Biol. Cybern. 54, 107–114 (1986)

Biological Cybernetics

A Mathematical Model of the Primary Visual Cortex and Hypercolumn

K. Okajima

Fundamental Research Laboratories, NEC Corporation, 1–1 Miyazaki 4-chome, Miyamae-ku, Kanagawa, 213 Japan

Abstract. A mathematical model of the primary visual cortex is presented. Basically, the model comprises two features. Firstly, in analogy with the principle of the computerized tomography (CT), it assumes that simple cells in each hypercolumn are not merely detecting line segments in images as features, but rather that they are as a whole representing the local image with a certain representation. Secondly, it assumes that each hypercolumn is performing spatial frequency analyses of local images using that representation, and that the resultant spectra are represented by complex cells. The model is analyzed using numerical simulations and its advantages are discussed from the viewpoint of visual information processing. It is shown that 1) the proposed processing is tolerant to shifts in position of input images, and that 2) spatial frequency filtering operations can be easily performed in the model.

1 Introduction

Owing to the work of Hubel and Wiesel (1962, 1977), and many other studies which followed it, today much is known about the primary visual cortex (area 17). The main points can be summarized:

1) Most cells in area 17 respond specifically to a line segment of certain orientation in the visual field.

2) Among them, simple cells respond to a line segment having a definite position, whereas complex cells (and hypercomplex cells) respond to a line segment, regardless of its precise position within their receptive fields, if it has the proper orientation (and usually if it moves in a proper direction).

3) These neurons are organized into a hypercolumn, the basic module of the cortical machinery. A hypercolumn, which is 1–2 mm^2 in size, contains a complete set of neurons which covers all orientations and both the ocular dominances.

4) Although the position of the receptive field slightly fluctuates from neuron to neuron in a hypercolumn, they all fall in a definite region in the visual field (aggregate field).

5) The primary visual cortex is constructed from a repeated structure of the basic modules (hypercolumns), and the whole visual field is fully covered by all the hypercolumns, each processing the local image fallen in its own aggregate field.

Now the problem is to elucidate whatever task this primary visual cortex is performing in the visual information processing pathway.

Some suppose that each cell in it acts as a feature (line segments of various orientations) detector. However this seems unlikely, since it is obvious that line segments alone are not sufficient as features to describe an image, though certainly they are very important ones. If cells in the primary visual cortex were merely detecting line segments, they would fail to deal with images having other features (e.g., blobs etc.).

Meanwhile, another important contribution to the understanding of the visual system is due to Campbell and Robson (1968). Mainly based on psycho-physical evidence, they proposed that there exists a certain neural mechanism which performs the spatial frequency analysis of images (see also Blakemore et al. 1969). If their proposition is correct, the problems to be elucidated are:

1) What neural mechanism could perform the spatial frequency analysis?

2) What role could such a neural mechanism play in the brain's image analysis function?

In this paper, I try to pose possible answers to these problems by proposing a model concerning the function of the primary visual cortex.

Basically, the model comprises two features. Firstly, it assumes that simple cells in a hypercolumn are not merely detecting line segments, but rather representing images by some kind of representation which I call "tomographic representation" in this paper. The details will be explained in Sect. 2.

Secondly, the model assumes that each hypercolumn is performing a Fourier analysis of its local image. Thus as a whole the primary visual cortex is performing local spatial Fourier analyses of the retinal images. It turns out that in the tomographic representation, two dimensional Fourier transforms of images can be obtained by performing only one dimensional Fourier transforms. It is assumed that complex cells are engaged in performing the Fourier transform, in accord with the electrophysiological work by Glezer et al. (1973).

The model will be described in Sect. 2. In Sect. 3 the function of the model will be analyzed and its advantages will be discussed from the viewpoint of visual information processing. In Sect. 4 a possible extension of the model to the binocular case will be proposed. Some experimental findings which can be explained by the model will be discussed in Sect. 5. Finally, in Sect. 6 the results will be summarized.

2 The Model

In this model, it is assumed that

1) Each local image is represented by the tomographic representation using all the simple cells in each hypercolumn.

2) Each hypercolumn performs a Fourier analysis of the local image using this representation.

In this section these two features will be described.

2.1 Tomographic Representation

Figure 1 shows a typical setup for computerized tomography (CT). For instance, in X-ray CT, one measures the attenuation of the beam across an object (see for instance Hounsfield 1976). The attenuation, $d(t, \theta)$ is proportional to the integral of density along the beam line, $l(l: r\cos(\phi-\theta)=t)$.

$$d(t,\theta)\propto\iint a(r,\phi)\delta(t-r\cos(\theta-\phi))\,|r|\,dr d\phi\,, \tag{1}$$

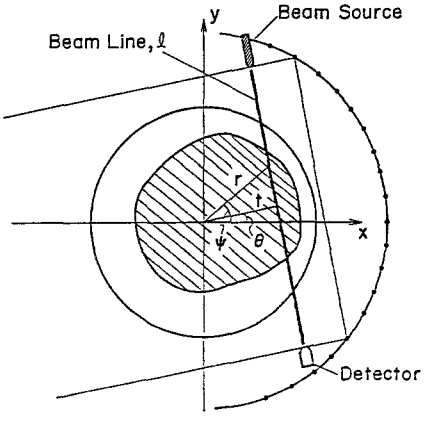

Fig. 1. A typical setup for computerized tomography

where $a(r, \phi)$ is the density distribution of the object and $\delta(x)$ is the δ-function.

If one measures the attenuation, $d(t, \theta)$ while scanning the beam to cover the whole object and rotating the beam axis within $0\leqq\theta\leqq\pi$, one can calculate the original density distribution, $a(r, \phi)$ from these data, $d(t, \theta)$ (Randon 1917; Hounsfield 1976).

$$\begin{aligned} a(r,\phi) &= \int_0^{\pi} g(r\cos(\phi-\theta),\theta)d\theta\,;\\ g(t,\theta) &= \int d(\tau,\theta)h(t-\tau)d\tau\,,\\ h(t) &= \int_{-K_c}^{K_c} |k|\exp(ikt)dk\,, \end{aligned} \tag{2}$$

where K_c is a cutoff frequency determined by the resolution.

By analogy with this, we define the tomographic representation $(d(t, \theta))$ of an image as

$$d(t,\theta)=\iint a(r,\phi)\delta(t-r\cos(\theta-\phi))\,|r|\,dr d\phi\,, \tag{3}$$

where $a(r, \phi)$ now denotes the light intensity of the image.

This representation is complete in the sense that one can reconstruct the original image, $a(r, \phi)$ using (2) if we know all the values of $d(t, \theta)$.

Now, if one assumes that each simple cell in a hypercolumn is representing this $d(t, \theta)$, then one can reconstruct the local image projected on the aggregate field of the hypercolumn from the activity pattern of all the simple cells.

Certainly, this representation is especially suited for dealing with line segments, since the bases are lines of various orientations. However it should be emphasized here that since it is complete, one can also deal with other features (such as blobs etc.) in this representation.

Figure 2 shows some examples of the tomographic representation.

2.2 Local Spatial Fourier Analysis

One advantage with representing images using the tomographic representation is that their two dimensional Fourier transforms or autocorrelation functions etc. can be obtained by performing the operation only on the argument t.

Suppose $d(t, \theta)$ is the tomographic representation of a local image, $a(r, \phi)$. Then its Fourier transform with respect to t is calculated as

$$\begin{aligned} D(k,\theta) &= \int d(t,\theta)\exp(-ikt)dt\\ &= \iiint a(r,\phi)\delta(t-r\cos(\theta-\phi))\,|r|\\ &\quad\cdot dr d\phi\exp(-ikt)dt\\ &= \iint a(r,\phi)\exp(-ikr\cos(\theta-\phi))\,|r|\,dr d\phi\\ &= \iint a(\mathbf{r})\exp(-i\mathbf{k}\cdot\mathbf{r})d\mathbf{r}=A(\mathbf{k})\,,\\ &\quad \mathbf{k}=(k\cos\theta,\ k\sin\theta)\,. \end{aligned} \tag{4}$$

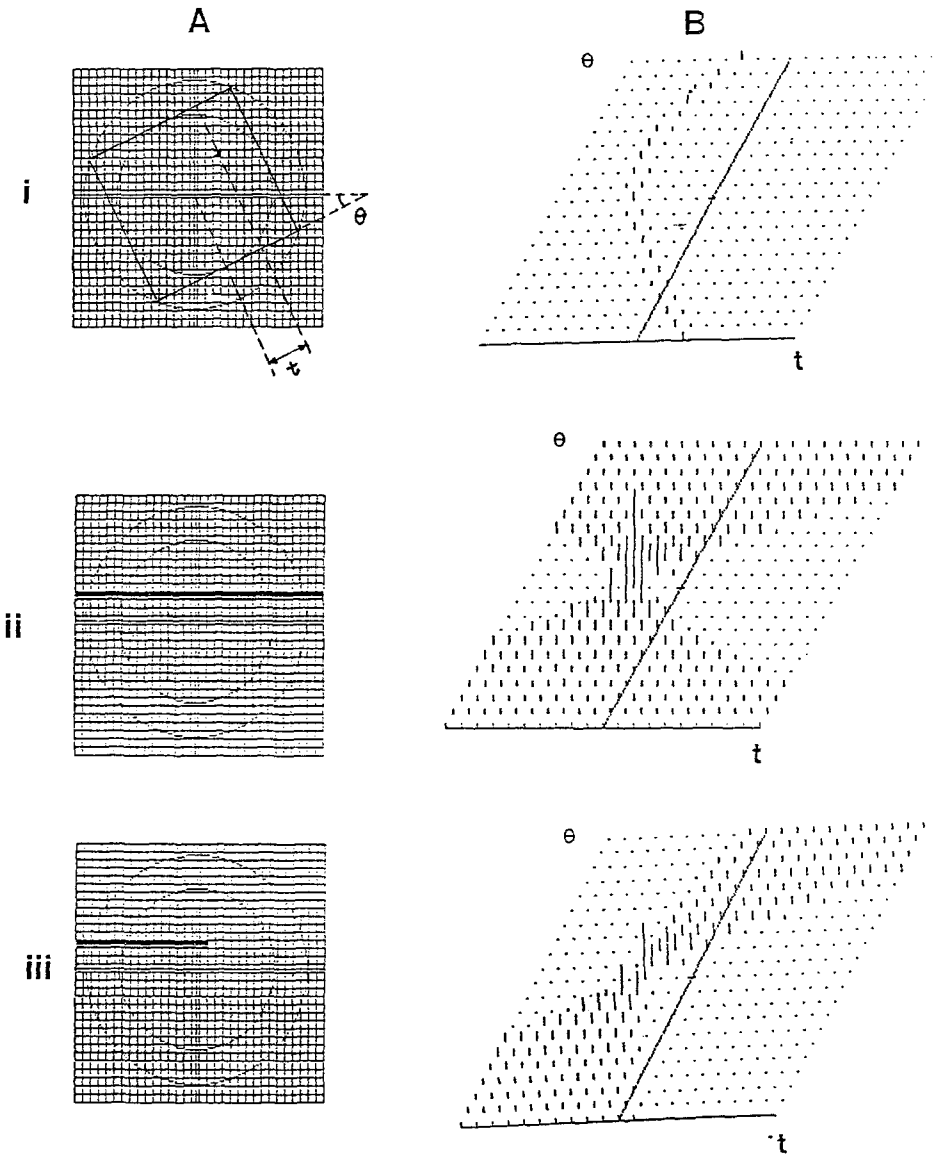

Fig. 2A and B. Examples of the tomographic representation. Original (local) images (**A** a point (*i*), a line (*ii*) and a line with a termination (*iii*)) and their tomographic representations (**B**). In **A**, the outer circles represent the local aggregate fields. The parameters, t and θ are defined as indicated in **A** (*i*)

Thus, $D(k, \theta)$, the Fourier transform of the tomographic representation ($d(t, \theta)$) with respect to t, is equal to $A(\mathbf{k})$, the Fourier transform of the original image $a(\mathbf{r})$ with respect to $\mathbf{r}$, expressed in polar coordinates (k, θ). In other words, one can obtain the two dimensional Fourier transform of the original image by performing merely a one dimensional Fourier transform on its tomographic representation.

In order to perform the Fourier transform, it is necessary to gather all the activities of simple cells having the same orientation selectivity (i.e., $d(t, \theta)$, $\theta = \text{const}$) to one cell. It is assumed in the model that some complex cells are playing this role.

Since activities of neurons are expressed in terms of pulse frequencies, basically they can not deal with negative quantities. Consequently, when one intends to realize the model with some neural network, one should assign certain neurons for dealing with negative quantities (e.g., on-center cells and off-center cells) in every processing stage of the model.

The features of the model are summarized in Fig. 3. In a hypercolumn, the local image fallen in its aggregate receptive field is

1) represented by all the simple cells with the tomographic representation, and

2) Fourier analyzed using that representation, the resultant spatial frequency spectra being represented by complex cells using the polar coordinate representation.

Thus, the whole input image is analyzed by all the hypercolumns with a local spatial Fourier analysis.

Figure 4 shows an example of the processed pattern produced by the model. The pattern consists of 13×10 local spatial Fourier transforms, each calculated using

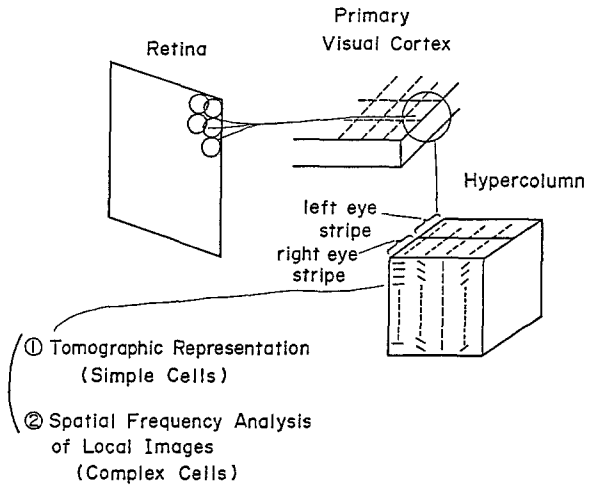

Fig. 3. Basic features of the model

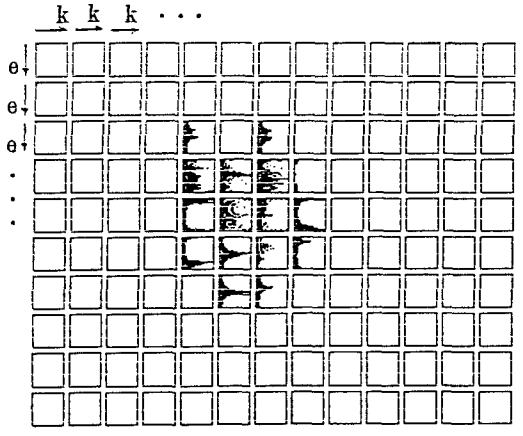

Fig. 4. An example of an output pattern of the model. The original image is the same as in Fig. 8B. The original image (256×200 pixels) is divided into overlapping 13×10 local images, and their Fourier transforms are calculated using the tomographic representation. To eliminate the discontinuity at boundaries, a Gaussian window ($\sigma = 10$) was employed in the Fourier transform calculation. In this figure, dots represent the spatial frequency components whose magnitudes are larger than a certain threshold

the tomographic representation of the corresponding local images (in Fig. 4, only the amplitudes squared of the Fourier transforms are shown).

3 The Function of the Model

3.1 Stability to Shifts in Position

The spatial frequency power spectrum of an image is invariant under a shift in its position. Accordingly, by local spatial Fourier analysis, one can extract separately a position-insensitive component (amplitudes of local Fourier transforms) and a position-sensitive component (their phases) from an input image. The former is stable with regard to displacements of the input image within the range of each aggregate field size.

An analogous stability would result if one took a moving average of images using some proper weighting function (e.g., Gaussian) by performing a convolution of the images with it. In this case, images become stable with respect to displacement as they are smoothed by the weighting function in the range of its width. However a decrease in resolution occurs, since in the averaging process the high spatial frequency components of the images are lost.

Meanwhile, with a local spatial Fourier analysis, one can obtain patterns, stable with regard to displacements in the original images, without any loss in resolution, since their high frequency components are also preserved throughout the process.

Figure 5a shows the overlap between an original image **I** and its displaced image **I′** [displaced by (dx, dy)] as a function of the displacement (dx, dy). The overlap $\mathbf{I} \cdot \mathbf{I}'/|\mathbf{I}| \cdot |\mathbf{I}'|$ steeply decreases when one displaces the image over a few digits.

On the other hand, the overlap between the output patterns (their amplitude components squared), **O** and **O′**, which are obtained by processing the images, **I** and **I′** with the model, maintains a rather high value when the displacement is within the range of the aggregate field size (see Fig. 5b).[1]

The same order of stability with respect to displacement can be achieved if one smoothes the image with a weighting function such as a Gaussian whose width (variance) is about 15 digits. However, in this case the resolution is restricted by the width of the weighting function.

Figure 6 shows an image of one line (2 digits width, Fig. 6a) and that of two lines (1 digit width and 1 digit spacing, Fig. 6b). These two images could not be distinguished if one smoothed the images by taking a moving average with a weighting function whose width is larger than about 1 digit. Meanwhile, one can distinguish these two from the high frequency components of the output patterns of the model (Fig. 7).

1 As in Fig. 4, **O** and **O′** consist of 13×10 local spatial frequency spectra. If we write **O** and **O′** as $P(k, \theta; X, Y)$ and $P'(k, \theta; X, Y)$, X and Y denoting the location of the local image, then the overlap, $\mathbf{O} \cdot \mathbf{O}'/|\mathbf{O}| \cdot |\mathbf{O}'|$ is calculated as

$$\sum_{\substack{k,\theta \\ X,Y}} P(k, \theta; X, Y)\, P'(k, \theta; X, Y) \Big/ \sqrt{\sum_{\substack{k,\theta \\ X,Y}} P(k, \theta; X, Y)^2} \cdot \sqrt{\sum_{\substack{k,\theta \\ X,Y}} P'(k, \theta; X, Y)^2}$$

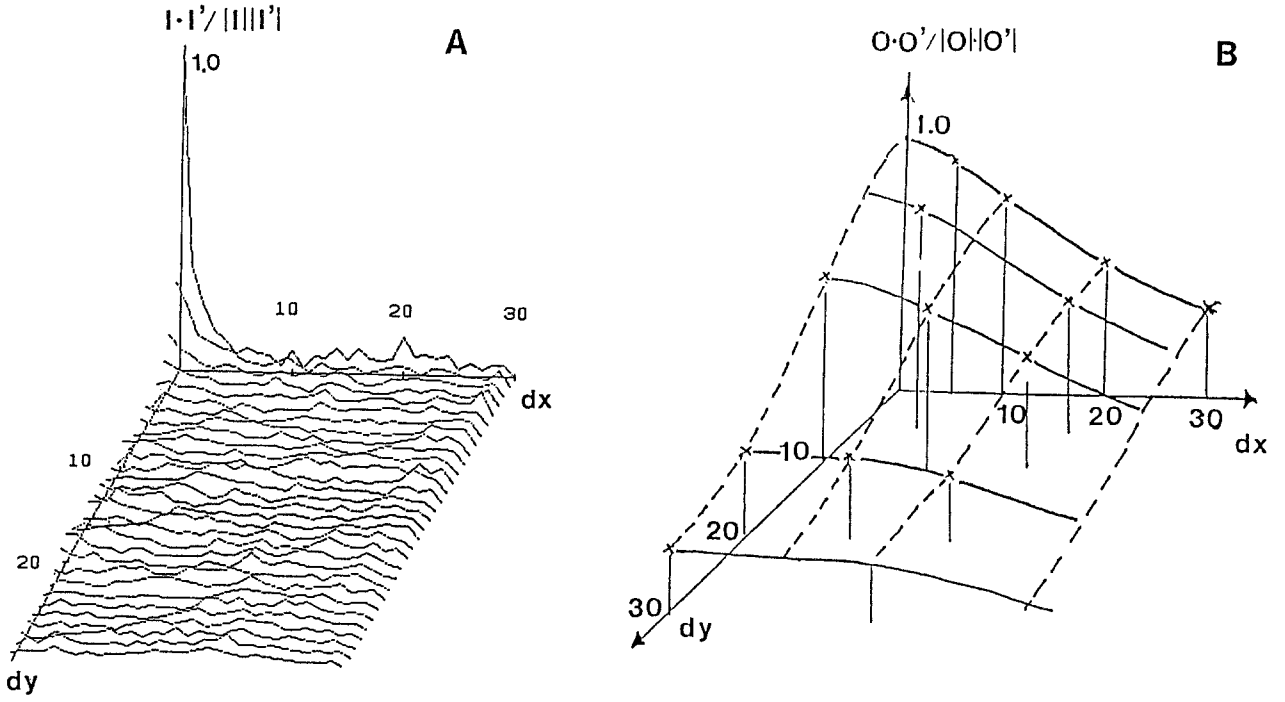

Fig. 5. A Overlap between an original image (shown in Fig. 8b) and the displaced image as a function of the displacement. **B** Overlap between output patterns (their amplitude components), each obtained by processing the original image and the displaced one with the model

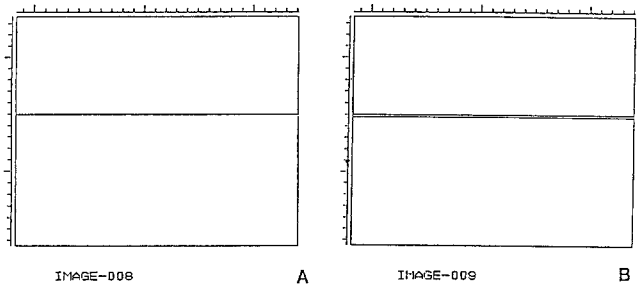

Fig. 6A and B. A line (width 2, **A**) and two lines (width 1, distance 1, **B**)

3.2 Spatial Frequency Filtering

With the model described in 2, one can easily perform a spatial frequency filtering operation by, for instance, inhibiting cells subserving certain frequency components.

Figure 8a shows the quantitized image (quantitization unit is 4×4 pixels) of an image shown in Fig. 8b.

The quantitized image and the original image are processed by the model, and the overlap between the two output patterns (their amplitude components squared) are calculated with proper weighting func-

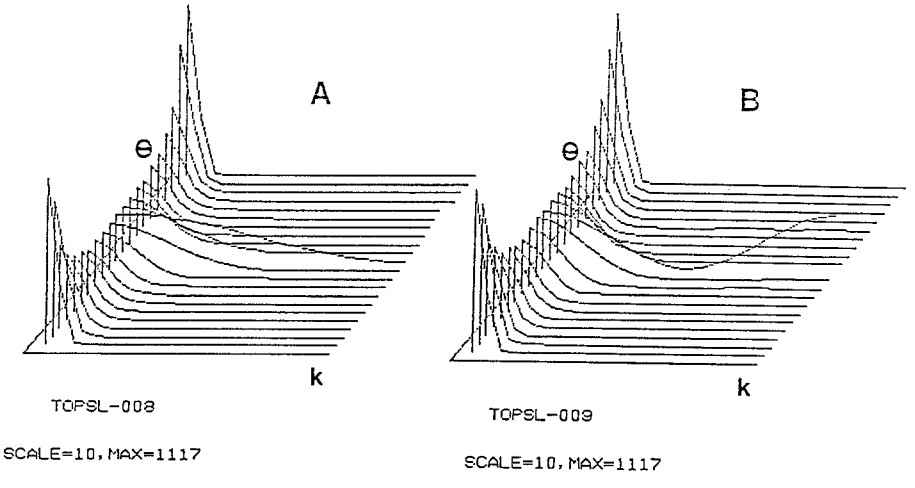

Fig. 7A and B. Output patterns of a hypercolumn which are, obtained by processing the images in Fig. 6. Here, only their amplitudes squared are shown

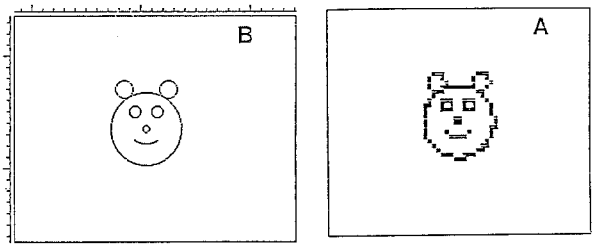

Fig. 8A and B. An image (**B**) and its quantitized image (**A**). Quantitization unit is 4×4 pixels

Table 1. Overlaps between output patterns obtained by processing the two images in Fig. 8. These are calculated using different weighting functions in spatial frequency space[a]

Weighting function	Case 1	Case 2
Overlap $\mathbf{O} \cdot \mathbf{O}'/\|\mathbf{O}\| \cdot \|\mathbf{O}'\|$	0.523	0.987

[a] In the model, spatial frequency spectra of local images are calculated and represented in polar coordinates ($P(k, \theta)$). Therefore, if one calculates the inner products (in spatial frequency space) by

$$\iint PP' dk d\theta$$

one underestimates the high frequency components (note that in polar coordinates, the surface element dS is given as $dS = k dk d\theta$). This occurs with Case 2.

In Case 1, in order to enhance the high frequency components, the inner products are calculated by

$$\iint PP' M^2 dk d\theta$$

where $M = k \exp(-0.125\, k^2)$ is a weighting function in spatial frequency space. In real space, this corresponds to some proper lateral inhibition among simple cells, $d(t, \theta)$

tions in spatial frequency space. The results are shown in Table 1. One can see that the model can either treat the two images in Fig. 8 as the same (case 2) or as different (case 1) by varying the weighting function in spatial frequency space.

4 A Possible Extension of the Model to the Binocular Case

In reality, hypercolumns get afferents from both eyes (Hubel and Wiesel 1962, 1977). If one intends to extend the model to the binocular case, it is desirable that it satisfies two points:

1) Stimulus equivalence between both eyes: a mechanism to evoke the same neuronal activities for the same stimulus, regardless of whether it is from the right or left eye (note that the model described so far satisfies quasi stimulus equivalence across the aggregate field).

2) A mechanism to detect the disparity between the right eye (local) image and the left eye one (as it is an important parameter for stereo vision) (Barlow et al. 1967; Marr et al. 1976).

In this section, I will suggest a possible extension of the model to the binocular case, which satisfies the above two points by introducing an additional Fourier transform with respect to the coordinate which specifies right eye (R) and left eye (L).

Here, we concentrate on a particular hypercolumn. Suppose the right eye image and the left eye image to that hypercolumn are $a_R(x, y)$ and $a_L(x, y)$ respectively. We denote their tomographic representations as $d_R(t, \theta)$ and $d_L(t, \theta)$. [Hereafter we abbreviate them as $d_R(t)$ and $d_L(t)$ since each θ component can be treated independently.]

First, we perform the Fourier transform with regard to coordinate t (we denote the Fourier transform of $d(t)$ as $D(k)$).

$$D_R(k) = \int d_R(t) \exp(-ikt) dt$$
$$D_L(k) = \int d_L(t) \exp(-ikt) dt\,. \tag{5}$$

Second, we perform the Fourier transform with regard to the coordinates R and L.

$$D_0(t) = D_R(t) + D_L(t)$$
$$D_1(t) = D_R(t) - D_L(t)\,. \tag{6}$$

Then their amplitudes squared ($P = |D|^2$) are given as

$$P_0(k) = P_R(k) + P_L(k) + (D_R^* D_L + D_R D_L^*)$$
$$P_1(k) = P_R(k) + P_L(k) - (D_R^* D_L + D_R D_L^*)\,. \tag{7}$$

In (7), the first and second terms are the power spectra of the right eye image and the left eye image respectively, and the third term is their interference term. Therefore if one expresses (7) as

$$\tilde{P}_0(k) = P_0(k) + P_1(k)$$
$$\tilde{P}_1(k) = P_0(k) - P_1(k) \tag{8}$$

then $\tilde{P}_0(k)$ satisfies the stimulus equivalence of right and left eyes, whereas $\tilde{P}_1(k)$ bears the information on the disparity between the two images.

For example, if $d_R(t)$ and $d_L(t)$ are

$$d_R(t) = f(t)$$
$$d_L(t) = f(t - \Delta) \tag{9}$$

then $\tilde{P}_0(k)$ and $\tilde{P}_1(k)$ are calculated as

$$\tilde{P}_0(k) = 4|F(k)|^2$$
$$\tilde{P}_1(k) = 4|F(k)|^2 \cos(\Delta \cdot k)\,, \tag{10}$$

where $F(k)$ is the Fourier transform of $f(t)$.

Thus $\tilde{P}_0(k)$ is proportional to the power spectrum of the input image ($f(t)$ or $f(t - \Delta)$), whereas $\tilde{P}_1(k)$ bears the information on the disparity, Δ.

Of course, it is also possible to detect disparities directly in real space by, for instance, calculating the mutual correlation function, $c(\mathbf{r})$ between both eyes local images, $a_R(\mathbf{r})$ and $a_L(\mathbf{r})$.

$$c(\mathbf{r}) = \sum_{\mathbf{r}'} a_R(\mathbf{r})\, a_L(\mathbf{r}+\mathbf{r}') \tag{11}$$

One can detect the disparity from the maximal point in $c(\mathbf{r})$.

In the tomographic representation, (11) becomes[2]

$$c(t, \theta) = \sum_{t'} d_R(t, \theta)\, d_L(t+t', \theta) \tag{12}$$

where $c(t, \theta)$ is the tomographic representation of $c(\mathbf{r})$, and $d_R(t, \theta)$ and $d_L(t, \theta)$ are those of $a_R(\mathbf{r})$ and $a_L(\mathbf{r})$ respectively.[3]

Basically, the model suggested in this section calculates the Fourier transform of $c(t, \theta)$ to extract the information on the disparity, i.e.

$$C(k, \theta) = D_R^*(k, \theta)\, D_L(k, \theta)\,. \tag{13}$$

Note that by comparing this with the calculation of the mutual correlation functions in real space [cf. (12)], one can immediately obtain $C(k, \theta)$ in spatial frequency space, if $D_R(k, \theta)$ and $D_L(k, \theta)$ have been calculated beforehand, which is one of the model's assumption made in Sect. 2.

Kaufman (1964) and Julez (1971) suggested that in the human visual system, some independent spatial frequency-tuned channels are involved in detecting disparities. However, at present it is not certain whether the disparity is dealt with in real space or in spatial frequency space at the stage of the primary visual cortex.

5 Discussion

The proposed model is consistent with some physiological or psychological findings.

1) Simple cells show maximal responses to line segments of various orientations and various positions (within the aggregate field), and a hypercolumn contains a complete set of simple cells covering all the orientations. These facts support the assumption that the local image fallen in the aggregate field is represented with the tomographic representation.

2) A hypercolumn also contains complex cells which respond to line segments regardless of their position within the receptive fields. This suggests that complex cells are performing some operation only on the argument t (with $\theta = \text{const}$) in the tomographic representation, such as the Fourier analysis on t assumed in the model. This assumption is consistent with the findings of Glezer et al. (1973).

3) Thus, in the model, the spatial frequency channels suggested by Campbell and Robson (1968) are assigned to some complex cells. However, in the model the channels are for local spatial frequencies.

In reality, simple cells have inhibitory regions (Hubel and Wiesel 1962) surrounding the excitatory region (or vice versa in case of off-center cells). Generally, such a lateral inhibition is known to enhance contrasts of an image. As a first approximation, one can treat this effect by replacing $d(t, \theta)$ in Sect. 2 by

$$d'(t, \theta) = \int d(\tau, \theta)\, w(\tau - t)\, d\tau\,, \tag{14}$$

where $w(\tau)$ is an adequate weighting function to represent the lateral inhibition. Then (4) becomes

$$D'(k, \theta) = D(k, \theta)\, W^*(k)\,, \tag{15}$$

where $W(k)$ is the Fourier transform of $w(\tau)$. Since $w(\tau)$ has surrounding antagonistic regions, $W(k)$ has small value around $k = 0$. This means, according to (15), that the high frequency components are relatively enhanced under the lateral inhibition (see also the footnote of Table 1).[4]

---

2 (12) can be derived as follows

$$\begin{aligned}
&\sum_{t'} d_R(t, \theta)\, d_L(t+t', \theta) \\
&= \sum_{t', \mathbf{r}_1, \mathbf{r}_2} a_R(\mathbf{r}_1)\, a_L(\mathbf{r}_2)\, \delta(t - |\mathbf{r}_1| \\
&\quad \cdot \cos(\theta - \phi_1))\, \delta(t + t' - |\mathbf{r}_2| \cos(\theta - \phi_2)) \\
&= \sum_{\mathbf{r}_1, \mathbf{r}_2} a_R(\mathbf{r}_1)\, a_L(\mathbf{r}_2)\, \delta(t + |\mathbf{r}_1 - \mathbf{r}_2| \cos(\theta - \phi_{1-2})) \\
&= \sum_{\mathbf{r}_1, \Delta} a_R(\mathbf{r}_1)\, a_L(\mathbf{r}_1 + \Delta)\, \delta(t - |\Delta| \cos(\theta - \phi_\Delta)) \\
&= \sum_{\Delta} c(\Delta)\, \delta(t - |\Delta| \cos(\theta - \phi_\Delta))
\end{aligned}$$

where $\Delta = \mathbf{r}_1 - \mathbf{r}_2$, and $\phi_1, \phi_2, \phi_{1-2} = \phi_\Delta$ are defined as

$\mathbf{r}_i = |\mathbf{r}_i|(\cos\phi_i, \sin\phi_i)$

and

$\mathbf{r}_1 - \mathbf{r}_2 = |\mathbf{r}_1 - \mathbf{r}_2|(\cos\phi_{1-2}, \sin\phi_{1-2})$

3 To detect the disparity, it is sufficient to calculate $c(\Delta)$ only in the region $(\Delta, 0)$; $|\Delta| \leqq \Delta_{\max}$ (i.e., along the horizontal direction), where $\Delta_{\max}$ is the maximal disparity to be detected. This means, in the tomographic representation, one can determine the disparity mainly from the value of $c(t, \theta)$ in the region $t \leqq \Delta_{\max} \cos\theta$. Therefore, for example, one need not calculate $c(t, \theta)$ along the vertical direction $(\theta = \pi/2)$

4 Even from $d'(t, \theta)$, one can reconstruct the original image $a(r, \phi)$ as follows.

$$a(r, \phi) = \int_0^\pi g'(r\cos(\phi - \theta), \theta)\, d\theta;$$

$$g'(t, \theta) = \int d'(\tau, \theta)\, h'(t - \tau)\, d\tau\,,$$

$$h'(t) = \int_{-K_c}^{K_c} |k| \exp(ikt)/W(k)\, dk$$

To achieve global visual information processing, one need to integrate the locally processed outputs. This integration process is left for the following stages. For example, in the proposed model, each hypercolumn can not deal with spatial changes whose periods are larger than their aggregate field sizes. In order to detect these frequencies, one needs to gather the outputs of hypercolumns with different locations.

6 Summary

To summarize the results,

1) A possible model was proposed concerning the function of the primary visual cortex and its basic module, the hypercolumn. The basic features of the model are:

a) First, hypercolumns transform local images which fall in their aggregate receptive fields into their tomographic representations.

b) Second, each hypercolumn performs Fourier analyses using that representation. Thus the whole input image is analyzed by local spatial Fourier analysis.

2) From the information processing viewpoint, the early stage visual information processing described above has the following advantages:

a) It is tolerant to shifts in position of the input images, preserving resolution.

b) One can easily perform filtering operations in spatial frequency space if necessary.

3) The model can be extended to the binocular case in a reasonable way.

Acknowledgements. The author thanks Mr. Fujiwara, S. for helpful discussions. He also thanks Dr. Fowler, M.T. for correcting the English manuscript.

References

Barlow HB, Blakemore C, Pettigrew JD (1967) The neural mechanism of binocular depth discrimination. J Physiol 193:327–342

Blakemore C, Campbell FW (1969) On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. J Physiol 203:237–260

Campbell FW, Robson JG (1968) Application of Fourier analysis to the visibility of gratings. J Physiol 197:551–556

Glezer VD, Ivanoff VA, Tscherbach TA (1973) Investigation of complex and hypercomplex receptive fields of visual cortex of the cat as spatial frequency filters. Vision Res 13:1875–1904

Hounsfield GN (1976) Computerized transverse axial scanning (tomography). Part 1. Description of the system. Br J Radiol 121:189–195

Hubel DH, Wiesel TN (1962) Receptive fields, binocular interactions and functional architecture in the cat's visual cortex. J Physiol 60:106–154

Hubel DH, Wiesel TN (1977) Functional architecture of macaque monkey visual cortex. Proc R Soc Lond Ser B 198:1–59

Julez B (1971) Foundations of cyclopean perception. University of Chicago Press

Kaufman L (1964) On the nature of binocular disparity. Am J Psychol 77:393–402

Marr D, Poggio T (1976) Cooperative computation of stereo disparity. Science 194:283–287

Radon J (1917) On the determination of functions from their integrals along certain manifolds. Ber Sächs Akad Wiss 69:262–277

Received: December 13, 1985

Dr. Kenji Okajima
Fundamental Research Laboratories
NEC Corporation
1–1 Miyazaki 4-chome
Miyamae-ku
Kanagawa
213 Japan