ERROR BOUND FOR MULTI-STAGE SYNTHESIS OF NARROW BANDWIDTH GABOR FILTERS

R. Neil Braithwaite

3241 N. Green Road
Oak Harbor, WA, 98277
Formerly at UC Riverside

Bir Bhanu

College of Engineering
University of California
Riverside, CA 92521

ABSTRACT

This paper develops an error bound for narrow bandwidth Gabor filters synthesized using multiple stages. It is shown that the error introduced by approximating narrow bandwidth Gabor kernels by a weighted sum of spatially offset, separable kernels is a function of the frequency offset and the reduction in bandwidth of the desired kernel compared to the basis values, as well as the spatial subsampling rate between filter stages. This error bound should prove useful in the design of a general basis filter set for multi-stage filtering because the maximum frequency offset is largely determined by the spacing of the basis filters.

1. MULTI-STAGE FILTERING

A direct spatial implementation of a narrow bandwidth filter requires a large size kernel. Multi-stage filtering makes it possible to synthesize a narrow bandwidth response in two or more stages, each comprising wide bandwidth filters. In this work, a two-stage implementation is used. The first stage performs the bulk of the image processing and is designed for computational efficiency. The second stage filter adjusts the frequency and bandwidth of the image response to the desired values. Since the bandwidth of the basis filter output is limited, the second stage kernel can be subsampled relative to the image, allowing for computational savings. The primary restriction in this approach is that the synthesized response must have a narrower bandwidth than the basis filters.

The primary benefits of two-stage filtering, over filtering with a single kernel, are *flexibility* and *efficiency*. The second stage provides *flexibility* in the sense that the center frequency and bandwidth of a given filter can be adjusted in a continuous manner. This allows individual basis filters (or a set of basis filters) to adjust to the image data, tuning Gabor filter responses to match certain narrow-bandwidth image features, such as periodic patterns. When the basis filters are adjusted together as a set, the effects of scale changes and image rotations, as well as the effect of foreshortening due to aspect changes, can be simulated (which could be exploited to simplify object and model matching for object recognition). The multi-stage filter is *efficient* because it allows the choice of a convenient set of basis filters. That is, the size and shape of the basis kernel can be selected independent of the desired two-stage response, thereby exploiting desirable implementation properties such as kernel separability. In addition, the basis filters can be selected for dual use, so that the basis output is used for detecting initial interesting features, as well as the first stage of the two-stage filter.

2. GABOR THEORY

The kernels of the basis set and the desired impulse response of the two-stage filter are 2D Gabor functions. The general form of the Gabor function $G_n(x, y)$ is given by [2] [3]

$$G_n(x, y) = \exp\left\{-\frac{1}{4\pi\sigma_n^2}\left[\acute{x}^2 + \alpha_n^{-2}\acute{y}^2\right]\right\} \quad (1)$$
$$\times \exp\left\{j\omega_n[x\cos\phi_n + y\sin\phi_n]\right\},$$

$$\begin{bmatrix} \acute{x} \\ \acute{y} \end{bmatrix} = \begin{bmatrix} \cos\phi_g & \sin\phi_g \\ -\sin\phi_g & \cos\phi_g \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad (2)$$

where x and y are the horizontal and vertical image coordinates, respectively; $\acute{x}$ and $\acute{y}$ denote the axes that are rotated by ϕ_g; ω_n and ϕ_n are the modulation frequency and orientation, respectively; and σ_n and α_n are the scale and the aspect ratio of the elliptical Gaussian window, respectively. The "bandwidths" of the Gabor function, measured along the principal axes of the Gaussian window, are defined as σ_n^{-1} and $\alpha_n^{-1}\sigma_n^{-1}$.

In this work, the basis filter set is comprised of spatially separable Gabor kernels with log-polar frequency spacing. Each quadrature pair of basis kernels is given

by

$$G_b(x, y, \omega_b, \phi_b) = \exp\left(-\frac{\lambda^2 \omega_b^2 (x^2 + y^2)}{4\pi}\right) \quad (3)$$
$$\times \quad \exp\{j\omega_b[x\cos\phi_b + y\sin\phi_b]\}$$

where $\lambda = (\sigma_b \omega_b)^{-1}$ and σ_b^{-1}, ω_b, and ϕ_b are the bandwidth, frequency, and orientation, respectively, of the basis kernel. Since, in this work, G_n is synthesized using a single basis channel (a sine-cosine pair), the frequency spacing of the basis channels determines the maximum frequency offset between the G_n and the "best" basis channel (with respect to minimum synthesis error). The relationship between the frequency offset and the error in synthesizing G_n is discussed in section 2.2. The computational efficiency associated with using separable basis kernels is discussed in section 2.1.

The spatially sampled output of the Gabor filter is referred to as a *Gabor coefficient*, and is given by

$$a_n = \iint I(x, y) G_n(x_n - x, y_n - y, \omega_n, \phi_n) dx\, dy, \quad (4)$$

where $I(x, y)$ is the input image. In a multi-stage filtering approach, a non-separable, narrow bandwidth kernel, G_n, is synthesized using a weighted combination of spatially offset basis kernels:

$$G_n(x, y) = \sum_i c_i G_b(x - x_i, y - y_i, \omega_b, \phi_b), \quad (5)$$

where c_i are complex weights. Instead of forming a new kernel $G_n(x, y)$ and then re-filtering the original image $I(x, y)$, it is possible to create the same filter response by a weighted summation of the basis filter outputs ($a_{b(i)}$):

$$a_n = \sum_{y_i} \sum_{x_i} c_i a_{b(i)}. \quad (6)$$

Since the bulk of the image processing is performed by the basis filters, the computational advantage of separable filtering is preserved.

The expansion coefficients $\bar{c} = [c_0 \cdots c_i]^T$, which are the complex weights for the second stage of the two-stage filter, are obtained by solving $\bar{b} = Q\bar{c}$, where $\bar{b}$ is a vector of inner products between G_i and $G_{b(i)}$, Q is a Grammian matrix (see [1]). The resulting least squared error is given by

$$J_{lms} = (\bar{b} - Q\bar{c})^T (\bar{b} - Q\bar{c}). \quad (7)$$

This error measure has limited value because $\bar{c}$ must be first calculated. It is not apparent how the error will change as the sampling density of the second stage, the parameters of the basis filter set, and the frequency offset between the basis kernel and the desired kernel are altered.

2.1. Computational Efficiency

Two-stage implementations require less computations than an equivalent single-stage filter if the outputs of the basis filters are subsampled before applying the second filter. The computational cost of a filter is proportional to the area of the kernel. Assume that a rectangular mask has been applied to truncate the elliptical Gaussian window at n_σ standard deviations in each direction along the principal axes. For the case of a single-stage filter, the cost of filtering an image of size N_x by N_y is

$$C_1 = 4n_\sigma^2 \sigma_n^2 \alpha_n N_x N_y. \quad (8)$$

If implemented as a two-stage filter, with the second stage filter subsampled by a factor S, we get

$$C_2 = n_{basis} \left[C_{basis} + \frac{C_1}{S^2}\right], \quad (9)$$

where C_{basis} is the computational cost of each basis filter, and n_{basis} is the number of basis filters used in the second stage filter. The cost of filtering with a separable basis kernel is given by

$$C_{basis} = 4n_\sigma \sigma_b N_x N_y. \quad (10)$$

The ratio of the two-stage and the single-stage costs, for a separable basis kernel implementation, is given by

$$\frac{C_2}{C_1} = n_{basis} \left[\left(\frac{\sigma_b}{\sigma_n}\right) \frac{1}{n_\sigma \sigma_n \alpha_n} + \frac{1}{S^2}\right]. \quad (11)$$

In this work, the sine and cosine responses from a single basis channel are used in the two-stage approach, which makes $n_{basis} = 2$. As an example, for $\omega_b = \frac{\pi}{8}$, $\lambda = \frac{\pi}{4}$, $\sigma_b = (\lambda \omega_b)^{-1}$, $\sigma_n = 2\sigma_b$, $\alpha_n = 1$, $S = 8$, $n_\sigma = 2$, the two-stage computational cost, for separable basis kernels, is reduced to 0.054 of the single-stage cost.

2.2. Predicting the Synthesis Error

The impulse response of the two-stage filter is not, in general, exactly the same as the desired Gabor response. The difference between the two responses is referred to as the "synthesis error." In this subsection, an expression for the upper bound of the synthesis error is provided. This upper bound considers the effects of frequency offsets and bandwidth reductions of the desired filter (compared to the basis filters), and the effects of subsampling before applying the second stage filter. It does not account for errors associated with the truncation of the kernel's Gaussian window.

Throughout this subsection, the responses of the basis filter, the second stage filter, and the desired filter are represented by their frequency domain transfer

functions, and are denoted by $\hat{G}_b$, $\hat{F}_s$, and $\hat{G}_n$, respectively. The term $\hat{G}$ indicates that the filter response in the frequency domain is Gaussian. When the second stage filter response is Gaussian, it will be denoted by $\hat{G}_s$.

The relationship between the basis function $\hat{G}_b$, the second stage filter $\hat{F}_s$, and the desired filter $\hat{G}_n$ is given, ideally, by

$$\hat{G}_n = \hat{F}_s \hat{G}_b. \qquad (12)$$

Both $\hat{G}_n$ and $\hat{G}_b$ are Gaussian responses in the frequency domain and are given by

$$\hat{G}_n = \exp(-\pi\sigma_n^2 \left[u^2 + \alpha_n^2 v^2\right]), \qquad (13)$$

$$\hat{G}_b = \exp(-\pi\sigma_b^2 \left[(u - \Delta u)^2 + (v - \Delta v)^2\right]), \qquad (14)$$

respectively, where u and v are the frequencies measured along the principal axes of the elliptical Gaussian $\hat{G}_n$, and $(\Delta u, \Delta v)$ are the frequency offset values measured relative to the basis frequency. The frequency offset values, or the frequency shifts to be introduced by the second stage filter $\hat{F}_s$, are given by

$$\Delta u = \omega_n \cos(\phi_n - \phi_g) - \omega_b \cos(\phi_b - \phi_g) \qquad (15)$$

$$\Delta v = \omega_n \sin(\phi_n - \phi_g) - \omega_b \sin(\phi_b - \phi_g). \qquad (16)$$

Although we know the basis response and the desired response; we need to know the shape of $\hat{F}_s$ in order to predict the synthesis error. In the ideal case, the shape of $\hat{F}_s$ is a Gaussian defined by $\hat{G}_n \hat{G}_b^{-1}$.

When the output of $\hat{G}_b$ is subsampled compared to the original image, an aliasing error arises. The sampling process causes the basis response $\hat{G}_b$ to be replicated throughout the spectral domain at intervals equal to the $\hat{x}$ and $\hat{y}$ sampling frequencies. The replicated basis responses are given by

$$\hat{G}_b(l, m) = \exp[-\pi\sigma_b^2(\omega_{\hat{x}} - \omega_b \cos\phi_b - l\,\omega_{s(x)})^2]$$
$$\times \exp[-\pi\sigma_b^2(\omega_{\hat{y}} - \omega_b \sin\phi_b - m\,\omega_{s(y)})^2]$$

where $\omega_{s(x)}$ and $\omega_{s(y)}$ are the sampling frequencies in the $\hat{x}$ and $\hat{y}$ directions, respectively, and l and m are integers. The replicated basis responses overlap with the original basis response. Thus, the filter $\hat{F}_s$ must shape the synthesized response $(\hat{F}_s \hat{G}_b)$ to best approximate $\hat{G}_n$ from a spectrum containing both $\hat{G}_b$ and $\hat{G}_a$, where

$$\hat{G}_a = \sum_{(l,m)\neq(0,0)} \hat{G}_b(l, m). \qquad (17)$$

If it assumed that the basis response $\hat{G}_b$ and the aliased response $\hat{G}_a$ are uncorrelated then the optimal $\hat{F}_s$ minimizes

$$J = \frac{\int\int |\hat{G}_n - \hat{F}_s \hat{G}_b|^2 + |\hat{F}_s \hat{G}_a|^2 dudv}{\int\int \hat{G}_n^2 dudv}, \qquad (18)$$

and is given by

$$\hat{F}_s = \hat{G}_n \hat{G}_b^{-1} \hat{F}_c \qquad (19)$$

where

$$\hat{F}_c = \frac{\hat{G}_b^2}{\hat{G}_b^2 + \hat{G}_a^2} \qquad (20)$$

The filter $\hat{F}_c$ attenuates the replicated (aliased) part of the spectrum.

Although optimal under the uncorrelated assumption, the above expression for $\hat{F}_s$ does not provide an analytical solution for the squared synthesis error J. To allow for a tractable solution, a sub-optimal window $\hat{G}_s$, which has a Gaussian shape, is used:

$$\hat{G}_s = \hat{G}_n \hat{G}_b^{-1} \hat{G}_c \qquad (21)$$

where

$$\hat{G}_c = \exp(-\pi\sigma_c^2 \left[u^2 + v^2\right]). \qquad (22)$$

Note that $\hat{G}_c$ is a circular Gaussian with bandwidth σ_c^{-1} and frequency (u, v) (same frequency as $\hat{G}_n$). The synthesis error for the Gaussian solution has the following form:

$$J = J_n + J_a = \frac{\int\int \hat{G}_n^2(1 - \hat{G}_c)^2 + \hat{G}_n^2 \hat{G}_c^2 \hat{G}_b^{-2} \hat{G}_a^2 dudv}{\int\int \hat{G}_n^2 dudv}, \qquad (23)$$

where J_n is the "basis-only" reconstruction error (no aliasing) and J_a is the aliasing error. These two error components are given by

$$J_n = 1 \quad - \quad 2\left[\frac{2\sigma_n^2}{2\sigma_n^2 + \sigma_c^2} \cdot \frac{2\alpha_n^2\sigma_n^2}{2\alpha_n^2\sigma_n^2 + \sigma_c^2}\right]^{0.5}$$
$$+ \quad \left[\frac{\sigma_n^2}{\sigma_n^2 + \sigma_c^2} \cdot \frac{\alpha_n^2\sigma_n^2}{\alpha_n^2\sigma_n^2 + \sigma_c^2}\right]^{0.5} \qquad (24)$$

$$J_a = c_\sigma \cdot h_\Delta(\Delta u, \Delta v) \cdot \sum_{(l,m)\neq(0,0)} h_s(l\,\omega_{s(x)}, m\,\omega_{s(y)}) \qquad (25)$$

where

$$c_\sigma = \left[\frac{\sigma_n^2}{\sigma_n^2 + \sigma_c^2} \cdot \frac{\alpha_n^2\sigma_n^2}{\alpha_n^2\sigma_n^2 + \sigma_c^2}\right]^{0.5} \qquad (26)$$

$$h_\Delta = \exp(2\pi\sigma_b^2 \left[\gamma_u^{-1}(\Delta u)^2 + \gamma_v^{-1}(\Delta v)^2\right]) \qquad (27)$$

$$h_s = \exp[-2\pi\sigma_b^2\gamma_u(\omega_{s(u)} - \gamma_u^{-1}\Delta u)^2]$$
$$\times \exp[-2\pi\sigma_b^2\gamma_v(\omega_{s(v)} - \gamma_v^{-1}\Delta v)^2]. \qquad (28)$$

$$\gamma_u = 1 - \frac{\sigma_b^2}{\sigma_n^2 + \sigma_c^2} \qquad (29)$$

$$\gamma_v = 1 - \frac{\sigma_b^2}{\alpha_n^2\sigma_n^2 + \sigma_c^2} \qquad (30)$$

$$\begin{bmatrix} \omega_{s(u)} \\ \omega_{s(v)} \end{bmatrix} = \begin{bmatrix} \cos(\phi_g) & \sin(\phi_g) \\ -\sin(\phi_g) & \cos(\phi_g) \end{bmatrix} \begin{bmatrix} l\ \omega_{s(x)} \\ m\ \omega_{s(y)} \end{bmatrix}. \quad (31)$$

For given values of σ_c and σ_n, the basis-only reconstruction error, J_n, is constant. The aliasing error, J_a, is a function of the frequency offset $(\Delta u, \Delta v)$ and the sampling frequencies $(\omega_{s(x)}, \omega_{s(y)})$, in addition to σ_c and σ_n. It can be seen, by substituting (29) and (30) into (27) and (28), that reducing the bandwidth of the desired filter $\hat{G}_n$ (that is, increasing σ_n), greatly reduces the aliasing error J_a. Similarly, the filter $\hat{G}_c$ is effective at reducing the aliasing error because it reduces the bandwidth of the synthesized response from σ_n^{-1} to $(\sigma_c^2 + \sigma_n^2)^{-0.5}$. However, a bandwidth reduction using $\hat{G}_c$ increases in the basis-only reconstruction error (J_n).

3. RESULTS AND CONCLUSIONS

Figure 1 shows the impulse responses of narrow bandwidth filters (cosine part) with various frequency and orientation offsets from the basis values. The synthesis error for each filter is listed in table 1. It should be noted that the predicted error values in table 1 measure the complex error (that is, both sine and cosine kernels), so it should overestimate the measured error (based on the cosine kernel, only). However, it can be seen in table 1 that this is only the case for figure 1 (f). The measured error includes the effects of truncation of the Gabor function (Gaussian window) and of the spatial lattice used in the second stage, none of which are accounted for in the predicted errors. The truncation errors in these examples are small, and hence are discernible only when the error is small. For figure 1 (f), the large orientation offset results in a 1.5 percent error whereas the prediction based on (7) is a 1.9 percent error. This overestimation, by a factor of 1.3, is close to the $\sqrt{2}$ factor that should appear on average. The prediction based on (23) is larger, a 3.3 percent error (factor of 2.2), because the shape of the optimal filter $\hat{F}_c$ becomes less like a Gaussian as Δu and/or Δv increases (relative to the second stage sampling frequencies, $\omega_{s(u)}$ and $\omega_{s(v)}$).

In conclusion, the synthesis error can be made arbitrarily small by reducing the maximum frequency offset (increasing the number and density of the basis channels) and/or increasing the second stage sampling frequencies. When the maximum allowable error is small, equation (23) is a good prediction, allowing one to predict how adjustments to the basis filter set and the second stage sampling frequencies will affect the synthesis error.

Acknowledgement: This work was supported by

(a)   (b)   (c)

(d)   (e)   (f)
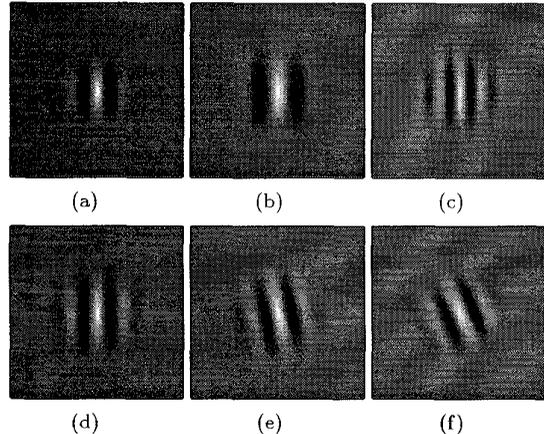
Figure 1: Cosine Gabor filters with $\alpha_n = 1$. (a) Basis kernel with bandwidth σ_b^{-1} and frequency, orientation $[\omega_b, 0]$. (b)-(f) Impulse response of synthesized filters with reduced bandwidth $(\sigma^{-1} = 0.7\sigma_b^{-1})$. (b) $[0.7\omega_b, 0]$. (c) $[1.3\omega_b, 0]$. (d) $[\omega_b, 0]$. (e) $[\omega_b, 0.20]$. (f) $[\omega_b, 0.39]$. Orientations measured in radians.

Table 1: Synthesis Errors for Impulse Responses in figure 1. Orientations, ϕ, are measured in radians.

Fig. no.	Freq. Offset (ω, ϕ)	Synth. Error	Predicted Errors Eq. (7)	Eq. (23)
1 (b)	$[0.7\omega_b, 0]$	0.0065	0.0060	0.0136
1 (c)	$[1.3\omega_b, 0]$	0.0064	0.0057	0.0136
1 (d)	$[\omega_b, 0]$	0.0030	0.0005	0.0001
1 (e)	$[\omega_b, 0.20]$	0.0036	0.0013	0.0015
1 (f)	$[\omega_b, 0.39]$	0.0150	0.0196	0.0334

4. REFERENCES

[1] R. N. Braithwaite and B. Bhanu, "Hierarchical Gabor filters for object detection in infrared images," *Proc., IEEE Conf. CVPR, June 1994*, pp. 628-631.

[2] J. G. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by 2-D visual cortical filters," *J. Opt. Soc. Am. A*, vol. 2, no. 7, pp. 1160-1169, 1985.

[3] D. Gabor, "Theory of communication," *J. Inst. Elec. Eng.*, vol. 93, pp. 429-457, 1946.