Proc. European Conference on Circuit Theory and Design, Special Session on Applications of Gabor Filters and Transforms in Image Processing, Helsinki, Finland, vol. 3, pp. 45-48, Aug. 2001.

Competitive Orientation Selective Arrays

Bertram E. Shi and Kwabena Boahen[1]

***Abstract* - We present an architecture for orientation selective image filtering that includes competitive interactions between different arrays tuned to different orientations. We prove that the presence of these interactions sharpens the spatial frequency tuning of the filters in comparison to a similar architecture which lacks these interactions.**

1 Introduction

Recent work has described circuit architectures implementing orientation selective image filters similar to Gabor filters[1]. The processing circuits are based upon the cellular neural network architecture[2], where analog continuous time circuits at each pixel are interconnected with identical circuits at neighboring pixels. By adjusting the interconnections via external bias voltages, we can tune the filter circuits to respond maximally to different scales and orientations.

In this paper, we describe an architecture which couples the outputs of a set of broadly tuned orientation selective layers so that they compete to best model the input. We prove theoretically that the introduction of this competitive coupling enhances the orientation tuning and illustrate this via system level simulations. One of the advantages of this formulation is that is extends previous work in a natural way, implying that similar circuit architectures as used in the past can be re-used in constructing this new architecture.

2 Uncoupled Orientation Selective Layers

In this section, we briefly review the previous work in orientation selective circuit arrays which the current work is based upon. This paper focuses on the system level description and the computational properties of the architecture. For information regarding circuit design and implementation, please refer to [3][4].

1. BES is with the Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology, Kowloon, Hong Kong. KB is with the Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, USA. This work was supported by the Research Grants Council of Hong Kong, China (Project no. HKUST6216/98E) and by the US National Science Foundation's LIS/KDI program (no. KDI-ECS98-74463). This work was inspired by discussions held at the 2000 Telluride Workshop on Neuromorphic Engineering.

The analog processing circuits implementing the orientation selectivity settle to a point which minimizes the following cost function:

$$
\begin{aligned}
E(\hat{v}) = {} & \frac{1}{2}\sum_m \sum_n \|v(m,n) - u(m,n)\|^2 \\
& + \frac{1}{2(\Delta\Omega_x)^2}\sum_m \sum_n \left\|v(m,n) - e^{-j\Omega_x} v(m+1,n)\right\|^2 \\
& + \frac{1}{2(\Delta\Omega_y)^2}\sum_m \sum_n \left\|v(m,n) - e^{-j\Omega_y} v(m,n+1)\right\|^2
\end{aligned} \tag{1}
$$

where $v(m,n)$ is the complex valued filter output and $u(m,n)$ is the real valued input image. This cost function is the sum of two terms: a data fidelity term that penalizes the difference between the filter output and input and a regularization term that is minimized if the output is a complex exponential waveform with spatial frequency (Ω_x, Ω_y). The amount each term contributes to the cost function is controlled by $\Delta\Omega_x$ and $\Delta\Omega_y$. It turns out that these parameters control the bandwidth of the filters (see (3) below).

By differentiating this cost function with respect to the real and imaginary parts of $v(m,n)$ ($v_r(m,n)$ and $v_i(m,n)$) and setting the results equal to zero, we obtain two equations for every pixel (m,n) which can be compactly expressed as the real and imaginary parts of the following:

$$
\begin{aligned}
0 = {} & u(m,n) \\
& - \left(1 + \frac{2}{(\Delta\Omega_x)^2} + \frac{2}{(\Delta\Omega_y)^2}\right) v(m,n) \\
& + \frac{e^{j\Omega_x}}{(\Delta\Omega_x)^2} v(m-1,n) + \frac{e^{-j\Omega_x}}{(\Delta\Omega_x)^2} v(m+1,n) \\
& + \frac{e^{j\Omega_y}}{(\Delta\Omega_y)^2} v(m,n-1) + \frac{e^{-j\Omega_y}}{(\Delta\Omega_y)^2} v(m,n+1)
\end{aligned} \tag{2}
$$

Assume an infinite array of pixels and define $U(\omega_x, \omega_y)$ to be the discrete space Fourier transform of $u(m,n)$,

$$U(\omega_x, \omega_y) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} u(m,n) e^{-j\omega_x m - j\omega_y n}$$

and $V(\omega_x, \omega_y)$ to be the discrete space Fourier transform of $v(m,n)$. Taking the Fourier transform of (2),

$$\frac{V(\omega_x, \omega_y)}{U(\omega_x, \omega_y)} = \frac{1}{1 + \frac{2 - 2\cos(\omega_x - \Omega_x)}{(\Delta\Omega_x)^2} + \frac{2 - 2\cos(\omega_y - \Omega_y)}{(\Delta\Omega_y)^2}} \tag{3}$$

For (ω_x, ω_y) close to (Ω_x, Ω_y),

$$\frac{V(\omega_x, \omega_y)}{U(\omega_x, \omega_y)} = \frac{1}{1 + \frac{(\omega_x - \Omega_x)^2}{(\Delta\Omega_x)^2} + \frac{(\omega_y - \Omega_y)^2}{(\Delta\Omega_y)^2}}$$

which drops to half its maximum value when

$$\frac{(\omega_x - \Omega_x)^2}{(\Delta\Omega_x)^2} + \frac{(\omega_y - \Omega_y)^2}{(\Delta\Omega_y)^2} = 1$$

This equation describes an ellipse centered at (Ω_x, Ω_y) whose axes are parallel to the ω_x and ω_y axes. The lengths of the axes are the 6dB bandwidths in x and y are $2\Delta\Omega_x$ and $2\Delta\Omega_y$. Define

$$Q = \sqrt{\left(\frac{\Omega_x}{2\Delta\Omega_x}\right)^2 + \left(\frac{\Omega_y}{2\Delta\Omega_y}\right)^2}$$

to be the quality factor, characterizing the sharpness of the tuning. Typically, $\Delta\Omega_x = \Delta\Omega_y = \Delta\Omega$, so that the frequency response is approximately circularly symmetric around the center frequency (Ω_x, Ω_y) and $Q = \Omega/(2\Delta\Omega)$ where $\Omega = \sqrt{\Omega_x^2 + \Omega_y^2}$.

The inverse Fourier transform gives us the complex valued impulse response:

$$h(m, n) = f(m, n)e^{j(\Omega_x m + \Omega_y n)}$$

where $f(m, n)$ is a real valued function which decays approximately circularly symmetrically away from the origin if $\Delta\Omega_x = \Delta\Omega_y$. These filters are similar to Gabor filters, in which the modulating function $f(m, n)$ is a Gaussian function.

3 Competitively Coupled Layers

In this section, we describe an extension of the previous architecture in which the outputs of that enables different outputs interact so that their sum best models the input. In this way, the approach is similar to that proposed by Daugman[5]. By introducing competition between different orientations, we enhance the orientation selectivity, especially in the case of multiple competing orientations. A major advantage of this approach is that by extending our previous approach in a logically consistent way, we can re-use much of the circuit architecture that we have already designed and tested.

Suppose we wish to analyze an input image at K different orientations $(\Omega_{xk},\Omega_{yk})$ for $k \in \{1, \ldots, K\}$. Let $v_k(m, n)$ be the orientation selective output for the kth orientation. Consider the following extension of the cost function:

$$\begin{aligned} E(\vec{v}) = {} & \frac{1}{2}\sum_m \sum_n \left\| \sum_k v_k(m, n) - u(m, n) \right\|^2 \\ & + \frac{1}{2}\sum_k \frac{1}{(\Delta\Omega_{xk})^2} \sum_m \sum_n \left\| v_k(m, n) - e^{-j\Omega_{xk}} v_k(m+1, n) \right\|^2 \\ & + \frac{1}{2}\sum_k \frac{1}{(\Delta\Omega_{yk})^2} \sum_m \sum_n \left\| v_k(m, n) - e^{-j\Omega_{yk}} v_k(m, n+1) \right\|^2 \end{aligned} \quad (4)$$

The critical difference between this cost function and (1) is in the data fidelity term. In (1), the data fidelity term penalizes the difference between the input image and a single orientation selective output. In (4), the data fidelity term penalizes the difference between the input image and the sum of the orientation selective outputs over all orientations. This introduces competitive couplings between the different orientations.

Differentiating the cost function with respect to the real and imaginary parts of $v_k(m, n)$ and setting the results equal to zero, we obtain two equations which can be compactly expressed as the real and imaginary parts of the following:

$$\begin{aligned} 0 = {} & u(m, n) - \sum_{j \neq k} v_j(m, n) \\ & - \left(1 + \frac{2}{(\Delta\Omega_{xk})^2} + \frac{2}{(\Delta\Omega_{yk})^2}\right) v_k(m, n) \\ & + \frac{e^{j\Omega_{xk}}}{(\Delta\Omega_{xk})^2} v_k(m-1, n) + \frac{e^{-j\Omega_{xk}}}{(\Delta\Omega_{xk})^2} v_k(m+1, n) \\ & + \frac{e^{j\Omega_{yk}}}{(\Delta\Omega_{yk})^2} v_k(m, n-1) + \frac{e^{-j\Omega_{yk}}}{(\Delta\Omega_{yk})^2} v_k(m, n+1) \end{aligned} \quad (5)$$

3.1 Frequency Response

Since relationship between the input and output of the coupled arrays in (5) is linear, we can find the frequency response of each array. In the following, we assume that center frequencies $(\Omega_{xk}, \Omega_{yk})$ are distinct for $k \in \{1, \ldots, K\}$.

Taking the Fourier transform of (5)

$$U(\omega_x, \omega_y) = \sum_j V_j(\omega_x, \omega_y) + D_k(\omega_x, \omega_y)V_k(\omega_x, \omega_y) \quad (6)$$

for $k \in \{1, \ldots, K\}$ where

$$D_k(\omega_x, \omega_y) = \frac{2 - 2\cos(\omega_x - \Omega_{xk})}{(\Delta\Omega_{xk})^2} + \frac{2 - 2\cos(\omega_y - \Omega_{yk})}{(\Delta\Omega_{yk})^2}$$

Note that $D_k(\omega_x, \omega_y) \geq 0$ with equality if and only of $\omega_x = \Omega_{xk}$ and $\omega_y = \Omega_{yk}$ where both equalities are modulo 2π. There are two cases which we must consider:

Case 1: $(\omega_x, \omega_y) \neq (\Omega_{xk}, \Omega_{yk})$ for all k

In this case, $D_k(\omega_x, \omega_y) > 0$ for all $k \in \{1, \ldots, K\}$. Equation (6) can be solved by noting that for all k,

$$V_k(\omega_x, \omega_y) = \frac{U(\omega_x, \omega_y) - \sum_j V_j(\omega_x, \omega_y)}{D_k(\omega_x, \omega_y)} \quad (7)$$

Summing over k,

$$\sum_k V_k(\omega_x, \omega_y) = \{U(\omega_x, \omega_y) - \sum_j V_j(\omega_x, \omega_y)\} \sum_k D_k(\omega_x, \omega_y)^{-1}$$

Thus,

$$\sum_k V_k(\omega_x, \omega_y) = \frac{\sum_k D_k(\omega_x, \omega_y)^{-1}}{1 + \sum_k D_k(\omega_x, \omega_y)^{-1}} U(\omega_x, \omega_y)$$

Substituting into (7),

$$H_k(\omega_x, \omega_y) = \frac{V_k(\omega_x, \omega_y)}{U(\omega_x, \omega_y)} = \frac{D_k(\omega_x, \omega_y)^{-1}}{1 + \sum_j D_j(\omega_x, \omega_y)^{-1}} \quad (8)$$

Case 2: $(\omega_x, \omega_y) = (\Omega_{xl}, \Omega_{yl})$ for some l.

Since $D_l(\omega_x, \omega_y) = 0$, evaluating (6) for $k = l$,

$$\sum_j V_j(\omega_x, \omega_y) = U(\omega_x, \omega_y) \quad (9)$$

For $k \neq l$, $D_k(\omega_x, \omega_y) > 0$. By (7), $V_k(\omega_x, \omega_y) = 0$. Substituting into (9), $V_l(\omega_x, \omega_y) = U(\omega_x, \omega_y)$. Thus,

$$H_k(\omega_x, \omega_y) = \frac{V_k(\omega_x, \omega_y)}{U(\omega_x, \omega_y)} = \begin{cases} 1 & l = k \\ 0 & l \neq k \end{cases} \quad (10)$$

This expression is the limit of the expression derived for Case 1 as we let $D_l(\omega_x, \omega_y) \to 0$.

The last equation shows that when the input spatial frequency component is exactly equal to tuned spatial frequency of one of the arrays, that spatial frequency is passed directly by the corresponding array and completely blocked by all others. This result is independent of the bandwidth of the spatial tuning.

3.2 Enhancement of Orientation Tuning

This system should exhibit a sharper orientation tuning in the presence of competing orientations for the coupled case. To see this intuitively, assume that the orientations are indexed in order of increasing angle, $\theta_1 < \theta_2 < \ldots < \theta_K$. and consider an input consisting of two lines oriented at angles θ_1 and θ_3. In the uncoupled case, there will be a strong response at the array tuned to the orientation θ_2 which lies between θ_1 and θ_3, since both lines will contribute to the output. On the other hand, in the coupled case, the different outputs compete with each other to see which best model the input. Thus, large responses at the outputs tuned to θ_1 and θ_3 will suppress the output tuned to the orientation θ_2.

The following theorem justifies this intuition mathematically. We compare the frequency responses of coupled and uncoupled orientation selective arrays tuned to the same sets of orientations. Suppose that at some spatial frequency (ω_x, ω_y) the response of the kth array is smaller than that at the lth array. The ratio

$$\frac{H_k(\omega_x, \omega_y)}{H_l(\omega_x, \omega_y)} < 1$$

indicates the sharpness of the tuning. A smaller ratio indicates sharper tuning, since the response of the kth array is more suppressed.

PROPOSITION 1:

Consider one coupled set and one uncoupled set of orientation selective arrays tuned to K distinct spatial frequencies $\{(\Omega_{xk}, \Omega_{yk})\}_{k=1}^{K}$. Define $H_k^c(\omega_x, \omega_y)$ and $H_k^u(\omega_x, \omega_y)$ to be the transfer functions from the input to the output tuned to spatial frequency $(\Omega_{xk}, \Omega_{yk})$ for the coupled and the uncoupled array. If for some k, l, ω_x and ω_y,

$$\frac{H_k^u(\omega_x, \omega_y)}{H_l^u(\omega_x, \omega_y)} < 1 \quad (11)$$

then

$$\frac{H_k^c(\omega_x, \omega_y)}{H_l^c(\omega_x, \omega_y)} < \frac{H_k^u(\omega_x, \omega_y)}{H_l^u(\omega_x, \omega_y)}$$

PROOF:

For the uncoupled arrays, we have from (3)

$$\frac{H_k^u}{H_l^u} = \frac{1 + D_l}{1 + D_k}$$

where we do not indicate the dependence on (ω_x, ω_y) to avoid clutter. If $D_l = 0$, then by (10),

$$\frac{H_k^c}{H_l^c} = 0 < \frac{H_k^u}{H_l^u}.$$

Otherwise, (11) implies that $0 < D_l < D_k$

$$D_l + D_k D_l < D_k + D_k D_l$$

$$D_l(1 + D_k) < D_k(1 + D_l)$$

$$\frac{H_k^c}{H_l^c} = \frac{D_l}{D_k} < \frac{1 + D_l}{1 + D_k} = \frac{H_k^u}{H_l^u}$$

where the left hand equality follows from (8).
QED

It is important to note that the sharpening exhibited by this system is robust in the sense that the theorem above does not depend upon any conditions on the selection of tuned spatial frequencies or in the bandwidths of the individual arrays, apart from the requirement that the tunings of the different arrays be distinct. Of course the amount of sharpening will depend critically on the choice of the spatial frequency tunings of the arrays and their bandwidths.

Figure 1 compares the responses of the uncoupled and the coupled arrays. The uncoupled arrays are broadly tuned with $Q = 0.64$. The introduction of competitive couplings between the arrays clearly enhances the orientation selective tuning. Examination of the impulse response indicates that this enhancement is primarily achieved by widening the impulse response in the direction parallel to the desired orientation. It is impossible to achieve such spreading by adjusting the bandwidths with the uncoupled arrays, which only adjust the widths in the ω_x and ω_y directions.

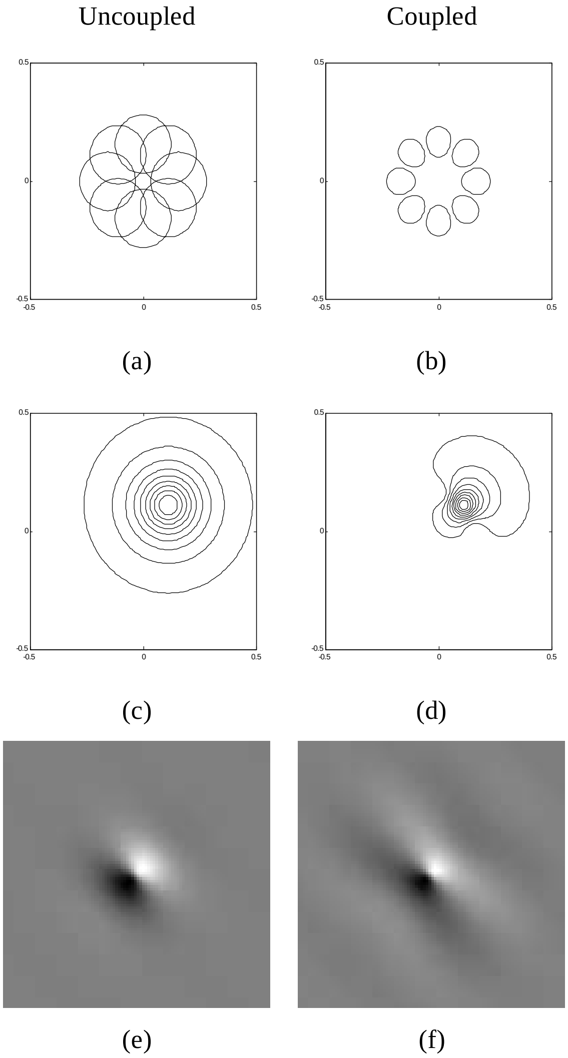

Fig. 1: Comparison of the responses from a set of eight uncoupled and coupled orientation selective layers. The orientation selective layers are tuned to respond to spatial frequencies with magnitude $\pi/20$ and orientations 0, $\pi/4$, $\pi/2$, $3\pi/4$, π, $5\pi/4$, $3\pi/2$, $7\pi/4$. The Q for the uncoupled arrays is 0.64. (a,b) Contour plots indicating where the frequency responses of the 8 arrays drops by 6dB. (c,d) Contour plots of the frequency responses of the arrays tuned to orientation $\pi/4$. (e,f). Impulse responses of the odd symmetric filters tuned to orientation $\pi/4$.

4 Conclusion

We have described an architecture for orientation selective filtering which includes competitive couplings between outputs tuned to different orientations. We have proven that this competition enhance orientation selectivity.

Our ongoing work in this area aims to implement these orientation selective circuits in CMOS VLSI. One of the advantages of this formulation is that the circuits performing the minimization are natural extensions of our previously designed circuits. Due to limitations of the 2D silicon substrate, we plan to split the processing among multiple chips implementing separate orientation selective layers which communicate with each other via the Address Event Representation (AER) communication protocol[6].

References

[1] B. E. Shi, “Gabor-type filtering in space and time with cellular neural networks,” *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, vol. 45, no. 2, pp. 121-132, Feb. 1998.

[2] L. O. Chua and T. Roska, “The CNN paradigm,” *IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications*, vol. 40, pp. 147-156, Mar. 1993.

[3] B. E. Shi, “Focal Plane Implementation of 2D Steerable and Scalable Cortical Filters,” *Journal of VLSI Signal Processing*, vol. 23, no. 2/3, pp. 319-334, Nov./Dec. 1999.

[4] B. E. Shi, “A low power orientation selective vision sensor,” *IEEE Trans. Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 47, no. 5, pp. 435-440, May 2000.

[5] J. G. Daugman, “Complete discrete 2-D gabor transforms by neural networks for image analysis and compression,” *IEEE Transactions on Acoustics*, *Speech and Signal Processing*, vol. 36, no. 7, pp. 1169-1179, Jul. 1988.

[6] K. A. Boahen, “Point-to-point connectivity between neuromorphic chips using address events,” *IEEE Transactions on Circuits and Systems-II: Analog and Digital Signal Processing*, vol. 47, no. 5, pp. 416-434, May 2000.