

RESEARCH ARTICLE

Open Access

Neural computation of visual imaging based on Kronecker product in the primary visual cortex

Zhao Songnian^{1†}, Zou Qi^{2†}, Jin Zhen³, Yao Guozheng⁴, Yao Li^{5*}

Abstract

Background: What kind of neural computation is actually performed by the primary visual cortex and how is this represented mathematically at the system level? It is an important problem in the visual information processing, but has not been well answered. In this paper, according to our understanding of retinal organization and parallel multi-channel topographical mapping between retina and primary visual cortex V1, we divide an image into orthogonal and orderly array of image primitives (or patches), in which each patch will evoke activities of simple cells in V1. From viewpoint of information processing, this activated process, essentially, involves optimal detection and optimal matching of receptive fields of simple cells with features contained in image patches. For the reconstruction of the visual image in the visual cortex V1 based on the principle of minimum mean squares error, it is natural to use the inner product expression in neural computation, which then is transformed into matrix form.

Results: The inner product is carried out by using Kronecker product between patches and function architecture (or functional column) in localized and oriented neural computing. Compared with Fourier Transform, the mathematical description of Kronecker product is simple and intuitive, so is the algorithm more suitable for neural computation of visual cortex V1. Results of computer simulation based on two-dimensional Gabor pyramid wavelets show that the theoretical analysis and the proposed model are reasonable.

Conclusions: Our results are:

1. The neural computation of the retinal image in cortex V1 can be expressed to Kronecker product operation and its matrix form, this algorithm is implemented by the inner operation between retinal image primitives and primary visual cortex's column. It has simple, efficient and robust features, which is, therefore, such a neural algorithm, which can be completed by biological vision.
2. It is more suitable that the function of cortical column in cortex V1 is considered as the basic unit of visual image processing (such unit can implement basic multiplication of visual primitives, such as contour, line, and edge), rather than a set of tiled array filter. Fourier Transformation is replaced with Kronecker product, which greatly reduces the computational complexity. The neurobiological basis of this idea is that a visual image can be represented as a linear combination of orderly orthogonal primitive image containing some local feature. In the visual pathway, the image patches are topographically mapped onto cortex V1 through parallel multi-channels and then are processed independently by functional columns. Clearly, the above new perspective has some reference significance to exploring the neural mechanisms on the human visual information processing.

* Correspondence: yaoli@bnu.edu.cn

† Contributed equally

⁵State Key Laboratory of Cognitive Neuroscience and Learning, School of Information Science and Technology, Beijing Normal University, Beijing 100875, China

Background

Human vision can be considered as a perfect image information processing device, it can easily recognize object's position, size, and orientation, pose in space, and so on. For a long time, visual scientists, computational neuroscientists, image processing experts and computer vision researchers make great effort to explore the neural mechanism of humans' remarkable visual abilities or how the retina image is represented in the primary visual cortex, which is related with the following two questions: what kind of neural computation is actually performed by the primary visual cortex, and how this is described mathematically.

It is well known that there is a one-to-one topographical mapping between retina and cortex V1, which determines projecting relations in visual space and represents some transformations from retina to cortex V1 [1-16]. Currently, it is believed that responses of neurons in cortex V1 can be simulated by a set of tiled spatio-temporal filters array. So the function of cortex V1 is to make a spatial local Fourier Transform. Theoretically, these filters involve many processes about spatial frequencies, orientations, motion and velocities (frequencies in temporal space) [17-23].

Is this notion consistent with the actual biological visual information processes? Research in neurobiology indicates that the metabolism and decay of neurons do not affect the visual function. Every neuron performs a simple ON-OFF function and transfers the information through spikes. So a dead neuron can easily be replaced by other nearby neurons [3,24,25]. In case of a complicated function, this replacement would be difficult. Therefore, the simplicity of algorithms not only reduces error rates to the minimum, but also guarantees repeatability and stability, i.e. robustness.

Complex computations can be carried out by parallel computations of neuronal groups with high efficiency. So we believe that the actual computations in retinal image must be simple, repeatable and robust and be performed by individual neurons at the system level and by neuronal groups with high efficiency. They are obvious requirements for neural computations in V1.

Then, how is the topographical mapping from retina to V1 be realized? Many neurobiological experiments and visual computational models show that when every primitive (edge, corner and contour) in the visual image finds matches in the receptive fields densely distributed on V1, only the neurons whose frequencies and orientations are similar to those of the primitive fire [26]. Therefore, the patterns of the fired neurons correspond to the primitives in the visual image, which may be represented by a topographical mapping, reflecting the adjustment of the visual image to neurons in V1, and

reflecting distributive and parallel visual information processing between retina and V1[27,28]. In this paper, we discuss the mathematical representation of this information processing and use the normalized matching measure (i.e. energy function) to measure the matching extent [29].

This paper proposes a model of neural information processing based on the topographical map in place of Fourier Transform. In this model, the functional columns are not considered as sets of tiled filters, but basis elements of the visual information, including orientation selectivity and feature matching [2,30-32]. The visual image carried by spike trains is processed by Kronecker product with functional columns in V1. So synchronously parallel computations on the whole image can be performed by receptive field-to-receptive field rather than by pixel-to-pixel, and be represented by Kronecker product between matrixes. Its complexity is greatly reduced as compared with Fourier transform and other matrix computations [33,34]. What is more, this algorithm can simulate stimulations of the elements in the visual image to cortical neurons as the embodiment of simple neural functions. The aggregative computation based on simple functions is one of plausible approaches of the visual cortex in realizing topographical mapping.

Numerical simulations are carried out to justify the above notions. In our experiments, receptive fields of neurons in V1 are simulated by hierarchical Gabor functions [35-40]. Visual image is the feature image of Lenna processed by pre-processing (filtered) in front of the pathway. Results of our algorithm are consistent with theoretical expectations.

Results

The visual pathway ('what' pathway [41]) from retina across LGN to V1 is modelled. The following discussions are focussed on: 1. the optimal detection in V1 of retinal image $R(x, y)$; 2. the optimal matching between $R(x, y)$ and firing pattern of neuronal groups in V1; 3. Kronecker product obtained by optimal detection and matching; 4. determination of kernel function $G(x, y)$; 5. realization of Kronecker product; 6. numerical experiments and discussions.

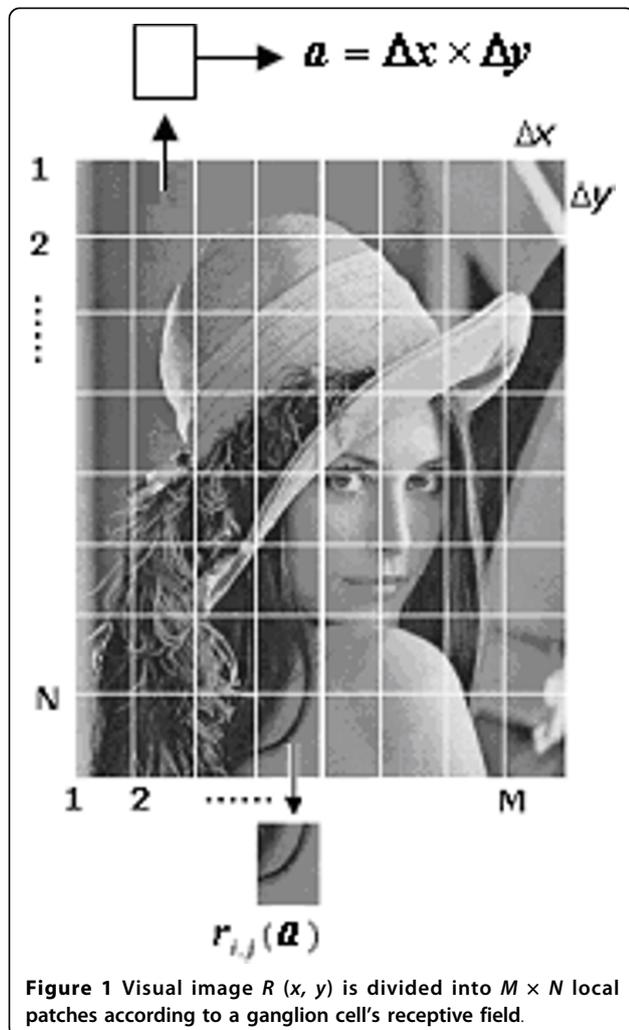
Image division and cortical response

As a first step of visual perception, external stimuli form a retinal image $R(x, y)$. In the retina, a ganglion cell receives inputs from about $10^3 \sim 10^4$ photo-receptor cells. If the size of a ganglion cell's receptive field is $a = \Delta x \times \Delta y$, the total imaging area of $R(x, y)$ in the retina is A , which is divided into $M \times N$ patches with the same size a as that of a ganglion cells receptive field, namely $A = (M \times N)a$. Here a is referred to as sub-

region, therefore, a indicates only its size; and the partially small image covered by the area a is called image unit or image primitive, and expressed by $r(a)$. In this way, a patch or the image primitive at the i line and the j column can be denoted as $r_{i,j}(a)$ ($i = 1, 2, \dots, M$; $j = 1, 2, \dots, N$), and the entire image included in the area A is also expressed by $R(A)$, as shown in Figure 1. That is to say, the whole image $R(A)$ can be also formed from $(M \times N)$ patches by means of orderly putting patches together, i.e. $R(A) = \sum_i^M \sum_j^N r_{i,j}(a)$.

Figure 1. Visual image $R(x, y)$ is divided into $M \times N$ local patches according to a ganglion cell's receptive field.

This division involves two aspects. First, every patch $r_{i,j}(a)$ contains local features (such as shapes of receptive fields and orientations) at (i, j) with area $(\Delta x \times \Delta y)$, while a pixel at (i, j) only contains intensity information. Second, in the parallel multi-channel vision system,



every channel only deals with a local patch. Obviously this division is consistent with parallel and multi-channel properties of a vision system, and in a topographical mapping any patch can be located in the retina, so the neural processing based on functional columns can be realized and the corresponding mathematical description is possible. A patch $r_{i,j}(a)$ will activate the corresponding ganglion cell and output a coded firing spike train. This is transferred across LGN into V1 as a topographical mapping. Then, the firing spike train is decoded and the image represented by $r_{i,j}(a)$ is restored. A Kronecker product of the restored image with functional columns $B_{k,l}(s)$ in V1 [40,41,33] leads to firing of neurons in receptive fields having similar orientation and bandwidth (that is so-called firing under a preferential stimulus). We denote the firing pattern of one singular neuron (or simple cell) as $\phi_{i,j}(b)$, where b denotes the area of the receptive field of the neuron. Since, A , the retinal image area, will be enlarged in the cortex V1 [42], for simplicity, with a magnification factor $z = 2^h$ ($h = 0, 1, 2, \dots$); if the area of image on the cortex V1 is denoted by B , we have $B = zA = 2^h A$, $b = za = 2^h a$. In this way, the spatial sum of all signals $\phi_{i,j}(b)$ in an orderly manner will form the overall firing pattern, that is $\Phi(B) = \sum_i^M \sum_j^N \phi_{i,j}(b)$, which represents a reconstruction of the retinal image $R(A)$ on V1.

Optimum detection in V1

For a single neuron in the functional columns, its receptive field consists of orientation, band-pass and spatial location [43,44], with a strong selectivity with respect to visual image $R(x, y)$ topologically mapped from the retina. When some patch $r_{i,j}(a)$ in visual image $R(x, y)$ shows properties at specific orientation and specific frequency similar to some receptive field $g_{i,j}(b)$, the corresponding neuron will respond strongly. In other words, when a local patch coincides with the receptive field of a neuron at its sensitive orientation and sensitive frequency, the neuron fires most strongly, which means a detection and matching of functional columns in V1 to local retinal patch's features $r_{i,j}(a)$, as a random process, in other words, which is a detection and matching in a random process. Therefore, this process can be expressed as [45]

$$g_{i,j}(b) = \alpha_{i,j} r_{i,j}(2^h a) + n_{i,j} \quad i, j = 1, 2, \dots \quad (1)$$

where $r_{i,j}(2^h a)$ is a patch from input image $R(A)$ containing basic features such as orientation, edge, and contour within the receptive field. Here $n_{i,j}$ is Gaussian white noise with zero mean and variance σ_n^2 . It is assumed here that the noise environment in different visual pathways is the same, i.e. $n_{i,j}$ is the same. $\alpha_{i,j}$ is

weight coefficient that can reflect excitation strength of an image primitive to neuron $g_{i,j}(b)$. We want to know what value of $\alpha_{i,j}$ can strongly activate neuron $g_{i,j}(b)$ to fire. Obviously, $\alpha_{i,j}$ is an unknown estimated coefficient. If we make K observations on the random process (1), in other words, we carry out K sampling on the random process (1), then formula (1) can be written in the following general form

$$[g_{i,j}(b)]_k = \alpha_{i,j} [r_{i,j}(2^h a)]_k + [n_{i,j}]_k \quad k = 1, 2, \dots, K \quad (2)$$

According to maximum likelihood estimation and least mean square error rule [45,46]

$$\begin{aligned} & \left. \frac{\partial}{\partial \alpha_{i,j}} \ln P(g_{i,j}(b) | \alpha_{i,j}) \right|_{\alpha_{i,j} = \hat{\alpha}_{ML}} \\ &= - \frac{\partial}{\partial \alpha_{i,j}} \sum_{k=1}^K \left[\frac{([g_{i,j}(b)]_k - \alpha_{i,j} [r_{i,j}(2^h a)]_k)^2}{2\sigma_n^2} \right] \bigg|_{\alpha_{i,j} = \hat{\alpha}_{ML}} \quad (3) \\ &= \frac{1}{\sigma_n^2} \sum_{k=1}^K ([g_{i,j}(b)]_k - \alpha_{i,j} [r_{i,j}(2^h a)]_k) [r_{i,j}(2^h a)]_k \bigg|_{\alpha_{i,j} = \hat{\alpha}_{ML}} = 0 \end{aligned}$$

The optimum estimation of $\alpha_{i,j}$ is

$$\hat{\alpha}_{i,j} = \frac{\sum_{k=1}^K [g_{i,j}(b)]_k [r_{i,j}(2^h a)]_k}{\sum_{k=1}^K [r_{i,j}(2^h a)]_k^2} = c \sum_{k=1}^K [g_{i,j}(b)]_k [r_{i,j}(2^h a)]_k \quad (4)$$

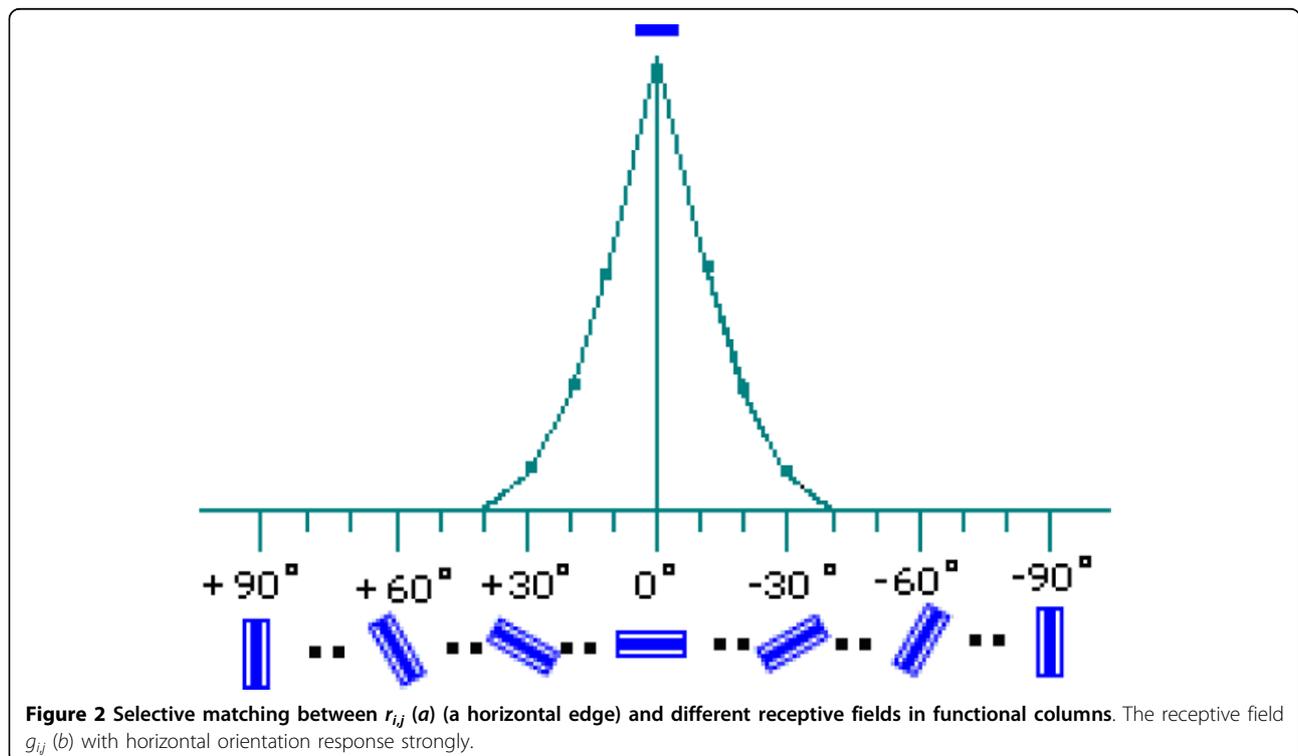
where $c = \frac{1}{\sum_{k=1}^K [r_{i,j}(2^h a)]_k^2}$ is a ratio coefficient.

We can see $\alpha_{i,j}$ reaches its optimum value $\hat{\alpha}_{i,j} = \hat{\alpha}_{ML} = 1$ when image patch $r_{i,j}(a)$ coincides with receptive fields feature $g_{i,j}(b)$, i.e. $\alpha_{i,j} r_{i,j}(2^h a) = g_{i,j}(b)$. So, $\hat{\alpha}_{i,j}$ can be taken as the measure for matching extent between neuronal receptive field $g_{i,j}(b)$ and patch $r_{i,j}(a)$ in image $R(A)$.

Above process is illustrated in Figure 2, where the local patch is a horizontal edge and an optimal matching is found in receptive fields of horizontal orientation with a strong response. No optimal matches are found in receptive fields of other orientations, so we have weak responses.

Figure 2. Selective matching between $r_{i,j}(a)$ (a horizontal edge) and different receptive fields in functional columns. The receptive field $g_{i,j}(b)$ with horizontal orientation responses strongly.

It is worthy noting that the multi-scale processing function of a visual pathway is to guarantee a clear image at a proper resolution in V1. So, the optimal matching is related with some resolution. When the scale or resolution changes, the optimal matching may concentrate on different extents of details, or may include or exclude some details, as is determined by the circumstances when the vision system is "observing" the world.



The whole matching between the retinal image and receptive field patterns in V1

In the previous section, we have described the local matching or detection (formula 3). Next we will analyze the image matching as a whole, to obtain a mathematical representation of the neural firing pattern in V1. It is understood that the topographical mapping from retinal image $R(x, y)$ to V1 is actually a stimulating process of $R(x, y)$ with respect to simple selective cells with features of orientation, bandwidth and so on. Receptive fields $G(x, y)$ of activated neurons are combined to form the global firing pattern $\Phi(x, y)$ in the cortex, as the responding process of simple cell groups to the visual image. So, the firing process can be regarded as a global matching between $G(x, y)$ and $R(x, y)$ at the system level, and eventually responding pattern $\Phi(x, y)$ will be formed. Many methods can be used to measure the extent of matching. However, in order to ensure a minimal reconstruction error, we adopt the following measure:

$$\varepsilon = \iint_B [G(x, y) - R(x, y)]^2 dx dy = \min \quad (5)$$

where B is the imaging region in V1.

$$\iint_B G(x, y)R(x, y) dx dy \leq \sqrt{\iint_B [G(x, y)]^2 dx dy \iint_B [R(x, y)]^2 dx dy} \quad (6)$$

Let $\lambda_{RG} = \iint_B R(x, y)G(x, y) dx dy$, $\lambda_{MAX} = \iint_B R(x, y)^2 dx dy$. In the optimal matching in region B , we will have $R(x, y) = G(x, y)$, as is the case when the equal operation is adopted in formula (5) and λ_{RG} reaches its maximum value λ_{MAX} . Therefore, we can define the normalized matching coefficient ρ_{rg} as follows:

$$\rho_{rg} = \frac{\lambda_{RG}}{\lambda_{MAX}} = \frac{\iint_B R(x, y)G(x, y) dx dy}{\iint_B [R(x, y)]^2 dx dy} \quad (7)$$

Obviously when $\rho_{rg} = 1$, $G(x, y)$ and $R(x, y)$ reaches a complete match. In other words, the receptive fields of all activated neurons in V1 are combined to form the same responding pattern $\Phi(x, y)$ as the whole visual image. It shown that this process can be mathematically described as follows by multiplication of $R(x, y)$ and $G(x, y)$

$$\Phi(x, y) = R(x, y) \otimes G(x, y) \quad (8)$$

It can be seen later, $R(x, y)$ and $G(x, y)$ can be expressed as matrix form, for this reason, the formula (8) is essentially an inner product operation, it is not only more elegant on the mathematical form, but also more clear on the neurobiological significance.

Additionally, we can see the normalized matching coefficient ρ_{rg} is equivalent to \hat{a} in the previous section.

Determination of integral kernel function

In formula (8), the role played by receptive fields $G(x, y)$ of cortical neurons is the same as the integral kernel in wavelet transform [46,47]. It can be described as oriented and bandwidth two-dimensional Gabor function $G(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$ [34-39]

$$G(x, y)_{\lambda, \sigma, \theta, \phi, \gamma} = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \phi\right) \quad (9)$$

$$x' = x \cos \Theta + y \sin \Theta$$

$$y' = -x \sin \Theta + y \cos \Theta$$

where γ is the ratio of the length in the major axis direction to that of in minor axis direction, usually set to a constant 0.5; σ is derivative of Gauss, determining the size of receptive fields; ϕ is the phase, when $\phi = 0; \pi$; $G(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$ is symmetric about the origin; when $\phi = -(\pi/2); (\pi/2)$, $G(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$ is anti-symmetric about the origin; Θ is the optimal orientation, and λ is the wavelength. These arguments should be determined by experimental results from morphology and biophysics, but the exact data are not available so far [48]. One plausible way is to set the arguments according to input image features in an input-driven topological mapping [2]. This will be explained in the last section.

Substituting (9) into (8) and considering the cortical responses to orientation and bandwidth properties, we replace $\Phi(x, y)$ with $\Phi(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$

$$\Phi(x, y)_{\lambda, \sigma, \theta, \phi, \gamma} = R(x, y) \otimes G(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$$

$$= R(x, y) \otimes \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \phi\right) \quad (10)$$

It is also the inner product, because the formula can be also expressed as matrix form. The formulae (8) and (10), as the inner product, which shows such an important neurobiological fact, that is, in the visual pathway topographical mapping indicates accurate positioning of retinal image in the visual cortex, therefore, these primitives can only respectively activate cells which are in the corresponding locations in visual cortex. Since it is a one to one excitation, scanning and convolution is no longer needed.

Comparison of inner product with convolution

For comparison, the convolution operation may be expressed as follows

$$\Phi(x, y)_{\lambda, \sigma, \theta, \phi, \gamma} = \iint R(u, v)G(x-u, y-v)_{\lambda, \sigma, \theta, \phi, \gamma} dudv \quad (11)$$

We know that convolution and cross-correlation operations are essentially filtering operations in the

frequency domain, which is not needed for V1, because such a filtering operation would lead to loss of high- and low-frequency information from the retinal picture. The second reason is that the scan process in such operations (convolution and cross-correlation) is a calculation with a high cost (see the section of discussion in this paper, for detail), in which, $G(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$ should be taken as the template to scan the whole image $R(x, y)$ from top to bottom and from left to right. Obviously, it is not an effective method, because this scanning will cause too many responses of corresponding cells and the energy cost is too great.

Mathematically, the discrete convolution of formula (11) can be expressed as

$$\Phi(x, y)_{\lambda, \sigma, \theta, \phi, \gamma} = \sum_k \sum_l R(k, l) G(x - k, y - l)_{\lambda, \sigma, \theta, \phi, \gamma} \quad (12)$$

For some point (x_0, y_0) in an image, $G(x_0 - k, y_0 - l)_{\lambda, \sigma, \theta, \phi, \gamma}$ is moved around on the image (by changing k and l) to realize an optimal match between $\Phi(x_k, y_k)$ and $R(x, y)$. The matched simple cell then is activated. Figure 3 shows an example, in which the visual image $R(x, y)$ is a small horizontal or vertical line. Obviously, a horizontal line in $R(x, y)$ excites numerous cells whose receptive fields have an orientation similar to a horizontal line, and a similar effect occurs for a vertical line. Such a calculation comes with a high costs in time and complexity.

Figure 3. Convolution operation for horizontal lines (A) and vertical lines (B) in image patches in $R(x, y)$.

The use of the inner product reveals that neuron firing caused by a visual stimulus is in fact a simple function. This is a simpler neural computation than the cross-correlation and convolution operations, because it needs only multiplication between corresponding pixels of $\Phi(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$ and $R(x_k, y_k)$. It is worth pointing out that the product $R(x_k, y_k) \Phi(x_k, y_k)_{\lambda, \sigma, \theta, \phi, \gamma}$ means that the retinal image $R(x_k, y_k)$ excites all cortical cells and forms a global

activity pattern $\Phi(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$ in V1. Different visual stimuli will excite and form different activity patterns corresponding to that stimulus; the differences in activity patterns occur only in a topographically connected weight coefficient of the pixels of $R(x_k, y_k)$ with the corresponding pixels of $\Phi(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$ in a fully mapping neural computation. Generally, the weight coefficients corresponding to detailed image information are much smaller than those corresponding to contour and edge information. The intensity of the spike firing of simple cells excited by the details of the stimulus is also weaker than the intensity corresponding to contours and edges.

The inner product $\sum_k \sum_l R(x_k, y_l) G(x_k, y_l)_{\lambda, \sigma, \theta, \phi, \gamma}$ in equation (8) reveals the collective calculation of a simple neuronal "on" or "off" function. From this, we can see that the calculation of the inner product is very well suitable to the visual system in that it satisfies the prerequisites of efficiency, simplicity, and robustness and also provides an optimal means of detection under the condition of least-mean-square-error reconstruction.

In fact, formula (9) reflects a specific wavelet transform on retinal image $R(x, y)$ by basis function $G(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$. This formula reflects the neural firing stimulated by the retinal image at the system level. Next we will discuss how to process visual images according to this formula. Two important problems will be discussed, that is, how to divide visual image $R(x, y)$ according to structures and functions of the visual pathway and how to express the orientation selectivity of functional columns in V1 by two-dimensional wavelet function $G(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$.

Kronecker product in V1

Usually, visual image $R(x, y)$ is independently transferred to LGN through 1 million ganglion cells, and then reaches the layers of $4C\alpha$ and $4C\beta$ in V1, and finally an image is reproduced in V1. Obviously, every image patch $r_{i,j}(a)$ is transferred through one channel. Suppose the number of

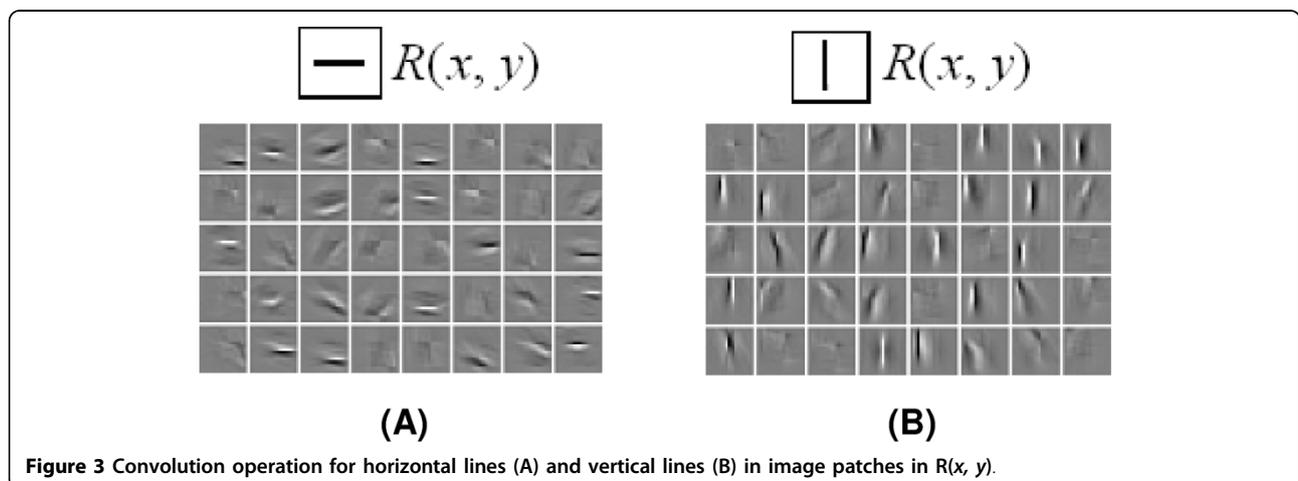


Figure 3 Convolution operation for horizontal lines (A) and vertical lines (B) in image patches in $R(x, y)$.

channels is $M \times N$, which means visual image $R(x, y)$ is divided into $M \times N$ units. As pointed out in previous section, each patch is assumed to have the same size as the receptive field of a ganglion cell, namely, $a = \Delta x \times \Delta y$. The area of the whole image is A . In every channel only one patch of $[R_{i,j}(a)]_{M \times N}$, with local features, is transferred. All $r_{i,j}(a)$ are added to form $[R_{i,j}(a)]_{M \times N}$.

$$[R_{i,j}(a)]_{M \times N} = \begin{bmatrix} r_{1,1}(a) & r_{1,2}(a) & \cdots & r_{1,N}(a) \\ r_{2,1}(a) & r_{2,2}(a) & \cdots & r_{2,N}(a) \\ \vdots & \vdots & \vdots & \vdots \\ r_{M,1}(a) & r_{M,2}(a) & \cdots & r_{M,N}(a) \end{bmatrix} \quad (13)$$

$$i = 1, 2, \dots, M, j = 1, 2, \dots, N.$$

The receptive field of a neuron distributed in V1 is $g_{i,j}(b)$. After being stimulated by $[R_{i,j}(a)]_{M \times N}$, a response pattern $[\Phi_{i,j}(b)]_{M \times N}$ (a $M \times N$ matrix) is formed from these receptive fields as:

$$[\Phi_{i,j}(b)]_{M \times N} = \begin{bmatrix} \phi_{1,1}(b) & \phi_{1,2}(b) & \cdots & \phi_{1,N}(b) \\ \phi_{2,1}(b) & \phi_{2,2}(b) & \cdots & \phi_{2,N}(b) \\ \vdots & \vdots & \cdots & \vdots \\ \phi_{M,1}(b) & \phi_{M,2}(b) & \cdots & \phi_{M,N}(b) \end{bmatrix} \quad (14)$$

$$i = 1, 2, \dots, M, j = 1, 2, \dots, N.$$

According to neurophysiology and neuroanatomy [43], cortical modules are densely distributed in V1, with approximately 10^3 modules; the area of each module is approximately $1.8 \text{ mm} \times 1.8 \text{ mm}$, containing two function columns for both left and right eyes. Thus, the area related with every function column $B_{k,l}(s)$ is $0.9 \text{ mm} \times 0.9 \text{ mm}$. At the system level, before adequate neurophysiological and neuroanatomical knowledge may be available, these function columns are assumed to have the same function and be composed of many receptive fields with different orientations and frequencies [3].

In this paper, the receptive fields of the function column can be represented as a matrix. As in Figure 4, each row of the matrix represents eighteen oriented receptive fields of the same type uniformly distributed from 0° to 180° . Each column of the matrix represents eight types of receptive fields (orthogonal Gabor of different frequencies) with a same orientation. So $[B_{k,l}(s)]_{K \times L}$ is made up of 144 elements $g_{i,j}(b)$ ($k = 1, 2, \dots, K; l = 1, 2, \dots, L$)

$$B_{k,l}(s) = [B_{k,l}(s)]_{K \times L} = \begin{bmatrix} g_{1,1}(b) & g_{1,2}(b) & \cdots & g_{1,18}(b) \\ g_{2,1}(b) & g_{2,2}(b) & \cdots & g_{2,18}(b) \\ \vdots & \vdots & \cdots & \vdots \\ g_{8,1}(b) & g_{8,2}(b) & \cdots & g_{8,18}(b) \end{bmatrix} \quad (15)$$

Figure 4. Functional columns as basic information processing units. (A) Eight representative types of receptive fields in function columns in V1; (B) Orientations range from 0° to 180° with a same interval of 10° ; (C) An example of receptive field calculated by formula (9).

Therefore, some edge or contour located at (i, j) in retinal image $[R_{i,j}(a)]_{M \times N}$ with area a will find a best match with the receptive field $g_{i,j}(b)$ of the same orientation and shape. When $\rho_{rg} = 1$, it means the patch at (i, j) in $[R_{i,j}(a)]_{M \times N}$ completely matches the cortical module $[B_{k,l}(s)]_{K \times L}$ with the specific orientation. Then the neuron is activated with the strongest response. The process of this image reconstruction is shown in figure 5.

Figure 5. Optimal matching between a patch (upper right corner of the hat) and receptive fields of specific orientations in cortical modules $[B_{k,l}(s)]_{K \times L}$.

When all the $M \times N$ patches in retinal image $[R_{i,j}(a)]_{M \times N}$ simultaneously (in a parallel manner) activate topographically corresponding neurons, response pattern $[\Phi_{i,j}(b)]_{M \times N}$ is formed in V1. At the system level, this process can be described by Kronecker Product.

Optimal responding of a neuron in V1 is actually reached through detection of cortical module $[B_{k,l}(s)]_{K \times L}$ with respect to a patch $r_{i,j}(a)$ in retinal image $[R_{i,j}(a)]_{M \times N}$, which can be represented as a product of them, i.e. $r_{i,j}(2^h a)[B_{k,l}(s)]_{K \times L}$

$$\begin{aligned} r_{i,j}(2^h a)[B_{k,l}(s)]_{K \times L} &= \\ &= [r_{i,j}(2^h a)g_{1,1}(b), r_{i,j}(2^h a)g_{1,2}(b), \dots, r_{i,j}(2^h a)g_{1,18}(b); \\ & r_{i,j}(2^h a)g_{2,1}(b), r_{i,j}(2^h a)g_{2,2}(b), \dots, r_{i,j}(2^h a)g_{2,18}(b); \dots; \\ & \dots r_{i,j}(2^h a)g_{8,1}(b), \dots, r_{i,j}(2^h a)g_{8,18}(b)] \max \\ & \text{for } \{r_{i,j}(2^h a)B_{k,l}(s) \mid k = 1, 2, \dots, 8; l = 1, 2, \dots, 18\} \end{aligned} \quad (16)$$

Where $|\max \text{ for } \{r_{i,j}(2^h a)B_{k,l}(s) \mid k = 1, 2, \dots, 8; l = 1, 2, \dots, 18\}$ means taking the maximum in $[r_{i,j}(2^h a)g_{1,1}(b), \dots, r_{i,j}(2^h a)g_{1,18}(b), \dots, r_{i,j}(2^h a)g_{8,18}(b)]$. When $i = 1, 2, \dots, M; j = 1, 2, \dots, N$, it is equal to Kronecker product between $[R_{i,j}(a)]_{M \times N}$ and $[B_{k,l}(s)]_{K \times L}$. The two matrixes are not necessarily of the same dimension. It can be represented as

$$\begin{aligned} [\Phi_{i,j}(b)]_{M \times N} &= [R_{i,j}(2^h a)]_{M \times N} \otimes B_{k,l}(s) = \\ &= \begin{bmatrix} r_{1,1}(2^h a) & r_{1,2}(2^h a) & \cdots & r_{1,N}(2^h a) \\ r_{2,1}(2^h a) & r_{2,2}(2^h a) & \cdots & r_{2,N}(2^h a) \\ \vdots & \vdots & \vdots & \vdots \\ r_{M,1}(2^h a) & r_{M,2}(2^h a) & \cdots & r_{M,N}(2^h a) \end{bmatrix} \otimes [B_{k,l}(s)] = \\ &= \begin{bmatrix} r_{1,1}(2^h a)B_{k,l}(s) & r_{1,2}(2^h a)B_{k,l}(s) & \cdots & r_{1,N}(2^h a)B_{k,l}(s) \\ r_{2,1}(2^h a)B_{k,l}(s) & r_{2,2}(2^h a)B_{k,l}(s) & \cdots & r_{2,N}(2^h a)B_{k,l}(s) \\ \vdots & \vdots & \vdots & \vdots \\ r_{M,1}(2^h a)B_{k,l}(s) & r_{M,2}(2^h a)B_{k,l}(s) & \cdots & r_{M,N}(2^h a)B_{k,l}(s) \end{bmatrix} \max \\ & \text{For } \begin{cases} \forall \{r_{i,j}(2^h a)B_{k,l}(s)\} \\ \tilde{a}_{ML} = 1, \text{ or } \rho_{\gamma\varphi} = 1 \end{cases} \end{aligned} \quad (17)$$

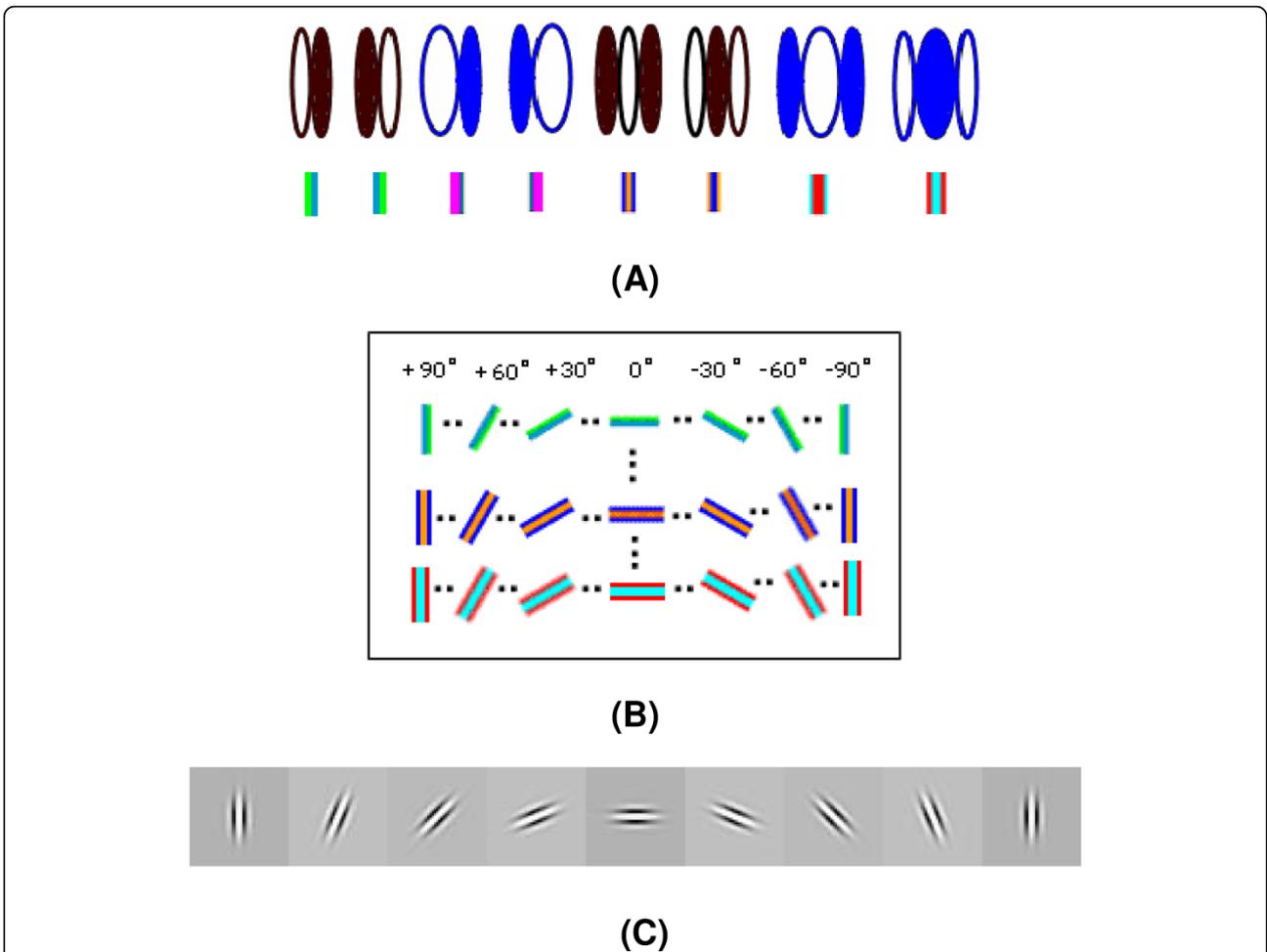


Figure 4 Functional columns as basic information processing units. (A) Eight representative types of receptive fields in function columns in V1; (B) Orientations range from 0° to 180° with a same interval of 10°; (C) An example of receptive field calculated by formula (9).

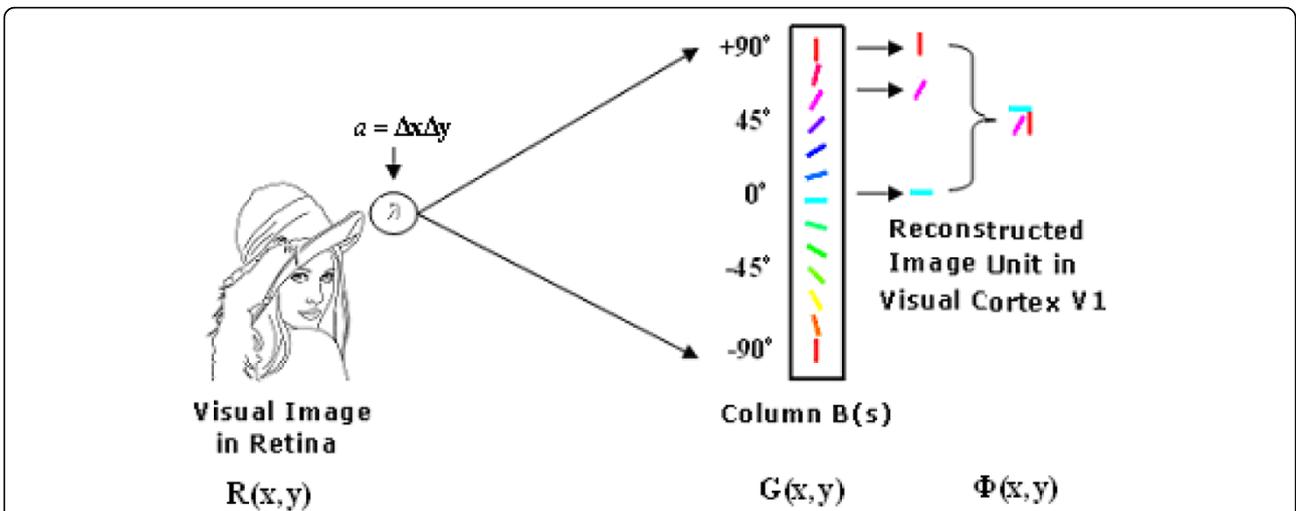


Figure 5 Optimal matching between a patch (upper right corner of the hat) and receptive fields of specific orientations in cortical modules $[B_{k,l}(s)]_{K \times L}$.

Where \otimes denotes Kronecker product, $|\max$ for $\forall\{r_{i,j}(2^h a)B_{k,l}(s)\}$, $\hat{a}_{ML} = 1$ or $\rho_{r\phi} = 1$ denotes the maximal of all products between $r_{i,j}(2^h a)$ and $B_{1,1}(s), B_{1,2}(s), \dots, B_{8,18}(s)$ (As the result of an improvement of signal-to-noise ratio, the noise is reduced). In a neurobiological sense, only stimuli with optimal orientation and frequency may activate the strongest response of simple cells in V1. Formula (16) represents the activated pattern corresponding to a typical image. This pattern can be represented as a matrix

$$[\Phi_{i,j}(b)]_{M \times N} = \begin{bmatrix} \left\{r_{1,1}(2^h a)B_{k,l}(s)\right\}_{\max} & \left\{r_{1,2}(2^h a)B_{k,l}(s)\right\}_{\max} & \dots & \left\{r_{1,N}(2^h a)B_{k,l}(s)\right\}_{\max} \\ \left\{r_{2,1}(2^h a)B_{k,l}(s)\right\}_{\max} & \left\{r_{2,2}(2^h a)B_{k,l}(s)\right\}_{\max} & \dots & \left\{r_{2,N}(2^h a)B_{k,l}(s)\right\}_{\max} \\ \vdots & \vdots & \ddots & \vdots \\ \left\{r_{M,1}(2^h a)B_{k,l}(s)\right\}_{\max} & \left\{r_{M,2}(2^h a)B_{k,l}(s)\right\}_{\max} & \dots & \left\{r_{M,N}(2^h a)B_{k,l}(s)\right\}_{\max} \end{bmatrix} \quad (18)$$

In formula (17), $[\Phi_{i,j}(b)]_{M \times N}$ is the representation of retinal image $[R_{i,j}(2^h a)]_{M \times N}$ in V1, which involves an essential difference with the traditional coding.

As is known, an image $I(x, y)$ can be represented as a linear combination of orthogonal basis functions $\psi_{i,j}(x, y)$

$$I(x, y) = \sum_{i=1}^M \sum_{j=1}^N a_{i,j} \psi_{i,j}(x, y) \quad (19)$$

where $a_{i,j}$ are weights. Obviously, the intersection of $a_{i,j} \psi_{i,j}(x, y)$ is not null, i.e.

$$\bigcap_{i=1, j=1}^{M, N} a_{i,j} \psi_{i,j}(x, y) \neq \emptyset \quad (20)$$

That is to say, the intensity at any location (x, y) in image $I(x, y)$ is contributed by all basis functions $\psi_{i,j}(x, y)$ ($i = 1, 2, \dots, M; j = 1, 2, \dots, N$), so the computation can be highly complicated. In our case, the orthogonal division of input images makes every patch $r_{i,j}(a)$ orthogonal as shown in Figure 6; at the same time, the following condition is satisfied:

$$\bigcap_{i=1, j=1}^{M, N} r_{i,j}(\Delta x, \Delta y) = \emptyset \quad (21)$$

Figure 6. Orthogonal division of a visual image.

Therefore, every patch $r_{i,j}(a)$ can be processed independently by functional columns. This is consistent with the neural mechanism of cortical information processing, and reduces computational complexity as well.

Discussion

Currently, it is widely believed that simple cells densely distributed in V1 function similarly as a tiled set of selective spatio-temporal filters, while V1 carries out

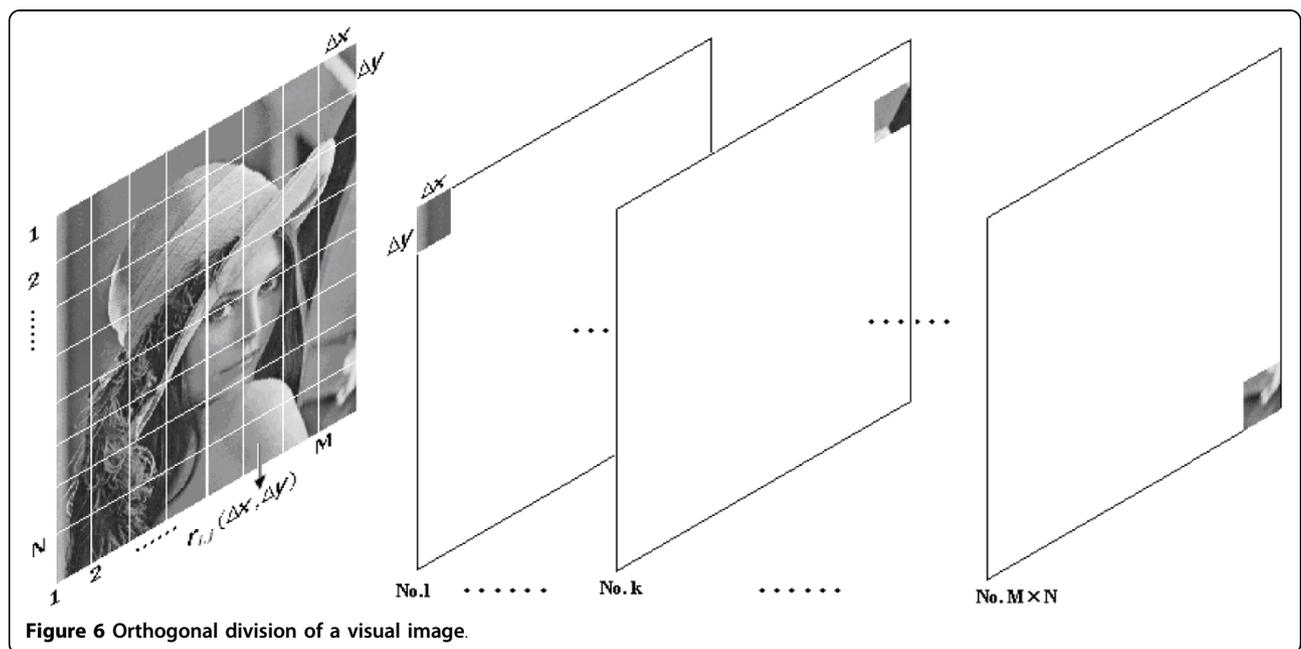


Figure 6 Orthogonal division of a visual image.

operations similar to the local complex Fourier transform. Theoretically, various kinds of neural processing about frequency, orientation, motion and other spatio-temporal operations can thus be performed [49,50].

In this way, the responding property $\Phi(x, y)_{\lambda, \sigma, \theta, \varphi, \gamma}$ in V1 is realized by a convolution between image $R(x, y)$ and receptive fields $G(x, y)_{\lambda, \sigma, \theta, \varphi, \gamma}$ [39].

$$\begin{aligned} \Phi(x, y)_{\lambda, \sigma, \theta, \varphi, \gamma} &= R(x, y) * G(x, y)_{\lambda, \sigma, \theta, \varphi, \gamma} \\ &= \iint R(u, v) G(x - u, y - v)_{\lambda, \sigma, \theta, \varphi, \gamma} dudv \end{aligned} \quad (22)$$

That is to say, $G(x, y)_{\lambda, \sigma, \theta, \varphi, \gamma}$ is taken as a template to scan the whole image $R(x, y)$ from above to bottom and from left to right. For example, if $G(x, y)_{\lambda, \sigma, \theta, \varphi, \gamma}$ is a horizontal orientation receptive field, it will match to many edges with a similar orientation in $R(x, y)$, so many cells in V1 are activated. The activated pattern $\Phi(x, y)_{\lambda, \sigma, \theta, \varphi, \gamma}$ is shown in Figure 3. A similar activated pattern corresponding to a vertical edge is shown in Figure 3(b). This is not an effective method for it stimulates too many responses of relative cells and costs a large amount of energy [51].

While in our case, in order to reconstruct retinal image $[R_{i,j}(a)]_{M \times N}$, we only calculate activated pattern $\varphi_{i,j}(b) = r_{i,j}(2^h a) B_{k,l}(s)$ of the receptive field stimulated by every patch $r_{i,j}(a)$ according to formula (15), and then the location of every patch is determined according to the topological mapping to V1 according to formulas (17) and (18). Finally, the whole activated pattern $[\Phi_{i,j}(b)]_{M \times N}$ stimulated by image $[R_{i,j}(a)]_{M \times N}$ is obtained. Obviously, the related computation is much less complicated, which thus is more consistent with the multi-channel parallel processing mechanism in biology vision.

In order to compare computational complexity of the two ways of computation, we discretize formula (22) as

$$\begin{aligned} \phi(i, j) &= \sum_{m=1}^M \sum_{n=1}^N r(m, n) g(i - m, j - n) \\ i &= 1, 2, \dots, M; j = 1, 2, \dots, N \end{aligned} \quad (23)$$

Every element $\varphi(i, j)$ in array $[\Phi_{i,j}(b)]_{M \times N}$ involves $M \times N$ times of calculations, making the total calculations for all elements as $M^2 \times N^2$.

While in our case, the main computation is $\varphi_{i,j}(b) = r_{i,j}(2^h a) B_{k,l}(s)$, so the total number of calculations are $M \times N \times K \times L$ ($K \ll M, L \ll N$). So the computation of Kronecker product is much less complicated than that of convolution.

We already noticed that a number of other researchers have developed linear-nonlinear models based on response properties of visual neurons [52,53], or on optimal nonlinear transformation [54].

In essence, they are a combination of linear filtering and divisive inhibition model; all of the models have been used to model the nonlinear responses of visual neurons and primary visual cortex. In terms of our proposed model, as already pointed out that theoretical analysis and simulated results show that at the system level, the inner product operator reflects the nature of the excitation of neurons in the cortex V1 by local characteristics of the external stimuli. This is also a plausible assumption for neural computation in the cortex V1. Therefore, it may have some reference value in investigations of neural mechanisms in visual information processing.

Conclusions

It is understood that the retinal image must be in one-to-one correspondence with cortex V1, for all subsequent processing will extract information from V1 and the information kept in V1 is vital. Only in this way, the brain can perceive a vision through the retinal image with high fidelity. Neurophysiology shows that when a retinal image topographically projects to the visual cortex, corresponding neurons will be activated. The whole activated pattern is a copy of the retinal image with high fidelity. In view of signal processing, product $r_{i,j}(2^h a) B_{k,l}(s)$ means that receptive fields $g_{i,j}(b)$ in V1 are activated when stimulated by retinal image $r_{i,j}(a)$. Therefore, this operator is consistent with this neurobiology mechanism. It involves both the simple function of a single neuron and the population function of neuronal groups. $\varphi_{i,j}(b)$ is the local activated pattern corresponding to patch $r_{i,j}(a)$. Different stimuli produce different activated patterns of neuronal groups. The signals activated by details in visual stimuli are much weaker than those activated by contours. According to our understanding of the precise reconstruction of retinal images in the visual process, and based on multi-channel parallel processing features of the visual pathway, a visual image is divided into basic image units (patches) or primitives, which are topographically mapped onto the visual cortex by a one-to-one correspondence, by means of multiplication computing, features contained in image's primitives can be extracted by thousands of visual cortical modules in parallel and synchronously, where only an inner product (or Kronecker product) is needed, then an image will be formed in the primary visual cortex. This algorithm is simple, efficient and in line with the current knowledge about the neural mechanism of visual information processing, the mathematical description is also appropriate to the visual neural computation.

Visual information processing that is actually carried out in V1 is very important, but so far our knowledge of it at the system level remains inadequate [55-57] apart

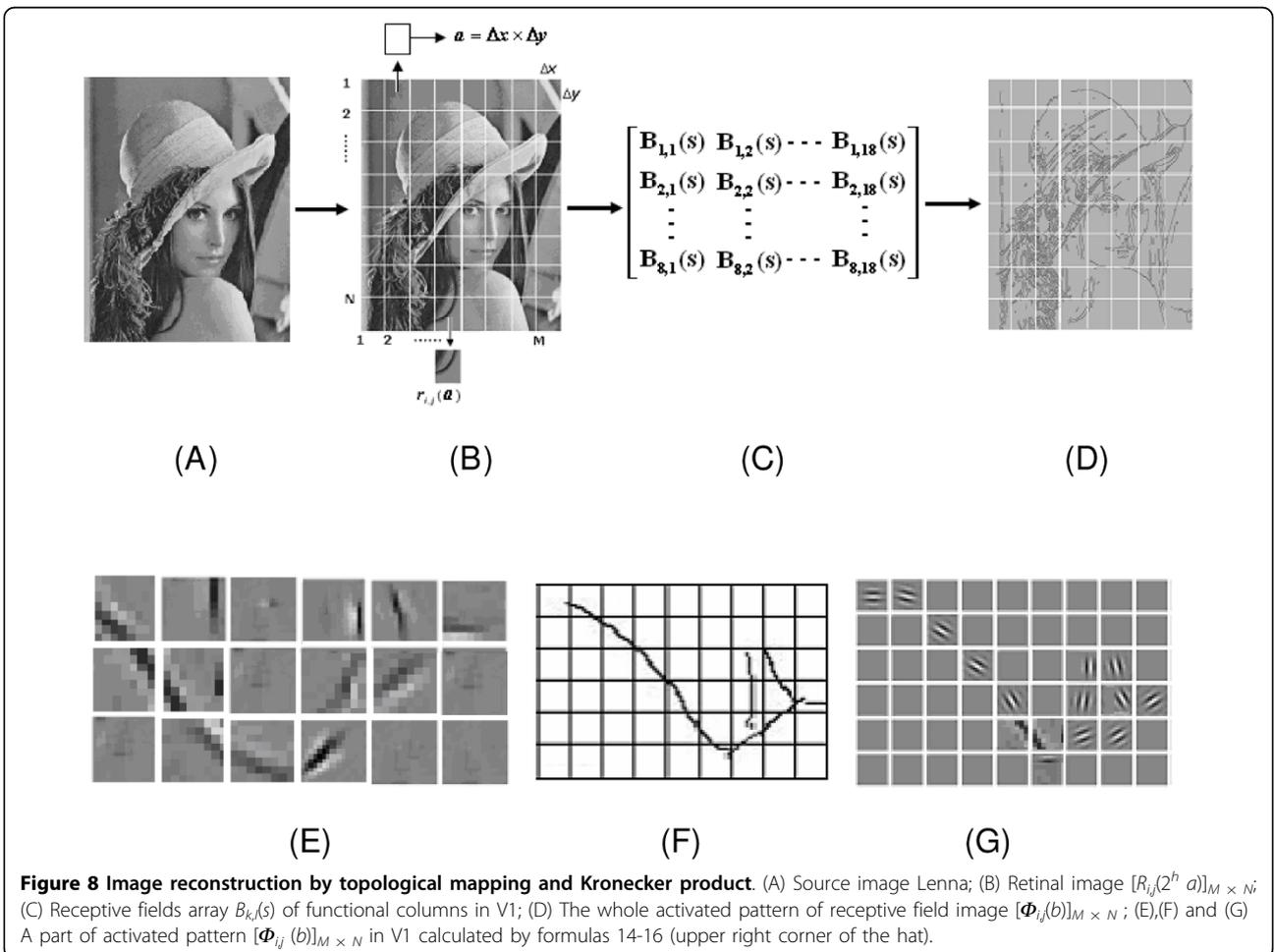
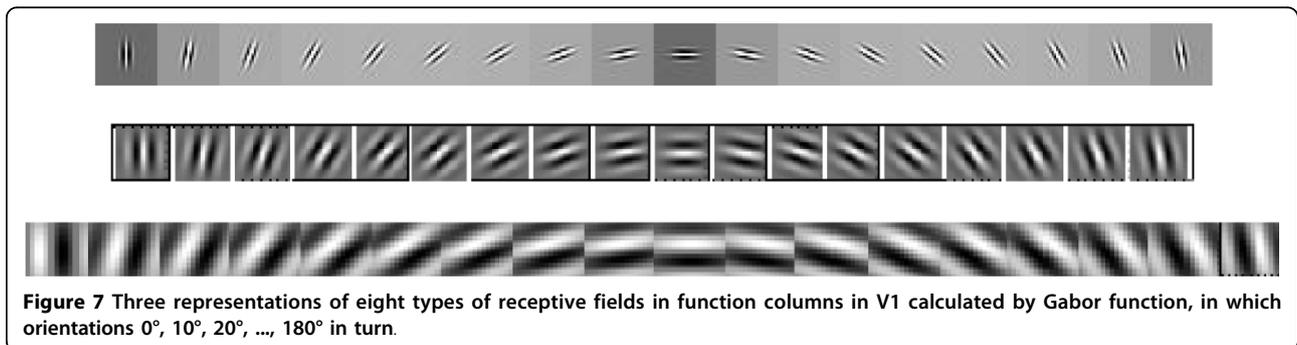
from Hubel and Wiesel's discovery [43] in the 1960s and Field and Olshausen's sparse coding theory [58] in the 1990s. Therefore, the neural computation model based on available knowledge about structure and function of V1 [59-65], presented in this paper may throw some light towards that direction, of course, will require further proof in neurobiology.

The next step of the studies on the visual information processing will be focussed on functional modules in cortex V1.

Methods

In numerical simulations, according to a resolution of $10^\circ/50 \mu\text{m}$, we divide orientations of receptive fields in function columns into 18 parts ranged from 0° to 180° with a same interval of 10° , i.e. $10^\circ, 20^\circ, \dots, 180^\circ$. Eight types of receptive fields are chosen in Gabor function $G(x, y)_{\lambda, \sigma, \theta, \phi, \gamma}$ according to formula (8); the result is shown in Figure 7.

Figure 7. Three representations of eight types of receptive fields in function columns in V1 calculated by



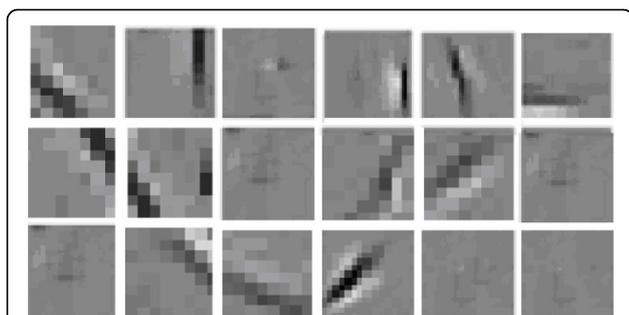


Figure 9 Result (only of the upper right part of the hat shown) corresponding to the damaged function columns at $i = 5$, $j = 9$; $i = 7$, $j = 2$; $i = 11$, $j = 15$.

Gabor function, in which orientations 0° , 10° , 20° , ..., 180° in turn.

And the arrays $B_{k,l}(s)$ are formed according to formula (13). Then the test image Lenna is divided into $M \times N$ patches according to formula (11). The activated pattern of every receptive field $\phi_{i,j}(b)$ stimulated by a patch $r_{i,j}(a)$ is calculated by formula (14). The whole activated pattern $[\Phi_{i,j}(b)]_{M \times N}$ stimulated by $[R_{i,j}(a)]_{M \times N}$ is processed by formulas (16) and (17). A simulated example of the whole activated pattern is shown in Figure 8.

Figure 8. Image reconstruction by topological mapping and Kronecker product (A)Source image Lenna; (B)Retinal image $[R_{i,j}(2^h a)]_{M \times N}$; (C) Receptive fields array $B_{k,l}(s)$ of functional columns in V1; (D) The whole activated pattern of receptive field image $[\Phi_{i,j}(b)]_{M \times N}$; (E), (F) and (G) A part of activated pattern $[\Phi_{i,j}(b)]_{M \times N}$ in V1 calculated by formulas 14-16 (upper right corner of the hat).

In numerical experiments, we also consider the case when a part of cortex is damaged, i.e. functional columns at these locations do not function as they should in the image processing. We assume that the damaged positions are those with $i = 5$, $j = 9$; $i = 7$, $j = 2$; $i = 11$, $j = 15$. The corresponding result is shown in Figure 9. It can be seen that the damaged functional columns do not affect the integrity of the image, for other columns have made compensation for the damaged ones.

Figure 9. Result (only of the upper right part of the hat shown) corresponding to the damaged function columns at $i = 5$, $j = 9$; $i = 7$, $j = 2$; $i = 11$, $j = 15$.

Acknowledgements

The authors appreciated help from Li Weige and Wang Keren for their assistance in English editing for this manuscript and would like to thank the anonymous reviewers for their detailed comments and helpful suggestions which improved the quality of our manuscript. This work was supported by the NSFC of China (Grant no. 60902058, 60371045, 60905063 & 60931003). Authors declare they have no conflict of interest.

Author details

¹LAPC, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China. ²Department of Computer Science, Beijing Jiaotong University, Beijing 100044, China. ³fMRI Center of Brain's function, Beijing 306 Hospital, Chinese People's Liberation Army, Beijing 100101, China. ⁴College of Information Science, Peking University, Beijing 100871, China. ⁵State Key Laboratory of Cognitive Neuroscience and Learning, School of Information Science and Technology, Beijing Normal University, Beijing 100875, China.

Authors' contributions

ZSN, ZQ took part in conceiving and designing the study, analyzed the data, and drafted the manuscript. JZ took part in designing the study and contributed to the data analysis. YGZ, YL took part in conceiving and designing the study. All authors have read and accepted the final version of the manuscript.

Received: 15 June 2009 Accepted: 26 March 2010

Published: 26 March 2010

References

1. Tusa RJ, Palmer LA, Rosenquist AC: The retinotopic organization of area 17 (striate cortex) in the cat. *J Comp Neurol* 1978, **177**:213-235.
2. Miikkulainen R, Bednar JA, Choe Y, Sirosh J: *Computational Maps in the Visual Cortex* Berlin: Springer Science +Business Media, Inc 2005.
3. Nicholls JG, Martin AR, Wallace BG, Fuchs PA: *From Neuron to Brain* Sinauer Associates, Inc, Fourth 2001.
4. Ringach DL: On the origin of the functional architecture of the cortex. *PLoS ONE* 2007, e251.
5. Ishai A, Ungerleider LG, Martin A, Cchouten JL, Haxby JV: Distributed representation of objects in the human ventral visual pathway. *Proc Natl Acad Sci USA* 1999, **96**:9379-9384.
6. Olshausen BA: Principles of Image Representation in Visual Cortex. *Visual Neurosciences* Massachusetts: Cambridge, The MIT PressChalupa LM, Werner JS 2004, 1603-1615.
7. Ferster D, Miller KD: Neural mechanisms of orientation selectivity in the visual cortex. *Annu Rev Neurosci* 2000, **23**:441-471.
8. Bosking WH, Crowley JC, Fitzpatrick D: Spatial coding of position and orientation in primary visual cortex. *Nat Neurosci* 2002, **5**:874-882.
9. Husson TR, Mallik AK, Zhang J, Issa NP: Functional imaging of primary visual cortex using flavoprotein autofluorescence. *J Neurosci* 2007, **27**:8665-8675.
10. Dayan P: Pattern formation and cortical maps. *Journal of Physiology-Paris* 2003, **97**:475-489.
11. Everson RM, Prashanth AK, Gabbay M, Knight BW, Sirovich L, Kaplan E: Representation of spatial frequency and orientation in the visual cortex. *Proc Natl Acad Sci USA* 1998, **83**:834-8338.
12. Khaytin I, Chen X, Royal DW, Ruiz O, Jermakowicz WJ, Siegel RM, Casagrande VA: Functional organization of temporal frequency selectivity in primate visual cortex. *Cereb Cortex* 2007, **3**, 10.1093/cercor/bhm.210.
13. Engel SA, Glover GH, Wandell BA: Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb Cortex* 1997, **7**:181-192.
14. Tootell RB, Hadjikhani NK, Vanduffel W, Liu AK, Mendola JD, Sereno MI, Dale AM: Functional analysis of primary visual cortex (V1) in humans. *Proc Natl Acad Sci USA* 1998, **95**:811-817.
15. Ohki K, Chung S, Kara P, Hubener M, Bonhoeffer T, Reid RC: Highly ordered arrangement of single neurons in orientation pinwheels. *Nature* 2006, **442**:925-928.
16. Olman C, Ronen I, Ugurbil K, Kim DS: Retinotopic mapping in cat visual cortex using high-field functional magnetic resonance imaging. *J Neurosci Methods* 2003, **131**:161-170.
17. Adelson EH, Bergen JR: Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A Opt Image Sci Vis* 1985, **2**:284-299.
18. Baker TI, Issa NP: Cortical maps of separable tuning properties predict population responses to complex visual stimuli. *J Neurophysiol* 2005, **94**:775-787.
19. Mante V, Carandini M: Mapping of stimulus energy in primary visual cortex. *J Neurophysiol* 2005, **94**:788-798.

20. Hyvarinen A, Hoyer PO: **A two-layer sparse coding model learn simple and complex cell receptive fields and topography from natural images.** *Vision Research* 2002, **41**(18):2413-2423.
21. van Hateren JH, Schaaf van der A: **Independent component filters of natural images compared with simple cells in primary visual cortex.** *Proc R Soc Lond B* 1998, **265**:359-366.
22. Horton JC, Adams DL: **The cortical column: a structure without a function.** *Philos Trans R Soc Lond B Biol Sci* 2005, **360**:837-862.
23. Martinez LM, Alonso JM: **Complex receptive fields in primary visual cortex.** *The neuroscientist* 2003, **9**(5):317-331.
24. Kandel ER, Schwartz JH, Jessell TM: *Principles of Neural Science* New York: McGraw-Hill, 4 2000.
25. Palmer SE: *Vision Science* Massachusetts: MIT Press 1999, 186-193, 579-560.
26. Fiorillo CD: **Towards a General Theory of Neural Computation Based on Prediction by Single Neurons.** *PLoS One* 2008, **3**(10):e3298.
27. Hubel DH, Wiesel TN: **Receptive fields and functional architecture of monkeys striate cortex.** *J Physical* 1968, **195**:215-243.
28. Felleman DJ, van Essen DC: **Distributed hierarchical processing in the primate cerebral cortex.** *Cereb Cortex* 1991, **1**:1-47.
29. McClelland JL, Rogers TT: **The parallel distributed processing approach to semantic cognition.** *Nature Reviews Neuroscience* 2003, **4**:1-14.
30. Snyder WE, Hairong Qi: *Machine Vision* Cambridge: Cambridge University Press 2004, 257-261.
31. Hubel DH: **Exploration of the primary visual cortex: 1955-1978.** *Nature* 1982, **299**:515-524.
32. Tinsley CJ, Webb BS, Barraclough NEM, Vincent CJ, Parer A, Derrington AM: **The nature of V1 neural responses to 2D moving patterns depends on receptive-field structure in the marmoset monkey.** *J Neurophysiol* 2003, **90**(2):930-937.
33. Bonheffer T, Grinvald A: **Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns.** *Nature* 1991, **353**:429-431.
34. Jian AK: *Fundamentals of Digital Image Processing* Prentice-Hill 1989.
35. Daugman JG: **Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters.** *J Opt Soc Am A Opt Image Sci Vis* 1985, **2**:1160-1169.
36. Daugman JG: **Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression.** *IEEE Trans Acoustics, Speech Signal Process* 1988, **37**(6):1160-1179.
37. Lee TS: **Image representation using 2D Gabor wavelets.** *IEEE Trans Pattern Anal* 1996, **18**:959-971.
38. Jones JP, Palmer LA: **The two-dimensional spatial structure of simple receptive fields in cat striate cortex.** *J Neurophysiol* 1987, **58**:1187-1211.
39. Feichtinger HG, Strohmer T: *Gabor analysis and algorithms: theory and application* Boston: Birkhauser/Feichtinger HG, Strohmer T 1998.
40. Grigorescu C, Petkov N, Westenberg MA: **Contour detection by band-limited noise and its relation to non-classical receptive field inhibition.** *IEEE Trans. On Image Processing* 2003, **12**(7):729-739.
41. Goodale M, Milner AD: **Separate visual pathways for perception and action.** *Trends in Neuroscience* 1992, **15**:20-25.
42. Teichert T, Wachtler T, Michler F, Gail A, Eckhorn R: **Scale-invariance of receptive field properties in primary visual cortex.** *BMC Neuroscience* 2007, **8**(38):1-16.
43. Hubel DH, Wiesel TN: **Receptive fields, binocular interaction and functional architecture in the cat's striate cortex.** *J Physical* 1962, **160**:106-154.
44. Hubel DH, Wiesel TN: **Ferrier lecture, Functional architecture of macaque monkey visual cortex.** *Proc R Soc Lond B Biol Sci* 1977, **198**:1-59.
45. Kay SM: *Fundamentals of statistical signal processing, Detection theory* Prentice Hall PTR 1998, 520-550.
46. Paninski L, Pillow JW, Simoncelli EP: **Maximum likelihood estimation of a stochastic integrate-and-fire neural encoding model.** *Neural Comput* 2004, **16**:2533-2561.
47. Daubechies I: *Ten Lecture on Wavelets* Philadelphia Pennsylvania: SIAM Press 1992.
48. Issa NP, Rosenberg A, Husson TR: **Models and measurements of functional maps in V1.** *J Neurophysiol* 2008, **99**:2745-2754.
49. Rosa MGP: **Visual maps in the adult primate cerebral cortex: some implication for brain development and evolution.** *Braz J Med Biol Res* 2002, **35**(12):1485-1498.
50. Lamme VA, Roelfsema PR: **The distinct modes of vision offered by feedforward and recurrent processing.** *Trends Neurosci* 2000, **23**(11):571-579.
51. Lennie P: **The cost of cortical computation.** *Current Biology* 2003, **38**:101-109.
52. Mante V, Bonin V, Carandini M: **Functional mechanisms shaping lateral geniculate responses to artificial and natural stimuli.** *Neuron* 2008, **58**:625-638.
53. Rust NC, Mante V, Simoncelli EP, Movshon JA: **How MT cells analyze the motion of visual patterns.** *Nature Neuroscience* 2006, **9**(11):1421-1431.
54. Siwei Lyu, Simoncelli EP: **Nonlinear extraction of independent components of natural image using radial Gaussianization.** *Neural Computation* 2009, **21**:1485-1519.
55. Roelfsema PR: **Cortical algorithms for perceptual grouping.** *Annu Rev Neurosci* 2006, **29**:203-27.
56. Carandini M, Demb JB, Mante V, Tolhurst DJ, Dan Y, Olshausen BA, Gallant JL, Rust NC: **Do we know what the early visual system does?** *J Neuroscience* 2005, **25**(46):10577-10597.
57. Olshausen BA, Field DJ: **How close are we to understanding V1?** *Neural Comput* 2005, **17**:1665-1699.
58. Olshausen BA, Field DJ: **Emergence of simple cell receptive field properties by learning a sparse code for natural images.** *Nature* 1996, **381**:607-609.
59. Somers DC, Todorov EV, Siapas AG, Toth LJ, Kim DS, Sur M: **A local circuit integration approach to understanding visual cortical receptive fields.** *Cerebral Cortex* 1998, **8**:204-217.
60. Troyer TW, Krukowski AE, Miller KD: **LGN input to simple cells and contrast-invariant orientation: an analysis.** *J Neurophysiol* 2002, **87**:2741-2752.
61. Swindale NV: **Feedback decoding of spatially structured population activity in cortical maps.** *Neural Comput* 2007, **20**(1):176-204.
62. Larsson J, Landy MS, Heeger DJ: **Orientation-selective adaptation to first- and second-order patterns in human visual cortex.** *J Neurophysiol* 2006, **95**:862-881.
63. Bednar JA, Miikkulainen R: **Joint maps for orientation, eye, and direction preference in a self-organizing model of V1.** *Neurocomputing* 2006, **69**:1272-1276.
64. Ringach DL: **Mapping receptive fields in primary visual cortex.** *J Physiol* 2004, **558**(3):717-728.
65. Singh G, Memoli F, Ishkhanov T, Sapiro G, Carsson G, Ringach DL: **Topological analysis of population activity in visual cortex.** *Journal of Vision* 2008, **8**(11):1-18.

doi:10.1186/1471-2202-11-43

Cite this article as: Songnian et al.: Neural computation of visual imaging based on Kronecker product in the primary visual cortex. *BMC Neuroscience* 2010 **11**:43.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

