Cogn Comput (2010) 2:297–302
DOI 10.1007/s12559-010-9047-2

Face Recognition with Quantum Associative Networks Using Overcomplete Gabor Wavelet

Nuo Wi Tay · Chu Kiong Loo · Mitja Peruš

Received: 19 April 2010 / Accepted: 10 May 2010 / Published online: 29 May 2010

Abstract Gabor wavelet is considered the best mathematical descriptor for receptive fields in the striate cortex. Besides, as a basis function, it is suitable to sparsely represent natural scenes due to its property in maximizing information. It is argued that Gabor-like receptive fields are emerged by sparseness-enforcing or infomax method. In this paper, we incorporate Gabor overcomplete representation into quantum holography for image recognition tasks, with suggestions in improvements through iterative method for reconstruction.

Keywords Quantum Holography · Gabor wavelet · Information maximization · Invariance · Face recognition

Introduction

Gabor wavelet is considered the best mathematical descriptor for striate cortex receptive fields [4, 17], which is selective to a particular orientation and spatial frequency. Careful comparison is done by [9], which supported this claim. It can be used to describe the orientation and spatial-frequency-tuned receptive fields of simple cells in V1 [13]. The design of biological cells fits the two 90 degree phase difference representations of Gabor wavelet [22]. Gabor wavelet representation is nearly affine invariant by manipulating its parameters [10]. Although non-orthogonal, Lee [11] derives conditions under which Gabor wavelets, which are generally non-orthogonal, behave as if they are orthogonal. Oversampling in primate's visual system achieves this almost tight-frame condition, enabling reconstruction of high-resolution images from the wavelets. This, of course, is a lot more convenient than finding biorthogonal function, which is quite difficult [1]. According to [7], this overcompleteness provides a robust representation that is able to be stored by low-precision neurons through redundancy. Besides, it provides a good medium for tasks like image segmentation [5, 11].

Normal images are generally highly self-correlated due to internal morphological consistency, which should be utilized and exploited in image recognition [5]. Field [6] shows that wavelet is the most suitable descriptor for natural images, which is relatively sparse compared to total spectral or spatial domain representation. Under the general cases, Gabor wavelet function can extract maximum information from an input image [14]. As shown by [12], 1993, the RF of neurons employs an information-theoretic method by maximizing mutual information. He showed that mutual information maximization is related to Hebbian learning rule for neural network connectivity [12]. Gabor-like receptive fields can emerge by sparseness-enforcing [15, 16] or by an infomax method [2, 3]. Both methods produce Gabor-like outputs as statistically independent scene-components. Peruš et al. [18, 19, 21] have argued that the sparseness-enforcing method is more biologically plausible than the infomax approach.

According to [23], there is a triple convolution for preliminary visual pathway from retina to lateral geniculate nucleus (LGN) to the striate cortex (V1), with the 1st and 2nd convoluted with a Difference-of-Gaussian function and the 3rd with the Gabor wavelet. Since this paper focuses on the Gabor transform in image recognition, the 1st and 2nd

N. W. Tay · C. K. Loo (✉)
Faculty of Information Science and Technology,
Multimedia University, Melaka, Malaysia
e-mail: ckloo@mmu.edu.my

M. Peruš
Laboratory for Cognitive Modeling, FRI, University
of Ljubljana, Ljubljana, Slovenia

stage of the visual pathway will not be touched upon. Images are convoluted with the Gabor-like receptive field, which is then conveyed to V1 hyper-column that is over-complete where every point of the column is selective to a particular orientation and ocular dominance [8]. According to [23], the spectral image in V1 is inversed to the spatial topologically correct image into V2. Both spatial and spectral representation can be processed by Hebbian-like association for image recognition and other tasks.

Quantum Associative Memory

A quantum implementable model has been proposed by [18–21] for associative memory for image recognition.

For the associative model, an n by m image ($n \times m = N$) is represented as a vector. Hebbian matrix for classical Hopfield network is defined as $vv^T = H$ with

$$v = \begin{bmatrix} p^1(1) & p^2(1) & \dots & p^k(1) \\ p^1(2) & p^2(2) & & \\ \vdots & & \ddots & \vdots \\ p^1(N) & & & p^k(N) \end{bmatrix}$$

Superscript T denotes the transpose of v. Reconstruction is achieved by $v^{\text{out}} = Hv^{\text{input}}$. For quantum holography, v is replaced by plane wave function $A_1 e^{i\phi 1}$, where pixel intensity is encoded in wave-phase and T represents the complex conjugate. Both approaches are equivalent as proven in [20]. The former representation (pixel intensity, v) will be used for ease of description. For a more detailed description of wave function in quantum holography, refer to [20, 21].

Image recognition is probably realized in V1 by wavelet interference and in its connections to V2 that is an inversion of Gabor transform [23]. Since images are being Gabor-transformed, the vector v contains, instead, the Gabor coefficients of the transform for every point, orientation and frequency. Recognition and reconstruction are described in Dirac notation as [21]:

$$|I_{V2}\rangle \cong \sum_k |v_k^{\text{V1}}\rangle\langle v_k^{\text{V1}}|I_{\text{gabor}}\rangle \tag{1}$$

where I_{gabor} is Gabor-filtered quantum-encoded retinal input (thus, the Gabor coefficients), which can be described as $I_{\text{gabor}} = \langle \psi_{f,\theta,t} | I_{\text{input}} \rangle$ where f, θ and t are different frequencies, orientation and spatial center points, respectively. The result is a reconstruction, which is projected to V2 secondary visual cortex. Just for a clarity of terminology, which will be used in section "Image recognition tests", $\langle v_k^{\text{V1}} | I_{\text{gabor}} \rangle$ from Eq. 1 is being termed as order parameter, which shows the magnitude of correlation between the input and the prototypes k.

Cross-talk will occur if $v^T v \neq I$, which is the identity matrix. This is generally the case due to non-orthogonality. Here, we assumed that v^k is normalized with zero mean. In order to ensure orthogonality, a matrix 'a' is multiplied, giving $av^T v = I$.

To orthogonalize, then, is to satisfy the following relation:

$$v^{k'}(v^k)^+ = \delta_{kk'}(\text{Kronec ker}) \tag{2}$$

with

$$(v^k)^+ = \sum_{k'} a_{kk'} \left(v^{k'}\right)^T \tag{3}$$

Substituting Eq. 3 to 2, we obtain $[a_{kk'}] = \text{inv}(v^T v)$, where inv( ) is the generalized Moore–Penrose inverse. Thus, from here onwards, v^T from $vv^T = H$ is replaced by $(v^k)^+$. But orthogonalization is not biologically plausible and quite artificial. Cross-talk for pixel level associative memory is severe due too much redundant similarities [20], causing the vectors of distinct objects to be inclined to each other. Though a winner-take-all method for recognition and reconstruction can be used, in order to abide to constraints of quantum holography [18], a more elegant method to ensure as much orthogonality as possible is needed (Fig. 1).

Gabor Wavelet Family

Our visual world is highly structured and predictable. It is not random but, on the contrary, follows a set of strict rules. Thus, our visual system evolved in order to exploit this redundancy to extract useful information with least processing cost. One of them is being scale-invariant, where statistical properties remain unchanged with magnification. Using sparse self-similar wavelet code, Field [6] reconstructed a synthetic waveform, which is scale-invariant. Gabor function is a self-similar basis function, which means that it is a band-passed localized function that varies with dilation, translation and rotation. Very few number of basis functions is required (sparse) for reconstruction as shown in Fig. 2, given the right coefficients.

We follow the biological Gabor family derived by [11] abiding to the 4 constraints:

- The aspect ratio of the elliptical Gaussian envelope is 2:1
- Wave's propagating direction along short axis of the envelope
- Bandwidth of 1.5 octaves. (Though for biological receptive fields, the bandwidth may span from 1 octave to 2 octaves)
- Zero mean $\sum_{x,y} \psi(x,y) = 0$

Fig. 1 **a** is the actual image used. **b** is a reconstructed image from continuous sampling. **c** is a reconstructed image for sampling outside minimum sampling rate, thus, the poorer quality. **d** shows reconstruction for discrete sampling. This is the condition we will be using for our experimental tests. **e** is a result from the summation of coefficients instead of Gabor elementary function. It resembles the original image and is computationally cheap

(a) Reconstruction before iteration **(b)** After iteration (50 iterations)

Fig. 2 **a** shows the reconstructed image from mere summation of Gabor function. Iteration is performed on **b** to produce sharper and better image. Images are sampled and reconstructed at low frequencies within 2 octaves

Where the Gabor wavelets are described by

$$\psi(x,y,\omega_0,\theta) = \frac{\omega_0^2}{8k^2}\left[e^{-\frac{\omega_0^2}{8k^2}(4(x\cos\theta+y\sin\theta)^2+(-x\sin\theta+y\cos\theta)^2)}\right] \cdot \left[e^{i(\omega_0 x\cos\theta+\omega_0 y\sin\theta)} - e^{-\frac{k^2}{2}}\right] \quad (4)$$

where (x, y) is the center point of the wavelet, ω_0 the unit spatial frequency and θ the orientation.

There are 8 orientations between a range of 0 to π rad. For spatial frequency, there are 3 octaves, with every 1 octave containing 3 frequencies.

$$k = \sqrt{2\ln 2}\left(\frac{2^{\varphi}+1}{2^{\varphi}-1}\right)$$

where ϕ is the bandwidth (1.5 octave).

The sampling rate for the wavelets is $\approx 0.4\lambda$ where λ is the wavelength of the respective wavelet.

Gabor Reconstruction Process

Reconstruction is performed through $\sum_n C_n\psi(n)$ where C_n is the coefficient of the particular wavelet and $\psi(n)$ the Gabor elementary function. n is the parameter vector containing information on the orientation, spatial frequency and center point of the function. For a particular orientation and frequency, since the functions are shift invariant, the reconstruction from that particular elementary function is performed through convolution as $f(\tau) = \int c(t)\psi(\tau - t)\partial t$. The coefficients determine the amplitude of the elementary functions, which is summed to obtain the output. The τ and t are 2 element vectors of the 2-D coordinate. c is the coefficient for a particular function shifted to t. $f(\tau)$ is the reconstruction result. In order to transform images to Gabor coefficients, the image is filtered through the Gabor-shaped receptive field. In contrast to reconstruction, the mathematical definition for transformation is as follows $\int f(\tau)\psi(\tau - t)\partial\tau$, where f is the input image and ψ the Gabor-like receptive field at a particular frequency and orientation. Given the definition for transformation and reconstruction, the path from retina to V2 (omitting Difference of Gaussian transformation that is performed at the ganglion cells and lateral geniculate nucleus LGN) can be described by:

$$f_{\text{output}}(t) = \iint f_{\text{input}}(a)\psi(a-\tau)\psi(t-\tau)\partial a\partial\tau$$

Reconstruction Improvement Under Iteration

Reconstruction is effected by sampling. Under low frequency sampling and low rate, iterative method can assist in reconstruction with the motivation that information is contained in neighboring Gabor wavelet due to redundancy, which is not exploited by mere superposition of Gabor function multiplied by its direct Gabor coefficients.

After iteration, the Gabor coefficients will converge to a state where they sufficiently describe the intended image. From Fig. 2, after iteration, the active points for that particular frequency and orientation reduced since those regions are sufficiently described by other elementary functions. For example, in Fig. 3 b1 and b2, the nose part is eliminated since the role of description is performed by other functions.

The iterative method is described by the dynamic equation:

$$\sum_{x,y}\left(I(x,y) - \sum_{n\neq i}\psi_n(x,y)C_n\right)\psi_i(x,y) - \alpha C_i = \frac{\partial C_i}{\partial t}$$

where $I(x,y)$ is the reconstruction result from Gabor elementary functions for the original image, ψ_n is the Gabor elementary function index n as described in Eq. 4 and C_n is the Gabor coefficients for Gabor function index n, which is

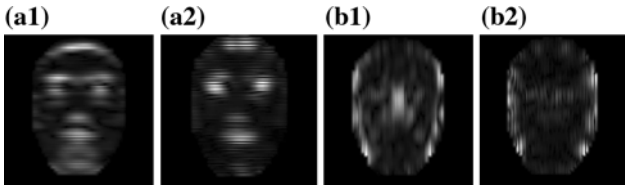

Fig. 3 Images show the active regions for a particular frequency and orientation Gabor coefficients before and after iteration (**a1** corresponds to **a2**, and **b1** to **b2**). Relative magnitude of activity is shown as pixel brightness value

an index number for different orientations, spatial frequencies and wavelet center points.

The system can be visualized as evaluating the contribution of a particular Gabor function after the superposition of all other functions in order to achieve maximum resemblance to the original image. Basically, the system will approach a state where there are just sufficient Gabor elementary function to describe the given image. The redundancy minimization effect needs to be explored further and with links to biologically plausible method.

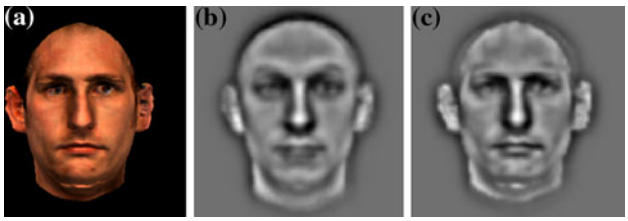

Fig. 6 Reconstruction result for facial recognition (where **a** is the input image) for different person using **b** all frequencies and **c** 1 optimum frequency which is 8 unit frequency. The test is performed with 8 prototype images

Image Recognition Tests

Test on Different Poses

Using Gabor coefficients as prototype vectors, image recognition is performed using quantum associative network described in Eq. 1 on a face with different poses (Fig. 4).

Figure 5c shows the recognition result for 7 poses each. Although result is better compared to pixel level correlation, cross-talk is still an issue even with only 7 faces. But

Fig. 4 Example of prototype images of same person with different poses

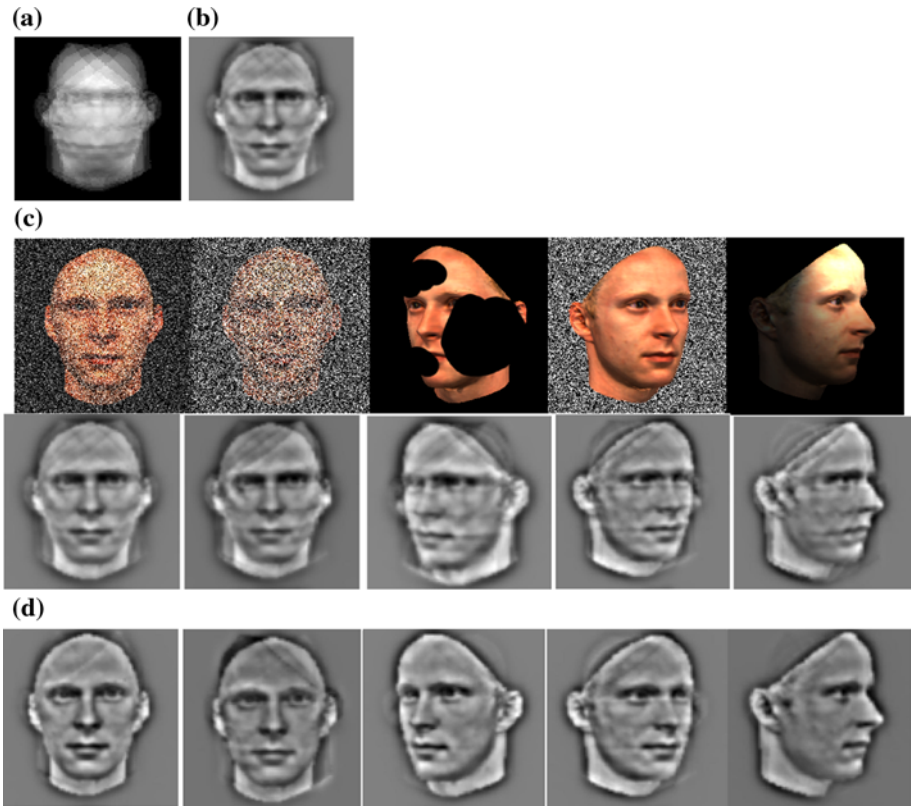

Fig. 5 **a** shows the pixel level recognition and reconstruction which shows severe cross-talk, whereas **b** shows result from Gabor wavelet sampled at all frequencies (4 octaves). For **c**, it shows the distorted input's (*upper*) Gabor recognition and reconstruction result (*lower*). Instead of correlating with coefficients at all frequencies, we choose coefficients from one optimum frequency, where the improved result is shown in **d**. Reconstruction is performed through mere summation of coefficients due to its high processing speed

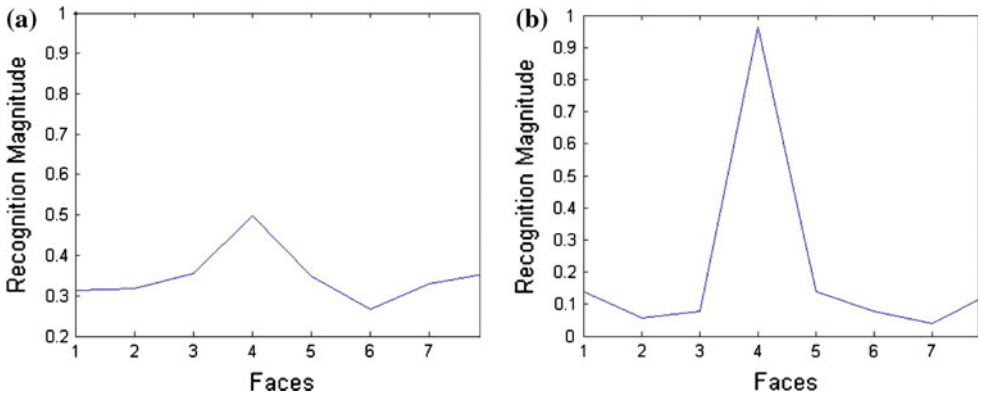

Fig. 7 The order parameter of the input image when sampled at every frequency (**a**) and with 1 optimum frequency (**b**). Order parameter means the magnitude of correlation between the input image with the images in the database, which there are 8 for the graphs

by recognition using coefficients from 1 optimum frequency (in our case is unit 8 frequency), the result is significantly improved. This is because for that particular frequency, information is most distributed, thus, more distinction between images.

For Face Recognition

Applying the same condition as the previous test, for face recognition (front-face pose), the optimum sampling frequency also performs reasonably well as shown in the result below (Figs. 6, 7).

Likewise, result for the optimum frequency is significantly better.

When sampled at 1 optimum frequency, the prototype faces seem nearly orthogonal, which produces better reconstruction due to less contribution from other erroneous database images. By choosing Gabor coefficients with least activity for correlation, the result is further improved. The points of least activity are determined from a sample of 99 faces. It is determined through the maximum absolute value of the normalized vector of the activities for a particular point.

$$\text{Inactivity magnitude}(x, y) = \max\left(\frac{V}{|V|}\right)$$

This is because if over a span of 99 samples, if only a few samples induce activity on a point, the maximum value of the normalized vector will be greater compared to those

Fig. 8 Example of points of LEAST activity for 3 orientations

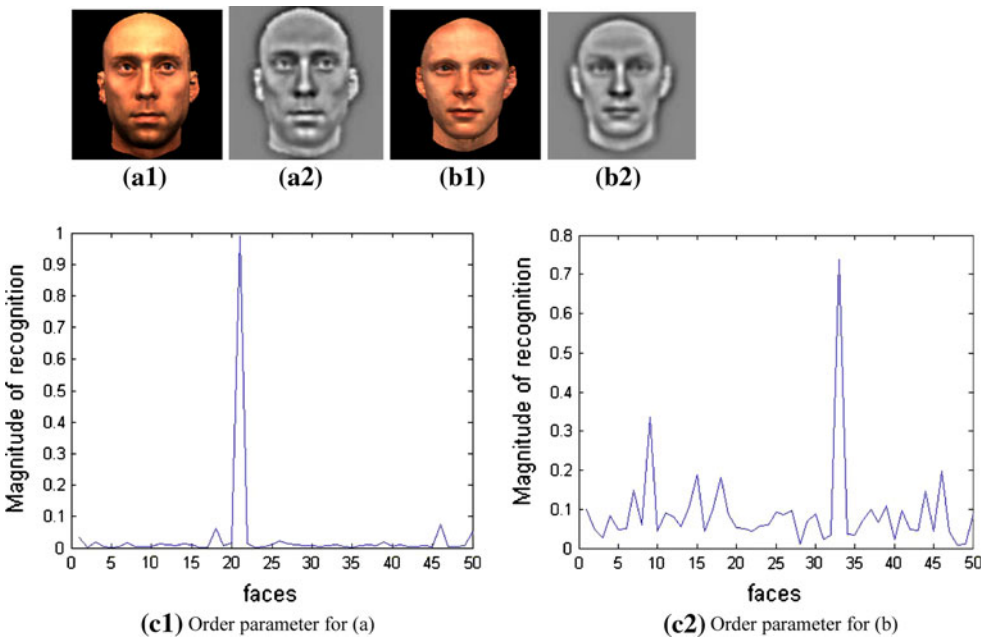

Fig. 9 Recognition using points of least activity. As shown in **c1** and **c2**, the order parameters are extremely distinctive in terms of the recognized face

(c1) Order parameter for (a)

(c2) Order parameter for (b)

with more activities. Points with inactivity magnitude above certain threshold (as shown in Fig. 8) are chosen for correlation in quantum associative network to find the order parameter (Fig. 9).

Conclusion

The paper reports on the application of Gabor transform in image recognition using the associative net-processing brain model from [18, 19]. Results show recognition that can provide acceptable distinction between different faces and also different poses using Gabor coefficients without artificial procedures like orthogonalization, which is superior to that of using mere pixel level correlation. Besides, Gabor coefficients provide room for exploitation for enhancement purposes, though analysis has yet to be performed here. For further research work, work will be done on the co-operation of Gabor functions to give rise to familiarity, which at the same time, exploiting the information maximization property of Gabor wavelet, which has yet to be touched upon yet for this paper.

References

1. Bastiaans M. Gabor's expansion of a signal into gaussian elementary signals. Proc IEEE. 1980;68:538–9.
2. Bell A, Sejnowski T. An information-maximization approach to blind separation and blind deconvolution. Neural Comput. 1995;7:1129–59.
3. Bell A, Sejnowski T. The "independent components" of natural scenes are edge filters. Vision Res. 1997;37:3327–38.
4. Buser P, Imbert M. Vision. Cambridge, MA: MIT Press; 1992.
5. Daugman J. Complete discrete 2-D gabor transforms by neural networks for image analysis and compression. IEEE Trans Acoust. 1988;36:1169–79.
6. Field DJ. Scale-invariance and self-similar 'wavelet' transforms: an analysis of natural scenes and mammalian visual systems. In: Wavelets, fractals and fourier transforms: new development and new applications. Oxford: Oxford University Press; 1993.
7. Grossmann A, Kronland-Martinet R, Morlet J. Reading and understanding continuous wavelet transforms. In: Combes JM, Grossman A, Tchamitchian P, editors. Wavelets. Berlin: Springer; 1989. p. 2–20.
8. Hubel D, Wiesel T. Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. J Physiol (Lond). 1962;160:106–54.
9. Jones J, Palmer L. An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. J Neurophysiol. 1987;58:1233–58.
10. Kyrki V, Kamarainen JK. Kälviäinen.: simple Gabor feature space for invariant object recognition. Pattern Recognit Lett. 2004;25:311–8.
11. Lee T. Image representation using 2D Gabor wavelets. IEEE Trans Pattern Anal Mach Intell. 1996;18(10):1–13.
12. Linsker R. Self-organization in a perceptual network. Computer. 1988;21(3):105–17.
13. Marcelja S. Mathematical description of the responses of simple cortical cells. J Opt Soc Am. 1980;70:1297–300.
14. Okajima K. The Gabor function extracts the maximum information from input local signals. Neural Netw. 1998;11:435–9.
15. Olshausen B, Field D. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. Nat Lett. 1996;381:607–9.
16. Olshausen B, Field D. Sparse coding with an overcomplete basis set: a strategy employed by V1? Vision Res. 1997;37:3311–25.
17. Palmer S. Vision science (photons to phenomenology). Cambridge, MA: MIT Press; 1999.
18. Peruš M. A synthesis of the Pribram Holonomic theory of vision with quantum associative nets after pre-processing using ICA and other computational models. Int J Comput Anticipatory Syst. 2001;10:352–67.
19. Peruš M. Multi-level synergetic computation in brain. Non-linear phenom. Complex Syst. 2001;4:157–93.
20. Peruš M, Bischof H, Caulfield HJ, Loo CK. Quantum-implementable selective reconstruction of high-resolution images. Appl Opt. 2004;43(33):6134–8.
21. Peruš M, Bischof H, Loo CK. Bio-computational model of object-recognition: quantum Hebbian processing with neurally shaped Gabor wavelets. Biosystems. 2005;82:116–26.
22. Pollen D, Ronner S. Phase relationships between adjacent simple cells in the visual cortex. Science. 1981;212:1409–11.
23. Pribram KH. Brain and perception (holonomy and structure in figural processing). Hillsdale, NJ: Lawrence Erlbaum Associates; 1991.