Tutorial on Gabor Filters

Javier R. Movellan

1 The Temporal (1-D) Gabor Filter

Gabor filters can serve as excellent band-pass filters for unidimensional signals (e.g., speech). A complex Gabor filter is defined as the product of a Gaussian kernel times a complex sinusoid, i.e.

$$g(t) = ke^{j\theta}\ w(at)s(t) \tag{1}$$

where

$$w(t) = e^{-\pi t^2} \tag{2}$$

$$s(t) = e^{j(2\pi f_o t)} \tag{3}$$

$$e^{j\theta}s(t)e^{j(2\pi f_o t+\theta)} = \Big(\sin(2\pi f_o t+\theta), j\cos(2\pi f_o t+\theta)\Big) \tag{4}$$

Here k, θ, f_o are filter parameters. We can think of the complex Gabor filter as two out of phase filters continently allocated in the real and complex part of a complex function, the real part holds the filter

$$g_r(t) = w(t)\sin(2\pi f_o t+\theta) \tag{5}$$

and the imaginary part holds the filter

$$g_i(t) = w(t)\cos(2\pi f_o t+\theta) \tag{6}$$

1.1 Frequency Response

Taking the Fourier transform

$$\hat{g}(f) = ke^{j\theta}\ \int_{-\infty}^{\infty} e^{-j2\pi ft}w(at)s(t)\ dt \qquad = ke^{j\theta}\int_{-\infty}^{\infty} e^{-j2\pi(f-f_o)t}w(at)dt \tag{7}$$

$$= \frac{k}{a}e^{j\theta}\hat{w}(\frac{f-f_o}{a}) \tag{8}$$

where

$$\hat{w}(f) = w(f) = e^{-\pi f^2} \tag{9}$$

1.2 Gabor Energy Filters

The real and imaginary components of a complex Gabor filter are phase sensitive, i.e., as a consequence their response to a sinusoid is another sinusoid (see Figure 1.2). By getting the magnitude of the output (square root of the sum of squared real and imaginary outputs) we can get a response that phase insensitive and thus unmodulated positive response to a target sinusoid input (see Figure 1.2). In some cases it is useful to compute the overall output of the two out of phase filters. One common way of doing so is to add the squared output (the energy) of each filter, equivalently we can get the magnitude. This corresponds to the magnitude (more precisely the squared magnitude) of the complex Gabor filter output. In the frequency domain, the magnitude of the response to a particular frequency is simply the magnitude of the complex Fourier transform, i.e.

$$\|g(f)\| = \frac{k}{a}\hat{w}(\frac{f-f_o}{a}) \tag{10}$$

Note this is a Gaussian function centered at f_0 and with width proportional to a.

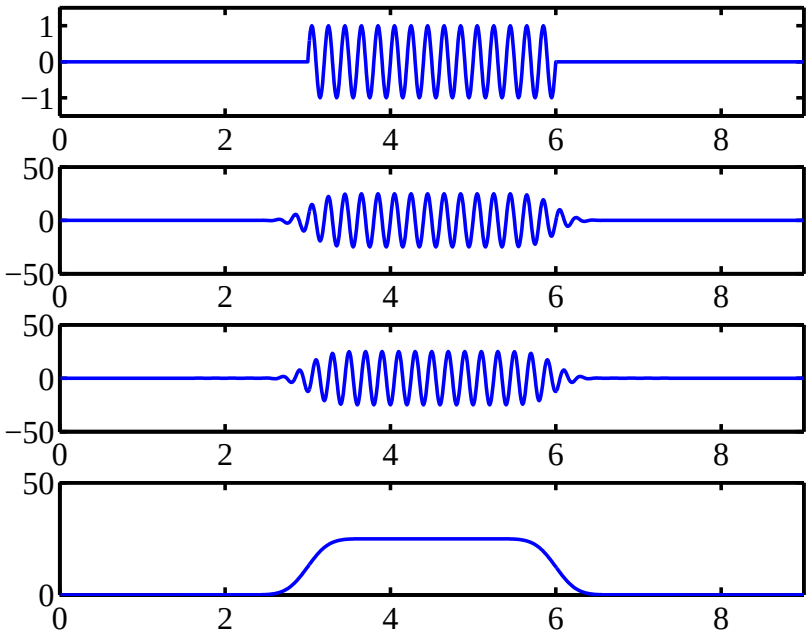

Figure 1: Top: An input signal. Second: Output of Gabor filter (cosine carrier). Third: Output of Gabor Filter in quadrature (sine carrier); Fourth: Output of Gabor Energy Filter

1.2.1 Bandwidth and Peak Response

Thus the peak filter response is at f_o. To get the half-magnitude bandwidth Δ_f note

$$\hat{w}(\frac{f - f_o}{a}) = e^{-\pi \frac{f-f_o}{a^2}} = 0.5 \tag{11}$$

Thus the half peak magnitude is achieved for

$$f - f_o \pm \sqrt{a^2 \log 2\pi} = 0.4697\, a \approx 0.5\, a \tag{12}$$

Thus the half-magnitude bandwidth is $(2)(0.4697)a)$ which is approximately equal to a. Thus a can be interpreted as the half-magnitude filter bandwidth.

1.3 Eliminating the DC response

Depending on the value of f_o and a the filter may have a large DC response. A popular approach to get a zero DC response is to subtract the output of a low-pass Gaussian filter,

$$h(t) = g(t) - c\, w(bt) = k e^{j\theta} w(at) s(t) - c\, w(bt) \tag{13}$$

Thus

$$\hat{h}(f) = \hat{g}(f) - \frac{c}{b} \hat{w}(\frac{f}{b}) \tag{14}$$

To get a zero DC response we need

$$\frac{c}{b}\hat{w}(0) = \hat{g}(0) \tag{15}$$

$$c = b\hat{g}(0) = b\frac{k}{a}\, e^{j\theta}\hat{w}(\frac{f_o}{a}) \tag{16}$$

where we used the fact that $\hat{w}(f_o) = \hat{w}(-f_o)$ Thus,

$$h(t) = g(t) - b\ \hat{g}(0) = ke^{j\theta}\Big(w(at)s(t) - \frac{b}{a}\hat{w}(\frac{f_o}{a})w(bt)\Big) \tag{17}$$

$$\hat{h}(f) = \frac{k}{a}e^{j\theta}\Big(\hat{w}(\frac{f-f_o}{a}) - \hat{w}(\frac{f_o}{a})\ \hat{w}(\frac{f}{b})\Big) \tag{18}$$

It is convenient, to let $b = a$, in which case

$$h(t) = ke^{j\theta}w(at)\Big(s(t) - \hat{w}(\frac{f_o}{a})\Big) \tag{19}$$

$$h(f) = \frac{k}{a}e^{j\theta}\Big(\hat{w}(\frac{f-f_o}{a}) - \hat{w}(\frac{f_o}{a})\ \hat{w}(\frac{f}{a})\Big) \tag{20}$$

2 The Spatial (2-D) Gabor Filter

Here is the formula of a complex Gabor function in space domain

$$g(x,y) = s(x,y)\ w_r(x,y) \tag{21}$$

where $s(x,y)$ is a complex sinusoid, known as the **carrier**, and $w_r(x,y)$ is a 2-D Gaussian-shaped function, known as the **envelope**.

2.1 The complex sinusoid carrier

The complex sinusoid is defined as follows,[1]

$$s(x,y) = \exp\left(j\ (2\pi(u_0\, x + v_0\, y) + P)\right) \tag{22}$$

where (u_0, v_0) and P define the spatial frequency and the phase of the sinusoid respectively.

We can think of this sinusoid as two separate real functions, conveniently allocated in the real and imaginary part of a complex function (see Figure 1).

[1] An offset constant parameter for $s(x,y)$ will be introduced later, to compensate the DC-component of this sinusoid. Refer to the appendix for detailed explanation.

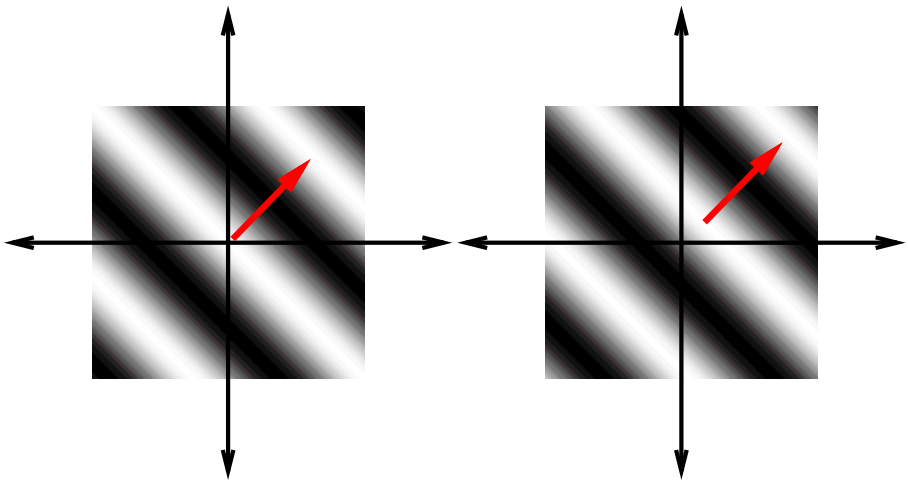

Figure 2: The real and imaginary parts of a complex sinusoid. The images are 128×128 pixels. The parameters are: $u_0 = v_0 = 1/80$ cycles/pixel, $P = 0$ deg.

The real part and the imaginary part of this sinusoid are

$$\begin{aligned} \mathrm{Re}\,(s(x,y)) &= \cos\left(2\pi(u_0\,x + v_0\,y) + P\right) \\ \mathrm{Im}\,(s(x,y)) &= \sin\left(2\pi(u_0\,x + v_0\,y) + P\right) \end{aligned} \tag{23}$$

The parameters u_0 and v_0 define the spatial frequency of the sinusoid in Cartesian coordinates. This spatial frequency can also be expressed in polar coordinates as magnitude F_0 and direction ω_0:

$$\begin{aligned} F_0 &= \sqrt{{u_0}^2 + {v_0}^2} \\ \omega_0 &= \tan^{-1}\left(\frac{v_0}{u_0}\right) \end{aligned} \tag{24}$$

i.e.

$$\begin{aligned} u_0 &= F_0\,\cos\,\omega_0 \\ v_0 &= F_0\,\sin\,\omega_0 \end{aligned} \tag{25}$$

Using this representation, the complex sinusoid is

$$s(x,y) = \exp\left(j\left(2\pi F_0\left(x\,\cos\,\omega_0 + y\,\sin\,\omega_0\right) + P\right)\right) \tag{26}$$

2.2 The Gaussian envelope

The Gaussian envelope looks as follows (see Figure 2):

$$w_r(x,y) = K\,\exp\left(-\pi\left(a^2\,{(x-x_0)_r}^2 + b^2\,{(y-y_0)_r}^2\right)\right) \tag{27}$$

where (x_0, y_0) is the peak of the function, a and b are scaling parameters[2] of the Gaussian, and the $_r$ subscript stands for a rotation operation[3] such that

$$\begin{aligned} (x-x_0)_r &= (x-x_0)\cos\theta + (y-y_0)\sin\theta \\ (y-y_0)_r &= -(x-x_0)\sin\theta + (y-y_0)\cos\theta \end{aligned} \tag{28}$$

[2]Note that the Gaussian gets smaller in the space domain, if a and b get larger.

[3]This rotation is clockwise, the inverse of the counterclockwise rotation of the ellipse.

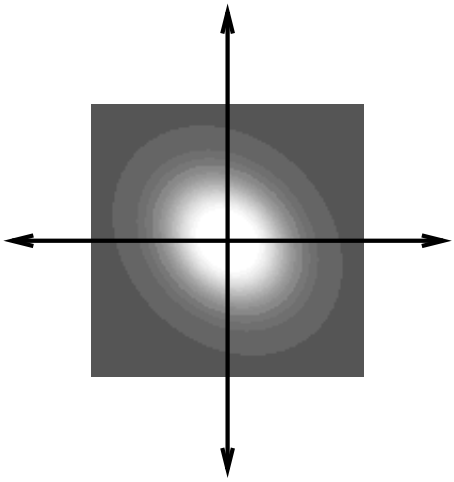

Figure 3: A Gaussian envelope. The image is 128×128 pixels. The parameters are as follows: $x_0 = y_0 = 0$. $a = 1/50$ pixels, $b = 1/40$ pixels, $\theta = -45$ deg.

2.3 The complex Gabor function

The complex Gabor function is defined by the following 9 parameters;

- K : Scales the magnitude of the Gaussian envelope.
- (a, b) : Scale the two axis of the Gaussian envelope.
- θ : Rotation angle of the Gaussian envelope.
- (x_0, y_0) : Location of the peak of the Gaussian envelope.
- (u_0, v_0) : Spatial frequencies of the sinusoid carrier in Cartesian coordinates.
  It can also be expressed in polar coordinates as (F_0, ω_0).
- P : Phase of the sinusoid carrier.

Each complex Gabor consists of two functions in quadrature (out of phase by 90 degrees), conveniently located in the real and imaginary parts of a complex function.

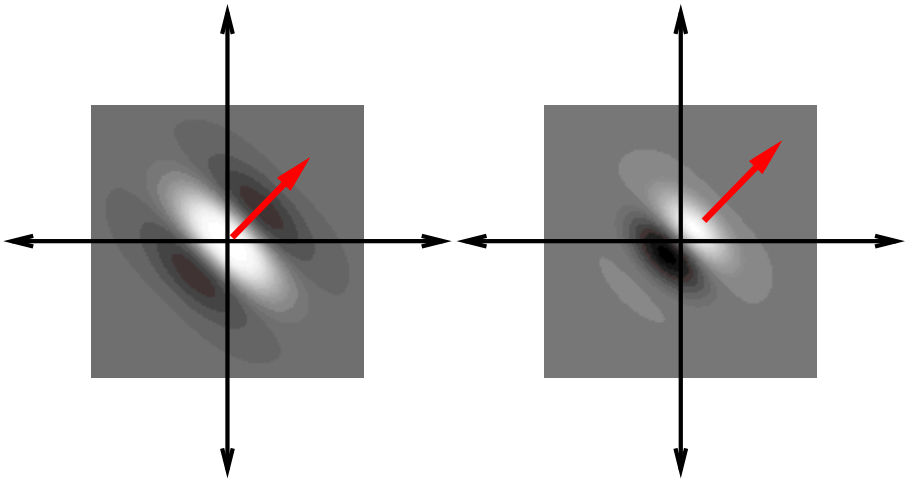

Figure 4: The real and imaginary parts of a complex Gabor function in space domain. The images are 128×128 pixels. The parameters are as follows: $x_0 = y_0 = 0$, $a = 1/50$ pixels, $b = 1/40$ pixels, $\theta = -45$ deg, $F_0 = \sqrt{2}/80$ cycles/pixel, $\omega_0 = 45$ deg, $P = 0$ deg.

Now we have the complex Gabor function in space domain[4] (see Figure 3):

$$\begin{aligned} g(x,y) = \quad & K \exp\left(-\pi\left(a^2 \left(x - x_0\right)_r^{\,2} + b^2 \left(y - y_0\right)_r^{\,2}\right)\right) \\ & \exp\left(j\left(2\pi(u_0\, x + v_0\, y) + P\right)\right) \end{aligned} \tag{29}$$

Or in polar coordinates,

$$\begin{aligned} g(x,y) = \quad & K \exp\left(-\pi\left(a^2 \left(x - x_0\right)_r^{\,2} + b^2 \left(y - y_0\right)_r^{\,2}\right)\right) \\ & \exp\left(j\left(2\pi F_0\left(x\, \cos\omega_0 + y\, \sin\omega_0\right) + P\right)\right) \end{aligned} \tag{30}$$

[4]In fact, there remains some DC component in this Gabor function. You have to compensate it to have the admissible Gabor function. Refer to the appendix.

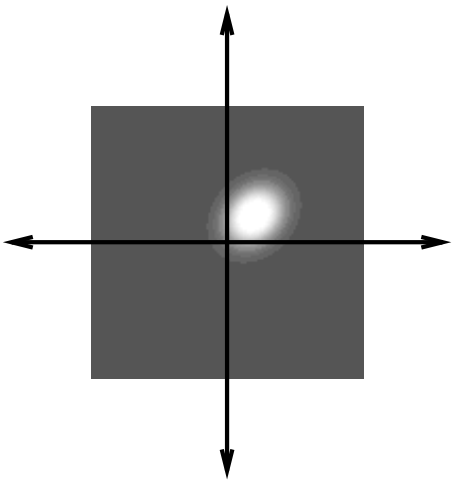

Figure 5: The Fourier transform of the Gabor filter. The peak response is at the spatial frequency of the complex sinusoid: $u_p = v_p = 1/80$ cycles/pixel. The parameters are as follows: $x_0 = y_0 = 0,\ a = 1/50$ pixels, $b = 1/40$ pixels, $\theta = -45$ deg, $F_0 = \sqrt{2}/80$ cycles/pixel, $\omega_0 = 45$ deg, $P = 0$ deg.

The 2-D Fourier transform of this Gabor[5] is as follows (see Figure 4):

$$\begin{aligned} \hat{g}(u,v) = \quad & \frac{K}{ab} \exp\left(j\left(-2\pi\left(x_0\left(u-u_0\right)+y_0\left(v-v_0\right)\right)+P\right)\right) \\ & \exp\left(-\pi\left(\frac{\left(u-u_0\right)_r{}^2}{a^2}+\frac{\left(v-v_0\right)_r{}^2}{b^2}\right)\right) \end{aligned} \tag{31}$$

Or in polar coordinates,

$$\begin{aligned} \text{Magnitude}\left(\hat{g}(u,v)\right) &= \frac{K}{ab} \exp\left(-\pi\left(\frac{\left(u-u_0\right)_r{}^2}{a^2}+\frac{\left(v-v_0\right)_r{}^2}{b^2}\right)\right) \\ \text{Phase}\left(\hat{g}(u,v)\right) &= -2\pi\left(x_0\left(u-u_0\right)+y_0\left(v-v_0\right)\right)+P \end{aligned} \tag{32}$$

[5]Refer to the appendix for detailed explanation.

3 Half-magnitude profile

The region of points, in frequency domain, with magnitude equal one-half the peak magnitude can be obtained as follows. Since the peak value is obtained for $(u, v) = (u_0, v_0)$, and the peak magnitude is K/ab, we just need to find the set of points (u, v) with magnitude $K/2ab$.

$$\frac{1}{2}\frac{K}{ab} = \frac{K}{ab}\exp\left(-\pi\left(\frac{(u-u_0)_r{}^2}{a^2} + \frac{(v-v_0)_r{}^2}{b^2}\right)\right) \tag{33}$$

or,

$$-\log 2 = -\pi\left(\frac{(u-u_0)_r{}^2}{a^2} + \frac{(v-v_0)_r{}^2}{b^2}\right) \tag{34}$$

or equivalently,

$$\begin{gathered}\left(\frac{(u-u_0)_r}{a\,C}\right)^2 + \left(\frac{(v-v_0)_r}{b\,C}\right)^2 = 1 \\ \text{where} \quad C = \sqrt{\frac{\log 2}{\pi}} = 0.46971864 \approx 0.5\end{gathered} \tag{35}$$

Equation 35 is an ellipse centered at (u_0, v_0) rotated with an angle θ with respect to the u axis. The main axis of the ellipse have length $2\,a\,C \approx a$ and $2\,b\,C \approx b$ respectively.

We will use the following convention: a is the length of the axis closer to ω_0, and b is the length of the axis perpendicular to the main axis[6] (See Figure 5).

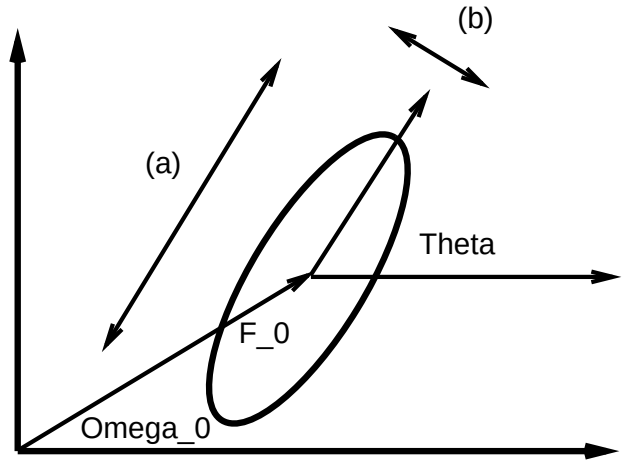

Figure 6: Parameters of the Gabor kernel as reflected in the half-magnitude elliptic profile. Note that this is a figure in frequency domain.

4 Half-magnitude frequency and orientation bandwidths

Frequency and orientation bandwidths of neurons are commonly measured in terms of the half-magnitude responses. Let u_0, v_0 the preferred spatial frequency of a neuron. In polar coordinates this spatial frequency can be expressed as F_0 and ω_0.

[6] More precisely a and b are 1.06 times the length of the respective axis

To find the half-magnitude frequency bandwidth, we probe the neuron with sinusoid images of orientation ω_0 and different spatial frequency magnitudes F. We increase F with respect to F_0 until the magnitude of the response is half the magnitude at (F_0, ω_0). Let's call that value $F_{\max}$. We then decrease F with respect to F_0 until the magnitude of the response is half the response at (F_0, ω_0). Call that $F_{\min}$. Half-magnitude frequency bandwidth is defined as follows:

$$\Delta F_{1/2} = F_{\max} - F_{\min} \tag{36}$$

or, when measured in octaves,[7]

$$\Delta F_{1/2} = \log_2 (F_{\max}/F_{\min}) \tag{37}$$

Half-magnitude orientation bandwidth is obtained following the same procedure but playing with the orientation ω instead of the frequency magnitude F.

$$\Delta \omega_{1/2} = \omega_{\max} - \omega_{\min} \tag{38}$$

In Gabor functions with $\theta_0 \approx \omega_0$ the frequency bandwidth can be obtained as follows (See Figure 6)

$$\Delta F_{1/2} = 2\, a\, C \approx a \tag{39}$$

and the orientation bandwidth can be approximated as follows (see Figure 6)

$$\Delta \omega_{1/2} \approx 2 \tan^{-1} \left(\frac{b\, C}{F_0} \right) \tag{40}$$

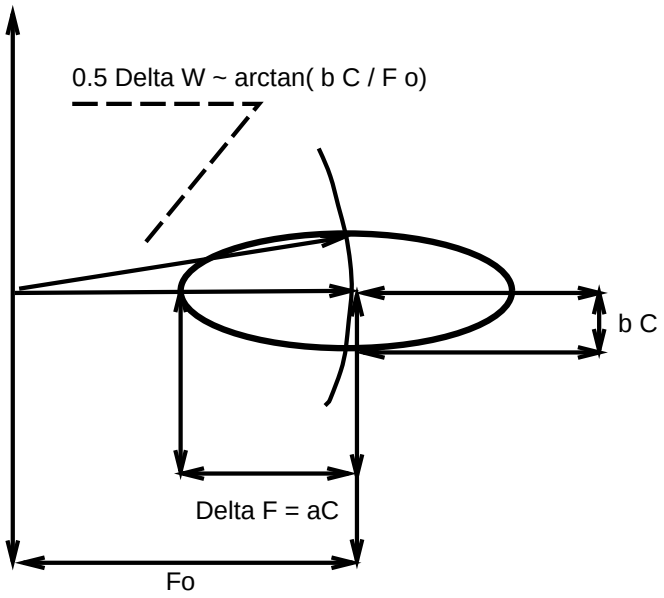

Figure 7: A half-magnitude profile and its relationship to the orientation and frequency bandwidths.

[7] Octave is a unit used for shown the ratio, as an index of 2. k octaves $= 2^k \times 100.0\%$

5 Effective spread and rms spread

The rms (which stands for root mean squares) length, rms width, and rms area of a 2-D function are defined in terms of their first and second moments:

The moments of a complex function $g(x, y)$ are defined by converting the function into a probability density (which must be always positive and integrates to 1.0) and then calculating the standard first and second moments.

A common way to achieve this is as follows:

From the function $g(x, y)$ we construct the following probability density

$$f(x, y) = \frac{1}{Z} |g(x, y)|^2 \tag{41}$$

where $|g(x, y)|^2$ is the squared magnitude of the signal, which is always positive, and Z guarantees that $f(x, y)$ integrates to 1.0, *i.e.*

$$Z = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |g(x, y)|^2 dxdy \tag{42}$$

Once we have defined a probability density function, the standard statistical measures of location and scale follow.

$$\mu_X = E_X(x) = \int \int f(x, y)\, x\, dxdy \tag{43}$$

$$\sigma_X^2 = E_X\left((x - \mu_X)^2\right) = \int \int f(x, y)\, (x - \mu_X)^2\, dxdy \tag{44}$$

with similar equations for μ_Y and σ_Y^2.

$$\mu_Y = E_Y(y) = \int \int f(x, y)\, y\, dxdy \tag{45}$$

$$\sigma_Y^2 = E_Y\left((y - \mu_Y)^2\right) = \int \int f(x, y)\, (y - \mu_Y)^2\, dxdy \tag{46}$$

And,

$$\sigma_{XY} = E_{XY}\left((x - \mu_X)\,(y - \mu_Y)\right) = \int \int f(x, y)\, (x - \mu_X)\, (y - \mu_Y)\, dxdy \tag{47}$$

The rms width and length are defined as the σ_X and σ_Y of a rotated version of $f(x, y)$ so that the covariance σ_{XY} of the rotated distribution be zero.

Let X_r, Y_r represent the rotated variables for which the covariance is zero, the rms length and width are

$$\Delta X_{\text{rms}} = \sqrt{\sigma_{X_r}^2} \tag{48}$$

$$\Delta Y_{\text{rms}} = \sqrt{\sigma_{Y_r}^2} \tag{49}$$

Similar definitions can be obtained also in the frequency domain, by working with the Fourier transform of the original complex function.

$$\Delta U_{\text{rms}} = \sqrt{\sigma_{U_r}^2} \tag{50}$$

$$\Delta V_{\text{rms}} = \sqrt{\sigma_{V_r}^2} \tag{51}$$

The rms area in the space and frequency domains are defined as follows:

$$\text{Area}\,(XY)_{\text{rms}} = (\Delta X_{\text{rms}})\,(\Delta Y_{\text{rms}}) \tag{52}$$

$$\text{Area}\,(UV)_{\text{rms}} = (\Delta U_{\text{rms}})\,(\Delta V_{\text{rms}}) \tag{53}$$

Some papers work with what are known as effective length, width and areas. They are simply the rms measures multiplied by $\sqrt{2\pi}$

$$\Delta X_{\text{eff}} = \sqrt{2\pi}\,\Delta X_{\text{rms}} \tag{54}$$

and so on.

It can be shown that the following relationships hold on any 2D function with finite moments

$$(\Delta X_{\text{rms}})\,(\Delta U_{\text{rms}}) \geq \frac{1}{4\pi} \tag{55}$$

$$(\Delta Y_{\text{rms}})\,(\Delta V_{\text{rms}}) \geq \frac{1}{4\pi} \tag{56}$$

and

$$\text{Area}\,(XY)_{\text{rms}}\ \text{Area}\,(UV)_{\text{rms}} \geq \frac{1}{16\pi^2} \tag{57}$$

It is easy to verify that the Gabor complex function achieves the lower limits of the uncertainty relations. For a given area in the space domain it provides the maximum possible resolution in the frequency domain, and vice-versa.

It can be shown that the rms width and lengths of Gabor functions are as follows:

$$\Delta U_{\text{rms}} = \frac{a}{2\sqrt{\pi}} \tag{58}$$

$$\Delta V_{\text{rms}} = \frac{b}{2\sqrt{\pi}} \tag{59}$$

Tow see why, simply consider that the probability density associated with the Gabor function $f(x,y) = \frac{1}{Z}\,|g(x,y)|^2$ is Gaussian

$$|\hat{g}(u,v)|^2 = \exp(-2\pi(\frac{u^2}{a^2} + \frac{v^2}{b^2})) \tag{60}$$

with variances equal to ΔU^2_{rms} and ΔV^2_{rms}.

Moreover, from the uncertainty relations,

$$\Delta X_{\text{rms}} = \frac{1}{2a\sqrt{\pi}} \tag{61}$$

$$\Delta Y_{\text{rms}} = \frac{1}{2b\sqrt{\pi}} \tag{62}$$

6 Gabor functions as models of simple cell receptive fields

Jones and Palmer (1987) showed that the real part of complex Gabor functions fit very well the receptive field weight functions found in simple cells in cat striate cortex.

Here are some useful pieces information for designing biologically inspired Gabor filters.

- To a first approximation the orientation of the Gaussian envelop ω_0 can be modeled as being equivalent to the orientation of the carrier.[8] $\theta_0 = \omega_0$. The actual absolute deviations between θ_0 and ω_0 have a Median of about 10 degrees (see Jones and Palmer, 1987, p. 1249).
- In macaque V1, most cells have a half magnitude spatial frequency bandwidth between 1 and 1.5 octaves. The median is about 1.4 octaves (see De Valois et al., 1982a, p. 551).
- In macaque V1, the range of half-magnitude orientation bandwidths among cells is very large: From 8 degrees to the most narrowly tuned. At the other end there were cells with no orientation selectivity at all. (see De Valois et al., 1982b, p. 535 and 541) reports the following statistics for the orientation bandwidth: mean = 65 degrees, median = 42 degrees, mode = 30 degrees. However they point out that others have reported significantly larger numbers. For example () reports a 71 % from max median bandwidth of 38.5 degrees. This would correspond to a median half magnitude bandwidth of 66 degrees. The median bandwidth of simple cells in the cat is a bit smaller than in the macaque, with a typical median half magnitude orientation bandwidth of 30 degrees (see De Valois et al., 1982b, p. 535 and 541).
- In macaque V1 the peak frequencies range from as low as 0.5 cycles per degree of visual angle, to as large as 15 cycles per degree of visual angle. Mean values are 2.7 cycles per degree for cells mapping into the parafoveal and 4.25 cycles per degree for cells mapping into the fovea.
- The spatial frequency bandwidth (in octaves) tends to be a bit larger for cells with low peak frequency than for cells with large peak frequency. For example, the median half magnitude bandwidth of cells tuned to frequencies higher than 5 cycles/degree is 1.2 octaves, whereas the median for cells tuned to frequencies smaller than 2 cycles/degree is 1.7 octaves. (see De Valois et al., 1982a, p. 552).
- Orientation selective simple cells in V1 show minimum response at about 30 to 40 degrees away from the optimal orientation, not at 90 degrees away from the optimal orientation (see De Valois et al., 1982b, p. 539).
- The spiking rate of simple cells neurons in macaque V1 is between close to 0 Hz, at rest, to about 120 Hz, when maximally excited (see De Valois et al., 1982a, p. 547).
- In the area mapping the fovea, there are more kernels oriented vertically and horizontally than oriented diagonally (about 3 to 2). (see De Valois et al., 1982b, p. 537).
- Pairs of adjacent simple cells in the visual cortex of the cat are in quadrature (Pollen and Ronner, 1981). We can then put these two cells in the real and imaginary parts of a complex function and treat them as a complex Gabor receptive field.

7 Gabor functions for spatial frequency filtering

Consider a massive set of simple cell neurons with Gabor kernel functions with equal parameters except for the location parameters (x_0, y_0). Let all these neurons be distributed uniformly about the foveal field. Each point in the foveal field contains

[8] domain. Note that the long axis in the frequency domain becomes the short axis in the space domain. Don't get confused!

at least two neurons in quadrature. We can model the operation of such a set of neurons as a convolution operation (assuming a continuous and uniform distribution of filters in all the foveal locations). Since convolution in space domain is product in frequency domain, the set of Gabor functions work as bandpass frequency filters of the foveal image. The peak frequency is controlled by the spatial frequency of the sinusoid carrier (u_0, v_0). The half-magnitude region is controlled by the rotation θ and scale parameters a, b, of the Gaussian envelope.

8 Energy filtering

A quadrature pair (or a Hilbert Transform pair) is a set of two linear operators with the same amplitude response but phase responses shifted by 90 degrees. Strictly speaking sine and cosine Gabor operators are not quadrature pairs because cosine phase Gabors have some DC response, whereas sine gabors do not. However, one can have quadrature Gabor pairs that look very much like sine/cosine pairs. Thus the sine and cosine Gabor pair is commonly refered to as a quadrature pair.

A system that sums the square of the outputs of a quadrature pair is called an energy mechanism (Adelson and Bergen, 1985). Energy mechanisms have unmodulated responses to drifting sinusoids.

Complex cells in V1 are commonly modeled as energy mechanisms since they are unmodulated by drifting sinusoids. Simple cells respond to a drifting sinusoid with a half-wave rectified analog of the signal, suggesting that the cells are linear up to rectification. Complex cells respond to a drifting sinusoid in an unmodulated way, as a maintained discharge. Movshon et al. (1987) showed that complex receptive fields are composed of subunits. The subunits of model complex cells are model simple cells with identical amplitude response. Emerson et al. (1992) have shown that behavior of complex cell to stimuli made of pairs of bars flashed in sequence is consist with an energy mechanism.

9 Contrast Normalization

Morrone et al. (1982) have shown that stimuli presented at orientations orthogonal to the optimal orientation inhibit simple cells activity. (De Valois et al., 1982a) have shown similar inhibitory effect between frequency bands. These inhibitory effects may play a serve as a gain control (or contrast normalization) mechanism. Heeger (1991) proposes the following model of gain control in complex cells: The amplitude response of each energy mechanism is divided by the total energy at all orientations and nearby spatial frequencies:

$$\bar{E_i} = \frac{E_i}{\kappa + \sum_j E_j} \tag{63}$$

where κ is a positive constant to avoid zero denominators.

10 Functional Interpretations

Section in preparation:

- minimizes number of neurons needed to achieve a desired frequency resolution.
- spatially and frequency localized.

- matched to "logons" likely to occur in images.
- for natural images the Gabor representation is more sparse than the δ (pixel) representation and than the DOG representation.

11 Constructing an idealized V1

Here we propose a way to construct a biologically inspired Gabor filter bank.

- The orientation of the complex sinusoid carrier and the Gaussian envelope are the same: $\omega_0 = \theta$. This is just an approximation. The actual median absolute deviation between ω_0 and θ is about 10 degrees.
- We will assume that the he half-magnitude frequency bandwidth, when measured in octaves, is constant and equal to 1.4 cycles per degree for all neurons. This is just an approximation. We know that neurons tuned to low spatial frequencies have larger bandwidth (median 1.7) and neurons tuned to high spatial frequencies have smaller bandwidth (median 1.2). In addition there is a significant range in bandwidths (bulk of the neurons have bandwidths between 1 and 1.5 octaves) that will not be addressed by the proposed model.
- We will assume that the half-magnitude orientation bandwidth is constant and equal to 40 degrees for all neurons (median value reported by (De Valois et al., 1982b). This is just an approximation since the actual range observed in simple cells is very large, going from 10 degrees to no orientation selectivity at all. Given this wide range in the distribution it is not surprising that other median bandwidth values have been reported in the literature, ranging from a reported median of about 30 degrees to a reported median of about 60 degrees (De Valois et al., 1982b).
- From the three constrained above, we will soon derive that all the filter kernels shall have a aspect ratio of about 1.24, i.e., $a/b \approx 1.24$.
- In addition, to facilitate the design we will design our filter bank so that the half-magnitude contour of a frequency band coincides with the lower contour of the next frequency band.

From these assumptions above, we can derive the relationship between the parameters F_0, a and b.

From equations 37 and 39, we know that the frequency bandwidth in octaves is

$$\begin{gathered} \Delta F_{1/2} = \log_2 \frac{F_0 + a\,C}{F_0 - a\,C} \\ \text{where} \quad C = \sqrt{\frac{\log 2}{\pi}} = 0.46971864 \approx 0.5 \end{gathered} \tag{64}$$

Thus,

$$a = F_0 \frac{K_a}{C} \tag{65}$$

where

$$K_a = \frac{2^{\Delta F} - 1}{2^{\Delta F} + 1} \tag{66}$$

With respect to the orientation bandwidth, equation 40 tells us that

$$\tan\left(\frac{1}{2}\Delta\omega\right) = \frac{b\,C}{F_0} \tag{67}$$

Thus,

$$b = F_0 \frac{K_b}{C} \tag{68}$$

where

$$K_b = \tan\left(\frac{1}{2}\Delta\omega\right) \tag{69}$$

Therefore, in this model the aspect ratio of a and b is constant:

$$\lambda = \frac{a}{b} = \frac{K_a}{K_b} \tag{70}$$

Moreover, from equation 39

$$\frac{1}{2}\Delta F = a\,C = F_0\,K_a \tag{71}$$

We can now locate our frequency peaks such that the upper half-magnitude contour of one channel coincides with the lower half-magnitude contour of the the next channel.

Let μ_i signify the peak frequency of the i^{th} band,
We know $F_{\max}$ for band i is

$$F^i_{\max} = \mu_i + \frac{1}{2}\Delta F_i = \mu_i + \mu_i\,K_a = \mu_i\,(1 + K_a) \tag{72}$$

and $F_{\min}$ for band $i+1$ is

$$F^{i+1}_{\max} = \mu_{i+1} - \frac{1}{2}\Delta F_{i+1} = \mu_{i+1} - \mu_{i+1}\,K_a = \mu_{i+1}\,(1 - K_a) \tag{73}$$

We want these two values to coincide, therefore

$$\mu_{i+1} = \mu_i \frac{1 + K_a}{1 - K_a} \tag{74}$$

Thus, the peak frequencies follow a geometric series

$$\mu_i = \mu_1 R^{i-1} \tag{75}$$

where

$$R = \frac{1 + K_a}{1 - K_a} \tag{76}$$

11.1 Example

If we use the standard values for simple cell median of the half magnitude bandwidths from macaque striate cortex:

- $\Delta F = 1.4$ octaves.
- $\Delta\omega = 40$ degrees.

Then,

$$K_a = \frac{2^{\Delta F} - 1}{2^{\Delta F} + 1} = 0.45040 \tag{77}$$

$$K_b = \tan\left(\frac{1}{2}\Delta\omega\right) = 0.36397 \tag{78}$$

$$a = \mu_i \frac{K_a}{C} = 0.9589\,\mu_i \tag{79}$$

$$\lambda = \frac{a}{b} = \frac{K_a}{K_b} = 1.23746 \tag{80}$$

$$b = \frac{a}{\lambda} = 1.1866\,\mu_i \tag{81}$$

$$R = \frac{1 + K_a}{1 - K_a} = 2.6390 \tag{82}$$

Suppose we want three frequency bands and we want the F_0 of the third band to be 0.25. Then,

$$\mu_1 = \frac{0.25}{2.6390^2} = 0.03589 \tag{83}$$

and

$$\frac{1}{2}\Delta F_1 = K_a\,\mu_1 = 0.01617 \tag{84}$$

Thus, the half magnitude interval[9] is $(0.01973, 0.05207)$

The second band peaks at

$$\mu_2 = \mu_1\,R = 0.09473 \tag{85}$$

and

$$\frac{1}{2}\Delta F_2 = K_a\,\mu_2 = 0.04267 \tag{86}$$

Thus, the half magnitude interval is $(0.05207, 0.1374)$

Finally, the third band peaks at

$$\mu_3 = \mu_2\,R = 0.2500 \tag{87}$$

and

$$\frac{1}{2}\Delta F_3 = K_a\,\mu_3 = 0.1126 \tag{88}$$

Thus, the half magnitude interval is $(0.1374, 0.3626)$

These three Gabors cover the frequency bands of $(0.01973, 0.3626)$

12 Appendix

12.1 Fourier transform of a Gaussian function

The Fourier transform of the simple 1-D Gaussian is

$$\begin{aligned}
&\int_{-\infty}^{\infty} \exp(-\pi x^2)\,\exp(-2\pi j f x)\,dx \\
&\quad = \int_{-\infty}^{\infty} \exp\left(-\pi(x + jf)^2 - \pi f^2\right)\,dx \\
&\quad = \exp\left(-\pi f^2\right)\int_{-\infty}^{\infty} \exp\left(-\pi x'^2\right)\,dx' \qquad (x' \equiv x + jf) \\
&\quad = \exp\left(-\pi f^2\right)
\end{aligned} \tag{89}$$

[9] The half magnitude interval here is the frequency coverage of that Gabor in terms of half-magnitude profile: $\left(\mu_i - \frac{1}{2}\Delta F_i, \mu_i + \frac{1}{2}\Delta F_i\right)$

In the same way, the Fourier transform of the simple 2-D Gaussian is

$$\begin{aligned}
&\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\pi(x^2+y^2)\right)\, \exp(-2\pi jux)\, \exp(-2\pi jvy)\, dxdy \\
&= \int_{-\infty}^{\infty} \exp(-\pi x^2)\, \exp(-2\pi jux)\, dx \int_{-\infty}^{\infty} \exp(-\pi y^2)\, \exp(-2\pi jvy)\, dy \\
&= \exp(-\pi u^2)\, \exp(-\pi v^2) \\
&= \exp\left(-\pi(u^2+v^2)\right)
\end{aligned} \tag{90}$$

and so on. More generally,

$$\int_{-\infty}^{\infty} \exp\left(-\pi\, x^T x\right)\, \exp\left(-2\pi j\, u^T x\right)\, dx = \exp\left(-\pi\, u^T u\right) \tag{91}$$

That is, the Fourier transform of an N-dimensional Gaussian is also an N-dimensional Gaussian.

12.2 Fourier transform of the Gabor function

Given a Gaussian envelope and sinusoid carrier:

$$w(x) = \exp(-\pi\, x^T x) \tag{92}$$

$$s(x) = \exp(j2\pi\, u_o^T x) \tag{93}$$

We define a Gabor function as follows

$$g(x) = K\, \exp(jP)\, w(A(x-x_o))\, s(x) \tag{94}$$

where K, P, A, u_o and x_o are function parameters. The Fourier transform of this function is as follows

$$\hat{g}(u) = \int_{-\infty}^{\infty} g(x)\, \exp(-2\pi j\, u^T x)\, dx \tag{95}$$

$$= K\, \exp(jP) \int_{-\infty}^{\infty} w(A(x-x_0))\, \exp(-2\pi j\, (u-u_0)^T x)\, dx \tag{96}$$

Letting $\tilde{x} = A(x-x_o)$ we get $x = A^{-1}\tilde{x} + x_o$, and[10] $d\tilde{x} = Adx$ and therefore

$$\hat{g}(u) = \frac{K}{\|A\|}\, \exp(jP) \int_{-\infty}^{\infty} w(\tilde{x})\, \exp(-2\pi j\, (u-u_0)^T (A^{-1}\tilde{x} - x_o))\, d\tilde{x} \tag{97}$$

$$= \frac{K}{\|A\|}\, \exp(jP)\, \exp((u-u_o)^T\, x_o) \tag{98}$$

$$\int_{-\infty}^{\infty} w(\tilde{x})\, \exp(-2\pi j\, (A^{-T}(u-u_0))^T \tilde{x})\, d\tilde{x} \tag{99}$$

Thus

$$\hat{g}(u) = \frac{K}{\|A\|}\, \exp(jP)\, \exp(-j2\pi\, (u-u_o)^T\, x_o)\, w(A^{-T}(u-u_o)) \tag{100}$$

where we used the fact that $\hat{w}(\cdot) = w(\cdot)$.

[10] Note the dx symbol in the integral stands for the product $dx_1 dx_2$

For the class of Gabor functions studied in the main section of this document we let $A = DV$, where D is a diagonal matrix and V is a rotation matrix such that

$$D = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \qquad V = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \tag{101}$$

Thus, since V is a rotation

$$A^{-1} = D^{-1}V^{-T} \tag{102}$$
$$A^{-T} = VD^{-1} = D^{-1}V \tag{103}$$
$$\|A\| = \|D\|\|V\| = ab \tag{104}$$

and therefore,

$$g(x,y) = K \exp\left(-\pi\left(a^2 {(x-x_0)_r}^2 + b^2 {(y-y_0)_r}^2\right)\right) \exp\left(j\left(2\pi\left(u_0x_0 + v_0y_0\right) + P\right)\right) \tag{105}$$

And its Fourier transform is:

$$\hat{g}(u,v) = \frac{K}{ab}\exp(jP)\,\exp(-2j\pi(x_0(u-u_0) + y_0(v-v_0))) \tag{106}$$
$$\exp(-\pi(\frac{(u-u_0)_r^2}{a^2} + \frac{(v-v_0)_r^2}{b^2}) \tag{107}$$

where

$$(x-x_0)_r = (x-x_0)\cos\theta + (y-y_0)\sin\theta \tag{108}$$
$$(y-y_0)_r = -(x-x_0)\sin\theta + (y-y_0)\cos\theta \tag{109}$$

12.3 Eliminating the DC response of Gabor Filters

The Gabor function as defined above may have a non-zero DC response

$$\hat{g}(0) = \frac{K}{\|A\|}\exp(jP)\,\exp(j2\pi\ u_o^T\ x_o)\ w(A^{-T}u_o) \tag{110}$$

where we used the fact that $w(x) = w(-x)$. In some cases it is useful to eliminate the DC response, for example, we may not want the filter to respond to the absolute intensity of an image. One approach to doing so is to subtract from the original filter the output of a low-pass filter.

$$h(x) = g(x) - Cf(x) \tag{111}$$

where C is a constant and $f(\cdot)$ is the low pass filter. A convenient and popular low pass filter is as follows

$$f(x) = \frac{K}{\|A\|}\,w(A(x-x_o)) \tag{112}$$

Note in this case

$$f(x) = \frac{K}{\|A\|}\,w(A(x-x_o))\Big(\exp(jP)\,\exp(j2\pi\ u_o^T\ x_o) - C\Big) \tag{113}$$

which corresponds to subtracting a complex constant from the complex sinusoid carrier.

Note f is a Gabor filter with zero phase and zero peak response. Therefore it has the following Fourier Transform

$$\hat{f}(u) = \frac{K}{\|A\|} \exp(-j2\pi\ u^T\ x_o)\ w(A^{-T}u) \tag{114}$$

Thus the DC response of the combined filter is as follows

$$\hat{h}(0) = \hat{g}(0) - C\hat{f}(0) = \frac{K}{\|A\|}\Big(\ \exp(jP)\ \exp(j2\pi\ u_o^T\ x_o)\ w(A^{-T}u_o) - C\Big) \tag{115}$$

and thus get a zero DC response we simply need to set C as follows

$$C = \exp(j(P + 2\pi\ u_o^T\ x_o))\ w(A^{-T}u_o) \tag{116}$$

12.4 Another formula of the Gabor function

In other papers, you may see another formula representation of the Gabor function. For example, in most papers, $x_0 = y_0 = 0, P = 0$. Then,

$$\begin{aligned} g(x,y) = K\ &\exp\left(-\pi\left(a^2\,{x_r}^2 + b^2\,{y_r}^2\right)\right) \\ &\left(\exp\left(2\pi j\left(u_0 x + v_0 y\right)\right) - \exp\left(-\pi\left(\frac{{u_{0r}}^2}{a^2} + \frac{{v_{0r}}^2}{b^2}\right)\right)\right) \end{aligned} \tag{117}$$

$$\begin{aligned} \hat{g}(u,v) = \frac{K}{ab}\ &\left(\exp\left(-\pi\left(\frac{{(u-u_0)_r}^2}{a^2} + \frac{{(v-v_0)_r}^2}{b^2}\right)\right)\right. \\ &\left. - \exp\left(-\pi\left(\frac{{u_{0r}}^2}{a^2} + \frac{{v_{0r}}^2}{b^2}\right)\right)\ \exp\left(-\pi\left(\frac{{u_r}^2}{a^2} + \frac{{v_r}^2}{b^2}\right)\right)\right) \end{aligned} \tag{118}$$

Moreover, $a = b \equiv \sigma$ in some paper. The rotation angle has no effect ($\theta = 0$) in this case.

$$\begin{aligned} g(x,y) = K\ &\exp\left(-\pi\sigma^2\left(x^2 + y^2\right)\right) \\ &\left(\exp\left(2\pi j\left(u_0 x + v_0 y\right)\right) - \exp\left(-\frac{\pi}{\sigma^2}\left({u_0}^2 + {v_0}^2\right)\right)\right) \end{aligned} \tag{119}$$

$$\begin{aligned} \hat{g}(u,v) = \frac{K}{\sigma^2}\ &\left(\exp\left(-\frac{\pi}{\sigma^2}\left((u-u_0)^2 + (v-v_0)^2\right)\right)\right. \\ &\left. - \exp\left(-\frac{\pi}{\sigma^2}\left({u_0}^2 + {v_0}^2\right)\right)\ \exp\left(-\frac{\pi}{\sigma^2}\left(u^2 + v^2\right)\right)\right) \end{aligned} \tag{120}$$

Then if you restrict the magnitude of spatial frequency of the sinusoid carrier F_0 to satisfy this equation:

$$F_0 = \sqrt{{u_0}^2 + {v_0}^2} = \frac{\sigma^2}{\sqrt{2\pi}} \tag{121}$$

the Gabor function will be

$$\begin{aligned} g(x,y) = K\ &\exp\left(-\pi\sigma^2\left(x^2 + y^2\right)\right) \\ &\left(\exp\left(j\sqrt{2\pi}\sigma^2\left(x\,\cos\,\omega_0 + y\,\sin\,\omega_0\right)\right) - \exp\left(-\frac{\sigma^2}{2}\right)\right) \end{aligned} \tag{122}$$

$$\hat{g}(u,v) = \frac{K}{\sigma^2}\left(\exp\left(-\frac{\pi}{\sigma^2}\left((u-u_0)^2+(v-v_0)^2\right)\right)\right.$$
$$\left. -\exp\left(-\frac{\sigma^2}{2}\right)\exp\left(-\frac{\pi}{\sigma^2}\left(u^2+v^2\right)\right)\right) \tag{123}$$

Finally if you use $K = 2\pi\sigma^2$,

$$g(x,y) = 2\pi\sigma^2\exp\left(-\pi\sigma^2\left(x^2+y^2\right)\right)$$
$$\left(\exp\left(j\sqrt{2\pi}\sigma^2\left(x\cos\omega_0 + y\sin\omega_0\right)\right) - \exp\left(-\frac{\sigma^2}{2}\right)\right) \tag{124}$$

$$\hat{g}(u,v) = 2\pi\left(\exp\left(-\frac{\pi}{\sigma^2}\left((u-u_0)^2+(v-v_0)^2\right)\right)\right.$$
$$\left. -\exp\left(-\frac{\sigma^2}{2}\right)\exp\left(-\frac{\pi}{\sigma^2}\left(u^2+v^2\right)\right)\right) \tag{125}$$

Additionally, you can use angular frequency (ν, ξ) instead of (u, v). Then,

$$\begin{aligned} g(x,y) &= 2\pi\sigma^2 \exp\left(-\pi\sigma^2\left(x^2+y^2\right)\right) \\ &\left(\exp\left(j\left(\nu_0 x+\xi_0 y\right)\right)-\exp\left(-\frac{\sigma^2}{2}\right)\right) \end{aligned} \tag{126}$$

$$\begin{aligned} \hat{g}(u,v) = 2\pi &\left(\exp\left(-\frac{1}{4\pi\sigma^2}\left(\left(\nu-\nu_0\right)^2+\left(\xi-\xi_0\right)^2\right)\right)\right. \\ &\left.-\exp\left(-\frac{\sigma^2}{2}\right)\exp\left(-\frac{1}{4\pi\sigma^2}\left(\nu^2+\xi^2\right)\right)\right) \end{aligned} \tag{127}$$

In fact, angular frequency representation can be seen in many papers. So it may be useful to have the quite general Gabor function[11] in that format:

$$\begin{aligned} g(x,y) &= K\exp\left(-\pi\left(a^2 {x_r}^2+b^2 {y_r}^2\right)\right) \\ &\left(\exp\left(j\left(\nu_0 x+\xi_0 y\right)\right)-\exp\left(-\frac{1}{4\pi}\left(\frac{{\nu_0}_r{}^2}{a^2}+\frac{{\xi_0}_r{}^2}{b^2}\right)\right)\right) \end{aligned} \tag{128}$$

$$\begin{aligned} \hat{g}(u,v) = \frac{K}{ab} &\left(\exp\left(-\frac{1}{4\pi}\left(\frac{{(\nu-\nu_0)}_r{}^2}{a^2}+\frac{{(\xi-\xi_0)}_r{}^2}{b^2}\right)\right)\right. \\ &\left.-\exp\left(-\frac{1}{4\pi}\left(\frac{{\nu_0}_r{}^2}{a^2}+\frac{{\xi_0}_r{}^2}{b^2}\right)\right)\exp\left(-\frac{1}{4\pi}\left(\frac{{\nu}_r{}^2}{a^2}+\frac{{\xi}_r{}^2}{b^2}\right)\right)\right) \end{aligned} \tag{129}$$

[11] Only $x_0 = y_0 = 0, P = 0$ are assumed.

13 History

- The first version of this document, which was 14 page long, was written by Javier R. Movellan in 1996.
- On September 3 2002 we added the changes made by Kenta Kawamoto. These included a 7 page Appendix with sections on the Fourier transform of the Gabor function, and an altenative formula for the Gabor function.
- Fall 2005. Georgios Britzolakis reported a bug on equation 60.
- Summer 2008. Javier Movellan added 1-D temporal Gabor Section, and polished the Appendix.

References

Adelson, E. H. and Bergen, J. R. (1985). Spationtemporal energy models for the perception of motion. *Journal of the optical society of america A*, 2:284–299.

De Valois, R. L., Albrecht, D. G., and Thorell, L. G. (1982a). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22:545–559.

De Valois, R. L., Yund, W., and Hepler, N. (1982b). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, 22:531–544.

Emerson, R. C., Bergen, J. R., and Adelson, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*, 32(2):203–218.

Heeger, D. (1991). Nonlinear model of neural responses in cat visual cortex. In Landy, M. and Movshon, J., editors, *Computational Models of Visual Processing*, pages 119–133. MIT Press, Cambridge, MA.

Jones, J. P. and Palmer, L. (1987). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58:1233–1258.

Morrone, M. C., Burr, D. C., and Maffei, L. (1982). Functional significance of cross-orientation inhibition, part I. *Neurophysiology. Proc. R. Soc. Lond. B*, 216:335–354.

Movshon, J. A., Thompson, I. D., and Tolhurst, D. J. (1987). Receptive field organization of complex cells in the cat's striate cortex. *Journal of Physiology (London)*, 283:53–77.

Pollen, D. A. and Ronner, S. F. (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science*, 212:1409–1411.