4x^2 + y^2)} e^{ikx}. \tag{8}$$

The Fourier transform of this function is equal to:

$$\hat{\psi}(\xi, v) = \sqrt{8\pi} e^{-\frac{1}{2}((\xi - k)^2 + 4v^2)}. \tag{9}$$

The self-similar filters are generated using the law:

$$\psi_{mn}(x, y) = (a_x a_y)^{-\frac{m}{2}} \psi(x'_{mn}, y'_{mn}), \tag{10}$$

where $x'_{mn} = a_x^{-m}(x \cos\theta_n + y \sin\theta_n)$, $y'_{mn} = a_y^{-m}(-x \sin\theta_n + y \cos\theta_n)$.

In the frequency domain we obtain:

$$\hat{\psi}_{mn}(\xi, v) = (a_x a_y)^{\frac{m}{2}} \hat{\psi}(a_x^m \xi'_{mn}, a_y^m v'_{mn}), \tag{11}$$

where $\xi'_{mn} = a_x^{-m}(\xi \cos\theta_n + v \sin\theta_n)$, $v'_{mn} = a_y^{-m}(-\xi \sin\theta_n + v \cos\theta_n)$, and $\theta_n = n\pi/r$.

Next, we apply the constraint that the λ-peak curves kiss. We follow the guidelines of Manjunath and Ma [15]. It is important to notice the difference in this situation from the Manjunath and Ma analysis: here the number of orientations depends on the scale. We denote, therefore, by $r(m)$ the number of orientations in the mth scale.
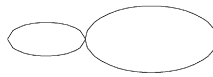
Fig. 1. The zero level scale filter, $\hat{\psi}_{00}$, and the first level scale filter, $\hat{\psi}_{10}$, both aligned to the same orientation, touch at λ-peak.
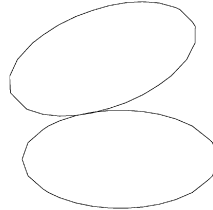

Fig. 2. The filter in the zero level orientation $\hat{\psi}_{00}$ is tangent to the first level orientation filter $\hat{\psi}_{01}$ at the point: $\nu = \tan\frac{\pi}{2r}$.

We would like the zero level scale filter, $\hat{\psi}_{00}$, and the first level scale filter, $\hat{\psi}_{10}$, both aligned to the same orientation, to touch at λ-peak as illustrated in Fig. 1. This constraint leads to the relation between the scaling factor a_x, the factor λ and the wave number k (see proof in Appendix A):

$$k = \frac{a_x + 1}{a_x - 1}\sqrt{-2\log\lambda}. \tag{12}$$

The second constraint is that the filter in the zero level orientation $\hat{\psi}_{00}$ touches the first level orientation filter $\hat{\psi}_{01}$ at λ-peak as can be seen in Fig. 2. This is the same as requiring that the line $\nu(\xi) = \tan(\frac{\pi}{2r})\xi$ is tangent to the λ-peak level set of both filters in the frequency domain. This means that we have to find ξ and ν that obey the tangent line equation, and that are located at the λ-peak of $\hat{\psi}_{00}(\xi, \nu)$. This means that there should be a single solution to the equation:

$$\hat{\psi}_{00}(\xi, \nu(\xi)) = \lambda \max|\hat{\psi}_{00}|.$$

Substituting $\hat{\psi}_{00}$ and the relationship $\nu = \tan\frac{\pi}{2r(0)}\xi$, we obtain (see proof in Appendix A):

$$\frac{\pi}{2r(0)} = \arctan\sqrt{\frac{-\log\lambda}{2k^2 + 4\log\lambda}}. \tag{13}$$

This is a constraint that relates the filter's basic parameter k to the number of orientations needed in order to obtain the λ-peak tangency relation. This derivation holds only for the zero level scale filters. Going up to the first level scale filters, and applying the same procedure, we obtain the following result:

$$\frac{\pi}{2r(1)} = \arctan\left(\frac{a_x}{a_y}\sqrt{\frac{-\log\lambda}{2k^2 + 4\log\lambda}}\right). \tag{14}$$

This is easily extrapolated to the m-level scale filters as follows:

$$\frac{\pi}{2r(m)} = \arctan\left(\left(\frac{a_x}{a_y}\right)^m\sqrt{\frac{-\log\lambda}{2k^2 + 4\log\lambda}}\right). \tag{15}$$

Thus, we have a mechanism to determine the number of orientations per scale so that there is a certain amount of overlap between the Gabor functions:

$$r(m) = \frac{\pi}{2\arctan\left(\left(\frac{a_x}{a_y}\right)^m\sqrt{\frac{-\log\lambda}{2k^2 + 4\log\lambda}}\right)}. \tag{16}$$

This can be related to the work of Daugman [4] which deals with the orientation selectivity of simple cells in the visual cortex and its relation to spatial frequency and spatial resolution. His analysis is within the framework of the uncertainty principle. In his work, Daugman has presented the relations between the width/length aspect ratios of the filter in the spatial domain and the resolution of the spatial frequency and orientation in the frequency domain. He has
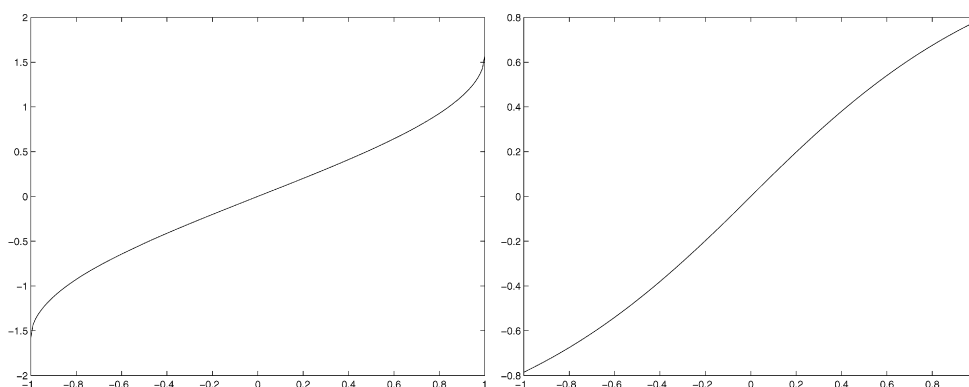
Fig. 3. Both the asin (left) and atan (right) functions have a monotonically increasing behavior in the range $[-1, 1]$. Thus, the dependence on the ratio $\frac{a_x}{a_y}$ is the same for the orientation bandwidth and the basic angle in our sampling scheme.

shown that elongating the filter in the direction parallel to the modulating wave sharpens the orientation bandwidth but has no effect on the spatial-frequency bandwidth. If the filter is elongated in the direction perpendicular to the modulating wave, then the spatial-frequency bandwidth sharpens, but there is no effect on the orientation bandwidth. Thus, sharp spatial resolution in the y direction can be obtained at the expense of the orientation selectivity, or sharp spatial resolution in the x direction can be obtained at the expense of the spatial-frequency selectivity. Daugman has found that the following relation exists between the orientation and the spatial-frequency bandwidths:

$$\Delta\theta_{\frac{1}{2}} = \arcsin\left(\beta \frac{2^{\Delta\omega} - 1}{2^{\Delta\omega} + 1}\right), \tag{17}$$

where $\Delta\theta_{\frac{1}{2}}$ is the orientation half-bandwidth, $\Delta\omega$ is the frequency bandwidth in octaves, and β is the aspect ratio between the x and y axes of the Gaussian. This relation was proved by Daugman in [4].

We extend the derivation of Daugman from the mother wavelet to the whole filter bank. Assuming that an affine transformation is applied to the mother wavelet such that it is scaled by some global scale m, while taking into account our freedom to have different scalings in the x and y directions. Consider a Gaussian modulated sine-wave function. Let $\Delta\theta_{\frac{1}{2}}$ be its orientation half-bandwidth, $\Delta\omega$ its frequency bandwidth in octaves, and β the aspect ratio between the x and y axes of the Gaussian. Then we obtain

Lemma 2.1. *The following relationship holds*:

$$\Delta\theta_{\frac{1}{2}} = \arcsin\left(\beta \left(\frac{a_x}{a_y}\right)^m \frac{2^{\Delta\omega} - 1}{2^{\Delta\omega} + 1}\right). \tag{18}$$

Proof. The proof is a straightforward extension of the proof in [4] to some scale m. We provide the detailed proof in Appendix B. $\square$

The arcsin function is monotonically increasing in the range $(-\pi, \pi)$. Therefore, the larger a_x is, for a specific scale, the larger is $\Delta\theta_{\frac{1}{2}}$. On the other hand, as a_y is larger, $\Delta\theta_{\frac{1}{2}}$ becomes smaller.

It is interesting to see that the two derivations, one coming from some reasonable pavement of the frequency plane and the other one coming from uncertainty considerations lead to the same dependence on the ratio $\frac{a_x}{a_y}$. This is because both the arctan and the arcsin functions have a monotonically increasing behavior in the range $(-\pi, \pi)$ (Fig. 3).

To conclude, we can derive filter banks that pave the frequency plane in various ways: constant number of orientations per scale, increasing number of orientations per scale and decreasing number of orientations pet scale, depending on the ratio between a_x and a_y, as can be seen in Fig. 4.

3. Frame bounds for Gabor filters for affine group parameters

In 1996 Lee has extended the frame criterion developed by Daubechies for one-dimensional wavelets to the two-dimensional situation [13]. He also computed the frame bounds for Gabor wavelets, and specifically has found a
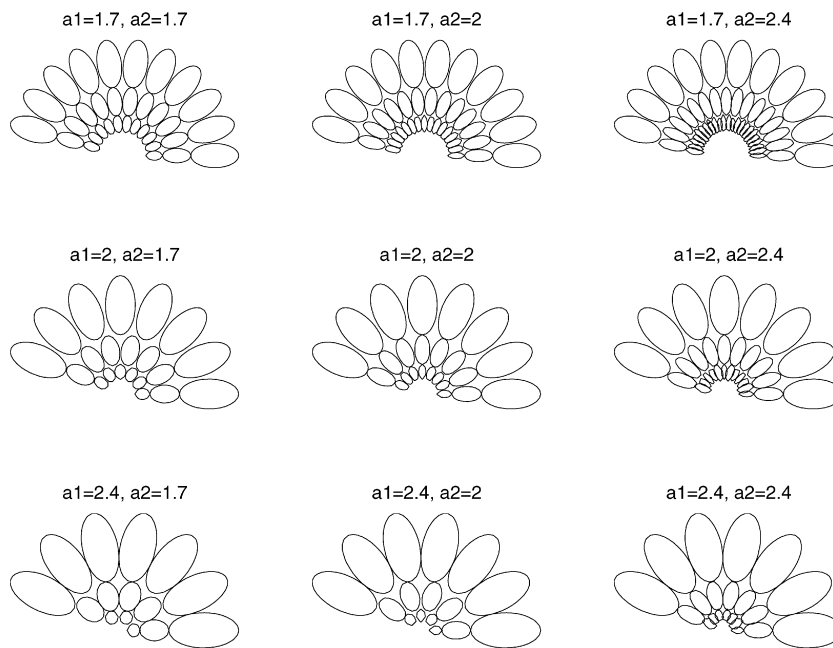
Fig. 4. In this figure the number of orientations depends on the scale m, and specifically on the ratio $(\frac{a_x}{a_y})^m$. As can be seen as this ratio is smaller, the number of orientations per scale becomes larger and vice versa.

parameterizations and a group sampling scheme which allowed stable reconstruction by summation, as if the Gabor wavelets would form an orthonormal basis. This happens when the parameterization and sampling selected lead to a setting of an almost tight frame, or a perceptually tight frame. By perceptually tight frame we mean a frame which leads to signal reconstruction that is *perceptually* indistinguishable from the original one. Thus, non-perceptible errors in the construction are allowed. This freedom enables to accept frames for which the frame bounds are not identical but such that the ration between the upper to the lower bound is sufficiently close to 1.

In his derivation, Lee has used the SIM(2) group operations to generate the family of self-similar Gabor wavelets. In this section we extend the derivation of the frame bounds to the affine group. This means that we may sample the x and y directions using different scaling parameters. We also look for parameterizations that enable a tight frame, so that the simple summation formula can be used for reconstruction.

Moreover, we would like to offer a link between the design of Gabor filters, as is manifested in the work of Manjunath and Ma, to the mathematical framework that Lee developed.

3.1. The design of Gabor wavelets according to the SIM(2) group

Once the Gabor mother wavelet is set, a whole family of Gabor wavelets is constructed by the group actions. Lee uses a constraint based on the HVS, which states that the half-peak of the filter, in the frequency domain, is between one to one and a half octaves along the optimal orientation. This constraint determines the attributes of the Gabor function, as it relates the sinewave frequency to the Gaussian's width, thus giving the Gabor functions a wavelet flavor. Moreover, it actually determines the degree of overlapping between the Gabor functions in the filter bank. Thus, the mother Gabor function is

$$\psi(x, y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{8}(4x^2 + y^2)} \left(e^{ikx} - e^{-\frac{k^2}{2}} \right), \tag{19}$$

where the term $e^{-\frac{k^2}{2}}$ is added to make the Gabor filter an admissible wavelet, i.e., a L_2 function with zero mean.

Next, Lee uses the SIM(2) group to generate the self-similar functions. He then samples the group parameters in various ways and investigates the frame bounds of the corresponding filter banks.

We generalize Lee's results to account for filter banks that are generated by more general transformations, e.g., affine group parameters.

3.2. The design of Gabor wavelets according to the affine group

In the SIM(2) group the scaling of the x and y directions is the same. If we allow different scaling in the x and y directions, we obtain, by sampling the group parameters, the following discrete representation of the affine Gabor wavelets in the spatial space:

$$\psi_{m,n,k,l}(x, y) = a_x^{-\frac{m}{2}} a_y^{-\frac{m}{2}} \psi_{\theta_l}\left(a_x^{-m}x - nb_x, a_y^{-m}x - kb_y\right), \tag{20}$$

where

$$\psi_{\theta_l}(x, y) = \psi\left(x\cos(l\theta_0) + y\sin(l\theta_0), -x\sin(l\theta_0) + y\cos(l\theta_0)\right), \tag{21}$$

and where θ_0 denotes the step size of angular rotation, and l indexes the number of rotation steps.

Thus, at the m level, the x direction is sampled by $a_x^m b_x$ steps, and the y direction is sampled by $a_y^m b_y$ steps.

In the frequency space the wavelets have the following form:

$$\hat{\psi}_{mnkl}(\xi, \nu) = a_x^{\frac{m}{2}} a_y^{\frac{m}{2}} \hat{\psi}_{\theta_l}\left(a_x^m \xi, a_y^m \nu\right) e^{-ia_x^m b_x n\xi} e^{-ia_y^m b_y k\nu}, \tag{22}$$

and the affine wavelets transform is given by

$$\left(T^{wav}f\right)(a_x, a_y, \theta, x_0, y_0) = \|a_x\|^{-\frac{1}{2}} \|a_y\|^{-\frac{1}{2}} \iint dx\, dy\, f(x, y)\psi_\theta\left(\frac{x - x_0}{a_x}, \frac{y - y_0}{a_y}\right). \tag{23}$$

The design of the filter bank amounts at this stage to a proper choice of the discretization of the group action. The group action is now parameterized by a_x, a_y, b_x, b_y and θ_0. The design criterion is the quality of the frame generated by the set of functions $\{\psi_{m,n,k,l}\}$. Before we state our result let us review briefly the notion of a frame.

Let I denote a set. A set of functions $\{\psi_i\}_{i\in I}$, in a vector space with a scalar product, is a frame if for any f in this vector space

$$0 < A\|f\|^2 \leqslant \sum_{i\in I}\left|\langle f, \psi_i\rangle\right|^2 \leqslant B\|f\|^2 < \infty,$$

where the scalars A and B are called the frame bounds. If A equals B the frame is called "tight." Otherwise the tightness of the frame is determined by the ration B/A. The closer this ratio is to 1 the tighter is the frame and a better reconstruction of the signal f from its projections $\langle f, \psi_i\rangle$ is obtained.

The frame bounds for the set $\{\psi_{m,n,k,l}\}$ depend on the choice of discretization. Namely, the frame bounds A and B depend on the parameters a_x, a_y, b_x, b_y and θ_0. This dependence is given in

Theorem 3.1. *The frame bounds for the set $\{\psi_{m,n,k,l}\}_{m,n,k\in\mathbb{Z}, l\in\mathbb{Z}_r}$ are given by*

$$A = \frac{1}{b_x b_y}\left\{\inf_{\xi,\nu}\sum_{m\in Z, l\in\mathbb{Z}/r\mathbb{Z}} F(\xi, \nu) - \sum_{(p,q)\in\mathbb{Z}^2\backslash(0,0)}\sqrt{\beta\left(\frac{2\pi p}{b_x}, \frac{2\pi q}{b_y}\right)\beta\left(-\frac{2\pi p}{b_x}, -\frac{2\pi q}{b_y}\right)}\right\},$$

$$B = \frac{1}{b_x b_y}\left\{\sup_{\xi,\nu}\sum_{m\in Z, l\in\mathbb{Z}/r\mathbb{Z}} F(\xi, \nu) + \sum_{(p,q)\in\mathbb{Z}^2\backslash(0,0)}\sqrt{\beta\left(\frac{2\pi p}{b_x}, \frac{2\pi q}{b_y}\right)\beta\left(-\frac{2\pi p}{b_x}, -\frac{2\pi q}{b_y}\right)}\right\},$$

where

$$F(\xi, \nu) = \left|\hat{\psi}_{\theta_l}\left(a_x^m \xi, a_y^m \nu\right)\right|^2$$

and

$$\beta(s, t) = \sup_{\xi,\nu}\sum_{m\in\mathbb{Z}}\left|\hat{\psi}_{\theta_l}\left(a_x^m \xi, a_y^m \nu\right)\right|\left|\hat{\psi}_{\theta_{l(m)}}\left(a_x^m \xi + s, a_y^m \nu + t\right)\right|.$$

Note that Eq. (16) gives a relation between m and r, namely a relation between the scale and the number of rotations.

Proof. See Appendix C. $\square$

4. Results

We present in this section several filter banks that are generated with various scalings and orientations. We compare the design criteria of the filter bank: the tessellation of the frequency domain and the degree of tightness of the generated frame. The filters' appearance in the frequency domain is shown for each selection of filter bank parameters and the ratio $\frac{B}{A}$ between the frame bounds obtained using these parameters is calculated. The reconstruction for each sampling scheme is also presented. The reconstruction is based on the linear summation formula:

$$\tilde{f} = \frac{2}{A+B} \sum_{m,n,k\in\mathbb{Z},\, l\in\mathbb{Z}/r\mathbb{Z}} \langle f, \psi_{mnkl}\rangle \psi_{mnkl}, \tag{24}$$

where $\tilde{f}$ is the approximation for the reconstructed image f. We use the linear summation formula rather than using the dual frame $\tilde{\psi}_{mnkl}$ for obtaining f:

$$f = \sum_{m,n,k\in\mathbb{Z},\, l\in\mathbb{Z}/r\mathbb{Z}} \langle f, \psi_{mnkl}\rangle \tilde{\psi}_{mnkl}. \tag{25}$$

The legitimacy of this procedure depends on the degree of the tightness of the frame. In principle, the tighter the frame, the better the result we obtain. Nevertheless, in order to use non-dyadic translations of the filters, we are forced to use some kind of interpolation to obtain sub pixel values. Thus, the interpolation itself may introduce numerical errors. To simplify the summation over points which do not lie on the image's grid, we make use of the fact that translations in the spatial space are equivalent to phase shifts in the frequency space:

$$\sum_{m,n,k\in\mathbb{Z},\, l\in\mathbb{Z}/r\mathbb{Z}} \langle f, \psi_{mnkl}\rangle \psi_{mnkl} = \sum_{m,n,k\in\mathbb{Z},\, l\in\mathbb{Z}/r\mathbb{Z}} \langle f, \psi_{mnkl}\rangle F^{-1}\big(F(\psi_{mnkl})\big)$$

$$= \sum_{m,n,k\in\mathbb{Z},\, l\in\mathbb{Z}/r\mathbb{Z}} \langle f, \psi_{mnkl}\rangle F^{-1}\big(F\big(\psi_{mn}\big(x - nb_x a_x^m, y - kb_y a_y^m\big)\big)\big)$$

$$= \sum_{m,n,k\in\mathbb{Z},\, l\in\mathbb{Z}/r\mathbb{Z}} \langle f, \psi_{mnkl}\rangle F^{-1}\big(F\big(\psi_{mn}(x, y)\big)e^{-i\xi nb_x a_x^m} e^{-i\nu kb_y a_y^m}\big),$$

where $F(\cdot)$ denotes the Fourier transform.

4.1. Number of orientations is constant through different scales

First, we address different scalings in the x and y directions, where the number of orientations is kept fixed, and the value of the wave number k is also kept constant and is equal to π. This means, according to Eq. (16), that the degree of overlap between the filters depends on the values of the scaling factors in the x and y directions.

4.1.1. The case of 8 orientations

In Fig. 5, the half-peak curves of the filters are presented for eight orientations, with different scaling parameters. In this case, the value of a_x is constant and is equal to 1.4 while a_y takes the values: 1.4, 1.8, 2, 2.2. We may observe that the larger a_y is, the more overlap the filters exhibit in the angular direction. The angular resolution is therefore lower as a_y increases.

The frame bounds A and B are calculated for the eight orientations, and the tightness of the frame is measured by the ratio $\frac{B}{A}$. These ratio values are shown in Table 1. The smaller the value of a_y, the tighter is the frame.

Next, in Fig. 6, we may see the appearance of the filter bank for eight orientations, where now the value of a_y is constant and is equal to 1.4 and a_x takes the values: 1.4, 1.8, 2, 2.2. We observe that the smaller a_x is, the denser is the sampling in the radial direction.

The tightness of the frame, measured by the ratio $\frac{B}{A}$, is shown in Table 2. As can be expected, the smaller the value of a_x, the tighter is the frame.

4.1.2. The case of 16 orientations

In Fig. 7, we may see the appearance of the filter bank for sixteen orientations, with different scaling parameters. The half-peak values of the filters are presented. Here, the value of a_x is constant and is equal to 1.4. a_y takes the
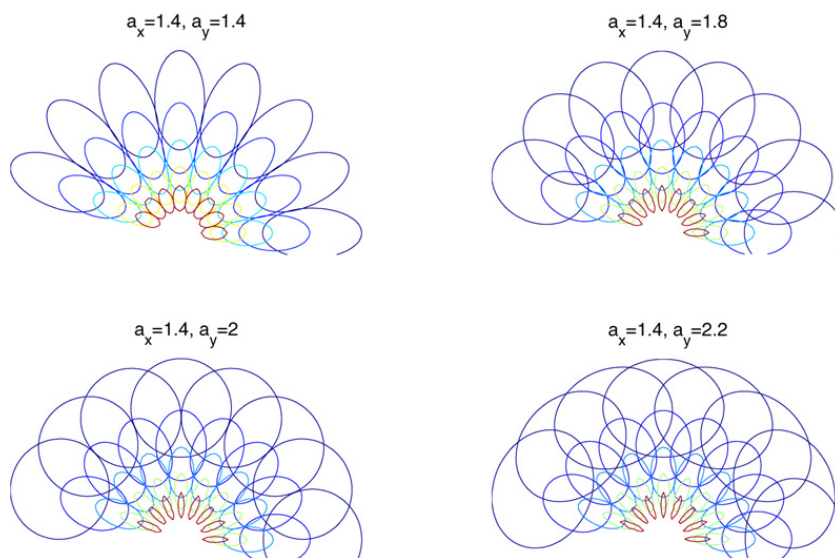
Fig. 5. The filter look for $N_\theta = 8$ for $a_x = 1.4$ and $a_y = 1.4, 1.8, 2, 2.2$. The larger a_y is, the more overlap the filters have in the orientation direction.
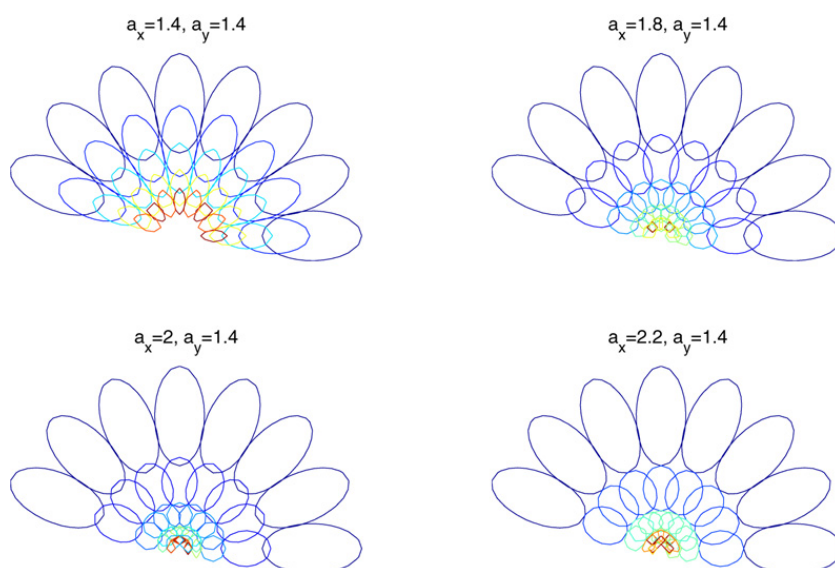

Fig. 6. The filter look for $N_\theta = 8$ for $a_y = 1.4$ and $a_x = 1.4, 1.8, 2, 2.2$.

Table 1
The ratio $\frac{B}{A}$ for eight orientations, where $a_x = 1.4$ and $a_y = 1.4, 1.8, 2, 2.2$

a_y	$\frac{B}{A}$
1.4	2.61
1.8	3.52
2.0	5.03
2.2	6.40

Table 2
The ratio $\frac{B}{A}$ for eight orientations, where $a_y = 1.4$ and $a_x = 1.4, 1.8, 2, 2.2$

a_x	$\frac{B}{A}$
1.4	2.61
1.8	3.12
2.0	3.37
2.2	3.77

values: 1.4, 1.8, 2, 2.2. The frame bounds A and B are calculated for the sixteen orientations, and the tightness of the frame is shown in Table 3.

The reconstruction results for this case are shown in Fig. 8.

In Fig. 9, we may see the appearance of the filter bank for sixteen orientations, where now the value of a_y is constant and is equal to 1.4 and a_x takes the values: 1.4, 1.8, 2.
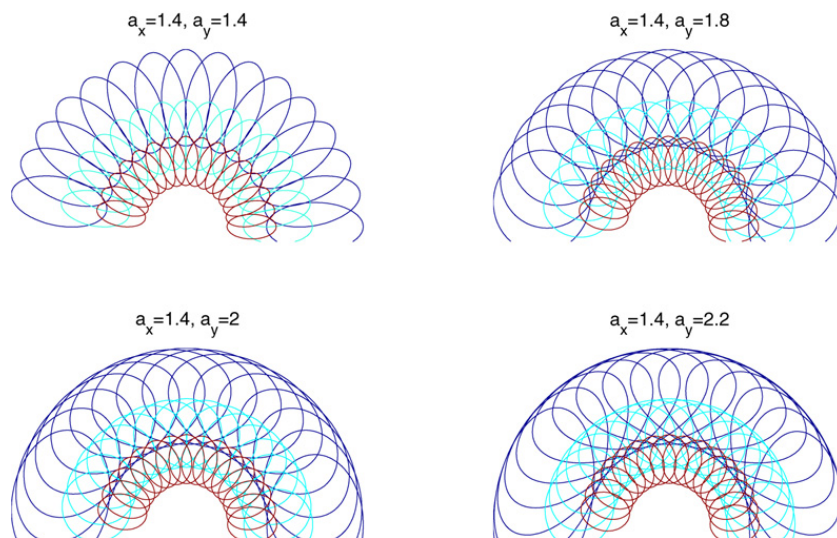
Fig. 7. The filter look for $N_\theta = 16$ for $a_x = 1.4$ and $a_y = 1.4, 1.8, 2, 2.2$.
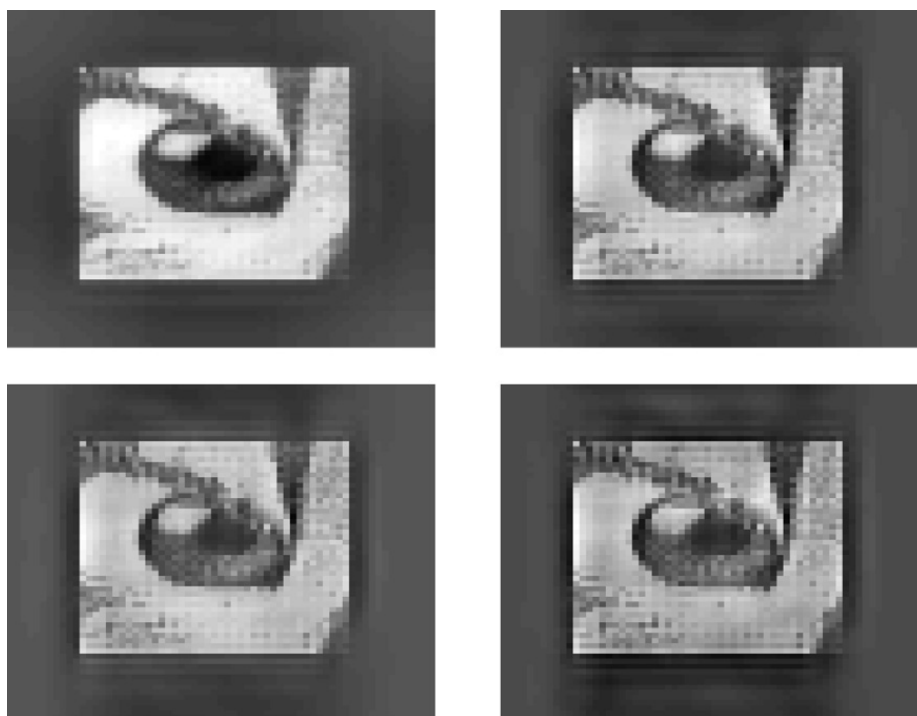

Fig. 8. The reconstruction results for $N_\theta = 16$ for $a_x = 1.4$ and $a_y = 1.4$ (top left), 1.8 (top right), 2 (bottom left), 2.2 (bottom right).

The tightness of the frame is measured by the ratios $\frac{B}{A}$, which are shown in Table 4.
The reconstruction results are given in Fig. 10.

4.2. Number of orientations depends on scale

Next, we do not pre-determine the number of orientations. We fix the overlap of the filters by selecting $\lambda = 0.5$. We explore the effect of changing the values of a_x and a_y where the number of orientations per scale is calculated using:

$$r(m) = \frac{\pi}{2 \arctan\left(\left(\frac{a_x}{a_y}\right)^m \sqrt{\frac{-\log \lambda}{2k^2 + 4 \log \lambda}}\right)}. \tag{26}$$
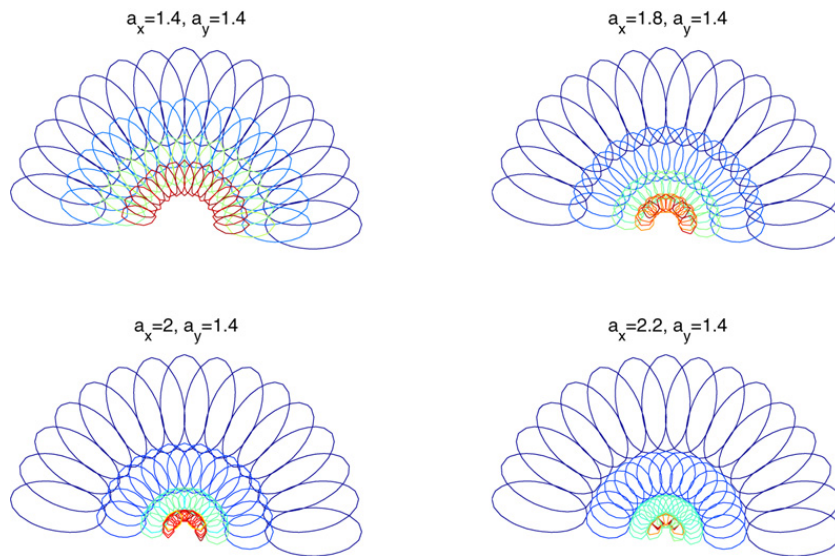
The filters are drawn using the half-peak values.

Fig. 9. The filter look for $N_\theta = 16$ for $a_y = 1.4$ and $a_x = 1.4, 1.8, 2$.

Table 3
The ratio $\frac{B}{A}$ for sixteen orientations, where $a_x = 1.4$ and $a_y = 1.4, 1.8, 2, 2.2$

a_y	$\frac{B}{A}$
1.4	1.04
1.8	1.78
2.0	2.27
2.2	4.00

Table 4
The ratio $\frac{B}{A}$ for sixteen orientations, where $a_y = 1.4$ and $a_x = 1.4, 1.8, 2$

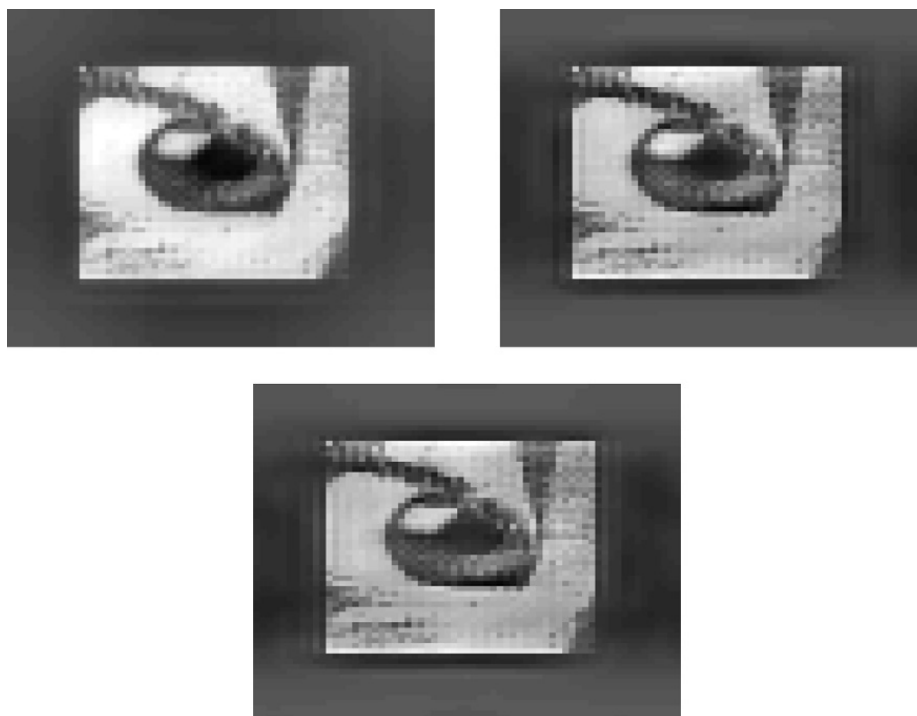| a_x	$\frac{B}{A}$
1.4	1.04
1.8	1.80
2.0	2.56

Fig. 10. The reconstruction results for $N_\theta = 16$ for $a_y = 1.4$ and $a_x = 1.4$ (top left), 1.8 (top right), 2 (bottom).
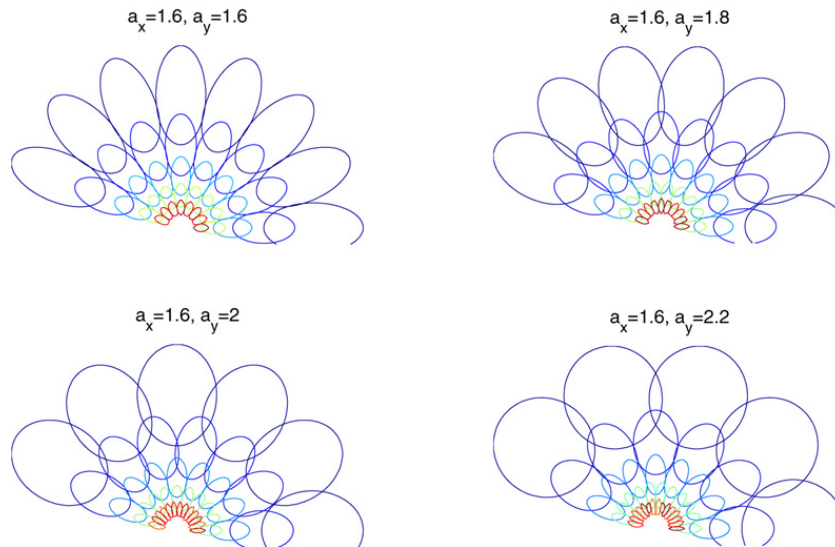
Fig. 11. The filter look for $a_x = 1.6$ and for $a_y = 1.6, 1.8, 2, 2.2$. The number of orientations increases as scale increases (meaning that frequency decreases, towards the origin).
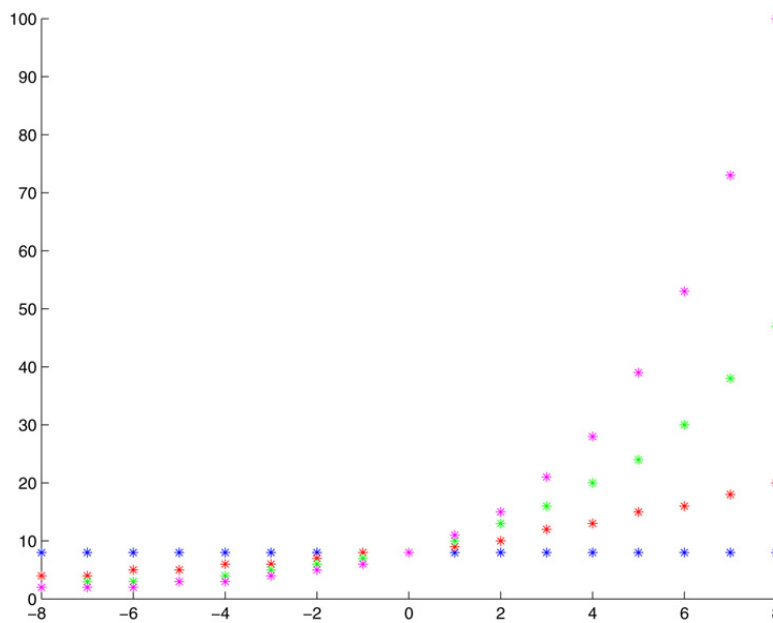

Fig. 12. The number of orientations as a function of scale, according to Eq. (16). In this case $a_x \leqslant a_y$, and the number of angles increases with scale for $m > 0$ and decreases with scale for $m < 0$.

In Fig. 11 we observe the appearance of the filter bank for $a_x = 1.6$ and $a_y = 1.6, 1.8, 2, 2.2$. Since for this case $\frac{a_x}{a_y} < 1$, we may see that the number of orientations is higher as the scale is larger (thus for lower frequencies). This means that for lower frequencies there are more orientations than for higher frequencies. As can be seen in Fig. 12, the number of angles increases with scale, when m is positive. When m is negative the number of angles decreases with scale. The frame bounds calculated for this case are shown in Table 5. The reconstruction results for $a_x = 1.6$ and $a_y = 1.6, 1.8, 2, 2.2$ are shown in Fig. 13.

In Fig. 14 we observe the appearance of the filter bank for $a_y = 1.6$ and $a_x = 1.6, 1.8, 2, 2.2$. Looking again at the dependence of the number of orientations on the scale:

$$r(m) = \frac{\pi}{2 \arctan\left(\left(\frac{a_x}{a_y}\right)^m \sqrt{\frac{-\log(\lambda_y)}{2*(k^2 + 2*\log(\lambda_y))}}\right)},$$
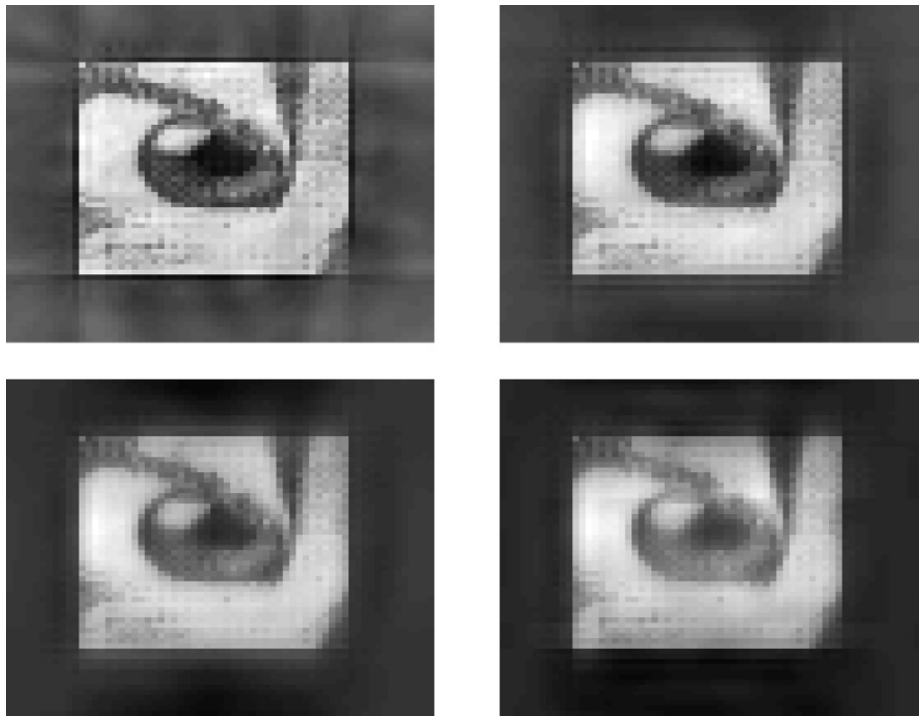
Fig. 13. The reconstruction results for $a_x = 1.6$ and $a_y = 1.6$ (top left), 1.8 (top right), 2 (bottom left), 2.2 (bottom right).
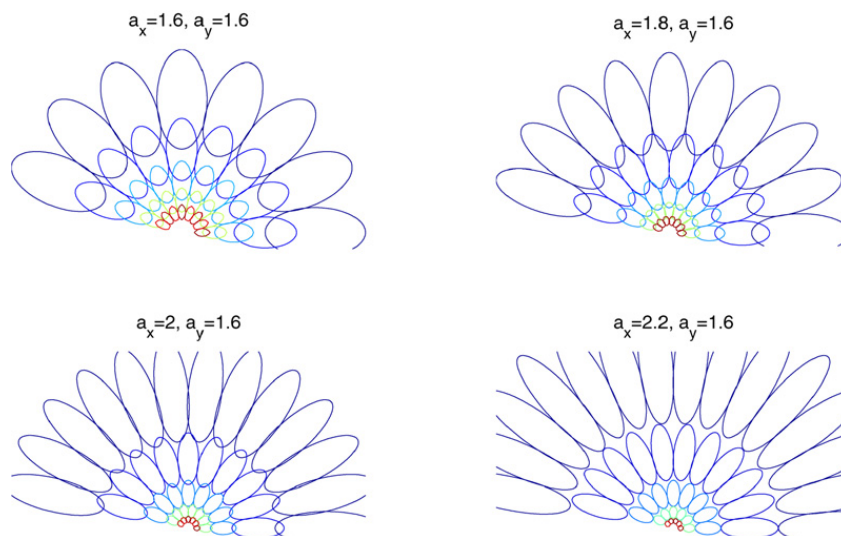

Fig. 14. The filter look for $a_y = 1.6$ and for $a_x = 1.6$, 1.8, 2, 2.2. The number of orientations decreases as scale increases (meaning that the frequency decreases, towards the origin).

we observe that for this case ($\frac{a_x}{a_y} > 1$) the number of orientations is higher as the scale is smaller (thus for higher frequencies). This means that for lower frequencies we obtain less orientations than for higher frequencies. As can be seen in Fig. 15, the number of angles deceases with scale, as the ratio $\frac{a_x}{a_y}$ is equal or larger than 1.

The frame bounds calculated for this case are shown in Table 6.

The reconstruction results for $a_y = 1.6$ and $a_x = 1.6$, 1.8, 2, 2.2 are shown in Fig. 16.

5. Discussion and conclusions

In this study we considered the issue of Gabor filter bank design where the dilations in the x and y directions are not necessarily the same. When the goal is to provide an adequate representation, it seems that the filter bank should
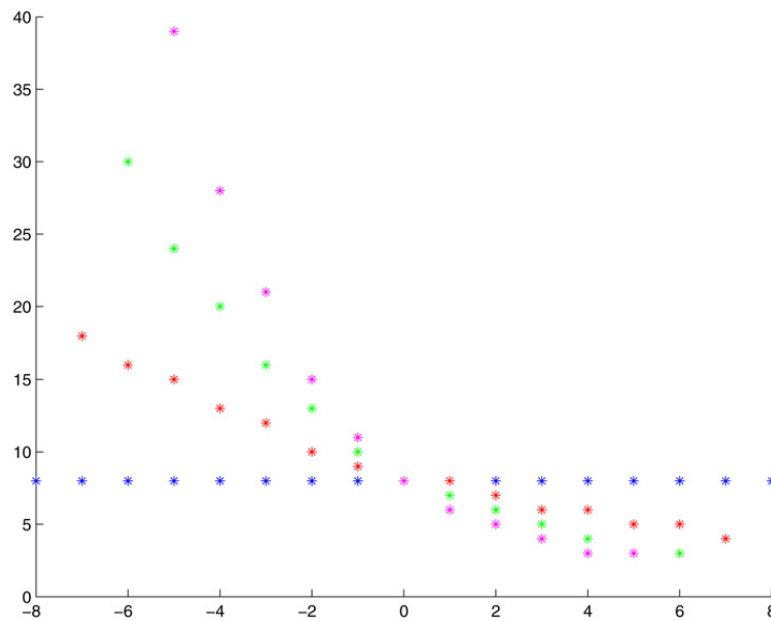
Fig. 15. The number of orientations as a function of scale according to Eq. (16). In this case $a_x \leqslant a_y$, and the number of angles decreases with scale.

Table 5
The ratio $\frac{B}{A}$ for the case $\lambda = 0.5$ and $a_x = 1.6$

a_y	$\frac{B}{A}$
1.6	1.092
1.8	2.11
2.0	3.72
2.2	4.99

Table 6
The ratio $\frac{B}{A}$ for the case $\lambda_x = \lambda_y = 0.5$ and $a_y = 1.6$

a_x	$\frac{B}{A}$
1.6	1.092
1.8	1.64
2.0	2.37
2.2	4.31

constitute a frame. We have extended the derivation of Lee for calculating the frame bounds to the case of different scalings in x and y, and have provided a tool to assess the filter design methodology of Manjunath and Ma.

When the number of orientations is kept constant, an increase in the number of orientations improves the frame bounds. When the number of orientations was set to sixteen, decreasing the values of either a_x or a_y results in a tighter frame. When the scaling in the y direction becomes finer, the overlapping between the filters is smaller, and a tighter frame, along with better reconstruction results are obtained. When the scaling in the x direction is smaller, the tessellation resolution in the radial direction increases, and the frame bounds, as well as the reconstruction results are better.

We have established the dependence of the number of orientations on the scale, in relation with the dilation factors a_x and a_y, and provided several different schemes for paving the phase space. We have also noted that the dependence of the basic orientation in the pavement scheme (i.e., $\frac{\pi}{2r(m)}$) on the ratio $\frac{a_x}{a_y}$ resembles the dependence of the orientation bandwidth on the same ratio in the context of uncertainty [4].

When the value of a_x is equal to a_y, the number of orientations is constant through scales. When we increase the value of a_y, and $\frac{a_x}{a_y} < 1$, the number of orientations decreases with increasing scale. When we increase the value of a_x, and $\frac{a_x}{a_y} > 1$, the number of orientations increases with increasing scale. However, we obtain a coarser sampling of the radial direction in the frequency space. Therefore, an increase in the values of either a_x or a_y results in a less tight frame, and poorer reconstruction results.

To conclude, we have established two results concerning a design of affine Gabor-based filter banks. The first result generalizes the analysis of Daugman and establishes a relationship between the spatial and angular resolutions. The second generalizes the spatial-frequency analysis of Manjunath and Ma, and the frame bound analysis of Lee and Daubechies. A new expression for the frame bound in non-equal scaling in the x and y directions is presented.
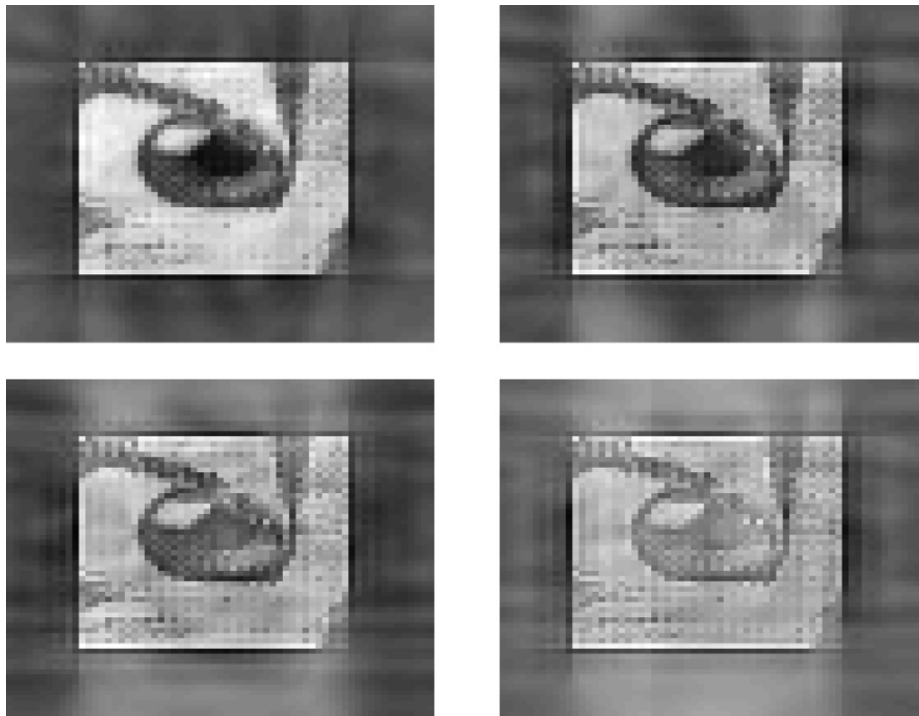
Fig. 16. The reconstruction results for $a_y = 1.6$ and $a_x = 1.6$ (top left), 1.8 (top right), 2 (bottom left), 2.2 (bottom right).

A frame tightness criterion for evaluating the design methodology of Gabor wavelets is then suggested and tested experimentally for several choices of group discretizations and therefore for several different filter banks. Few filter banks form an almost tight frame and the image reconstruction that is based on these frames is perceptually good.

Appendix A. Calculating the relations among the Gabor wavelets parameters

In this appendix we explicitly calculate the relationships among the parameters of the Gabor wavelets. This analysis is based on the constraint that the functions touch each other at λ-peak in the frequency space. We denote by K_m the number of orientations in the mth scale.

We prove in this appendix two relations. The first, relates the scaling factor a_x, the factor λ and the wave number k:

$$k = \frac{a_x + 1}{a_x - 1}\sqrt{-2\log\lambda}. \tag{A.1}$$

The second presents the dependency of the number of orientations on the scale:

$$\frac{\pi}{2r(0)} = \arctan\sqrt{\frac{-\log\lambda}{2k^2 + 4\log\lambda}}. \tag{A.2}$$

First, we would like the zero-level scale filter, $\hat{\psi}_{00}$, and the first-level scale filter, $\hat{\psi}_{10}$, both aligned to the same orientation, to touch at λ-peak. Let us choose, for simplicity and without loss of generality the x-axis, where $\nu = 0$. Given the explicit form of the two functions:

$$\hat{\psi}_{00}(\xi, \nu) = \sqrt{8\pi} \exp\left(-\frac{(\xi - k)^2 + 4\nu^2}{2}\right),$$

$$\hat{\psi}_{10}(\xi, \nu) = \sqrt{8\pi} \exp\left(-\frac{(a_x\xi - k)^2 + 4(a_y\nu)^2}{2}\right), \tag{A.3}$$

we substitute $\nu = 0$ to obtain

$$\hat{\psi}_{00}(\xi, 0) = \sqrt{8\pi} \exp\left(-\frac{(\xi - k)^2}{2}\right), \qquad \hat{\psi}_{10}(\xi, 0) = \sqrt{8\pi} \exp\left(-\frac{(a_x\xi - k)^2}{2}\right). \tag{A.4}$$

Next, the maximal value of $\hat{\psi}_{00}$ is given for $\xi = k$ and $\nu = 0$, thus $\max|\hat{\psi}_{00}| = \sqrt{8\pi}$. We simply search for the value of ξ that will give us $\frac{\sqrt{8pi}}{\lambda}$. First, we look at the λ-peak of $\hat{\psi}_{00}$:

$$\lambda = \exp\left(-\frac{(\xi - k)^2}{2}\right). \tag{A.5}$$

Solving this equation, we obtain

$$\xi = k \pm \sqrt{-2\log(\lambda)}, \tag{A.6}$$

where the relevant solution is only

$$\xi = k - \sqrt{-2\log(\lambda)}. \tag{A.7}$$

Next, we do the same calculation for $\hat{\psi}_{10}$ to obtain

$$\xi = \frac{k}{a_x} + \frac{\sqrt{-2\log(\lambda)}}{a_x}. \tag{A.8}$$

As we require that $\hat{\psi}_{00}$ and $\hat{\psi}_{10}$ kiss each other at this λ-peak, we obtain

$$k = \frac{a_x + 1}{a_x - 1}\sqrt{-2\log(\lambda)}, \tag{A.9}$$

as depicted in Eq. (A.1).

The second constraint is that the filter in the zero-level orientation $\hat{\psi}_{00}$ is tangent to the first-level orientation filter $\hat{\psi}_{01}$ at point $\nu(\xi) = \alpha\xi$ with $\alpha = \tan\frac{\pi}{2r(0)}$. Substituting this relation in the following expression:

$$\hat{\psi}_{00}(\xi, \nu) = \sqrt{8\pi} \exp\left(-\frac{(\xi - k)^2 + 4\nu^2}{2}\right),$$

and solving

$$\hat{\psi}_{00}(\xi, \nu(\xi)) = \lambda \max_{\xi}|\hat{\psi}_{00}|,$$

we obtain the constraint

$$\exp\left(-\frac{(\xi - k)^2 + 4\alpha^2\xi^2}{2}\right) = \lambda. \tag{A.10}$$

Solving for ξ yields the quadratic equation

$$(4\alpha^2 + 1)\xi^2 - 2k\xi + k^2 + 2\log\lambda = 0. \tag{A.11}$$

The kissing requirement means that we must have only single solution for this equation. This implies that the discriminant is zero. This gives us Eq. (A.2):

$$\frac{\pi}{2r(0)} = \arctan\sqrt{\frac{-\log\lambda}{2k^2 + 4\log\lambda}}. \tag{A.12}$$

Appendix B. Generalization of Daugman's uncertainty relationship

In this appendix, we aim at proving the following relationship:

$$\Delta\theta_{\frac{1}{2}} = \arcsin\left(\beta\left(\frac{a_x}{a_y}\right)^m \frac{2^{\Delta\omega} - 1}{2^{\Delta\omega} + 1}\right). \tag{B.1}$$

Let β be defined as the aspect ratio between the Δx and Δy of the 2D Gabor function. Thus, if $\beta = \frac{\Delta x}{\Delta y}$, it is also true that $\beta = \frac{\Delta v}{\Delta u}$, where Δu denotes the frequency bandwidth in the x-direction, and Δv denotes the frequency bandwidth in the y-direction. If the center frequency value of our filter is ω_0, then we may write the height of the filter in the following way: $\Delta v = 2\omega_0 \sin(\Delta\theta_{\frac{1}{2}})$.

Also, the full bandwidth on the x-axis $\Delta\omega$ can be described in octaves to be

$$\Delta\omega = \log_2\left[\frac{\omega_0 + \frac{\Delta u}{2}}{\omega_0 - \frac{\Delta u}{2}}\right],$$

and so

$$\Delta u = 2\omega_0\left[\frac{2^{\Delta\omega} - 1}{2^{\Delta\omega} + 1}\right].$$

Next, we can use these relations to derive a relationship between the filter's spatial aspect ratio β, the orientation half-bandwidth $\Delta\theta_{\frac{1}{2}}$ and the spatial-frequency bandwidth in octaves, $\Delta\omega$ as follows:

$$\Delta\theta_{\frac{1}{2}} = \arcsin\left[\beta\frac{2^{\Delta\omega} - 1}{2^{\Delta\omega} + 1}\right]. \tag{B.2}$$

When the scaling is the same, the aspect ratio between the x and y dimensions of the filter is constant through scales and is equal to β. Our generalization of this relationship involves accounting for the different scaling parameters in the x and y directions, a_x and a_y, respectively. The aspect ration between the height and width of the filter for the zero-level scale is equal to β. However, moving to the next scale level the width of the filter changes by a factor a_x, and its height changes by a factor a_y. When we account for this change in the aspect ratio through the scales, we obtain the general relationship:

$$\Delta\theta_{\frac{1}{2}} = \arcsin\left(\beta\left(\frac{a_x}{a_y}\right)^m\frac{2^{\Delta\omega} - 1}{2^{\Delta\omega} + 1}\right). \tag{B.3}$$

Appendix C. Calculating the frame bounds

In this appendix we provide a detailed calculation of the frame bounds, when accounting for the extension of the similitude (2) group by an independent scaling in the x and y directions.

We prove here Theorem 3.1: The frame bounds for the set $\{\psi_{m,n,k,l}\}_{m,n,k\in\mathbb{Z},l\in\mathbb{Z}_r}$ are given by

$$A = \frac{1}{b_x b_y}\left\{\inf_{\xi,\nu}\sum_{m\in Z,l\in\mathbb{Z}/r\mathbb{Z}}F(\xi,\nu) - \sum_{(p,q)\in\mathbb{Z}^2\backslash(0,0)}\sqrt{\beta\left(\frac{2\pi p}{b_x},\frac{2\pi q}{b_y}\right)\beta\left(-\frac{2\pi p}{b_x},-\frac{2\pi q}{b_y}\right)}\right\},$$

$$B = \frac{1}{b_x b_y}\left\{\sup_{\xi,\nu}\sum_{m\in Z,l\in\mathbb{Z}/r\mathbb{Z}}F(\xi,\nu) + \sum_{(p,q)\in\mathbb{Z}^2\backslash(0,0)}\sqrt{\beta\left(\frac{2\pi p}{b_x},\frac{2\pi q}{b_y}\right)\beta\left(-\frac{2\pi p}{b_x},-\frac{2\pi q}{b_y}\right)}\right\},$$

where

$$F(\xi,\nu) = \left|\hat{\psi}_{\theta_l}\left(a_x^m\xi,a_y^m\nu\right)\right|^2$$

and

$$\beta(s,t) = \sup_{\xi,\nu}\sum_{m\in\mathbb{Z}}\left|\hat{\psi}_{\theta_l}\left(a_x^m\xi,a_y^m\nu\right)\right|\left|\hat{\psi}_{\theta_{l(m)}}\left(a_x^m\xi + s,a_y^m\nu + t\right)\right|.$$

We have to bound the sum of all the contributions of the wavelet responses to the image:

$$C = \sum_{m,n,k\in\mathbb{Z},l\in\mathbb{Z}/r\mathbb{Z}}\left|\langle f,\psi_{m,n,k,l}\rangle\right|^2,$$

from below and from above. Using Parseval's theorem we obtain

$$4\pi\langle f,\psi_{m,n,k,l}\rangle = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}d\xi\,d\nu\,\hat{f}(\xi,\nu)a_x^{\frac{m}{2}}a_y^{\frac{m}{2}}\overline{\hat{\psi}_{\theta_l}\left(a_x^m\xi,a_y^m\nu\right)}e^{ib_x a_x^m n\xi}e^{ib_y a_y^m k\nu}$$

$$= \sqrt{a_x^m a_y^m}\int_0^{\frac{2\pi}{a_x^m b_x}}\int_0^{\frac{2\pi}{a_y^m b_y}}d\xi\,d\nu\,e^{ib_x a_x^m n\xi}e^{ib_y a_y^m k\nu}G(\xi,\nu), \tag{C.1}$$

where

$$G(\xi, \nu) = \sum_{h,j \in \mathbb{Z}} g\left(\xi + h\frac{2\pi}{a_x^m b_x}, \nu + j\frac{2\pi}{a_y^m b_y}\right)$$

and $g(\xi, \nu) = \hat{f}(\xi, \nu)\overline{\hat{\psi}_{\theta_l}(a_x^m \xi, a_y^m \nu)}$ and where we have assumed that $b_x, b_y \neq 0$.

Using Parseval's formula for periodic functions we obtain:

$$
\begin{aligned}
C &= \sum_{m \in \mathbb{Z}, l \in \mathbb{Z}/r\mathbb{Z}} a_x^m a_y^m \left(\frac{2\pi}{a_x^m b_x}\right)^2 \left(\frac{2\pi}{a_y^m b_y}\right)^2 \left[\frac{a_x^m b_x}{2\pi} \int\limits_0^{\frac{2\pi}{b_x a_x^m}} d\xi \frac{a_y^m b_y}{2\pi} \int\limits_0^{\frac{2\pi}{b_y a_y^m}} d\nu |G(\xi, \nu)|^2\right] \\
&= \sum_{h_2, j_2 \in \mathbb{Z}} \int d\xi \int d\nu\, g(\xi, \nu)\overline{g\left(\xi + \frac{2\pi h_2}{a_x^m b_x}, \nu + \frac{2\pi j_2}{a_y^m b_y}\right)} \\
&= \frac{1}{b_x b_y} \sum_{m,n,k \in \mathbb{Z}, l \in \mathbb{Z}} \iint d\xi\, d\nu\, \hat{f}(\xi, \nu)\overline{\hat{f}\left(\xi + \frac{2\pi}{b_x}n, \nu + \frac{2\pi}{b_y}k\right)} \\
&\quad \times \hat{\psi}_{\theta_l}\left(a_x^m \xi, a_y^m \nu\right)\overline{\hat{\psi}_{\theta_l}\left(a_x^m \xi + \frac{2\pi}{b_x}n, a_y^m \nu + \frac{2\pi}{b_y}k\right)},
\end{aligned}
$$

where the first equality is by definition, the second comes from the change of variables $\xi' = \xi + h_1\frac{2\pi}{a_x^m b_x}$ and $\nu' = \nu + j_1\frac{2\pi}{a_y^m b_y}$ and the third is substitution of the expression for $g(\xi, \nu)$.

First we look at the contributions of this expression for $n = 0$ and $k = 0$. This is the main contribution and we follow the notations of Lee who defined this contribution as P:

$$P = \frac{1}{b_x b_y} \iint d\xi\, d\nu \|\hat{f}\|^2_{L^2(R^2)} \sum_{m \in \mathbb{Z}, l \in \mathbb{Z}/r\mathbb{Z}} \left|\psi_{\theta_l}\left(a_x^m \xi, a_y^m \nu\right)\right|^2. \tag{C.2}$$

If we use the definition $F(\xi, \nu) = \sum_{m \in \mathbb{Z}, l \in \mathbb{Z}/r\mathbb{Z}} F_{ml}(\xi, \nu) = \sum_{m \in \mathbb{Z}, l \in \mathbb{Z}/r\mathbb{Z}} |\psi_{\theta_l}(a_x^m \xi, a_y^m \nu)|^2$ we finally obtain:

$$P = \frac{1}{b_x b_y} \int d\xi \int d\nu \|\hat{f}\|^2_{L^2(R^2)} F(\xi, \nu). \tag{C.3}$$

As we have created a family of self-similar wavelets, the term $F(\xi, \nu)$ should also be self-similar. Lee [13] had already noted that the various components of the formulas obtained have a self-similar behavior. Especially, he refers to the function $F(\xi, \nu)$ defined in the previous section, which is periodic with the scaling parameter along the radial direction and with the basic angle along the orientation direction. When the scaling parameters in the x and y directions are not the same the basic sector may vary in appearance according to the shear operation implied in this different scaling. See Fig. 17 for a visualization of F. Thus, in order to evaluate its maximal and minimal values it is sufficient to evaluate its minimal and maximal values in a basic sector which depends on the a_x, a_y selected.

The second component which accounts for the contributions of all other values of n, k measures the coupling between the wavelets due to non-orthogonality. It is composed of the sum of all cross products of each wavelet transform and its spectrally displaced versions. This term, denoted as the residue term, R, is given by

$$
\begin{aligned}
R = \Bigg| &\frac{1}{b_x b_y} \sum_{m,n,k \in \mathbb{Z}, l \in \mathbb{Z}/r\mathbb{Z}} \int\limits_{-\infty}^{\infty} d\xi \int\limits_{-\infty}^{\infty} d\nu\, \hat{f}(\xi, \nu)\overline{\hat{f}\left(\xi + \frac{2\pi n}{a_x^m b_x}, \nu + \frac{2\pi k}{a_y^m b_y}\right)} \\
&\times \hat{\psi}_{\theta_l(m)}\left(a_x^m \xi, a_y^m \nu\right)\overline{\hat{\psi}\left(a_x^m \xi + \frac{2\pi n}{b_x}, a_y^m \nu + \frac{2\pi k}{b_y}\right)}\Bigg|,
\end{aligned}
$$

where the sun does not include the point $(n, k) \neq (0, 0)$. Under the change of variables $\tilde{\xi} = \xi + \frac{2\pi n}{a_x^m b_x}$, $\tilde{\nu} = \nu + \frac{2\pi k}{a_y^m b_y}$ we obtain:
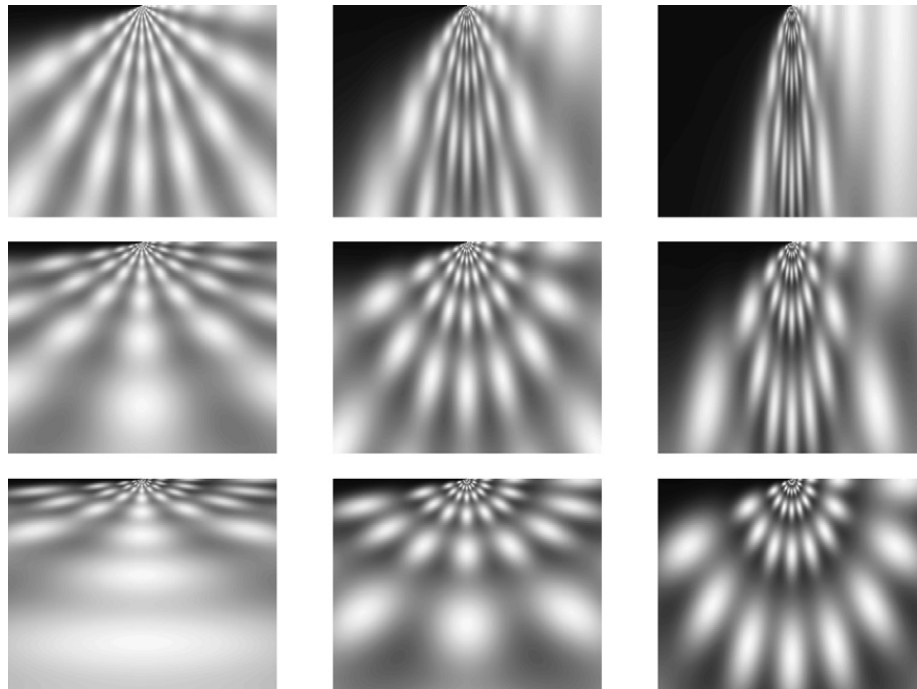
Fig. 17. The function $F(\xi, v) = \sum |\psi_{\theta_l(m)}(a_x^m \xi, a_y^m v)|^2$ for eight orientations. a_x grows from left to right and a_y grows from above to bottom.

$$R \leqslant \frac{1}{b_x b_y} \sum_{m,n,k \in \mathbb{Z}, \, l \in \mathbb{Z}/r\mathbb{Z}} \sqrt{\int d\xi \int dv \left| \hat{f}(\xi, v) \right|^2 \left| \hat{\psi}_{\theta_l}(a_x^m \xi, a_y^m v) \right| \left| \hat{\psi}_{\theta_l}\left(a_x^m \xi + \frac{2\pi n}{b_x}, a_y^m v + \frac{2\pi k}{b_y} \right) \right|}$$

$$\times \sqrt{\int d\tilde{\xi} \int d\tilde{v} \left| \hat{f}(\tilde{\xi}, \tilde{v}) \right|^2 \left| \hat{\psi}_{\theta_l}(a_x^m \tilde{\xi}, a_y^m \tilde{v}) \right| \left| \hat{\psi}_{\theta_l}\left(a_x^m \tilde{\xi} - \frac{2\pi n}{b_x}, a_y^m \tilde{v} - \frac{2\pi k}{b_y} \right) \right|}.$$

Using the Cauchy–Schwartz inequality, we get:

$$R \leqslant \frac{1}{b_x b_y} \sum_{n,k \in \mathbb{Z}} \left\{ \int d\xi \int dv \left| \hat{f}(\xi, v) \right|^2 \sum_{m \in \mathbb{Z}, \, l \in \mathbb{Z}/r\mathbb{Z}} \hat{\psi}_{\theta_l}(a_x^m \xi, a_y^m v) \hat{\psi}_{\theta_l}\left(a_x^m \xi + \frac{2\pi}{b_x} n, a_y^m v + \frac{2\pi}{b_y} k \right) \right\}^{\frac{1}{2}}$$

$$\times \left\{ \int d\tilde{\xi} \int d\tilde{v} \left| \hat{f}(\tilde{\xi}, \tilde{v}) \right|^2 \sum_{n,k \in \mathbb{Z}} \hat{\psi}_{\theta_l}(a_x^m \tilde{\xi}, a_y^m \tilde{v}) \hat{\psi}_{\theta_l}\left(a_x^m \tilde{\xi} - \frac{2\pi}{b_x} n, a_y^m \tilde{v} - \frac{2\pi}{b_y} k \right) \right\}^{\frac{1}{2}}$$

$$\leqslant \frac{1}{b_x b_y} \int d\xi \int dv \left| f(\xi, v) \right|^2 \sum_{n,k \in \mathbb{Z}} \left[\beta\left(\frac{2\pi}{b_x} n, \frac{2\pi}{b_y} k \right) \beta\left(-\frac{2\pi}{b_x} n, -\frac{2\pi}{b_y} k \right) \right]^{\frac{1}{2}}, \tag{C.4}$$

where

$$\beta(s, t) = \sup_{\xi, v} \sum_{m \in \mathbb{Z}, \, l \in \mathbb{Z}/r\mathbb{Z}} \left| \hat{\psi}_{\theta_l(m)}(a_x^m \xi, a_y^m v) \right| \left| \hat{\psi}_{\theta_l(m)}(a_x^m \xi + s, a_y^m v + t) \right|. \tag{C.5}$$

The term β demonstrates the same self-similarity of F and therefore can be calculated on the basic section which is defined per a pair (a_x, a_y).

Finally, the frame bounds A and B are obtained using:

$$A = \frac{1}{b_x b_y} \left\{ \inf_{\xi, v} \sum_{m \in Z, \, l \in \mathbb{Z}/r\mathbb{Z}} F(\xi, v) - \sum_{(p,q) \in Z^2 \backslash (0,0)} \sqrt{\beta\left(\frac{2\pi p}{b_x}, \frac{2\pi q}{b_y} \right) \beta\left(-\frac{2\pi p}{b_x}, -\frac{2\pi q}{b_y} \right)} \right\},$$

$$B = \frac{1}{b_x b_y} \left\{ \sup_{\xi, v} \sum_{m \in Z, \, l \in \mathbb{Z}/r\mathbb{Z}} F(\xi, v) + \sum_{(p,q) \in Z^2 \backslash (0,0)} \sqrt{\beta\left(\frac{2\pi p}{b_x}, \frac{2\pi q}{b_y} \right) \beta\left(-\frac{2\pi p}{b_x}, -\frac{2\pi q}{b_y} \right)} \right\}.$$

References

[1] S.T. Ali, J.P. Antoine, J.P. Gazeau, Coherent States, Wavelets and Their Generalizations, Springer-Verlag, 2000.

[2] I. Daubechies, The wavelet transform, time–frequency localization and signal analysis, IEEE Trans. Inform. Theory 36 (1990) 961–1005.

[3] Emmanuel J. Candes, David L. Donoho, Continuous curvelet transform. II. Discretization and frames, Appl. Comput. Harmon. Anal. 19 (2) (2005) 198–222.

[4] J.G. Daugman, Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, J. Opt. Soc. Am. 2 (7) (1985) 1160–1169.

[5] J.G. Daugman, Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression, IEEE Trans. Acoust. Speech Signal Process. 36 (7) (1988) 1169–1179.

[6] S. Dahlke, P. Maass, The affine uncertainty principle in one and two dimensions, Comput. Math. Appl. 30 (3–6) (1995) 293–305.

[7] M.N. Do, M. Vetterli, The contourlet transform: An efficient directional multiresolution image representation, IEEE Trans. Image Process. 14 (12) (2005) 2091–2106.

[8] H. Feichtinger, T. Strohmer, Gabor Analysis and Algorithms: Theory and Applications, Birkhäuser, 1998.

[9] D. Gabor, Theory of communication, J. IEEE 93 (1946) 429–459.

[10] K. Guo, G. Kutyniok, D. Labate, Sparse multidimensional representations using anisotropic dilation and shear operators, in: Wavelets and Splines, Athens, GA, 2005, Nashboro Press, Nashville, TN, 2006, pp. 189–201.

[11] K. Kanatani, Group Theoretical Methods in Image Understanding, Springer-Verlag, New York, 1990.

[12] G. Kutyniok, D. Labate, Resolution of the wavefront set using continuous shearlets, preprint, 31 pp., 2006, http://www4.ncsu.edu/~dlabate/wavefront.pdf.

[13] T.S. Lee, Image representation using 2D Gabor-wavelets, IEEE Trans. PAMI 18 (10) (1996) 959–971.

[14] R. Lenz, Group Theoretic Methods in Image Processing, Springer-Verlag, New York, 1990.

[15] B.S. Manjunath, W.Y. Ma, Texture features browsing and retrieval of image data, IEEE Trans. PAMI 18 (8) (1996) 837–842.

[16] S. Marcelja, Mathematical description of the response of simple cortical cells, J. Opt. Soc. Am. 70 (1980) 1297–1300.

[17] M. Porat, Y.Y. Zeevi, The generalized Gabor scheme of image representation in biological and machine vision, IEEE Trans. PAMI 10 (4) (1998) 452–468.

[18] C. Sagiv, Segmentation, uncertainty and phase-space pavement according to Gabor, PhD thesis, Tel Aviv, Israel, 2006.

[19] C. Sagiv, N.A. Sochen, Y.Y. Zeevi, Scale-space generation via uncertainty principles, in: R. Kimmel, N. Sochen, J. Weickert (Eds.), Proceedings of Scale Space and PDE Methods in Computer Vision 2005, Hofgeismar, Germany, 2005, pp. 351–362.

[20] J. Segman, Y.Y. Zeevi, Image analysis by wavelet-type transform: Group theoretic approach, J. Math. Imag. Vis. 3 (1993) 51–75.

[21] M. Zibulski, Y.Y. Zeevi, Analysis of multi-window Gabor-type schemes by frame methods, Appl. Comput. Harmon. Anal. 4 (2) (1997) 188–221.