Image Representation
Using 2D Gabor Wavelets

Tai Sing Lee

Abstract—This paper extends to two dimensions the frame criterion developed by Daubechies for one-dimensional wavelets, and it computes the frame bounds for the particular case of 2D Gabor wavelets. Completeness criteria for 2D Gabor image representations are important because of their increasing role in many computer vision applications and also in modeling biological vision, since recent neurophysiological evidence from the visual cortex of mammalian brains suggests that the filter response profiles of the main class of linearly-responding cortical neurons (called simple cells) are best modeled as a family of self-similar 2D Gabor wavelets. We therefore derive the conditions under which a set of continuous 2D Gabor wavelets will provide a complete representation of any image, and we also find self-similar wavelet parameterizations which allow stable reconstruction by summation as though the wavelets formed an orthonormal basis. Approximating a "tight frame" generates redundancy which allows low-resolution neural responses to represent high-resolution images, as we illustrate by image reconstructions with severely quantized 2D Gabor coefficients.

Index Terms—Gabor wavelets, coarse coding, image representation, visual cortex, image reconstruction.

---------------------- ❖ ----------------------

1 INTRODUCTION

SINCE Hubel and Wiesel's [16] discovery of the crystalline organization of the primary visual cortex in mammalian brains some thirty years ago, an enormous amount of experimental and theoretical research has greatly advanced our understanding of this area and the response properties of its cells. On the theoretical side, an important insight has been advanced by Marcelja [25] and Daugman [6], [7] that simple cells in the visual cortex can be modeled by Gabor functions. The 2D Gabor functions proposed by Daugman are local spatial bandpass filters that achieve the theoretical limit for conjoint resolution of information in the 2D spatial and 2D Fourier domains.

Gabor [13] showed that there exists a "quantum principle" for information: the conjoint time-frequency domain for 1D signals must necessarily be quantized so that no signal or filter can occupy less than a certain minimal area in it. This minimal area, which reflects the inevitable trade-off between time resolution and frequency resolution, has a lower bound in their product, analogous to Heisenberg's uncertainty principle in physics. He discovered that Gaussian-modulated complex exponentials provide the best trade-off. The original Gabor elementary functions, in the form proposed by Gabor [13], are generated with a fixed Gaussian while the frequency of the modulating wave varies. These are equivalent to a family of "canonical" coherent states generated by the Weyl-Heisenberg group. Examples of a set of these coherent states or elementary functions are presented in Fig. 1. A signal can be encoded by its projection onto these elementary functions. This decomposition is equivalent to the Gaussian-windowed Fourier transform.
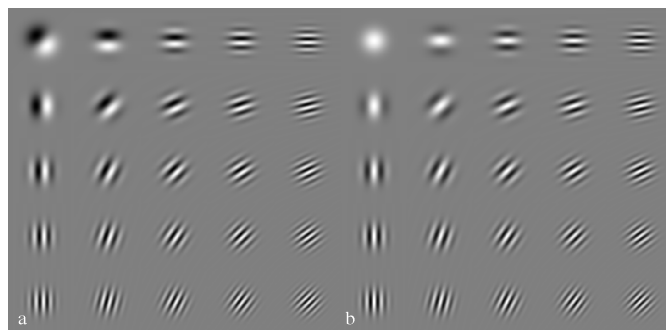

Fig. 1. An ensemble of odd (a) and even (b) Gabor filters or Weyl-Heisenberg coherent states.

Daugman [6], [7] generalized the Gabor function to the following 2D form to model the receptive fields of the orientation-selective simple cells:

$$G(x,y) = \frac{1}{2\pi\sigma\beta} e^{-\pi\left[\frac{(x-x_o)^2}{\sigma^2}+\frac{(y-y_o)^2}{\beta^2}\right]} e^{i[\xi_o x + \nu_0 y]} \qquad (1)$$

where (x_o, y_o) is the center of the receptive field in the spatial domain and (ξ_o, ν_o) is the optimal spatial frequency of the filter in the frequency domain. σ and β are the standard deviations of the elliptical Gaussian along x and y. The 2D Gabor function is thus a product of an elliptical Gaussian and a complex plane wave. The careful mapping of the receptive fields of the simple cells by Jones and Palmer [17] confirmed the validity of this model. Mathematically, the 2D Gabor function achieves the resolution limit in the conjoint space only in its complex form. Since a complex-valued 2D Gabor function contains in quadrature projection

an even-symmetric cosine component and an odd-symmetric sine component, Pollen and Ronner's [33] finding that simple cells exist in quadrature-phase pairs therefore showed that the design of the cells might indeed be optimal. The fact that the visual cortical cell has evolved to an optimal design for information encoding has caused a considerable amount of excitement not only in the neuroscience community but in the computer science community as well. Gabor filters, rediscovered and generalized to 2D, are now being used extensively in various computer vision applications [3], [20].

Recent neurophysiological evidence [11] suggests that the spatial structure of the receptive fields of simple cells having different sizes is virtually invariant. Daugman [8] and others [3], [34] have proposed that an ensemble of simple cells is best modeled as a family of 2D Gabor wavelets sampling the frequency domain in a log-polar manner. This class is equivalent to a family of affine coherent states generated by the affine group. The decomposition of an image f into these states is called the *wavelet* transform of the image:

$$\left(T^{wav}f\right)(a,\theta,x_o,y_o)=\|a\|^{-1}\int\int dxdyf(x,y)\psi_\theta\left(\frac{x-x_o}{a},\frac{y-y_o}{a}\right) \quad (2)$$

where a is the dilation parameter, related to σ and β, x_o, and y_o the spatial translation parameters, θ the orientation parameter of the wavelet. $\psi_\theta(a,x,y,x_o,y_o)=\|a\|^{-1}\psi_\theta\left(\frac{x-x_o}{a},\frac{y-y_o}{a}\right)$ is the 2D wavelet elementary function, rotated by θ. A particular Gabor elementary function can be used as the mother wavelet to generate a whole family of Gabor wavelets. Examples of a particular class of 2D Gabor wavelets, to be derived in the next section, are presented in Fig. 2.
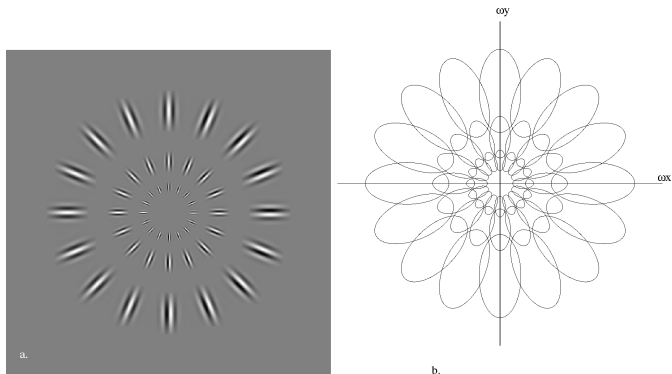

Fig. 2. An ensemble of Gabor wavelets (1.5 octave bandwidth) (a) and their coverage of the spatial frequency plane (b). Each ellipse shows the half-amplitude bandwidth contour dilated by a factor of 2, covering almost the complete support of a wavelet.

Despite this recognition, many questions, originally posed by Daugman [7], about how the degrees-of-freedom which create the "coding budget" in the visual cortex are allocated and constrained, such as the trade-off between orientation sampling and spatial sampling, have remained unanswered. We here address and propose answers to those questions.

In this paper, we first derive a class of 2D Gabor wavelets, with their parameters properly constrained by

neurophysiological data on simple cells and by the wavelet theory. We extend Daubechies' completeness criteria on 1D wavelets to 2D and apply such criteria to study the physiologically relevant family of 2D Gabor wavelets. By numerically computing the frame bounds for this family of wavelets in different phase space sampling schemes, we elucidate the conditions under which they form a tight frame. We find that the phase space sampling density provided by the simple cells in the primary visual cortex is sufficient to form an almost tight frame that allows stable reconstruction of the image by linear superposition of the Gabor wavelets with their own projection coefficients, and provides representation of high resolution images using coarse neuronal responses. Finally, we demonstrate these theoretical insights with results from image reconstruction experiments.

2 DERIVATION OF THE 2D GABOR WAVELETS

In this section, we will derive the following family of 2D Gabor wavelets that satisfies the wavelet theory and the neurophysiological constraints for simple cells,

$$\psi(x,y,\omega_o,\theta)=\frac{\omega_o}{\sqrt{2\pi}\kappa}e^{-\frac{\omega_o^2}{8\kappa^2}\left(4\left(x\cos\theta+y\sin\theta\right)^2+\left(-x\sin\theta+y\cos\theta\right)^2\right)}$$

$$\cdot\left[e^{i\left(\omega_o x\cos\theta+\omega_o y\sin\theta\right)}-e^{-\frac{\kappa^2}{2}}\right] \quad (3)$$

where ω_o is the radial frequency in radians per unit length and θ is the wavelet orientation in radians. The Gabor wavelet is centered at $(x=0, y=0)$ and the normalization factor is such that $<\psi,\psi>=1$, i.e., normalized by L^2 norm. κ is a constant, with $\kappa\approx\pi$ for a frequency bandwidth of one octave and $\kappa\approx 2.5$ for a frequency bandwidth of 1.5 octaves. A discrete ensemble of the 1.5 octave bandwidth family of Gabor wavelets is shown in Fig. 2.

Each member of this family of Gabor wavelets models the spatial receptive field structure of a simple cell. The response of a simple cell is the projection of an image onto a Gabor wavelet, which is the inner product of the image f with the receptive field centered at (x_o, y_o),

$$R_s=<\psi\left(x_o,y_o,\omega_o,\theta_o\right),f>$$

$$=\int_x\int_y\psi\left(x_o-x,y_o-y:\omega_o,\theta_o\right)f(x,y)dxdy \quad (4)$$

where $<,>$ denotes the inner product. Because the response of a simple cell is half-wave rectified, each projection is in fact represented by the responses of *two* simple cells, the response of an on-center cell (R_{s+}) and the response of an off-center cell (R_{s-}), where

$$R_{s+}(x_o, y_o, \omega_o, \theta_o)=[<\psi(x_o, y_o, \omega_o, \theta_o),f>]^+ \quad (5)$$

and

$$R_{s-}(x_o, y_o, \omega_o, \theta_o)=[<\psi(x_o, y_o, \omega_o, \theta_o),f>]^- \quad (6)$$

respectively, and $[\alpha]^+=\{\alpha$ if $\alpha\geq 0\}$, $[\alpha]^+=\{0$ if $\alpha<0\}$, $[\alpha]^-=\{-\alpha$ if $\alpha\leq 0\}$ and $[\alpha]^-=\{0$ if $\alpha>0\}$ are the half-wave rectifications of the function α.

Let us start our derivation with the most general 2D complex Gabor function, normalized so that its L^2 norm $<\psi, \psi> = 1$.

$$\psi(x, y, \xi_o, v_o, x_o, y_o, \rho, \theta, \sigma, \beta) =$$

$$\frac{1}{\sqrt{\pi\sigma\beta}} e^{-\left(\frac{((x-x_o)\cos\theta+(y-y_o)\sin\theta)^2}{2\sigma^2}+\frac{(-(x-x_o)\sin\theta+(y-y_o)\cos\theta)^2}{2\beta^2}\right)}$$

$$\cdot e^{i(\xi_o(x-x_o)+v_o(y-y_o)+\rho)} \quad (7)$$

The filter is centered at $(x = x_o, y = y_o)$ in the spatial domain, and at $(\xi = \xi_o, v = v_o)$ in the spatial frequency domain. σ and β are the standard deviations of an elliptical Gaussian along the x and y axes. θ is the orientation of the filter, rotated counter-clockwise around the origin. ρ is the absolute phase of an individual filter. There are, therefore, eight degrees of freedom in the general Gabor function: ξ_o, v_o, θ, ρ, σ, β, x_o, and y_o.

Simple cells are often not strictly even or odd symmetric, but the quadrature-phase relationship between a pair of simple cells is generally preserved, whatever the absolute phase may be [33], [7], [9]. Since only the power-modulus enters into our later calculation, we can simplify the Gabor filter by setting the spatial location of the filter's center $(x_o = 0, y_o = 0)$ and the absolute phase ρ of the filter to 0. The Fourier Transform of the simplified complex-valued Gabor function is

$$\hat{\psi}(\xi, v, \xi_o, v_o, \theta, \sigma, \beta) =$$

$$2\sqrt{\pi\sigma\beta}e^{-\frac{1}{2}\left[((\xi-\xi_0)\cos\theta+(v-\omega_o)\sin\theta)^2\sigma^2+(-(\xi-\xi_0)\sin\theta+(v-v_o)\cos\theta)^2\beta^2\right]} \quad (8)$$

where ξ and v are the spatial frequencies in radians per unit length along x and y.[1]

We can further reduce the number of degrees of freedom of the filter by noting that σ, β, and θ are constrained by ξ_o and v_o according to the following physiological findings [7]. The spatial frequency bandwidths of the simple and complex cells have been found to range from 0.5 to 2.5 octaves, clustering around 1.2 octaves [10], [18], and 1.5 octaves [38]. The Gaussian envelope is usually elliptical, with an aspect ratio of 1.5-2.0 [38], [7] and with the plane wave's propagating direction along the short axis of the elliptical Gaussian (Fig. 2).

CONSTRAINT #1. *The aspect ratio $\frac{\beta}{\sigma}$ of the elliptical Gaussian envelope is 2:1 [7].*

CONSTRAINT #2. *The plane wave with frequency (ξ_o, v_o) tends to have its "propagating direction" along the short axis of the elliptical Gaussian.*

When the plane wave rotates, the elliptical Gaussian rotates correspondingly. This implies the center frequency

(ξ_o, v_o) of the filter is related to the rotation angle θ of the modulating Gaussian by $\xi_o = \omega_o \cos\theta$ and $v_o = \omega_o \sin\theta$, where the radial frequency $\omega_o = \sqrt{\xi_o^2 + v_o^2}$.

With these two constraints, the Gabor filter becomes

$$\psi(x, y, \xi_o, v_o, \sigma) =$$

$$\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{8\sigma^2}\left[4\left(x\frac{\xi_o}{|\omega_0|}+y\frac{v_o}{|\omega_o|}\right)^2+\left(-x\frac{v_o}{|\omega_o|}+y\frac{\xi_o}{|\omega_o|}\right)^2\right]} \cdot e^{i(\xi_o x+v_o y)}. \quad (9)$$

The orientation of the filter is aligned with the long axis of the elliptical Gaussian.

CONSTRAINT #3. *The half-amplitude bandwidth of the frequency response is about 1 to 1.5 octaves along the optimal orientation [18], [38], [7].*

The relationship between σ and ω_o can be derived to be,

$$\sigma = \frac{\kappa}{\omega_o} \quad \text{where} \quad \kappa = \sqrt{2\ln 2}\left(\frac{2^\phi + 1}{2^\phi - 1}\right) \quad (10)$$

where ϕ is the bandwidth in octaves. For $\phi = 1$ octave, $\sigma \approx \pi/\omega_o = \lambda/2$. For $\phi = 1.5$ octave, $\sigma \approx 2.5/\omega_o$.

Imposing this constraint, we obtain the following family of self-similar Gabor filters depending on four *continuous variables*,

$$\psi(x, y, \omega_o, \theta) = \frac{\omega_o}{\sqrt{2\pi}\kappa} e^{-\frac{\omega_o^2}{8\kappa^2}\left(4(x\cos\theta+y\sin\theta)^2+(-x\sin\theta+y\cos\theta)^2\right)}$$

$$\cdot e^{i(\omega_o x\cos\theta+\omega_o y\sin\theta)} \quad (11)$$

where $\theta = \arctan\frac{v_o}{\xi_o}$, and κ is fixed for Gabor wavelets of a particular bandwidth. The whole family can be translated to any spatial position (x_o, y_o). In order to make the Gabor filters into admissible wavelets, we need to introduce the following constraint.

CONSTRAINT #4. *Admissible wavelets are functions having zero mean.*

Why is this important? Consider a map $\psi(f) : L^2(R^2) \rightarrow L^2(R^4)$, the Gabor wavelet is admissible only if its norm is finite, i.e.,

$$\|\psi(f)\|_{L^2(R^4)} = \int_R \int_R \int_0^{2\pi} \int_0^\infty \left|\psi(f(\omega, \theta, x, y))\right|^2 \frac{d\omega}{\omega} d\theta dx dy$$

$$= \|\hat{f}\|_{L^2(R^2)} \cdot \|\hat{\psi}\|_{L^2\left(\frac{d\omega}{\omega}d\theta\right)} < \infty. \quad (12)$$

This implies $\|\hat{\psi}(\omega = 0)\| = 0$ is a necessary condition; otherwise the norm will become infinite in the measure of $\frac{d\omega}{\omega}$ as $\omega \rightarrow 0$.

The sine component of the complex-valued Gabor filter has zero mean, but its cosine component has nonzero mean (d.c. response). The d.c. response can be computed from its Fourier transform, with $\xi = 0$ and $v = 0$,

$$\hat{\psi}(\xi = 0, v = 0 : \xi_o, v_o) = \sqrt{8\pi}\sigma e^{-\frac{\kappa^2}{2}}. \quad (13)$$

---

1. In this paper, the convention of the Fourier transform is,

$$\hat{f}(\xi, v) = \int\int_{R^2} f(x, y)e^{-i\xi x}e^{-ivy}dxdy$$

$$f(x, y) = \frac{1}{4\pi^2}\int\int_{R^2} \hat{f}(\xi, v)e^{i\xi x}e^{ivy}d\xi dv$$

A family of admissible 2D Gabor wavelets can be obtained by subtracting this d.c. response from the Gabor filter,

$$\psi(x, y, \omega_o, \theta) = \frac{\omega_o}{\sqrt{2\pi}\kappa}\left[e^{-\frac{\omega_o^2}{8\kappa^2}\left(4(x\cos\theta+y\sin\theta)^2+(-x\sin\theta+y\cos\theta)^2\right)}\right]$$

$$\cdot \left[e^{i(\omega_o x\cos\theta+\omega_o y\sin\theta)} - e^{-\frac{\kappa^2}{2}}\right] \quad (14)$$

with its Fourier transform equal to

$$\hat{\psi}(\xi, \nu, \xi_o, \nu_o) = \frac{\sqrt{8\pi}\kappa}{\omega_o}[e^{-\frac{\kappa^2}{2\omega_o^2}\left[((\xi-\xi_o)\cos\theta+(\nu-\nu_o)\sin\theta)^2+4(-(\xi-\xi_o)\sin\theta+(\nu-\nu_o)\cos\theta)^2\right]}$$

$$-e^{-\frac{\kappa^2}{2\omega_o^2}\left[(\xi\cos\theta+\nu\sin\theta)^2+4(-\xi\sin\theta+\nu\cos\theta)^2+\omega_o^2\right]}] \quad (15)$$

where

$$\xi\xi_o + \nu\nu_o = 0 \quad (16)$$

specifies the line of zeroes ($\hat{\psi}(\xi, \nu, \xi_o, \nu_o) = 0$) on the frequency plane.

Each of these two families of Gabor wavelets can be generated by rotation and dilation (affine group) from the following mother Gabor wavelet,

$$\psi(x, y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{8}(4x^2+y^2)}\cdot\left[e^{i\kappa x} - e^{-\frac{\kappa^2}{2}}\right] \quad (17)$$

whose Fourier transform is

$$\hat{\psi}(\xi, \nu) = \sqrt{8\pi}\left\{e^{-\frac{1}{2}\left[(\xi-\kappa)^2+4\nu^2\right]} - e^{-\frac{1}{2}\left[\xi^2+4\nu^2+\kappa^2\right]}\right\} \quad (18)$$

3 NONORTHOGONAL WAVELETS AND FRAMES

Now we have obtained the properly parameterized Gabor wavelet equation, we can address the issue of complete representation. A transform is said to provide a complete representation if we can reconstruct f in a numerically stable way from the transform of f, or alternatively, if any function f can be written as a superposition of the transform's elementary functions.

Recall that the continuous wavelet transform is given by

$$\left(T^{wav}f\right)(a, \theta, x_o, y_o) = \|a\|^{-1}\int\int dxdy f(x, y)\psi_\theta\left(\frac{x-x_o}{a}, \frac{y-y_o}{a}\right) \quad (19)$$

where a is the dilation parameter, x_o and y_o the spatial translation parameters, θ the orientation parameter of the wavelet. $\psi_\theta(a, x, y, x_o, y_o) = \|a\|^{-1}\psi_\theta(\frac{x-x_o}{a}, \frac{y-y_o}{a})$ is the 2D wavelet elementary function, rotated by θ.

A function can always be reconstructed from its continuous wavelet transform by means of the following *resolution of identity* formula, provided that the wavelets are admissible [5],

$$f = C_\psi^{-1}\int_0^\infty \frac{da}{a^3}\int_{-\infty}^\infty\int_{-\infty}^\infty dxdy\int_0^{2\pi} d\theta < f, \psi(a, x, y, \theta) > \psi(a, x, y, \theta) \quad (20)$$

where $< , >$ denotes the L^2 inner product and

$$C_\psi = 4\pi^2\int_0^\infty \frac{d\omega}{|\omega|}\int_0^{2\pi} d\theta|\hat{\psi}(\omega\cos\theta, \omega\sin\theta)|^2 \quad (21)$$

where $\omega = \sqrt{\xi^2 + \nu^2}$, and $\psi \in L^1(R)$, $\hat{\psi}$ is continuous and $C_\psi < \infty$ only if $\hat{\psi}(0, 0) = 0$. This is because it can be shown that [5]

$$\int_0^\infty \frac{da}{a^3}\int_{-\infty}^\infty\int_{-\infty}^\infty dxdy\int_0^{2\pi} d\theta T^{wav}f(a, x, y, \theta)T^{wav}g(a, x, y, \theta) =$$

$$C_\psi < f, g >. \quad (22)$$

When a, x, y, θ take on discrete values (i.e., $a = a_o^m$, $a_o > 1$, $\theta = \theta_l = l\theta_o$, $x = nb_o a_o^m$, and $y = kb_o a_o^m$, and m, n, l, k are integers ranging over Z), the resulting transform

$$\left(T_{m,n,k,l}^{wav}f\right) = a_o^{-m}\int\int dxdy f(x, y)\psi_{\theta_l}\left(a_o^{-m}x - nb_o, a_o^{-m}y - kb_o\right) \quad (23)$$

is called the *discrete wavelet transform*. ψ_{θ_l} is a rotated version of the mother wavelet $\psi(x, y)$. $a_o > 1$, $\theta_o > 0$, and $b_o > 0$ are fixed. In this scheme, a spatially narrower wavelet translates by finer steps, and a wider wavelet translated by larger steps. As a result, the discretized wavelets at each m level "cover" the spatial domain in the same way as $\psi(x - nb_o, y - kb_o)$ at the $m = 0$ level.

The wavelets are given by

$$\psi_{m,n,k,l}(x, y) = a_o^{-m}\psi_{\theta_l}\left(a_o^{-m}\left(x - nb_o a_o^m\right), a_o^{-m}\left(y - kb_o a_o^m\right)\right)$$

$$= a_o^{-m}\psi_{\theta_l}\left(a_o^{-m}x - nb_o, a_o^{-m}y - kb_o\right). \quad (24)$$

Note that although the parameterization is discrete, each wavelet elementary function is a function of the continuous variables x and y.

Certain discrete 1D wavelets, for examples, the one-dimensional Harr bases, Meyer wavelets, and Battle-Lemarie wavelets, form orthonormal bases. In this case, the function can be reconstructed by a linear superposition of the bases weighed by the wavelet coefficients,

$$f = \sum_{m,n} < f, \psi_{m,n} > \psi_{m,n} \quad (25)$$

where $<\alpha_1, \alpha_2>$ denotes the inner product of α_1 and α_2.

However, 1D or 2D Gabor wavelets do not form orthonormal bases. They are called *nonorthogonal wavelets*. In this case, we are interested in knowing whether a discrete set of the nonorthogonal wavelets form a frame. If the wavelets form a frame, then one can completely characterize a function f by knowing $T^{wav}(f)$ and can reconstruct f in a numerically stable way from $T^{wav}(f)$ with the dual frame $\tilde{\psi}$ [5],

$$f = \sum_{m,n,k,l} < f, \psi_{m,n,k,l} > \tilde{\psi}_{m,n,k,l} = \sum_{m,n,k,l} < f, \tilde{\psi}_{m,n,k,l} > \psi_{m,n,k,l} \quad (26)$$

The concept of frames was first introduced by Duffin and Schaeffer [12] in the context of non-harmonic Fourier series. Their definition is as follows:

DEFINITION. *A family of functions $\{\psi_{m,n}; m, n \in Z\}$ in a Hilbert Space H is called a* frame *if there exist $A > 0$, $B < \infty$ so that, for all f in H,*

$$A\|f\|^2 \le \sum_{m,n\in Z} \left|<f,\psi_{m,n}>\right|^2 \le B\|f\|^2 \qquad (27)$$

where $\|f\| = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|f(x,y)|^2 dxdy$. A and B are called frame bounds.

When a discrete family of wavelet forms a frame, they provide a complete representation of the input functions [2], [5].

When $0 < A \le B < 2$, the family of wavelets can be treated as if it were an orthonormal basis, and the function can be recovered with good approximation using the following inversion formula [26],

$$f \approx \frac{2}{A+B}\sum_{m,n,k,l}<\psi_{m,n,k,l},f>\psi_{m,n,k,l} \qquad (28)$$

$(A + B)/2$ is a measure of the redundancy of the frame whereas B/A is a measure of the tightness of the frame. When $B = A$, the frame is called a tight frame and the reconstruction by linear superposition using the above inversion formula is exact, analogous to the resolution of identity equation [5]. Therefore, a reasonably tight frame behaves in a way similar to the continuous wavelet transform and can be treated as though it were an orthonormal basis in image decomposition and reconstruction.

Readers should consult [5], [26], and [15] for a careful mathematical elucidation and review of frames. The frame bounds can be simply interpreted as the bounds on the gain of the wavelet transform, with $\|f\|^2$ as the total power of the input function and $\sum_{m,n}\left|<f,\psi_{m,n}>\right|^2$ the total output power of the transform. Therefore $A > 0$ means that the output will not be null if the input signal is nonzero and $B < \infty$ means that given an input of finite power, the output power of the transform has to be finite also. A tight frame provides a uniform gain for the frequency response of the input functions.

4   2D GENERALIZATION OF DAUBECHIES'S FRAME CRITERION

How dense should the phase space or the conjoint spatial/spatial frequency domain be sampled so that the 2D Gabor wavelets will provide a complete representation of any image? To answer this question, we need to generalize to 2D the following frame criterion for 1D wavelets of Daubechies [5]:

Daubechies's Frame Criterion: Let $\psi \in L^1(R)$ be a function of zero mean. It defines a sequence $\beta(k)$, $k \in Z$, by the formula,

$$\beta(k) = \sup_{1\le|\xi|\le 2}\sum_{-\infty}^{\infty}\left|\hat{\psi}(2^j\xi)\right|\left|\hat{\psi}(2^j\xi + 2k\pi)\right|. \qquad (29)$$

Now suppose that

$$\sup_{1\le|\xi|\le 2}\sum_{-\infty}^{\infty}\left|\hat{\psi}(2^j\xi)\right|^2 \le C < \infty \qquad (30)$$

and introduce the constant

$$W = \sup_{1\le|\xi|\le 2}\sum_{-\infty}^{\infty}\left|\hat{\psi}(2^j\xi)\right|^2 \qquad (31)$$

Then if

$$W > 2\sum_{1}^{\infty}\left(\beta(k)\beta(-k)\right)^{1/2} \qquad (32)$$

the collection $2^{j/2}\psi(2^jx - k)$, $j, k \in Z$, forms a frame for $L^2(R)$.

At the outset, it is important to note that here we are studying the frame formed by a discrete set of continuous wavelets, i.e., the discretization is in the phase space sampling but not in the description of the wavelet function.

Let us consider the following 2D admissible wavelet family

$$\psi_{m,n,k,l}(x,y) = a_o^{-m}\psi_{\theta_l}\left(a_o^{-m}x - nb_o, a_o^{-m}y - kb_o\right) \qquad (33)$$

where

$$\psi_{\theta_l}(x,y) = \psi(x\cos(l\theta_o) + y\sin(l\theta_o), -x\sin(l\theta_o) + y\cos(l\theta_o));$$

ψ is the mother wavelet as defined in (17); θ_o denotes the step size of each angular rotation, l the index of rotation steps, b_o the unit spatial interval, and a_o^m the dilation in scale. The (x, y) domain is sampled by each wavelet in a scale-dependent spatial lattice with horizontal and vertical translation intervals given by $a_o^m b_o$ at each m level.

The Fourier transform of the 2D wavelet is given by,

$$\hat{\psi}_{m,n,k,l}(\xi,v) = a_o^m\hat{\psi}_{\theta_l}\left(a_o^m\xi, a_o^m v\right)e^{-ia_o^m b_o n\xi}e^{-ia_o^m b_o kv}, \qquad (34)$$

where

$$\hat{\psi}_{\theta_l}(\xi,v) = \hat{\psi}(\xi\cos(l\theta_o) + v\sin(l\theta_o), -\xi\sin(l\theta_o) + v\cos(l\theta_0)),$$

and $\hat{\psi}$ is the Fourier transform of the mother wavelet as defined in (18).

To evaluate the frame bounds A and B as in

$$0 < A\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|\hat{f}(\xi,v)\right|^2 d\xi dv \le \sum_{m,n,k,l}\left|<f(x,y),\psi_{m,n,k,l}(x,y)>\right|^2$$

$$\le B\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|\hat{f}(\xi,v)\right|^2 d\xi dv < \infty, \qquad (35)$$

the basic idea is to sum up the power of all the wavelet responses to an image. This can be done in the spectral domain using Parseval's theorem:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|f(x,y)\right|^2 dxdy = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|\hat{f}(\xi,v)\right|^2 d\xi dv \qquad (36)$$

Let a_o be the scaling step and b_o be the unit translation step, then Daubechies's derivation of frame bounds for 1D wavelets [5] can be generalized for 2D wavelets as follows:

$$\sum_{m,n,k\in Z, l\in Q} \left| <f, \psi_{m,n,k,l}> \right|^2$$

$$= \frac{1}{4\pi^2} \sum_{m,n,k,l} | \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\xi d\nu \hat{f}(\xi,\nu) a_o^m \overline{\hat{\psi}_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right)} e^{ib_o a_o^m n\xi} e^{ib_o a_o^m k\nu}|^2$$

$$= \frac{1}{4\pi^2} \sum_{m,n,k,l} a_o^{2m} | \int_0^{\frac{2\pi}{b_o a_o^m}} \int_0^{\frac{2\pi}{b_o a_o^m}} d\xi d\nu e^{ib_o a_o^m n\xi} e^{ib_o a_o^m k\nu}$$

$$\cdot \sum_{h,j\in Z} \hat{f}\left(\xi + 2\pi h a_o^{-m} b_o^{-1}, \nu + 2\pi j a_o^{-m} b_o^{-1}\right) \overline{\hat{\psi}_{\theta_l}\left(a_o^m \xi + 2\pi h b_o^{-1}, a_o^m \nu + 2\pi j b_o^{-1}\right)}|^2$$

$$= \frac{1}{b_o^2} \sum_{m,l} \int_0^{\frac{2\pi}{b_o a_o^m}} \int_0^{\frac{2\pi}{b_o a_o^m}} d\xi d\nu \mid \sum_{h,j\in Z} [\hat{f}\left(\xi + 2\pi h b_o^{-1}, \nu + 2\pi j b_o^{-1}\right)$$

$$\cdot \overline{\hat{\psi}_{\theta_l}\left(a_o^m \xi + 2\pi h a_o^{-m} b_o^{-1}, a_o^m \nu + 2\pi j a_o^{-m} b_o^{-1}\right)}] \mid^2$$

$$= \frac{1}{b_o^2} \sum_{m,p,q\in Z, l\in Q} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\xi d\nu [\hat{f}(\xi,\nu) \hat{f}\left(\xi + 2\pi p a_o^{-m} b_o^{-1}, \nu + 2\pi q a_o^{-m} b_o^{-1}\right)$$

$$\cdot \overline{\hat{\psi}_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right)} \hat{\psi}_{\theta_l}\left(a_o^m \xi + 2\pi p b_o^{-1}, a_o^m \nu + 2\pi q b_o^{-1}\right)]$$

$$= P + R \qquad (37)$$

where the set $Q = \{0, 1, 2, 3 \dots K - 1\}$ and K is the number of sampling orientations, $\overline{\psi}$ is the complex conjugate of ψ.

The first principal term P (i.e., the power for $(p = 0, q = 0)$) is the product between the power of the input function and the sum of the spectral powers of all the wavelets.

$$P = \frac{1}{b_o^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\xi d\nu \left| \hat{f}(\xi,\nu) \right|^2 \sum_{l\in Q} \sum_{m\in Z} \left| \hat{\psi}_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right) \right|^2 \quad (38)$$

$$= \frac{1}{b_o^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\xi d\nu \left| \hat{f}(\xi,\nu) \right|^2 F(\xi,\nu) \qquad (39)$$

Because of the self-similarity of the wavelets in both the spectral and spatial domains, the spectral power of

$$F(\xi,\nu) = \sum_{m\in Z, l\in Q} \left| \hat{\psi}_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right) \right|^2$$

is also self-similar, i.e.,

$$F\left(a_o^m \xi, a_o^m \nu\right) = F(\xi,\nu) \qquad (40)$$

and

$$F(\xi \cos\theta_o + \nu \sin\theta_o, -\xi \sin\theta_o + \nu \cos\theta_o) = F(\xi,\nu). \quad (41)$$

Therefore, the infimum and supremum of F over the whole frequency domain is equivalent to the infimum and supremum of F over a fundamental sector S (see Figs. 3 and 4). S is defined by $1 < \sqrt{\xi^2 + \nu^2} < a_o$ and $0 < \tan(\frac{\nu}{\xi}) < \frac{\pi}{K}$, where K is the number of sampling orientations, and a_o is the dilation scaling factor. As a result, the product of the infimum and supremum of F over S with the spectral power of the input function provides the lower and upper bounds of the first term, respectively,

$$\left(\inf_{\xi,\nu\in S} F \right) \|f\|^2 \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\xi d\nu \left| \hat{f}(\xi,\nu) \right|^2 \sum_{l\in Q} \sum_{m\in Z} \left| \hat{\psi}_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right) \right|^2$$

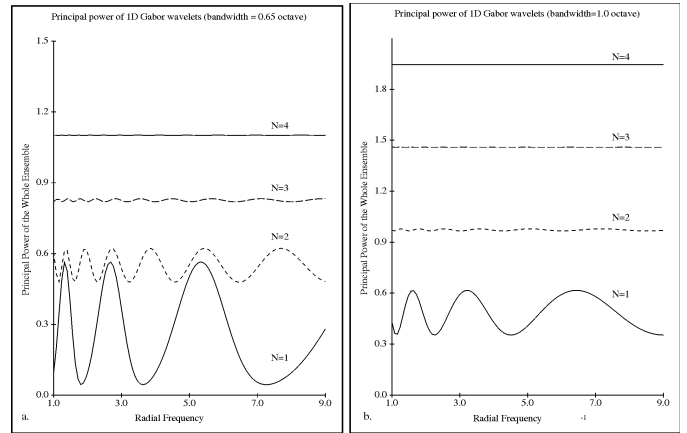$$< \left(\sup_{\xi,\nu\in S} F \right) \|f\|^2 \qquad (42)$$

Fig. 3. The principal power $F(\xi)$ of 1D Gabor wavelets of 0.65 octave bandwidth (a) and 1.0 octave bandwidth (b) are shown to illustrate its self-similarity in the frequency spectrum. The principal term becomes smoother as the number of voices (N) increases, corresponding to the increase in the tightness of the frame. The progression to tightness is faster for the wavelets with broader bandwidths (e.g., 1.0 octave bandwidth).
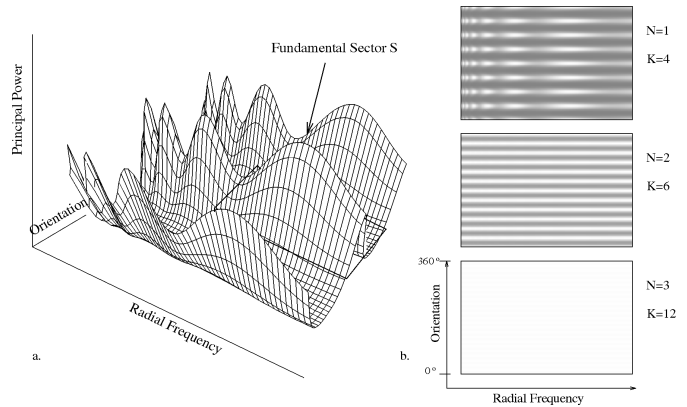

Fig. 4. (a) The principal power $F(\xi, \nu)$ for the 2D Gabor wavelet. It is periodic along the orientation axis, and exhibits dilating oscillation along the radial frequency axis. The fundamental sector S (enclosed by the black rectangle) is repeated in its dilated and translated versions to cover the whole frequency domain. (b) Three examples of $F(\xi, \nu)$ with different sampling densities are shown. The gray level indicates the value of the spectral power. Higher value is shown with lighter color. Generally, as the number of orientations and voices (i.e., frequency steps per octave) increases, the power spectrum becomes more uniform in absolute term.

The second residue term R captures the interference effect between the wavelets in a family due to their nonorthogonality. It is computed by summing all the cross products of each wavelet transform and the transform of its spectrally displaced (scaled and rotated) version, for all $(p, q) \in Z^2$ except at $(p = 0, q = 0)$,

$$R = | \frac{1}{b_o^2} \sum_{m\in Z, l\in Q} \sum_{(p,q)\in Z^2 \setminus \{(0,0)\}} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\xi d\nu \hat{f}(\xi,\nu) \overline{\hat{\psi}_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right)}$$

$$\overline{\hat{f}\left(\xi + 2\pi p a_o^{-m} b_o^{-1}, \nu + 2\pi q a_o^{-m} b_o^{-1}\right)} \hat{\psi}_{\theta_l}\left(a_o^m \xi + 2\pi p b_o^{-1}, a_o^m \nu + 2\pi q b_o^{-1}\right)| \quad (43)$$

$$\leq \frac{1}{b_o^2} \sum_{m \in Z, l \in Q} \sum_{(p,q) \in Z^2 \setminus \{0,0\}}$$

$$\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\xi d\nu \left| \hat{f}(\xi, \nu) \right|^2 \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right| \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi + 2\pi p b_o^{-1}, a_o^m \nu + 2\pi q b_o^{-1} \right) \right| \right]^{\frac{1}{2}}$$

$$\cdot \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\tilde{\xi} d\tilde{\nu} \left| \hat{f}(\tilde{\xi}, \tilde{\nu}) \right|^2 \left| \hat{\psi}_{\theta_l} \left(a_o^m \tilde{\xi} - 2\pi p b_o^{-1}, a_o^m \tilde{\nu} - 2\pi q b_o^{-1} \right) \right| \left| \hat{\psi}_{\theta_l} \left(a_o^m \tilde{\xi}, a_o^m \tilde{\nu} \right) \right| \right]^{\frac{1}{2}} \quad (44)$$

(with change of variables $\tilde{\xi} = \xi + 2\pi p b_o^{-1} a_o^m$, $\tilde{\nu} = \nu + 2\pi q b_o^{-1} a_o^{-m}$ in the second factor)

$$\leq \frac{1}{b_o^2} \sum_{(p,q) \in Z^2 \setminus \{0,0\}}$$

$$\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\xi d\nu \left| \hat{f}(\xi, \nu) \right|^2 \sum_{m,l} \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right| \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi + 2\pi p b_o^{-1}, a_o^m \nu + 2\pi q b_o^{-1} \right) \right| \right]^{\frac{1}{2}}$$

$$\cdot \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\tilde{\xi} d\tilde{\nu} \left| \hat{f}(\tilde{\xi}, \tilde{\nu}) \right|^2 \sum_{m,l} \left| \hat{\psi}_{\theta_l} \left(a_o^m \tilde{\xi}, a_o^m \tilde{\nu} \right) \right| \left| \hat{\psi}_{\theta_l} \left(a_o^m \tilde{\xi} - 2\pi p b_o^{-1}, a_o^m \tilde{\nu} - 2\pi q b_o^{-1} \right) \right| \right]^{\frac{1}{2}} \quad (45)$$

(use Cauchy-Schwarz on the sum over m and l)

$$\leq \frac{1}{b_o^2} \left\| f(\xi, \nu) \right\|^2 \sum_{(p,q) \in Z^2 \setminus \{0,0\}} \left[\beta \left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o} \right) \beta \left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o} \right) \right]^{1/2} \quad (46)$$

where

$$\beta(s,t) = \sup_{\xi, \nu} \sum_{m \in Z, l \in Q} \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right| \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi + s, a_o^m \nu + t \right) \right| \quad (47)$$

and $s = \frac{2\pi p}{b_o}$ and $t = \frac{2\pi q}{b_o}$.

Since the "cross-product" term,

$$\sum_{m \in Z, l \in Q} \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right| \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi + s, a_o^m \nu + t \right) \right|, \quad (48)$$

is also self-similar. Its supremum and infimum over the whole spectrum are also equal to its supremum and infimum over the fundamental sector S. Therefore, we can write,

$$\beta(s,t) = \sup_{\xi, \nu \in S} \sum_{m \in Z, l \in Q} \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right| \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi + s, a_o^m \nu + t \right) \right| \quad (49)$$

Thus, the infimum and supremum of the power output is given by

$$\inf_{f \in \mathcal{H}, f \neq 0} \left\| f \right\|^{-2} \sum_{m,n,k,l} \left| < f, \psi_{m,n,k,l} > \right|^2 \geq \frac{1}{b_o^2} \left\{ \inf_{\xi, \nu \in S} \sum_{m \in Z, l \in Q} \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right|^2 \right.$$

$$\left. - \sum_{(p,q) \in Z^2 \setminus \{0,0\}} \left[\beta \left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o} \right) \beta \left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o} \right) \right]^{1/2} \right\} \quad (50)$$

$$\sup_{f \in \mathcal{H}, f \neq 0} \left\| f \right\|^{-2} \sum_{m,n,k,l} \left| < f, \psi_{m,n,k,l} > \right|^2 \leq \frac{1}{b_o^2} \left\{ \sup_{\xi, \nu \in S} \sum_{m \in Z, l \in Q} \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right|^2 \right.$$

$$\left. + \sum_{(p,q) \in Z^2 \setminus \{0,0\}} \left[\beta \left(\frac{2\pi p}{b_o}, \frac{2\pi 1}{b_o} \right) \beta \left(-\frac{2\pi p}{b_o}, -\frac{2\pi 1}{b_o} \right) \right]^{1/2} \right\} \quad (51)$$

and the frame bounds A and B for 2D wavelets are given by,

$$A = \frac{1}{b_o^2} \left\{ \inf_{\xi, \nu \in S} \sum_{m \in Z, l \in Q} \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right|^2 - \right.$$

$$\left. \sum_{(p,q) \in Z^2 \setminus \{0,0\}} \left[\beta \left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o} \right) \beta \left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_0} \right) \right]^{1/2} \right\} \quad (52)$$

$$B = \frac{1}{b_o^2} \left\{ \sup_{\xi, \nu \in S} \sum_{m \in Z, l \in Q} \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right|^2 + \right.$$

$$\left. \sum_{(p,q) \in Z^2 \setminus \{0,0\}} \left[\beta \left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o} \right) \beta \left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o} \right) \right]^{1/2} \right\} \quad (53)$$

The Gabor wavelet is a complex-valued function. If only real signals f are represented and reconstructed by means of Gabor wavelet responses $< \psi_{m,n,k,l}, f >$, then the complex wavelet really consists of two wavelets, Re(ψ) and Im(ψ). Therefore, the frame bounds for 2D Gabor wavelets are given as follows,

$$A = \frac{1}{b_o^2} \left\{ \inf_{\xi, \nu \in S} \sum_{m \in Z} \sum_{l \in Q} \frac{1}{2} \left[\left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right|^2 + \left| \hat{\psi}_{\theta_l} \left(-a_o^m \xi, -a_o^m \nu \right) \right|^2 \right] - R \right\} \quad (54)$$

$$B = \frac{1}{b_o^2} \left\{ \sup_{\xi, \nu \in S} \sum_{m \in Z} \sum_{l \in Q} \frac{1}{2} \left[\left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) \right|^2 + \left| \hat{\psi}_{\theta_l} \left(-a_o^m \xi, -a_o^m \nu \right) \right|^2 \right] + R \right\} \quad (55)$$

where

$$R = \sum_{\varepsilon = +, -} \sum_{(p,q) \in Z^2 \setminus \{0,0\}} \left[\beta_\varepsilon \left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o} \right) \beta_\varepsilon \left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o} \right) \right]^{1/2} \quad (56)$$

and

$$\beta_\varepsilon(s,t) = \frac{1}{4} \sup_{\xi, \nu \in S} \sum_{m \in Z, l \in Q} \left| \hat{\psi}_{\theta_l} \left(a_o^m \xi, a_o^m \nu \right) + \varepsilon \hat{\psi}_{\theta_l} \left(-a_o^m \xi, -a_o^m \nu \right) \right|$$

$$\left| \hat{\psi}_{\theta_l} \left(a_o^m \xi + s, a_o^m \nu + t \right) + \varepsilon \hat{\psi}_{\theta_l} \left(-a_o^m \xi - s, -a_o^m \nu - t \right) \right| \quad (57)$$

When the frame bounds $A > 0$ and $B < \infty$ for a discrete family of Gabor wavelets, this family of wavelets is said to provide a complete representation.

5 FRAME BOUNDS FOR FRACTIONALLY DILATED 2D WAVELETS

In the above derivation, $a_o = 2$, the frequency space is sampled every octave. However, it is known that in the visual cortex, the frequency space is sampled every half or every third of an octave [32], [36]. What is the advantage of suboctave sampling? Grossmann, Kronland-Martinet, and Morlet [14] have found that when the frequency space is sampled suboctavely, the frame will become tighter, i.e., $A \to B$. They proposed to construct a frame based on "fractionally" dilated versions of a single 1D wavelet ψ:

$$\psi^\eta(x) = 2^{-\eta/N} \psi(2^{-\eta/N} x) \quad (58)$$

N is the number of sampling frequency steps per octave

and η is the index of the frequency step per octave. Within each octave, the scale of the wavelet is changed without the corresponding change in its spatial sampling intervals. Daubechies [1989] remarked that the ψ^η, constructed this way, do not have the same L^2-normalization, $\| \psi^\eta \|_2 = 2^{-\eta/2N} \| \psi \|_2$. This change in normalization compensates for the fact that the phase space lattice, which is a superposition of N lattices, is "denser" than the lattice whose translation interval is strictly proportional to the scale of the wavelet.

The 2D version of the fractionally dilated wavelets can be generalized to be

$$\psi^\eta_{\theta_l}(x,y) = 2^{-2\eta/N} \psi_{\theta_l}\left(2^{-\eta/N}x,\ 2^{-\eta/N}y\right), \quad \eta = 0, \dots N-1, \text{ (59)}$$

with its Fourier transform equal to

$$\hat{\psi}^\eta_{\theta_l}(\xi, \nu) = \hat{\psi}_{\theta_l}\left(2^{\eta/N}\xi,\ 2^{\nu/N}\nu\right), \quad \eta = 0, \dots N-1, \text{ (60)}$$

The general frame conditions for the fractionally dilated 2D wavelets are given below.

Define

$$m(a_o, \theta_o, N) = \inf_{\xi,\nu \in S} \sum_{\eta=0}^{N-1} \sum_{m \in Z, l \in Q} \left| \hat{\psi}^\eta_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right) \right|^2 \quad \text{(61)}$$

$$M(a_o, \theta_o, N) = \sup_{\xi,\nu \in S} \sum_{\eta=0}^{N-1} \sum_{m \in Z, l \in Q} \left| \hat{\psi}^\eta_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right) \right|^2 \quad \text{(62)}$$

$$\beta^\eta(s,t) = \sup_{\xi,\nu \in S} \sum_{m \in Z, l \in Q} \left| \hat{\psi}^\eta_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right) \right| \left| \hat{\psi}^\eta_{\theta_l}\left(a_o^m \xi + s, a_o^m \nu + t\right) \right| \quad \text{(63)}$$

where $\theta_o = \pi/K$, K is the number of sampling orientations, $Q = \{0, 1, 2 \dots K-1\}$; N is the number of voices or frequency steps per octave (i.e., $N = 4$ implies sampling at $2^{-1/4}$ intervals). S is the fundamental sector defined earlier.

Choose b_o and θ_o such that

$$\sum_{\eta=0}^{N-1} \sum_{(p,q) \in Z^2 \setminus \{(0,0)\}} \left[\beta^\eta\left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o}\right) \beta^\eta\left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o}\right) \right]^{1/2}$$

$$< m(a_o, \theta_o, N), \quad \text{(64)}$$

then

$$\{\psi^\eta_{m,n,k,l};\ m, n, k \in Z, l = 0, \dots, K-1, \eta = 0, \dots, N-1\}$$

constitute a frame with the following frame bounds,

$$A = \frac{1}{b_o^2} \left\{ m(a_o, \theta_o, N) - \sum_{\eta=0}^{N-1} \sum_{(p,q) \in Z^2 \setminus \{(0,0)\}} \left[\beta^\eta\left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o}\right) \beta^\eta\left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o}\right) \right]^{1/2} \right\}$$

$$\text{(65)}$$

$$B = \frac{1}{b_o^2} \left\{ M(a_o, \theta_o, N) + \sum_{\eta=0}^{N-1} \sum_{(p,q) \in Z^2 \setminus \{(0,0)\}} \left[\beta^\eta\left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_0}\right) \beta^\eta\left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o}\right) \right]^{1/2} \right\}.$$

$$\text{(66)}$$

The proof is a simple variant of the proof for the one-octave sampling case (see [5]).

When $A > 0$ and $B < \infty$, $\psi_{m,n,k,l}$ constitute a frame. This is valid only if the following conditions are true, with the definitions in (61), (62), and (63),

For b_o sufficiently small, $m(\psi, a_o, \theta_o, N) > 0$, $M(\psi, a_o, \theta_o, N) < \infty$, $\beta^\eta(s, t)$ decays at least as fast as $(1 + s^2 + t^2)^{-(1+\varepsilon)}$,

with $\epsilon > 0$, or put in another way,

$$\sup_{s,t \in R} \left[\left(1 + s^2 + t^2\right)^{(1+\varepsilon)} \beta(s,t) \right] = C_\varepsilon < \infty \quad \text{for some } \varepsilon > 0 \text{ (67)}$$

where s and t are defined as in $\beta^\eta(s,t)$.

The condition on β^η and the bounds are satisfied if $\left| \hat{\psi}^\eta(\xi, \nu) \right|$ decays exponentially as in our case and that $\hat{\psi}_{\theta_l}(a_o^m \xi, a_o^m \nu) = 0$ when $(\xi = 0, \nu = 0)$, for all m, l.

For 2D multivoice Gabor wavelets, the frame bounds are as follows,

$$A = \frac{1}{b_o^2} \left\{ \inf_{\xi,\nu \in S} \sum_{\eta=0}^{N-1} \sum_{m \in Z} \sum_{l \in Q} \frac{1}{2} \left[\left| \hat{\psi}^\eta_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right) \right|^2 + \left| \hat{\psi}^\eta_{\theta_l}\left(-a_o^m \xi, -a_o^m \nu\right) \right|^2 \right] - R \right\}$$

$$\text{(68)}$$

$$B = \frac{1}{b_o^2} \left\{ \sup_{\xi,\nu \in S} \sum_{\eta=0}^{N-1} \sum_{m \in Z} \sum_{l \in Q} \frac{1}{2} \left[\left| \hat{\psi}^\eta_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right) \right|^2 + \left| \hat{\psi}^\eta_{\theta_l}\left(-a_o^m \xi, -a_o^m \nu\right) \right|^2 \right] + R \right\}$$

$$\text{(69)}$$

where

$$R = \sum_{\varepsilon=+,-} \sum_{\eta=0}^{N-1} \sum_{(p,q) \in Z^2 \setminus \{(0,0)\}} \left[\beta^\eta_\varepsilon\left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o}\right) \beta^\eta_\varepsilon\left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o}\right) \right]^{1/2} \text{(70)}$$

and

$$\beta^\eta_\varepsilon(s,t) = \frac{1}{4} \sup_{\xi,\nu \in S} \sum_{m \in Z} \sum_{l=1}^{K} \left| \hat{\psi}^\eta_{\theta_l}\left(a_o^m \xi, a_o^m \nu\right) + \varepsilon \hat{\psi}^\eta_{\theta_l}\left(-a_o^m \xi, -a_o^m \nu\right) \right|$$

$$\left| \hat{\psi}^\eta_{\theta_l}\left(a_o^m \xi + s, a_o^m \nu + t\right) + \varepsilon \hat{\psi}^\eta_{\theta_l}\left(-a_o^m \xi - s, -a_o^m \nu - t\right) \right| \text{ (71)}$$

These frame bound equations for Gabor wavelets with suboctave, multi-orientation phase space sampling allow us to study the implications of the cortical sampling strategies.

6 FRAME BOUNDS FOR VARIOUS SAMPLING SCHEMES

Given ψ is admissible, i.e., it has reasonable decay in space and spatial frequency domains, with zero d.c. response ($\hat{\psi}(0, 0) = 0$), then there exists a range of a_o, θ_o, b_o for which a family of ψ constitute a frame. The lower bounds for these parameters are $a_o > 1$, $b_o > 0$, $\theta_o > 0$ and the upper bounds are given by the critical values a_o^c, θ_o^c, and b_o^c.

$$b_o^c = \inf\{b_o \text{ s.t.}$$

$$\sum_{\eta=0}^{N-1} \sum_{(p,q) \in Z^2 \setminus \{(0,0)\}} \left[\beta^\eta\left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o}\right) \beta^\eta\left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o}\right) \right] \geq m(\psi, a_o, \theta_o)\} \text{ (72)}$$

$$\theta_o^c = \inf\{\theta_o \text{ s.t.}$$

$$\sum_{\eta}^{N-1} \sum_{(p,q) \in Z^2 \setminus \{(0,0)\}} \left[\beta^\eta\left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o}\right) \beta^\eta\left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o}\right) \right] \geq m(\psi, a_o b_o)\} \text{ (73)}$$

$$a_o^c = \inf\{a_o \text{ s. t.}$$

$$\sum_{\eta=0}^{N-1} \sum_{(p,q)\in Z^2\setminus\{(0,0)\}} \left[\beta^\eta\left(\frac{2\pi p}{b_o}, \frac{2\pi q}{b_o}\right)\beta^\eta\left(-\frac{2\pi p}{b_o}, -\frac{2\pi q}{b_o}\right)\right] \geq m(\psi, \theta_o, b_o)\} \quad (74)$$

Because there is a line of zeroes in the frequency plane for each 2D Gabor wavelet (16),

$$\xi\xi_o + \nu\nu_o = 0 \quad (75)$$

at least two sampling orientations are required to cover the frequency domain, except at $(\xi = 0, \nu = 0)$. This is because wavelets of the same orientation have the same line of zeroes, since ν_o/ξ_o is fixed for each orientation. Therefore, the upper bound of the rotation step size θ_o^c, cannot exceed $\pi/2$.

The upper bound of the unit spatial interval b_o is 1 for the single voice wavelets. This can be deduced from the Nyquist sampling theorem: the maximum sampling intervals (max $\triangle x$) before aliasing occurs is $\lambda/2$, where λ is the wavelength of the filter. Since $\triangle x = a_o b_o$, where $a_o = \sigma = \lambda/2$, the upper bound of b_o is 1 for the one-octave sampling scheme. However, for fractionally dilated wavelets, the upper bound of b_o could be pushed above 1 because wavelets of the fractional scales are imposed on denser lattices.

To determine whether the 2D Gabor wavelets form a frame and how tight that frame is, we computed their frame bounds using the frame bound (68) and (69), the mother wavelet (18), and the fractional wavelet scaling (60). The frame bounds for two families of the 2D Gabor wavelets with bandwidths of 1.0 octave and 1.5 octaves
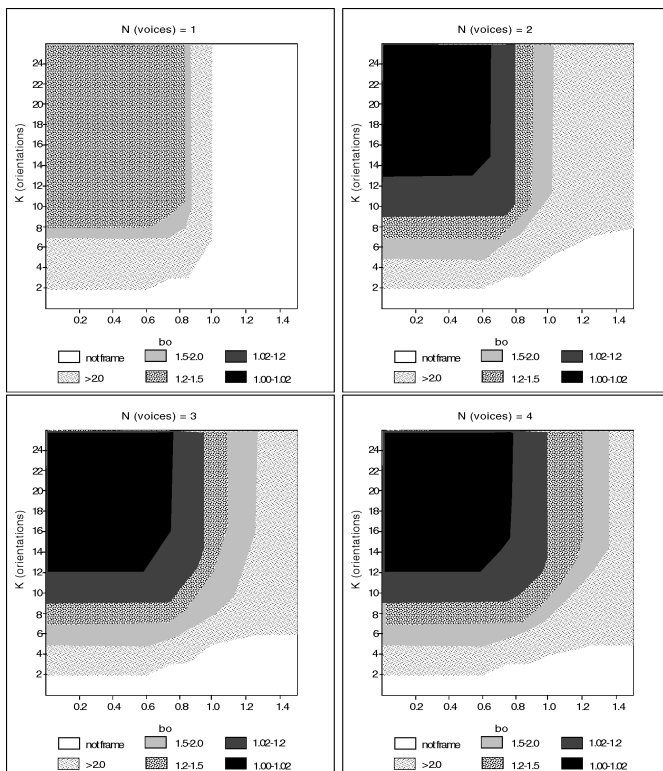

Fig. 5. The figures depict regions of tightness in the parametric space of sampling density. The frames become tighter with the increase in phase space sampling density.

TABLE 1
FRAME BOUNDS FOR THE 2D GABOR WAVELETS WITH 1.5 OCTAVE BANDWIDTH FOR DIFFERENT NUMBERS OF SAMPLING ORIENTATIONS (K), WITH THE NUMBER OF FREQUENCY STEPS PER OCTAVE N = 1 AND N = 3, AND THE UNIT SPATIAL SAMPLING INTERVAL $b_o = \Delta x/a_o = 0.8$

N = 1			
K	A	B	B/A
4	2.502	54.388	21.739
6	17.131	56.965	3.325
8	32.107	62.141	1.935
12	57.150	81.810	1.432
16	78.494	107.516	1.370
20	98.499	134.178	1.362

N = 3			
K	A	B	B/A
4	19.273	136.424	7.079
6	67.908	144.066	2.122
8	118.664	161.044	1.357
12	201.665	217.529	1.079
16	272.766	286.155	1.049
20	341.454	357.199	1.046

TABLE 2
FRAME BOUNDS FOR THE 2D GABOR WAVELETS WITH 1.5 OCTAVE BANDWIDTH FOR DIFFERENT NUMBERS OF SAMPLING ORIENTATIONS (K), WITH THE NUMBER OF FREQUENCY STEPS PER OCTAVE N = 1 AND 3, AND THE UNIT SPATIAL SAMPLING INTERVAL $b_o = 1$

N = 1			
K	A	B	B/A
4	-	-	-
6	-	-	-
8	5.114	55.205	10.795
12	18.168	70.766	3.895
16	26.255	92.792	3.534
20	33.033	115.880	3.508

N = 3			
K	A	B	B/A
4	-	-	-
6	25.440	110.223	4.333
8	57.460	121.554	2.115
12	107.492	160.792	1.496
16	146.924	210.786	1.435
20	184.056	263.082	1.429

The notation '-' implies not guaranteed to form a frame.

respectively were computed for different chosen sampling densities. The results are shown in Tables 1 to 5.

We also computed the frame bounds for a set of K and b_o for $N = 1$ to 4. Fig. 5 shows the frame regions of different degrees of tightness for each N. The tightness of the frame is measured by the ratio B/A and is classified into five degrees, each indicated by a different gray level in the figure. It is found that when $N = 1$, there is no tight frame. As N increases, the frame region expands and the frame becomes tighter and tighter.

Several general observations can be made from these results. First, the frame becomes tighter with the increase in the number of orientation, frequency and spatial sampling steps. Second, at each phase space sampling density, the 1.5 octave bandwidth Gabor wavelets provide a tighter frame than the 1.0 octave bandwidth Gabor wavelets.

Finally, with a phase space sampling density of 20 ori-

TABLE 3
FRAME BOUNDS FOR THE 2D GABOR WAVELETS WITH
1.5 OCTAVE BANDWIDTH FOR DIFFERENT UNIT SPATIAL
SAMPLING INTERVALS BO, WITH THE NUMBER OF
SAMPLING ORIENTATIONS K = 20 AND THE NUMBER OF
FREQUENCY STEPS PER OCTAVE N = 1 AND N = 3

N = 1

b_o	A	B	B/A
0.25	1087.822	1294.789	1.190
0.50	271.955	323.697	1.190
0.75	117.261	147.473	1.258
0.80	98.499	134.178	1.362
1.00	33.033	115.880	3.508
1.25	-	-	-

N = 3

b_o	A	B	B/A
0.25	3576.876	3577.334	1.000
0.50	894.219	894.333	1.000
0.75	393.801	401.112	1.019
0.80	341.454	357.200	1.046
1.00	184.056	263.082	1.429
1.25	65.683	220.456	3.357

TABLE 4
FRAME BOUNDS FOR THE 2D GABOR WAVELETS WITH
1.0 OCTAVE BANDWIDTH FOR DIFFERENT NUMBERS OF
SAMPLING ORIENTATIONS (K), WITH THE NUMBER OF
FREQUENCY STEPS PER OCTAVE N = 1 AND 3, AND THE
UNIT SPATIAL SAMPLING INTERVAL b_o = 0.8

N = 1

K	A	B	B/A
4	-	-	-
6	-	-	-
8	-	-	-
12	7.266	74.224	10.215
16	15.765	92.388	5.860
20	21.068	114.444	5.432

N = 3

K	A	B	B/A
4	-	-	-
6	2.727	125.798	46.137
8	34.536	129.686	3.755
12	92.217	152.605	1.655
16	134.362	192.045	1.429
20	170.087	237.924	1.398

TABLE 5
FRAME BOUNDS FOR 2D GABOR WAVELETS OF
1.0 OCTAVE BANDWIDTH FOR DIFFERENT UNIT SPATIAL
SAMPLING INTERVALS BO, WITH THE NUMBER OF SAMPLING
ORIENTATIONS K = 20 AND THE NUMBER OF
FREQUENCY STEPS PER OCTAVE N = 1 AND 3

N = 1

b_o	A	B	B/A
0.25	543.289	844.349	1.554
0.50	135.817	211.093	1.554
0.75	39.490	114.692	2.904
0.80	21.067	114.444	5.432
1.00	-	-	-
1.25	-	-	-

N = 3

b_o	A	B	B/A
0.25	2087.619	2090.410	1.001
0.50	521.899	522.608	1.001
0.75	210.553	253.672	1.205
0.80	170.087	237.924	1.399
1.00	57.254	203.873	3.561
1.25	1.550	165.572	106.846

These findings suggest that the cortical cells in the primary visual cortex not only provide a complete representation of the image within a three to five octave frequency band at each eccentricity, but also provide an almost tight frame representation through oversampling. In the next section, we will explore the implications of such a representation.

7 ILLUSTRATIVE EXPERIMENTAL RESULTS

The numerical results of the last section indicate that at the phase space sampling density of $K = 20$, $N = 3$, $b_o < 0.8$, the Gabor wavelets can form a very tight frame. In fact, even at $K = 8$, $N = 1$, $b_o = 0.8$, the frame is reasonably tight (B/A = 2) so that any image can be reconstructed approximately by a linear superposition of the Gabor wavelet elementary functions, as follows,

$$f_{\text{approx}} = \frac{2}{A + B} \sum_{m,n,k,l} < \psi_{m,n,k,l}, f > \psi_{m,n,k,l} \qquad (76)$$

without the use of a dual frame [5]. This is because as the redundancy increases, the frame becomes tighter. Both the lower and upper bounds of the spectral power of the transform (A, B in (68) and (69)) will increase and converge toward each other, rendering insignificant the interference due to the nonorthogonality of the Gabor wavelets (R in (70)). On the other hand, if the frame is not tight, the residues due to the nonorthogonality of the Gabor wavelets will introduce spurious effect to the reconstructed image as shown in Fig. 7a. In this case, coefficients from the projection onto the dual frame $\{\tilde{\psi}_{m,n,k,l}\}$ are needed if f is to be reconstructed by a linear superposition of the wavelet elementary functions as in (26). Dual frames are often difficult to obtain. But the projection coefficients to the dual frame can be obtained numerically by an iterative error minimization procedure [8].

In this iterative method, which we will use to evaluate empirically the completeness of a frame, the set of projec-

entations per cycle and 3 scale steps per octave, the 1.5 octave bandwidth Gabor wavelets family forms a tight frame when $\triangle x = 0.4\lambda$, ($b_o = 0.8$). There is physiological evidence suggesting that the phase space sampling density of the cortical cells in the brain's primary visual cortex is close to this tight frame sampling density. Hubel and Wiesel's [16] data suggested that at least 16 to 20 orientations are being sampled per hypercolumn, and the receptive fields of the cortical cells in one hypercolumn overlap half of the receptive fields of the cells in the adjacent hypercolumns. Campbell and Robson [4], De Valois et al. [11], [10], found that at each eccentricity the human visual system is sensitive to a spatial frequency range of three to five octaves. Pollen and Feldon [32] found that frequency is sampled at 1/2 octave intervals at cat visual cortex, while the data of Silverman et al. [36] suggested the sampling interval in spatial frequency is about 1/3 octave in monkeys.

tion coefficients to the dual frame $c_{m,n,k,l} = < f, \tilde{\psi}_{m,n,k,l} >$ can be obtained by minimizing the following cost function,

$$E = \left| f - \sum_{m,n,k,l} c_{m,n,k,l} \cdot \psi_{m,n,k,l} \right|^2 \tag{77}$$

where f is the original image, and ψ is a family of 2D Gabor filters as specified in (33). Then the image can be reconstructed by

$$f = \sum_{m,n,k,l} c_{m,n,k,l} \cdot \psi_{m,n,k,l} \tag{78}$$

However, the cortical sampling rate approximates a tight frame. Therefore, we should be able to reconstruct the image using the following *linear summation formula*,

$$f(x,y) = \frac{2}{A+B} \sum_{n,m,k,l} < f(x,y), \psi_{m,n,k,l}(x,y) > \psi_{m,n,k,l}(x,y) \tag{79}$$

This reconstruction formula implies the use of a pyramid sampling scheme, with spatial sampling interval $\Delta x = a_o^m b_o$ at each m level. We conducted a set of experiments to test these theoretical insights.

Fig. 6 presents the reconstruction results using both the linear summation method as well as the iterative method. In this reconstruction, only eight orientations and seven scales (one octave apart) are used ($K = 8$, $N = 1$, $b_o = 0.8$). The reasonable reconstruction result using the iterative method shows the completeness of the reconstruction (Fig. 6c), establishing that the wavelets indeed form a frame. The result using the linear summation method provides a good approximation for the original image also (Fig. 6b), suggesting that, even at this relatively coarse sampling density, the frame is reasonably tight. As the number of orientations and the number of voices increase toward the sampling rate of the visual cortex, the frame will theoretically become tighter, and the result using this linear summation method should be a better approximation of the original image.

a. Original image: Lena.   b. Simple summation.   c. Iterative method.

Fig. 6. Reconstruction of image 'Lena' (a) using the linear summation method (b) and the iterative error minimization method (c). The example is computed with a relatively tight frame ($K = 8$, $N = 1$, $b_o = 0.8$). Wavelets with seven scales are used in a pyramid scheme. Even at this coarse sampling density, the approximation to the original image is reasonable.

Fig. 7 presents the reconstruction results obtained when only three orientations are used ($K = 3$, $N = 1$, $b_o = 0.8$). The linear summation results indicate that the frame is very loose, and therefore linear summation is not a good approximation (Fig. 7a). The iterative method, however, demonstrates that even three orientations can provide a com-

plete representation of the image (Fig. 7b). On the other hand, when $b_o = 3.0$, the reconstruction ($K = 8$, $N = 1$, $b_o = 3.0$), even with the iterative method, is no longer complete because this family of Gabor wavelets no longer forms a frame (Fig. 7c).

a. Simple summation.   b. Iterative method.   c. Non-frame reconstruction.

Fig. 7. (a) Reconstruction of image 'Lena' using a complete by loose frame ($k = 3$, $N = 1$, $b_o = 0.8$) with the linear summation method shows that this method will introduce spurious results when the frame is not tight. However, the iterative error minimization method yields a reasonable reconstruction (b), indicating that the frame is complete. When a frame is not complete, even the iterative method cannot eliminate the spurious effect due to incompleteness. (c) shows an example of the reconstruction using an incomplete frame ($K = 8$, $N = 1$, $b_o = 3$).

The results above show that two or three orientations are sufficient for complete representation of the image. Why, then, does the brain construct a tight frame? Is there any advantage for sampling at such a large number of orientation and frequency steps?

Grossman et al. [14] have proposed that the continuous wavelet transform, through its redundancy, might provide a robust representation in the sense that the wavelet coefficients for a function can be computed and stored with low precision and still can be used to reconstruct or represent the original function with much higher precision. Could the redundancy provided by a tight frame reduce the "resolution burden" of the cortical cells? Suppose a neuron's firing rate is limited in resolution, probably to three bits or four bits. Can an eight bit resolution image be represented, for example, using cells with four bits in resolution?

To test this idea, we conducted another set of experiments in which the coefficients of the Gabor wavelets (i.e., the responses of the simple cells) are severely quantized. The phase space sampling density is the same as in the previous experiment, i.e., $K = 8$, $N = 1$, $b_o = 0.8$. Fig. 8 shows the reconstruction of an image from the 2D Gabor wavelet transform with its coefficients quantized to four bits, three bits, and two bits, respectively. The quantization is done independently on each scale, and the range of the response at each scale is stored and used subsequently in the reconstruction. We observe that even at such a coarse sampling density, the original image can be reasonably represented when the resolution of the coefficients or neuronal responses is reduced to four bits or even three bits (Fig. 8b). The degradation of the image representation is graceful with the decrease in the resolution of the representational elements (Fig. 8c).

8 CONCLUSION

In this paper, we have derived, based on physiological constraints and the wavelet theory, a family of 2D Gabor

a. 4 bit resolution.   b. 3 bit resolution.   c. 2 bit resolution.

Fig. 8. Reconstruction of image 'Lena' using the linear summation method with coefficients quantized to four bits, three bits, and two bits, respectively. The example is computed with a relatively tight frame ($K = 8$, $N = 1$, $b_o = 0.8$).

wavelets which model the receptive fields of the simple cells in the brain's primary visual cortex. By generalizing Daubechies's [5] frame criteria to 2D, we established the conditions under which a discrete class of continuous Gabor wavelets will provide complete representation of any image. We computed the frame bounds for the 2D Gabor wavelet families with 1.0 octave frequency bandwidth and 1.5 octave frequency bandwidth, and found that the latter family provides a tighter frame at identical phase space sampling densities. For the 1.5 octave family, the frame is tightened significantly when the number of voices (i.e., the number of wavelet frequencies per octave) is increased from one to two, but not much more with further increase in the number of voices (see Fig. 5). Even at one voice and eight orientations, the family forms a reasonably tight frame at $b_o = 0.8$ (i.e., spatial sampling interval $\triangle x = 0.4\lambda$). Hence, image reconstruction can be achieved by linear superposition of the Gabor wavelets using the wavelet projection coefficients. This finding suggests that such a parsimonious sampling of the phase space is sufficient for an approximation of the continuous wavelet transform—an insight that is important for computer vision applications and brain modeling which involve continuous wavelet transform formulations [20].

Physiological data suggest that the cortical sampling density is far greater than the parsimonious sampling density required for complete representation. In fact, it is dense enough to form a tight frame within a three to five octaves frequency band at each eccentricity, providing an over-complete and redundant representation of the retinal image within that frequency band. We have demonstrated there are at least two advantages to such a redundant representation: first, an image can be represented and easily reconstructed as a linear superposition of the receptive field structures of the simple cells weighed by their firing rates; second, high precision information can be computed and stored by a population of low-precision neurons. It has been suggested by information theoretic analysis that individual neurons can signal at most three to four bits of information in their instantaneous responses [39], [31], [37]. Earlier work on Gabor wavelets focused on the issues in image compression and compact representation [8]. Our work suggests that precision in resolution achieved through redundancy may be a more relevant issue in brain modeling. Furthermore, the high orientation and frequency sampling densities allow orientation and scale variables to be represented *precisely* and *explicitly* in the primary visual

cortex. As a result, only a small number of cells need to be activated at a time to signal precisely the structures in natural images [30]. Finally, the visual cortex is primarily concerned with extracting and computing perceptual information such as segmenting a scene [19], [22], [23], rather than re-presenting simply the retinal image. The simple cells, modeled by Gabor wavelets, with the redundancy provided by a tight frame, facilitate these computations by providing an ideal medium for representing surface texture and surface boundary with high resolution.

ACKNOWLEDGMENTS

REFERENCES

[1] Z. Li and J.J. Atick, "Towards a Theory of the Striate Cortex," *Neural Computation*, vol. 6, pp. 127-146, 1994.

[2] V. Bargmann, P. Butera, L. Girardello, and J.R. Klauder, "On the Completeness of Coherent States," *Rep. Math. Phys.*, vol. 2, pp. 221-228, 1989.

[3] A.C. Bovik, M. Clark, and W.S. Geisler, "Multichannel Texture Analysis Using Localized Spatial Filters," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 55-73, Jan. 1990.

[4] F.W. Campbell and J.G. Robson, "Application of Fourier Analysis to the Visibility of Gratings," *J. Physiology*, vol. 197, pp. 551-566, 1968.

[5] I. Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Trans. Information Theory*, vol. 36, no. 5, pp. 961-1004, 1990.

[6] J.G. Daugman, "Two-Dimensional Spectral Analysis of Cortical Receptive Field Profile," *Vision Research*, vol. 20, pp. 847-856, 1980.

[7] J.G. Daugman, "Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters," *J. Optical Soc. Amer.*, vol. 2, no. 7, pp. 1,160-1,169, 1985.

[8] J.G. Daugman, "Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 36, no. 7, pp. 1,169-1,179, 1988.

[9] J.G. Daugman, "Quadrature-Phase Simple-Cell Pairs Are Appropriately Described in Complex Analytic Form," *J. Optical Soc. Am.*, vol. 10, no. 7, pp. 375-377, 1993.

[10] R.L. De Valois, D.G. Albrecht, and L.G. Thorell, "Spatial Frequency Selectivity of Cells in Macaque Visual Cortex," *Vision Research*, vol. 22, pp. 545-559, 1982.

[11] R.L. De Valois and K.K. De Valois, *Spatial Vision*. New York: Oxford Univ. Press, 1988.

[12] R.J. Duffin and A.C. Schaeffer, "A Class of Nonharmonic Fourier Series," *Trans. Am. Math. Soc.*, vol. 72, pp. 341-366, 1952.

[13] D. Gabor, "Theory of Communication," *J. IEE*, vol. 93, pp. 429-459, 1946.

[14] A. Grossmann, R. Kronland-Martinet, and J. Morlet, "Reading and Understanding Continuous Wavelet Transforms," *Wavelets*, J.M. Combes, A. Grossmann, and P. Tchamitchian, eds., pp. 2-20. Berlin: Springer-Verlag, 1989.

[15] C.E. Heil and D.F. Walnut, "Continuous and Discrete Wavelet Transforms," *SIAM Review*, vol. 31, no. 4, pp. 628-666, 1989.

[16] D.H. Hubel and T.N. Wiesel, "Functional Architecture of Macaque Monkey Visual Cortex," *Proc. Royal Soc. B (London)*, vol. 198, pp. 1-59, 1978.

[17] J.P. Jones and L.A. Palmer, "An Evaluation of the Two-Dimensional Gabor Filter Model of Simple Receptive Fields in the Cat Striate Cortex," *J. Neurophysiology*, vol. 58, pp. 1,233-1,258, 1987.

[18] J.J. Kulikowski and P.O. Bishop, "Fourier Analysis and Spatial Representation in the Visual Cortex," *Experientia*, vol. 37, pp. 160-163, 1981.

[19] V.A.F. Lamme, "The Neurophysiology of Figure-Ground Segregation in Primary Visual Cortex," *J. Neuroscience*, vol. 10, no. 2, pp. 649-669, 1995.

[20] T.S. Lee, D. Mumford, and A.L. Yuille, "Texture Segmentation by Minimizing Vector-Valued Energy Functionals: The Couple-Membrane Model," *Lecture Notes in Computer Science, 588*, G. Sandini, ed., Computer Vision ECCV '92, pp. 165-173. Springer-Verlag, 1992.

[21] T.S. Lee, "A Bayesian Framework for Understanding the Texture Segmentation in the Primary Visual Cortex," *Vision Research*, vol. 35, no. 18, pp. 2,643-2,657, 1995.

[22] T.S. Lee, D. Mumford, and P.H. Schiller, "Neural Correlates of Boundary and Medial Axis Representations in Primate Visual Cortex." *Invest. Opth. Vis. Sci*, vol. 36, p. 477, 1995.

[23] T.S. Lee, V.A.F. Lamme, and D. Mumford, "The Role of V1 in Scene Segmentation and Shape Representation," unpublished manuscript.

[24] S. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 31, pp. 674-693, 1989.

[25] S. Marcelja, "Mathematical Description of the Responses of Simple Cortical Cells," *J. Optical Soc. Am.*, vol. 70, pp. 1,297-1,300, 1980.

[26] Y. Meyer, *Wavelets and Operators*. New York: Cambridge Univ. Press, 1992.

[27] J. Morlet, G. Arens, E. Fourgeau, and D. Giard, "Wave Propagation and Sampling Theory—Part II: Sampling Theory and Complex Waves," *Geophysics*, vol. 47, no. 2, pp. 222-236, 1982.

[28] J.A. Movshon and D.J. Tolhurst, "On the Response Linearity of Neurons in Cat Visual Cortex," *J. Physiology (London)*, vol. 249, pp. 56P-57P, 1975.

[29] D. Mumford, "Neuronal Architectures for Pattern-Theoretic Problems," *Large Scale Neuronal Theories of the Brain*, C. Koch, ed. MIT Press, 1993.

[30] B.A. Olshausen and D.J. Field, "Sparse Coding of Natural Images Produces Localized, Oriented, Bandpass Receptive Fields," unpublished manuscript.

[31] L.M. Optican and B.J. Richmond, "Temporal Encoding of Two-Dimensional Patterns by Single Units in Primate Inferior Temporal Cortex. III. Information Theoretic Analysis," *J. Neurophyisology*, vol. 57, pp. 162-178, 1987.

[32] D.A. Pollen and S.E. Feldon, "Spatial Periodicities of Periodic Complex Cells in the Visual Cortex Cluster at One-Half Octave Intervals," *Invest. Ophth. Vis. Sci*, pp. 429-434, Apr. 1979.

[33] D.A. Pollen and S.F. Ronner, "Phase Relationship Between Adjacent Simple Cells in the Visual Cortex," *Science*, vol. 212, pp. 1,409-1,411, 1981.

[34] M. Porat and Y. Zeevi, "The Generalized Gabor Scheme of Image Representation in Biological and Machine Vision," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 452-468, 1988.

[35] J. Segman and Y.Y. Zeevi, "A Wavelet-Type Approach to Image Analysis and Vision," Brussels: *NATO Book on Wavelets*, Aug. 1992.

[36] M.S. Silverman, D.H. Grosof, R.L. De Valois, and S.D. Elfar, "Spatial-Frequency Organization in Primate Striate Cortex," *Proc. Nat'l Academy of Science U.S.A.*, vol. 86, Jan. 1989.

[37] D.J. Tolhurst, "The Amount of Information Transmitted About Contrast by Neurons in the Cat's Visual Cortex," *Visual Neuroscience*, vol. 2, pp. 409-413, 1989.

[38] M.A. Webster and R.L. De Valois, "Relationship Between Spatial Frequency and Orientation Tuning of Striate Cortex Cells," *J. Optical Soc. Am.*, vol. A2, no. 7, July 1985.

[39] G. Werner and V.B. Mountcastle, "Neural Activity in Mechano-Receptive Cutaneous Afferents: Stimulus-Response Relations, Weber Functions, and Information Transmission," *J. Neurophysiology*, vol. 28, pp. 359-397, 1965.

Tai Sing Lee received the SB degree in 1986 and the PhD degree in engineering sciences in 1993, both from Harvard University, Cambridge, Massachusetts. He also graduated from the PhD program in medical engineering and medical physics from the Harvard-MIT Division of Health Science and Technology in 1993. From 1993 to 1996, he was a postdoctoral fellow in neurophysiology in the Department of Brain and Cognitive Sciences at MIT, and a postdoctoral fellow in biomedical physics in the Division of Applied Sciences, Harvard University.

Dr. Lee is currently an assistant professor in the Department of Computer Science at Carnegie Mellon University and in the CMU-University of Pittsburgh Center of Neural Basis of Cognition. His research involves the application of mathematical, computational, and neurophysiological techniques to uncover the neural mechanisms underlying visual perception in the brain.