Electronics and Communications in Japan, Part 2, Vol. 86, No. 3, 2003
Translated from Denshi Joho Tsushin Gakkai Ronbunshi, Vol. J84-D-II, No. 3, March 2001, pp. 559–570

Probing the Neural Mechanism of Binocular Information Processing with VEPs

Ryusuke Hayashi, Yoichi Miyawaki, Taro Maeda, and Susumu Tachi

Graduate School of Engineering, The University of Tokyo, Tokyo 113-8656 Japan

SUMMARY

We analyze the binocular stereo-vision process, and the visual evoked potential (VEP) is examined for the case in which a random-dot stereogram (RDS) is presented. In this study, three experimental parameters are considered, namely, stimulus presentation position, disparity, and correlation between binocular images. The analysis is based on the change of VEP latency, and the relation between the binocular stereo-vision process and the binocular competition process is investigated. The VEP waveform for RDS presentation is divided into three main components. The first component is the initial vision response, observed to be localized in the occipital area, which is considered to reflect the local disparity detection process. The second and third components are responses with middle to long latency spreading from the occipital area to the frontal areas, and the latency greatly depends on the presentation position and the disparity. The latency is shorter when the presentation is at the center than at the periphery of the view field, and is shorter for crossed disparity than for uncrossed disparity. The difference of latency seems to reflect the processing mechanisms for the presentation position and the disparity. In an experiment using the RDS with the contrast reversed between two eyes (anti-RDS), the latency is longer for crossed disparity than for uncrossed disparity, which is the reverse of the case of RDS. This phenomenon can be accounted for by the response of the disparity-selective neuron in V1, which suggests that both binocular stereo vision and binocular competition are based on the local disparity detection mechanism as the neural basis. The response of the disparity-selective neuron is simulated using the binocular energy model, and the difference between the two kinds of binocular vision processing. A mechanism that detects the unpairedness between the two eyes is proposed. © 2003 Wiley Periodicals, Inc. Electron Comm Jpn Pt 2, 86(3): 47–60, 2003; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/ecjb.10135

Key words: binocular stereo vision; binocular competition; interocular unpaired region; random-dot stereogram; visual evoked response; biomedical applications.

1. Introduction

As an experiment for the analysis of the human binocular stereo-vision process, the examination of the visual evoked potential (VEP) using a random-dot stereogram (RDS) [1] is useful. By examining the VEP, the electrophysiological response in the brain corresponding to the perception process can be evaluated noninvasively with high time resolution.

There have been many experiments in which the VEP is measured using the RDS [2–5]. Many of these report negative activity with a latency of 200 ms in the occipital area when the RDS is analyzed, which is considered as characteristic of the disparity detection process. The time

course of the VEP, as dependent on the stimulus parameters, however, has not been examined.

In many experiments the stimulus is presented at the center of the view field and the response is analyzed, but the characteristics of vision depend in general on the position of the stimulus in the view field. There is a study [3] in which the relation between the position of disparity stimulus presentation and the VEP waveform is investigated, but a sufficient view field for presentation is not provided.

In this context, we wish in this study to enlarge the view angle of presentation, and the VEP is observed by presenting the stimulus with a vertical deviation. By enlarging the view field of presentation and presenting the stimulus at deviated positions, the response time is elongated, making it easier to analyze the effect of the experimental parameters. As other parameters, the disparity of the RDS and the correlation between binocular images are also considered, and the neural basis and its time characteristics are investigated in relation to the whole binocular vision mechanism, including the binocular competition process.

A simulation is performed in which the disparity-selective neuron is modeled by the binocular energy neuron. A new mechanism is proposed involving interocularly unpaired regions, which are considered to play the important role in both binocular stereo vision and binocular competition.

2. Binocular Stereo Vision

2.1. Mechanism of binocular stereo vision

Binocular stereo-vision processing determines the depth of the object from the binocular disparity. For this purpose it is necessary to match (pair) the image elements between the left and the right scenes (called the pairing problem). Binocular stereo-vision processing in the brain is generally considered as composed of two main steps. One is local disparity detection in which the difference of the retinal images between the eyes is detected, including the case of incorrect pairing, from only local information. The other, called global stereo-vision processing, is a high-level form of processing in which the pairing problem is solved by matching of the patterns of the whole images [6].

Many disparity-selective neurons, which provide the neural basis for binocular stereo vision, have been identified in the vision-related areas located in the posterior pathways, such as V1, V2, V3, V3A, VP, MT and MST [7, 8]. The local detection of disparity is an initial form of visual information processing, which is considered to be performed by the disparity-selective neurons in V1 and V2 [6]. On the other hand, electrical neuron stimulation experiments suggest that the high-level visual area including the MT is related to conscious depth perception [8].

2.2. Marr's computation theory

As a theoretical approach to solving the matching problem, Marr and Poggio presented the constraints specified by the physical environment (which are pairedness, uniqueness, and continuity). They proposed a model in which the integration of disparity detectors is defined on the basis of these three constraints, and the solution is output uniquely by iterative computation [9]. Although Marr's computation theory has been used as the basis for many binocular stereo-vision models, many problems are encountered. In particular, it is assumed that the depth of the object of vision changes continuously, and that depth discontinuity, which is directly related to the object shape, is ignored. An extension of Marr's model defines as discontinuous the process of including a certain change of disparity by introducing, for example, line processing. The problem still remains that the method cannot handle a small but discontinuous change or a large and continuous change [10].

2.3. Interocularly unpaired region

One of the useful approaches to handling depth discontinuity is a method based on information concerning the interocularly unpaired region [11]. When there exist two surfaces with a depth difference, and the near surface covers the background plane, there always exists a region in the depth discontinuous part which is observed by only one of the eyes, and there exists a region in the scene observed by the other eye for which pairing is impossible.

The interocularly unpaired region is considered to provide depth discontinuity information and to contribute to the production of the edge sensation at the boundary [12]. In the processing of the interocularly unpaired region, it is considered important to specify the exact position on the retina where the V1 neuron has a role. The detailed neural mechanism, however, is not known.

2.4. Binocular competition (view field conflict)

When the images of the eyes differ greatly and are unpaired, binocular competition (also called view field conflict) occurs. The inhibition of the view field by binocular competition is produced locally in the initial vision processing. Since the perceptions of the left and right eyes compete, it is generally considered that binocular competition is produced by mutual inhibition of the left and right monocular neurons [13].

Phenomena which cannot be accounted for by mutual inhibition of the monocular neurons, such that the competition elements are not the left and right eyes, have been found [14]. Rather, it is suggested by recent neurophysi-

ological investigations that binocular competition originates from the V1 binocular neuron and is produced by inhibition of the eye superiority column [15].

It is not clear, however, what binocular neuron is related to the inhibition, nor has a model of binocular competition based on the binocular neuron been presented. Furthermore, both binocular stereo vision and binocular competition are produced when there is a difference between the retinal images of the two eyes. In other words, the two forms of perception are closely related, inhibiting each other [16]. Still, no binocular vision model to represent the two phenomena in a systematic way has been presented. Consequently, the experiments in this study also consider the difference of the neural mechanisms for binocular stereo vision and binocular competition.

3. Method of Experiment

3.1. Subjects

The measurements were performed on four adult males with vision corrected to normal binocular vision (40" or more by the stereo test of the Stereo Optical Co.).

3.2. Experimental setup

Figure 1 outlines the experimental setup. The subject enters a shielded room and observes the image presented on a screen parallel to the forehead at a distance of 100 cm. The chin of the subject is fixed using a chin stage. A liquid crystal projector (Sharp Co. XV-E550) is placed outside the shielded room, and the image is projected through a glass window with metal mesh.

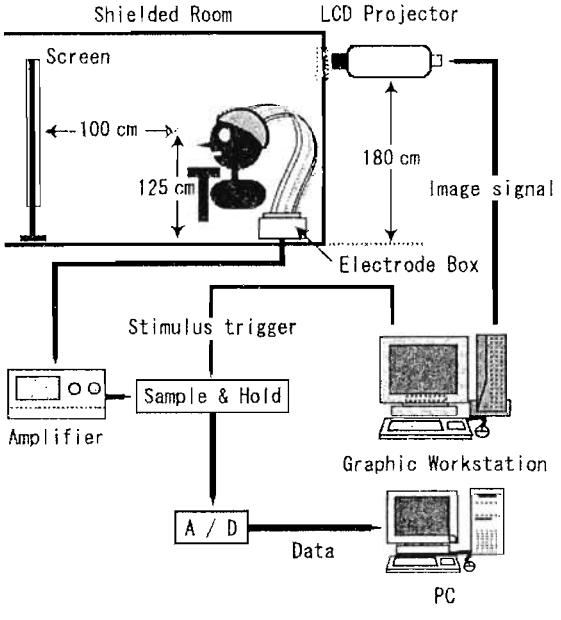

Fig. 1. The experimental system.

3.3. Method of measurement and recording

The electrodes are prepared following the international 10–20 placement. Nineteen channels are prepared on the skull, or 8 channels are prepared only on the occiput. In addition, 1 channel is prepared on each lower eyelid, to monitor eye movement and blinking. The impedance of each electrode is adjusted to 5 to 30 kΩ. The reference electrode is placed on the ear lobe, and the ground electrode is placed on the nose. Thus, monopolar leads are used. A multichannel biological signal amplifier (Nihon Kohden Co. MME-3132) is used in the measurement. The gain is set as 5 μV/V, and the overall frequency band is set as 0.53 to 100 Hz. The analog output from the EEG amplifier is fed to a sample-and-hold board, and the signal is sampled by an A-D board with a sampling frequency of 1 kHz.

The measured data are stored in an IBM PC/AT-compatible machine. After the experiment, averaging is applied to 50 trial data. The image is generated by a dedicated graphics work station (SGI Co. Indigo2). The trigger signal is output using the channel option function, synchronized to the image signal. The trigger signal is also recorded, to be used as the origin of synchronization in averaging.

3.4. Presentation and stimulus conditions

The stimulus is a dynamic RDS, which is obtained by switching the RDS image at 30 Hz, and is presented by the red-green anaglyph method. The random point pattern to be used as the control (binocularly correlated random-dot stereopattern) is also presented dynamically.

Figure 2 shows the image presentation procedure. First, the CRD is presented for 100 ms. Then, one of the stimulus images with binocular disparity or CRD is selected at random, and is shown for 700 ms. Lastly, the CRD is presented for 600 ms. The above procedure is defined as a set. For each stimulus, data of 25 sets are acquired in a session. Between the sets, the CRD is presented for a randomly set time of 1 to 2 s, avoiding situations where the stimulus is given periodically. The above sequence is applied twice with a recess of approximately 3 minutes in between. Thus, data for 50 sets are acquired for each stimulus. The subject is instructed to gaze at the point at the center of the screen during the experiment.

As the stimulus image with disparity, images are prepared in which a rectangular region with disparity is placed in the upper 10°, upper 5°, center 0°, lower 5°, and

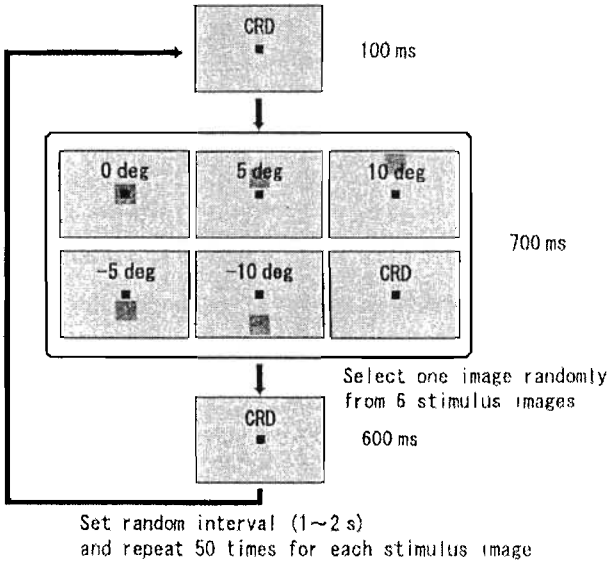

Fig. 2. The sequence of events in a single trial of stimulus-presentation.

lower 10° positions, respectively. The view angle of the whole screen is 31.3° × 39.6°. The size of the gaze point is 1.2° × 1.2°. The dot density is 50%. The size of a pixel is 12.1′ × 12.1′. The resolution on the screen is 213 × 160 pixels. The average brightness of the screen is 41 cd/m^2.

RDS of cross disparity −36.3′ and uncross disparity of 36.3′ are mostly used in the experiments. The size of the stimulus region is 6.9° × 6.9° for the crossed image, and 5.2° × 5.2° for the uncrossed image. In this study, constant size is emphasized, since the response with middle to long latency is examined.

In other words, let the disparity be θ [rad], the distance from the observer to the screen be d_{scr} [m], and the distance between the eyes be d_{eye} [m]. Then, the depth d m of the object plane from the screen is approximated as

$$d \approx \frac{d_{scr}^2 \theta}{d_{scr}\theta - d_{eye}} \qquad (1)$$

The disparity region is determined so that the edge length of the stimulus region is kept constant for the depth d. The effect of the size of the stimulus region is discussed later (Section 5.1).

4. Experimental Results

4.1. Effect of presentation position

Examining the topography when the RDS is presented, we find that a negative potential is produced first in the occipital area, and then the response spreads to the frontal area [5, 17]. Figure 3 is an example of a VEP waveform derived from the electrodes at the right occipital area (O2) and the postparietal area (Pz) when an RDS with cross disparity (−36.3′) is presented. The reason for using Pz is that the visual area along the posterior pathway is considered to be related to binocular stereo vision (Section 2.1). The vertical axis is the amplitude, with the upper direction being negative. The horizontal axis is the time, and the data are shown only up to 800 ms from stimulus presentation.

The waveforms in the figure are drawn for each presentation position of the disparity region. The stimulus is shifted downward, to the upper 10°, upper 5°, center 0°, lower 5°, and lower 10° positions, and the respective VEPs

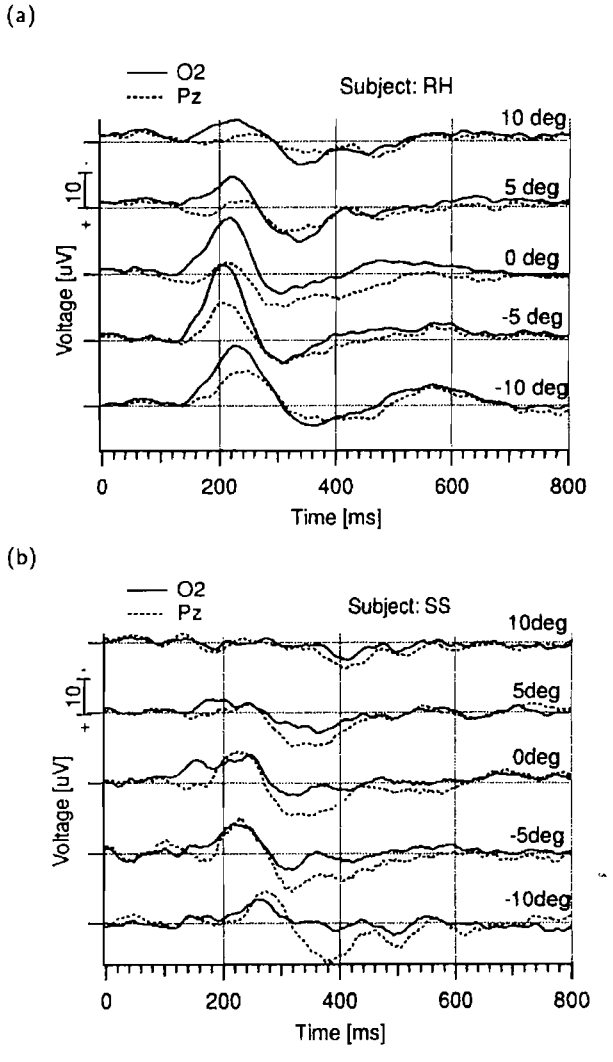

Fig. 3. VEPs to RDSs of a crossed disparity for subject RH (a) and SS (b), recorded at scalp sites O2 and Pz.

are shown. When the disparity stimulus is presented at the center 0°, a large negative peak is produced in the occipital area with a latency of 200 ms. This agrees with past research reports.

Except for the case of lower 5° presentation, the latency is seen to be increased as the stimulus is presented in the peripheral area. A positive peak follows the above negative peak, and a similar increase of the latency is observed. The reason for the longer latency in the periphery may be that the number of stimulated disparity-selective neurons depends on the view field position, due to the difference of the cortex magnification ratio.

However, on presentation of a stimulus that could be perceived by the monocular queue (a rectangular image drawn with a gray level between the bright dot and the dark dot, with a size of 6.0° × 6.0°, and observed by both eyes), the latency did not vary significantly with the presentation position (Figs. 4 and 5 are the results for the average peak latency of the O2 waveform as a function of the presentation position for four subjects).

When the stimulus is presented to the upper view field, a large positive peak appears first. Consequently, only the case of presentation to the lower view field is shown in Fig. 4. This suggests that the effect of the presentation position in the view field (especially the position in the vertical direction) is more marked in binocular stereo-vision processing.

In the case of RDS with crossed disparity, the latency tends to become shorter when the stimulus is presented at the lower 5° position than at the center. Consequently, it is conjectured that even if the number of disparity-selective neurons has an effect on the latency, it is not true that the

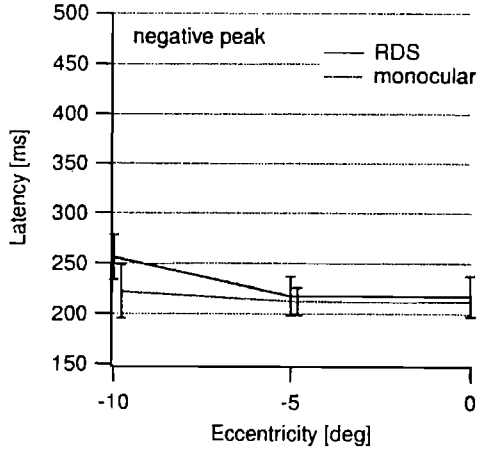

Fig. 4. The latencies of the negative peaks to RDSs and monocular-cue stimuli.

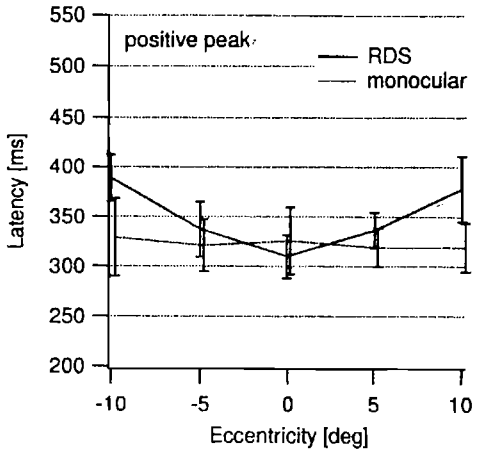

Fig. 5. The latencies of the positive peaks to RDSs and monocular-cue stimuli.

latency is simply determined by the cortex magnification ratio.

According to past studies [5, 17], the VEP for the RDS exhibits the following process. A negative potential is produced first, being localized in the occipital area, then expands to the frontal area including the occipital area. Then the potential changes to positive polarity. The tendency is also observed in Fig. 3, except for when the presented stimulus is greatly deviated toward the periphery (i.e., upper 10° and lower 10°). Thus, the VEP waveform when the RDS is presented is divided into three main components. The first is the negative potential localized in the occipital area, the second is the negative potential spreading from the occipital to the frontal area, and the third is the positive potential spreading from the occipital area to the frontal area.

In this experiment, the dynamic CRD was switched to the dynamic RDS. Consequently, clues except for binocular disparity were excluded in the perception of the stimulus appearance. Consequently, the detection of local disparity is the first form of visual information processing. It is known in neurophysiology that the detection of local disparity is performed by low-level visual areas such as V1 and V2 (Section 2.1).

Consequently, it is estimated that the first component, that is, the first visual response, localized in the occipital area, is a reflection of the local disparity detection process. It is also well known from past studies that this negative potential is induced in common by the images with binocular retinal image difference [2, 4]. When the presented stimulus is greatly deviated toward the periphery (such as upper 10° and lower 10°), the first component becomes obscure. The reason seems to be that the projection of the

lower peripheral visual field onto V1 is the inside of the upper calcarine sulcus, and the projection of the upper peripheral visual field is the inside of the lower calcarine sulcus. Consequently, the activity of this area as seen from the potential distribution on the cortex is no longer localized in the occipital area around O2.

4.2. Effect of disparity

Next, we examine the effect of RDS disparity on the VEP, and a similar experiment is performed using the uncrossed RDS (36.3′). It is also seen in the case of uncrossed disparity that the latencies of the second and third components are increased when the stimulus is presented to the peripheral view field (Fig. 6). Remarkably, in the case of the uncrossed disparity the latency is found to be increased at any presentation position compared to the case of crossed disparity.

Comparing the measurement results for RDS with disparity −12.1′ (stimulus region 6.3° × 6.3°) and disparity 12.1′ (stimulus region 5.8° × 5.8°), the time zone for the peak latency is separated for the crossed and uncrossed cases. Note that Figs. 7 and 8 give the average peak latencies for four subjects. Thus, it seems that the sign of the disparity has a great effect on the latency.

4.3. Effect of correlation between binocular images

4.3.1. Anti-RDS

In order to examine the difference in neural mechanisms between binocular stereo vision and binocular competition, an experiment was performed using anti-correlated RDS (anti-RDS). Anti-RDS is a stimulus in which, as shown in Fig. 9, the sign of the contrast is reversed between the left and right eyes in the area corresponding to the RDS disparity region. When this stimulus is presented to the two eyes, binocular competition is produced, since there is no pairing between the two scenes. In anti-RDS, the deviation of the contrast-inverted region between the two images is defined as the disparity. The presentation conditions in the experiment were the same as in the case of RDS, except that the stimulus was anti-RDS.

4.3.2. Uncorrelated stimulus

In addition to anti-RDS, a binocularly uncorrelated random-dot stereo-pattern (unCRD) was presented and the

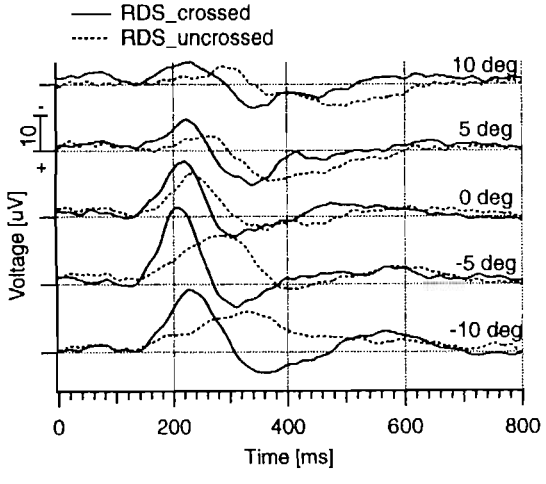

Fig. 6. VEPs to RDSs of crossed and uncrossed disparities for subject RH, recorded at site O2.

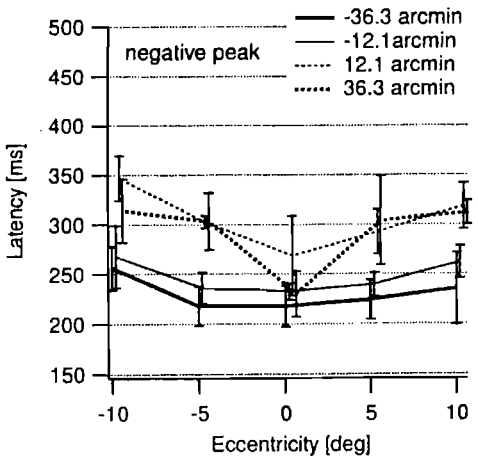

Fig. 7. The latencies of the negative peaks to RDSs.

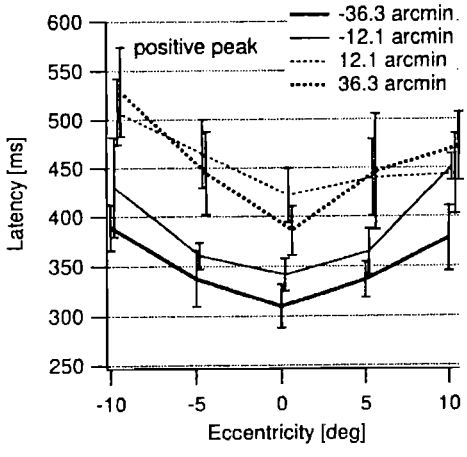

Fig. 8. The latencies of the positive peaks to RDSs.

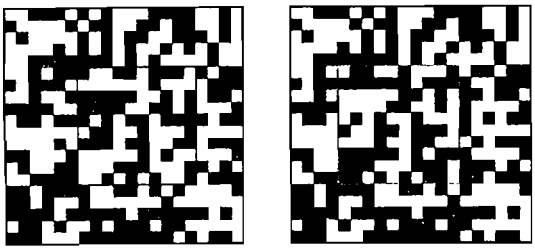

Fig. 9. An example of anti-RDS.

VEP was examined in order to see more details of the effect of unpairedness between two eyes. unCRD is a stimulus in which there is no correlation between the point placements of the left and right images. As in the case of anti-RDS, binocular competition is recognized. In the experiment, the size of the rectangular region without correlation between the two eyes is set as 6.0° × 6.0°.

4.3.3. Results

Figure 10 shows waveform examples derived from O2 when anti-RDS with the crossed disparity (−36.3′) and uncrossed disparity (36.3′) are presented. The same waveform change as in the case of RDS is observed, with the peak latency changing greatly according to the presentation position of the stimulus, and the VEP waveform consists of three components. The latencies of the second and third components, however, are always shorter for the uncrossed disparity than for the crossed disparity, which is a tendency opposite to the case of RDS.

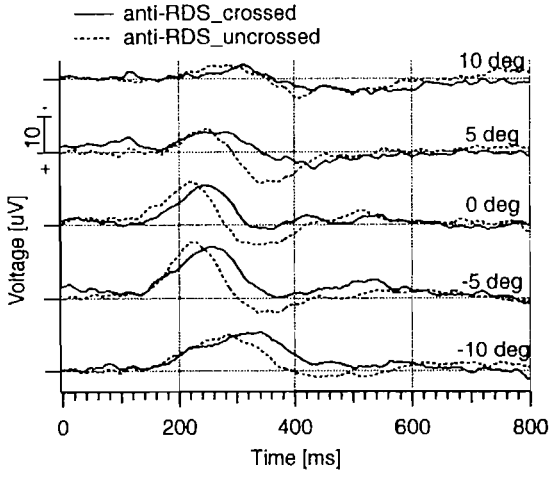

Fig. 10. VEPs to anti-RDSs of crossed and uncrossed disparities for subject RH, recorded at site O2.

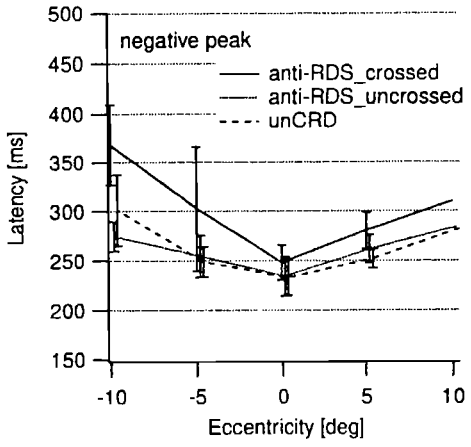

Fig. 11. The latencies of the negative peaks to anti-RDSs and unCRDs.

Figures 11 and 12 show the average peak latency for four subjects, for anti-RDS and unCRD. Depending on the subject, it sometimes happened that, when the interocularly unpaired stimulus was presented at the periphery, the amplitude of VEP was reduced, making it difficult to determine the peak (the number of subjects used in averaging was 1 for upper 10°, 2 for upper 5°, 3 for lower 5°, and 2 for lower 10°). On comparing the results of anti-RDS, the above reversed tendency of the latency change depending on the disparity is observed.

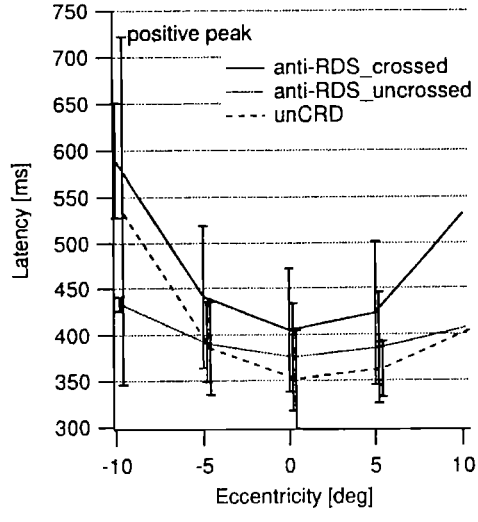

Fig. 12. The latencies of the positive peaks to anti-RDSs and unCRDs.

Comparing further the results for unCRD and anti-RDS, only the peak latency for anti-RDS with crossed disparity is increased, although the data exhibit a large variation.

5. Discussion

5.1. Difference of processing mechanisms between crossed and uncrossed disparities

The latencies of the second and third components are always longer when the RDS of uncrossed disparity is presented than when the RDS of crossed disparity is presented. The reason may be the different sizes of the stimulus region. When the stimulus region is modified from 6.9° × 6.9° to 4.6° × 4.6° and the RDS with cross disparity (–36.3′) is presented, there is an increase of latency due to size reduction of the stimulus region. Compared to the presentation of RDS with the stimulus region 5.2° × 5.2° and uncrossed disparity (36.3′), however, the latency is always smaller (Fig. 13).

In the case of anti-RDS, the size of the stimulus region is smaller in the uncrossed case than in the crossed case, but the latency is always shorter in the uncrossed case. This suggests that the size of the stimulus region has an effect on the difference in VEP latency, but that the sign of the disparity has a larger effect.

By past psychological experiments, it has been suggested that there exists a functional difference between the crossed disparity and the uncrossed disparity. Richards

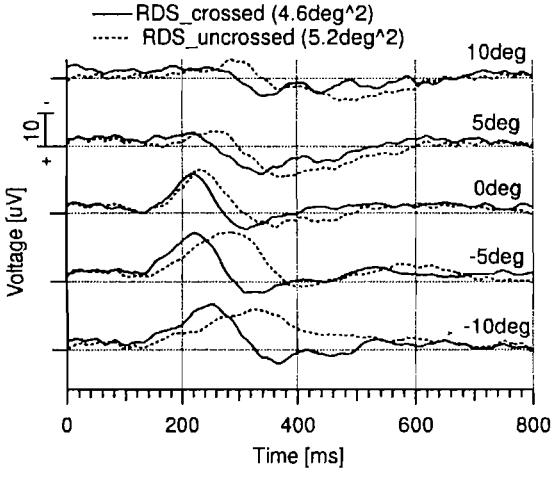

Fig. 13. VEPs to RDSs of crossed and uncrossed disparities for subject RH, recorded at site O2. The effect of stimulus size.

reported that a solid vision blind can occur selectively for either the crossed or uncrossed disparity [18]. According to another study, the discrimination ability is higher and the threshold for the presentation time needed for detection is lower in general for crossed disparity than for uncrossed disparity [19].

Thus, it is estimated that there is a difference in processing mechanism between the crossed and the uncrossed disparities, and that the capability of the function for the uncrossed disparity is lower. The difference in mechanism may involve the number of disparity-selective neurons and the pattern of connection to the high-level perception process.

5.2. Effect of eye movement

One of the reasons for the change of the latencies of the second and third components due to the presentation position and the disparity may be the effect of eye movement. The increase in latency on presenting the disparity stimulus to the peripheral view field can be considered as reflecting the saccade motion time until the stimulus is captured.

So far as the measurement of the eye potential is concerned, saccade motion was not observed during the experiments. Even if the subject was instructed to perform saccade motion in response to the stimulus, the start time of the eye movement changed depending on the stimulus presentation position. In addition, when a monocular stimulus was presented at the same position, no remarkable change in latency was observed. Thus, saccade is unlikely to be the reason for the increase in latency.

It is reported that the disparity stimulus induces congestive motion and that the disparity is adjusted to stay within the merge range. Thus, it is conceivable that the difference between the crossed and uncrossed disparities is produced by different mechanisms of motion for switching between the congestion and release. In the series of experiments in this study, the stimulus presentation time was sufficiently long (700 ms) to allow congestive motion to be induced.

Among the disparities used in the experiment, the disparity of ±36.3′ is larger than Panum's merge range, and it is possible that congestive motion was produced during the experiment. It should be noted, however, that the latency difference is observed for the disparity (±12.1′), which is within Panum's merge range, and that even if the same two kinds of disparities are compared, the latency difference of the two changes with the presentation position.

Thus, it is seen that even if congestive motion affects the latency difference, the onset time of the motion is important, rather than the eye movement duration. In other words, the major factor is not the mechanism of motion,

such as eye muscle, but the neural mechanism. In particular, it is estimated that fast congestive motion with a latency of 60 to 80 ms is related to the response of the disparity-selective neuron in V1 [20]. In other words, the reason for the latency difference between the crossed and uncrossed disparities appears to extend down to the disparity-selective neuron in V1.

5.3. Processing of anti-RDS

The disparity-selective neuron in V1 responds to the RDS with a particular disparity. Cumming and Parker reported that when an anti-RDS with the same disparity is presented to the above neuron, the response may be inhibited [6]. The phenomenon of a change in peak latency depending on the disparity, reversing the tendency in the case of anti-RDS, corresponds to the verification of the report by Cumming and Parker as a macroscopic electrical activity. It is strongly suggested that the anti-RDS inhibits the disparity-selective neuron. This is very interesting.

The binocular competition mechanism appears likely to be related to the perception of anti-RDS. The change in the latency of VEP for anti-RDS depending on the sign of the disparity, and the reversal of the dependency on the disparity between anti-RDS and RDS, correspond to the response characteristics of the disparity-selective neuron in V1. Consequently, the perception process of the binocular competition should also be based on the local disparity detection mechanism. Binocular competition is closely related to binocular stereo vision (Section 2.4). Furthermore, the VEP for anti-RDS exhibits the same waveform as in the case of RDS (Section 4.3.3). Consequently, it is inferred that the two mechanisms are based on a common neural mechanism and interact with each other.

As a mechanism that discriminates competition from binocular stereo vision, it will be adequate to assume a mechanism for detecting an unpairedness match between the two scenes. As was discussed in Section 2.3, the detection of unpairedness between the scenes of the two eyes appears to be related to ordinary stereo-vision experiences, such as the extraction of depth discontinuity information. In order to investigate the neural mechanism that detects interocular unpaired regions, a simulation experiment with the following model was performed.

5.4. Binocular energy model

As a model to describe the response of the disparity-selective neuron in V1, Ohzawa and colleagues proposed the binocular energy model [21]. It is a model which agrees well with the response of the disparity-selective neuron derived by neurophysiological experiments. It can also account for the reversal of the disparity response tendency of the neuron that is reversed in the case of antiRDS.

Figure 14 shows an outline of the binocular energy model. In this model, the neuron response in the receptor field is described by means of a Gabor function [which is Eq. (2) when only the one-dimensional space is considered]. x is a continuous variable representing the position in space or on the retina. σ_i is the standard deviation of the Gauss function, which represents the size (in pixels) of the window function. ω_i is the center spatial frequency (cycles/pixel). ϕ is the phase (rad) of the Gabor function.

For the spatial positions of the left and right scenes, let the image intensity be $I_l(x)$ and $I_r(x)$, respectively. Then, the responses (E_l and E_r) of the monocular neurons of the left and the right images are given by the convolution integral as in Eq. (3). x_l and x_r represent the positions (pixels) of the centers of the receptor fields in the left and the right images, respectively:

$$g(x,\phi,i)=\frac{1}{\sqrt{2}\pi^{\frac{3}{2}}\sigma_i}\exp\left(-\frac{x^2}{2\sigma_i^2}\right)\times\sin(2\pi\omega_i x+\phi) \tag{2}$$

$$E_l(x_l,\phi,i)=\int I_l(x)g(x-x_l,\phi,i)dx$$
$$E_r(x_r,\phi,i)=\int I_r(x)g(x-x_r,\phi,i)dx \tag{3}$$

Among the binocular neurons with the disparity-selective property, the output of a simple cell is given by the linear

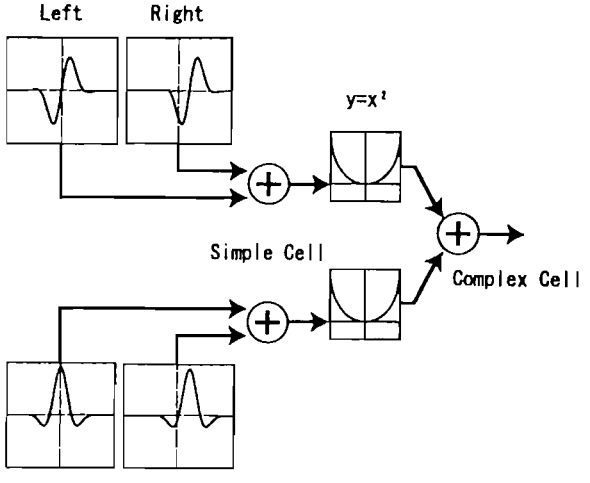

Fig. 14. A binocular energy neuron that prefers nonzero disparity.

sum [Eq. (4)] of the left and right monocular neurons. The output of the complex cell is given by the square-sum of the outputs from two simple cells with orthogonal phases [Eq. (5)]. In this study, the two images have the same receptor field response, given by the Gabor function, and the deviation of the centers of the receptor fields of two images is the coding of the binocular disparity.

The following $S(x_l, x_r, \phi, i)$ and $C(x_l, x_r, \phi, i)$ represents the selectivity for the disparity $(x_r - x_l)$:

$$S(x_l, x_r, \phi, i) = E_l(x_l, \phi, i) + E_r(x_r, \phi, i) \quad (4)$$

$$C(x_l, x_r, \phi, i) = S^2(x_l, x_r, \phi, i) + S^2\left(x_l, x_r, \phi + \frac{\pi}{2}, i\right) \quad (5)$$

Various neurons are assumed, which have different phases and spatial frequencies of the Gabor function, and the sum of the outputs from the complex cells for the same disparity is formed [Eq. (6)]. Then, the disparity can be detected almost exactly for RDS without iteration [22]:

$$Out(x_l, x_r) = \sum_{\phi=0}^{\frac{\pi}{2}} \sum_{i} C(x_l, x_r, \phi, i) \quad (6)$$

5.5. Result of processing by binocular energy model

Using the binocular energy model, the responses of the disparity-selective neuron to the RDS, anti-RDS, and unCRD stimuli were simulated. The number of pixels was 50. The size of the stimulus region was 17 pixels. The disparity was set as 4 pixels for the crossed case. A total of 20 Gabor function phases, spaced uniformly in the range $[0, \pi)$, were prepared. Five spatial frequencies were prepared for each octave. The sizes of the window function were set inversely proportional to the spatial frequency: $((\sigma_i, \omega_i) = \{(1, \frac{1}{4}), (2, \frac{1}{8}), (4, \frac{1}{16}), (8, \frac{1}{32}), (16, \frac{1}{64})\})$. The final output was the weighted average of the five trial computations. It was normalized as follows:

$$\widehat{Out}(x_l, x_r) = \frac{Out(x_l, x_r)}{\max_{x_l}(Out(x_l, x_r))} \times \frac{Out(x_l, x_r)}{\max_{x_r}(Out(x_l, x_r))} \quad (7)$$

Figure 15 shows the result of processing for RDS. The horizontal axis is x_l and the vertical axis is x_r. The pixels represent the intensity of $\widehat{Out}(x_l, x_r)$. It is seen that in the central disparity region, the neuron is strongly activated selectively to the corresponding disparity, correctly detecting the disparity. In the case of anti-RDS, the neuron

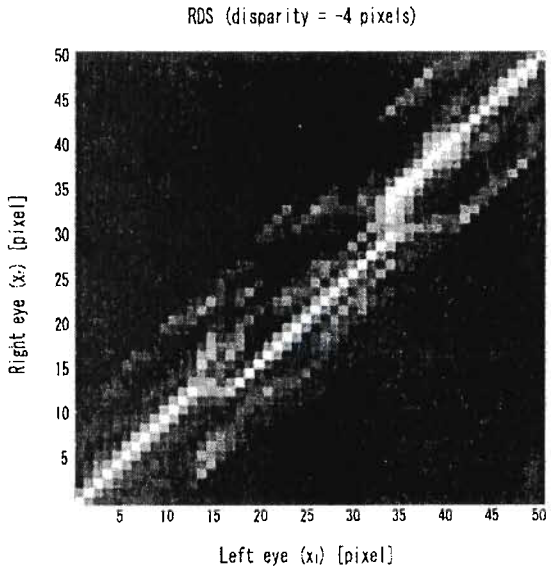

Fig. 15. The responses of binocular energy neurons to an RDS.

selective to the disparity in the stimulus region is strongly inhibited (Fig. 16).

In the case of the unCRD, there is no particular disparity in the stimulus region, and various neurons are simply activated as a whole (Fig. 17). Comparing these results, however, it is seen that there is a common aspect

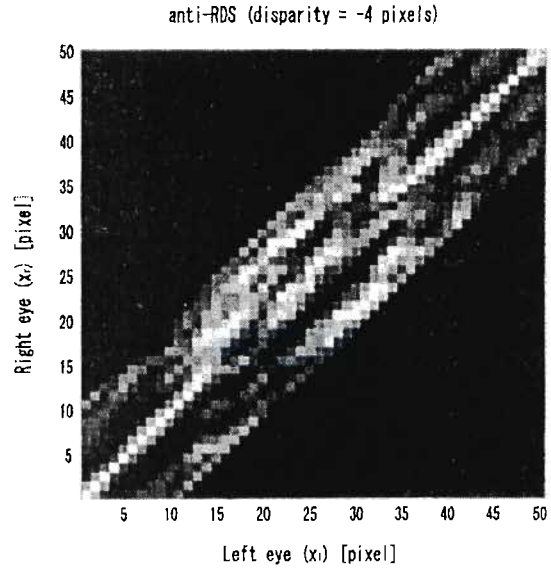

Fig. 16. The responses of binocular energy neurons to an anti-RDS.

between the activity pattern of the disparity-selective neuron in the interocularly unpaired region of RDS, and the activity pattern of the disparity-selective neuron in the case of anti-RDS and the uncorrelated stimulus. In any case, considering the position on the retina, it is not true that neurons selective to a particular disparity are highly activated; instead various disparity-selective neurons are uniformly activated. Consequently, if various disparity-selective neurons at a particular position on the retina are activated, it is inferred that the region is unpaired between the two images.

5.6. Interocularly unpaired detector

Based on the above results, the authors propose an interocularly unpaired detection mechanism shown in Fig. 18, as a mechanism for detecting unpairedness between images. This mechanism is prepared for each position on the retina. When the various disparity selection neurons are uniformly activated at the considered retinal position, the position is detected as an interocularly unpaired region, and it is assumed that binocular competition is induced. The merge process in binocular stereo vision, on the other hand, is defined as the case in which a particular disparity-selective neuron exhibits high activity at the retinal position.

Since binocular competition can occur for each local region of the view field, it is satisfactory to prepare the interocularly unpairedness detection mechanism for each retinal position. The detection mechanism is prepared for both the left and right eyes. The reason is that binocular

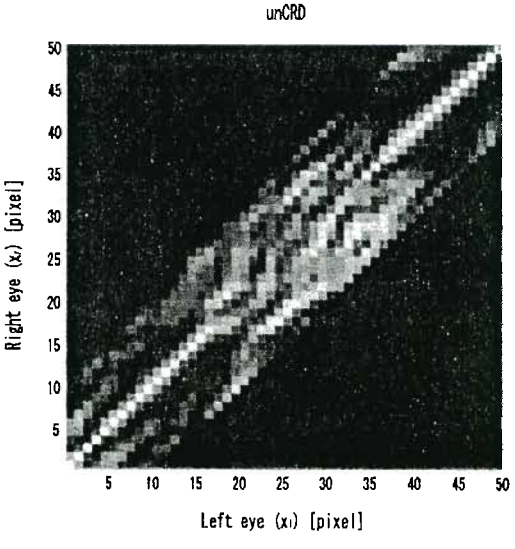

Fig. 17. The responses of binocular energy neurons to an unCRD.

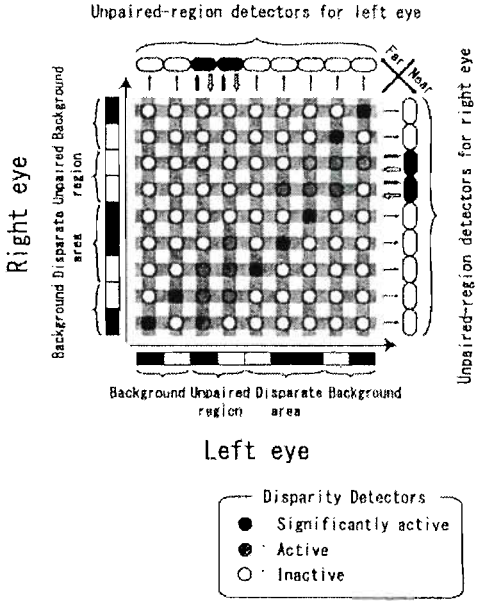

Fig. 18. A model for the detection mechanism of interocularly unpaired regions.

competition mostly occurs between the left and right eyes (Section 2.4), and the perception of the interocularly unpaired region depends on the eye used for the input [12].

The proposed model is one that realizes binocular competition based on binocular neurons, as is suggested by recent research [15]. In other words, interocular unpairedness can be handled integrally with the binocular stereo vision process. It is estimated that the interocularly unpaired detection mechanism is composed of common disparity-selection neurons, which receive inputs from the eye-priority column. It is believed that the inhibition of binocular competition [1] and view field conflicts is induced by inhibiting the above neurons.

McLoughlin and Grossberg have proposed a binocular stereo-vision model considering the interocularly unpaired region [11]. Their model is derived by formulating a problem related to depth discontinuity processing in binocular stereo vision. The model does not consider the processing of general interocularly unpaired stimuli, such as anti-RDS that induces binocular competition. In addition, their model assumes the monocular neuron as the neural basis for detection of the interocularly unpaired region. This is inconsistent with the neurophysiological finding that binocular competition originates from the binocular neuron [15].

Assuming the proposed interocularly unpaired detection mechanism, it is possible to account qualitatively for the tendency that the latency is increased only by the

anti-RDS with crossed disparity, in contrast to anti-RDS with uncrossed disparity and unCRD (Section 4.3.3). Based on the experiment using RDS, it is judged that the processing ability is higher for crossed disparity than for uncrossed disparity, and that there exists a difference in the number of disparity-selective neurons and their activities (Section 5.1).

The following interpretation is considered. When an unCRD or anti-RDS with uncrossed disparity is presented, a large number of disparity-selective neurons are activated, including the major disparity-selective neurons. But when an anti-RDS with crossed disparity is presented, the major crossed parity-selective neurons are inhibited, and the interocularly unpaired detection processing is delayed, increasing the latency.

5.7. Framework of binocular vision mechanism

Based on the above discussion, a framework for the neural mechanism is proposed, including the whole of binocular vision (Fig. 19). The visual inputs from both eyes are projected to the occipital area and are processed first by the local disparity detection mechanism. When the stimulus can be paired between the two eyes, only a particular disparity-selective neuron is activated, and the binocular disparity is determined. When the stimulus is unpaired between the eyes, various disparity-selective neurons are activated, and interocular unpairedness is detected. The interocularly unpaired detection mechanism is related to the binocular stereo-vision processing in extracting the depth discontinuity information. The local disparity information and the interocularly unpaired information are sent to the higher-level visual area through the posterior pathway, and generate the depth sensation and binocular competition perception.

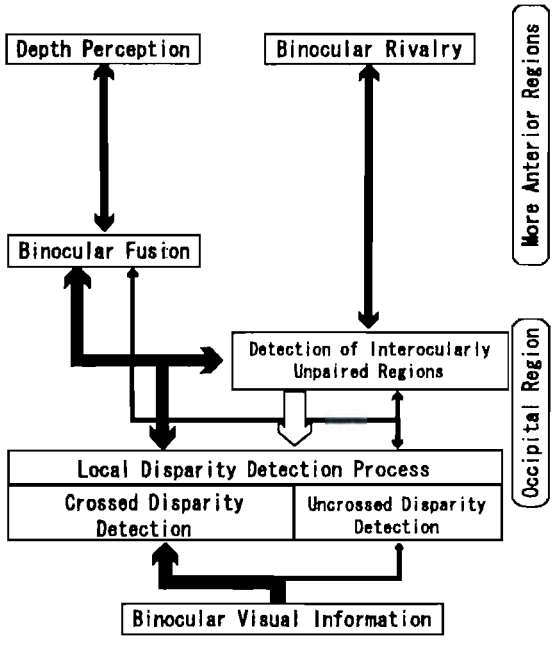

Fig. 19. A model framework for stereoscopic depth perception and binocular rivalry.

It is suggested that the first component is a reflection of the local disparity detection process. The second and third components, on the other hand, are the responses which propagate from the occipital area to the frontal areas. They seem to reflect binocular visual information processing at a higher level, such as binocular visual information processing, including global stereo-vision processing and binocular competition processing. The present investigation, however, casts little light on their roles in the detailed perception process.

In this study, the post-parietal electrode Pz is used as a typical position, and the data are compared to the data from the occipital O2 electrode. It is not true, however, that binocular vision processing following local disparity detection is restricted to the post-parietal area. There is a study that emphasizes the role of the post-temporal area rather than that of the post-parietal area [17].

6. Conclusions

The VEP for RDS presentation is examined, and it is shown that the peak latency is greatly affected by the stimulus presentation position, the disparity, and the correlation between the images of the two eyes. It is inferred from experimental results that both binocular competition and binocular stereo vision use the local disparity detection mechanism as a common neural basis. The interocularly unpaired detection mechanism is used to discriminate these. Based on simulation of the response of the disparity-selective neuron in the binocular energy model, a mechanism is proposed in which the position is detected as the interocularly unpaired region, based on the activities of various disparity-selective neurons for each position on the retina.

REFERENCES

1. Julesz B. Binocular depth perception without familiarity cues. Science 1964;145:356–362.
2. Julesz B, Kropfl W, Petrig B. Large evoked potentials to dynamic random-dot correlograms and stereograms permit quick determination of stereopsis. Proc Natl Acad Sci USA 1980;77:2348–2351.

3. Fenelon B, Neill RA, White CT. Evoked potentials to dynamic random dot stereograms in upper, center and lower fields. Doc Ophthalmol 1986;63:151–156.
4. Ohmoto T, Hatsukawa K, Murai Y. VEP for disparity stimulus using dynamic random-dot pattern. J Jpn Soc Ophthalmol 1984;88:559–564.
5. Miyawaki Y, Yanagida Y, Maeda T, Tachi S. Properties of two-peak wave in visual evoked response in depth perception. Trans IEICE 1999;J82-D-II:961–972.
6. Cumming BG, Parker AJ. Responses of primary visual cortical neurons to binocular disparity without depth perception. Nature 1997;389:280–283.
7. Poggio GF, Gonzalez F, Krause F. Stereoscopic mechanisms in monkey visual cortex: Binocular correlation and disparity selectivity. J Neurosci 1988;8: 4531–4550.
8. DeAngelis GC, Cumming BG, Newsome WT. Cortical area MT and the perception of stereoscopic depth. Nature 1998;394:677–680.
9. Marr D (Inui T, Ando H, translators). Vision—Computation theory of vision and representation in brain. Sangyo Tosho; 1987.
10. Inui T (editor). Perceptive psychology 1—Perception and motion. University of Tokyo Press; 1995. p 15–48.
11. McLoughlin N, Grossberg S. Cortical computation of stereo disparity. Vision Res 1998;38:91–99.
12. Nakayama K, Shimojo S. Da Vinci stereopsis: Depth and subjective occluding contours from unpaired image points. Vision Res 1990;30:1811–1825.
13. Blake R. A neural theory of binocular rivalry. Psychol Rev 1989;96:145–167.
14. Logothetis NK, Leopold DA, Sheinberg DL. What is rivalling during binocular rivalry? Nature 1996;380: 621–624.
15. Sengpiel F, Blakemore C, Harrad R. Interocular suppression in the primary visual cortex: A possible neural basis of binocular rivalry. Vision Res 1995;35:179–195.
16. Blake R, Yang Y, Wilson HR. On the coexistence of stereopsis and binocular rivalry. Vision Res 1991;31:1191–1203.
17. Neill RA, Fenelon B. Scalp response topography to dynamic random dot stereograms. Electroencephalogr Clin Neurophysiol 1988;69:209–217.
18. Richards W. Anomalous stereoscopic depth perception. J Opt Soc Am 1971;61:410–414.
19. Patterson R, Cayko R, Short GL, Flanagan R, Moe L, Taylor E, Day P. Temporal integration differences between crossed and uncrossed stereoscopic mechanisms. Percept Psychophys 1995;57:891–897.
20. Masson GS, Busettini C, Miles FA. Vergence eye movements in response to binocular disparity without depth perception. Nature 1997;389:283–286.
21. Ohzawa I, DeAngelis GC, Freeman RD. Stereoscopic depth discrimination in the visual cortex: Neurons ideally suited as disparity detectors. Science 1990;249:1037–1041.
22. Fleet DJ, Wagner H, Heeger DJ. Neural encoding of binocular disparity: Energy models, position shifts and phase shifts. Vision Res 1996;36:1839–1857.

AUTHORS

Ryusuke Hayashi (student member) received his B.S. degree from the Department of Physical Instrumentation of the University of Tokyo in 1997 and completed the M.E. program (physical instrumentation) in 1999. He is now a doctoral candidate, engaged in research on brain wave measurement and visual information processing models. He is a student member of ARVO, IEEE, the Society of Instrument and Control Engineers, the Japan Vision Society, the Japan Neural Network Society, and the Japan Virtual Reality Society.

AUTHORS (continued) (from left to right)



Yoichi Miyawaki (student member) received his B.S. degree from the Department of Applied Physics of Osaka University in 1996 and completed the M.E. program (physical instrumentation) in 1998. He is now a doctoral candidate (interdisciplinary engineering), engaged in research on biological signal measurement and visual information processing. He received an Encouragement Award from the Society of Instrument and Control Engineers in 1999. He is a student member of SFN, the Society of Instrument and Control Engineers, the Japan Vision Society, and the Japan Virtual Reality Society.

Taro Maeda (member) received his B.S. degree from the Department of Physical Instrumentation of the University of Tokyo in 1987 and joined the Ministry of International Trade and Industry, Research Institute of Mechanical Technology. He became a research associate at the Advanced Science and Technology Research Center of the University of Tokyo in 1992. He was appointed a research associate (1994) and lecturer (1997) in the Graduate School of Engineering. He received an adjunct appointment as lecturer, Science Group, Information Science, in 2001. He holds a D.Eng. degree (University of Tokyo). His research interests are human perception characteristics and neural network model, man–machine interface, and tele-existence. He received a Paper Award (1990) and Encouragement Award (1997) from the Society of Instrument and Control Engineers, and a Technology Award (1991) from the Robotics Society of Japan. He is a member of ARVO, IEEE, SFN, the Society of Instrument and Control Engineers, the Japan Vision Society, the Japan Neural Network Society, the Japan Virtual Reality Society, and the Robotics Society of Japan.

Susumu Tachi (member) received his B.S. degree from the Department of Physical Instrumentation of the University of Tokyo in 1968 and completed the doctoral program in 1973. He holds a D.Eng. degree. He became a research associate at the University of Tokyo in 1973. Since 1975, he was affiliated with the Ministry of International Trade and Industry, Research Institute of Mechanical Technology, successively as researcher, chief researcher, chief of Remote Control Laboratory, and chief of Robotics Laboratory. He was a visiting researcher at Massachusetts Institute of Technology. He was appointed an associate professor (1989) and professor (1992) at the Advanced Science Technology Research Center of the University of Tokyo. He has been a professor (physical instrumentation) at the Graduate School of Engineering since 1994. His research interests are seeing-eye dog robots, tele-existence, and virtual reality. He has received an IEEE/EMBS Award, MITI Minister's Award, and IMEKO Special Achievement Award. He served as chair of the IMEKO Robotics Conference; SICE Fellow and President, Japan Virtual Reality Society.