Vol. 39 No. 1 Transactions of Information Processing Society of Japan Jan. 1998

Regular Paper

Neural Net Pattern Recognition Equation for Stereoscopic Vision

YASUNARI YOSHITOMI,† TETU KANDA,†,☆ TETURO KITAZOE†
and TOMOHIRO SHII†

An algorithm for stereoscopic vision was developed on the basis of recent pattern recognition equations describing a dynamic self-organizing process with competition and cooperation. The algorithm includes a potential model and a simulated annealing process, and can thus be used to simulate stereoscopic depth perception from any initial situation of binocular neurons. It has been shown that simulated annealing is very useful in allowing a disparity map to develop by overcoming the local minimum in the potential.

1. Introduction

In stereoscopic vision, a three-dimensional scene is imaged from two different points, and the acquired two-dimensional images are then reconstructed in order to measure the disparity and estimate the depth. Our brains seem to be able to carry out this complicated procedure by means of a kind of neural network. It has been shown that it is possible for human to extract a depth map by using the disparity of corresponding image features 1). Various algorithms for solving the problem of stereoscopic vision have been reported 2)~7). Some of them have also been mentioned in reviews of stereoscopic vision 8),9). In recent investigations, a mathematical model in which the potential of a neuron is treated as a variable in the equation for cooperation and competition has been proposed as a theory of neural fields 2). The main stream of algorithms for solving the problem of stereoscopic vision has been based on finding the global minimum of the energy or of the potential function 4)~6).

On the other hand, an algorithm that is not based on the energy or the potential function has also recently been developed 7). In this algorithm, a variable, called the activity of binocular neurons, is assigned to each possible disparity, and some procedures such as competition, cooperation, and self-organizing processes also seem to be used for simulating the neural networks in our brains. However, our investigation of the algorithm has shown that special selection of the initial conditions is indispensable for a disparity map to develop as a self-organizing process. The initial condition of binocular neurons in our brains does not seem to be limited in such a way without some special reason.

As a means of analyzing the problem, we propose a new algorithm for stereoscopic vision, based on the modified pattern recognition equations. The algorithm includes a potential model and a simulated annealing process, and can thus be applied to any initial condition of binocular neurons by finding the global minimum of the energy or of the potential function.

2. Algorithm for Stereoscopic Vision

2.1 Coupled Pattern Recognition Equations

Since the present study is based on the Coupled Pattern Recognition (CPR) equations developed by Reimann and Haken7), the CPR equations are briefly described below.

In the CPR equations, there are three constraints, namely, the search area, the continuity, and the uniqueness. Starting from a feature in the left image, the feature in the right image which is the most similar to that of the left image is found by means of the CPR equations.

The function $S(u,v,a,b)$ for the similarity between the feature L of an image point (u,v) in the left image and the feature R of an image point $(u+a, v+b)$ in the right image is given as follows:

$$S(u,v,a,b) = E - \frac{\iint |L(x,y,u,v) - R(x+a,y+b,u+a,v+b)|dxdy}{\iint f(x,y)dxdy}, \tag{1}$$

where

$$L(x,y,u,v) = f(x-u,y-v)g_L(x,y) - \overline{g}_L(u,v). \tag{2}$$

---

† Department of Computer Science and Systems Engineering, Faculty of Engineering, Miyazaki University
☆ Presently with Nintendo, Kyoto

$R(x, y, u, v)$ is defined correspondingly. E is a positive constant defined in such a way that the similarity measure has a positive maximum for the highest similarity. $g_L(x, y)$ and $g_R(x, y)$ are the left and right stereo input data at the point (x, y). The area that defines the features is given by the window function $f(x, y)$. $\overline{g}_L(u, v)$ and $\overline{g}_R(u, v)$ are the local mean gray values of the features corresponding to the image points (u, v).

Since the disparity can vary only between certain limits, the search has to be performed in between these limits. The disparity search area (DSA) is defined as $-a_s \leq a \leq a_s$, $-b_s \leq b \leq b_s$. In the definition of $S(u, v, a, b)$, a and b are limited in those ranges.

A time-dependent activity of binocular neurons $\xi^{uv}_{ab}(t)$ is introduced. The activity will obey a dynamic process and have the value of the similarity measure as initial value. If the continuity constraint is ignored for a moment, in the course of the process the activity of the binocular neuron with the highest similarity measure in the DSA should exceed the activity of the others, while the others should vanish. Such a winner-takes-all system is given by the following equations for CPR processes [7]:

$$\begin{aligned}\dot{\xi}^{uv}_{ab}(t) &= \lambda^{uv}_{ab}\xi^{uv}_{ab}(t) \\ &\quad - (B + C)\sum_{a'b'}^{\prime} \xi^{uv}_{a'b'}(t)^2 \xi^{uv}_{ab}(t) \\ &\quad - C\xi^{uv}_{ab}(t)^3 + D\sum_{u'v'}^{\prime} \xi^{u'v'}_{ab}(t)\xi^{uv}_{ab}(t)\end{aligned} \tag{3}$$

The first term represents an exponential growth of the amplitude $\xi^{uv}_{ab}(t)$, the second term denotes the effect of competition among all amplitudes, the third term restricts the growth of an amplitude, and the fourth term denotes the cooperative coupling with the different pattern recognition processes. B, C, and D are positive constants that have to be chosen appropriately. Here, one possibility is to choose λ^{uv}_{ab} as a deviation of the similarity measure from its mean value in the DSA. In $\sum'_{a'b'}$ the summation indices $a'b'$ run over the DSA with the restrictions $a' \neq a$ and $b' \neq b$. In $\sum'_{u'v'}$ the summation indices u', v' run over the cooperation area (CA) defined as $u - l \leq u' \leq u + l$, $v - l \leq v' \leq v + l$ with the restrictions $u' \neq u$ and $v' \neq v$.

To solve the correspondence problem of stereoscopic vision, the CPR equations are used. The CPR equations have the specific feature that they cannot be derived from a potential or energy function, because of the fourth term. The values of the positive semi-definite similarity function, which are adjusted by E, are chosen as the initial values of the corresponding binocular neurons $\xi^{uv}_{ab}(t)$.

In the algorithm, a variable, called the activity of binocular neurons, is assigned to each possible disparity. Moreover, procedures such as competition, cooperation, and self-organizing processes also seem to be used for simulating neural networks in our brains. However, our simulation with the CPR equations showed that special selection of a positive constant E in the initial conditions of the activity of binocular neurons $\xi^{uv}_{ab}(t)$ is indispensable for a disparity map to develop as a self-organizing process, and that the activity of binocular neurons that will win changes to a negative value at equilibrium in the CPR process if a negative constant E is chosen. That is to say, in the similarity function, the constant E should be positive and defined in such a way that the similarity measure will have a positive maximum for the highest similarity. Since there is no special reason for the initial condition of binocular neurons in our brains to be constrained in such a way as described above, the selection of the initial condition of $\xi^{uv}_{ab}(t)$ does not seem to be reasonable. For this reason, the CPR equations are modified with a potential model in order to analyze more clearly the structure of the equations, especially, the role of the constant E in the similarity function.

2.2 Potential Model and Simulated Annealing

The CPR equations cannot be derived from a potential or energy function, because of the fourth term $+D\sum'_{u'v'} \xi^{u'v'}_{ab}(t)\xi^{uv}_{ab}(t)$. In general, however, in order to analyze the kinetics of $\xi^{uv}_{ab}(t)$, it is very helpful to introduce a potential as in the following equations:

$$\dot{\xi}^{uv}_{ab}(t) = -\frac{\partial U}{\partial \xi^{uv}_{ab}(t)} \tag{4}$$

The potential is introduced by modifying the CPR equations (3) as follows:

$$\begin{aligned}U(\xi^{uv}_{ab}(t)) &= -\frac{1}{2}\lambda^{uv}_{ab}\xi^{uv}_{ab}(t)^2 \\ &\quad + \frac{1}{2}(B + C)\sum_{a'b'}^{\prime} \xi^{uv}_{a'b'}(t)^2 \xi^{uv}_{ab}(t)^2\end{aligned}$$

$$+ \frac{C}{4}\xi_{ab}^{uv}(t)^4 - \frac{D}{2}\sideset{}{'}\sum_{u'v'}\xi_{ab}^{u'v'}(t)^2\xi_{ab}^{uv}(t)^2, \quad (5)$$

where $U(\xi_{ab}^{uv}(t))$ is the non-vanishing part of U under the derivative in Eq. (4).

It should be noticed that the potential function introduced above consists of terms of $\xi_{ab}^{uv}(t)^2$ and $\xi_{ab}^{uv}(t)^4$, and that the potential function can thus have two minimum points with the same potential value in both positive and negative regions of $\xi_{ab}^{uv}(t)$. Accordingly, the potential function can produce two equilibrium points, in both positive and negative regions of $\xi_{ab}^{uv}(t)$, as in the previous study without a potential model[7].

Since the equilibrium value of $\xi_{ab}^{uv}(t)$ for the winning binocular neuron should be unique and positive, we add a fifth term, $-(E/3)\xi_{ab}^{uv}(t)^3$, in the potential function to break the left-right symmetry, where E is positive. Consequently, the final potential function, given below, can have a global minimum in the positive region of $\xi_{ab}^{uv}(t)$ and a local minimum in its negative region, as shown by the example in **Fig. 1**, where $\xi_{ab}^{uv}(t)$ is denoted by ξ.

$$\begin{aligned} U(\xi_{ab}^{uv}(t)) = & -\frac{1}{2}\lambda_{ab}^{uv}\xi_{ab}^{uv}(t)^2 \\ & + \frac{1}{2}(B+C)\sideset{}{'}\sum_{a'b'}\xi_{a'b'}^{uv}(t)^2\xi_{ab}^{uv}(t)^2 \\ & + \frac{C}{4}\xi_{ab}^{uv}(t)^4 \\ & - \frac{D}{2}\sideset{}{'}\sum_{u'v'}\xi_{ab}^{u'v'}(t)^2\xi_{ab}^{uv}(t)^2 \\ & - \frac{E}{3}\xi_{ab}^{uv}(t)^3 \qquad (6) \end{aligned}$$

Therefore, the initial value of $\xi_{ab}^{uv}(t)$ can have a strong influence on the CPR processes. The structure of the potential function is discussed here. It is rewritten in the following form:

$$\begin{aligned} U(\xi_{ab}^{uv}(t)) = & \frac{\alpha}{2}\xi_{ab}^{uv}(t)^2 - \frac{E}{3}\xi_{ab}^{uv}(t)^3 \\ & + \frac{C}{4}\xi_{ab}^{uv}(t)^4 \qquad (7) \end{aligned}$$

$$\begin{aligned} \alpha = & -\lambda_{ab}^{uv} + (B+C)\sideset{}{'}\sum_{a'b'}\xi_{a'b'}^{uv}(t)^2 \\ & - D\sideset{}{'}\sum_{u'v'}\xi_{ab}^{u'v'}(t)^2 \qquad (8) \end{aligned}$$

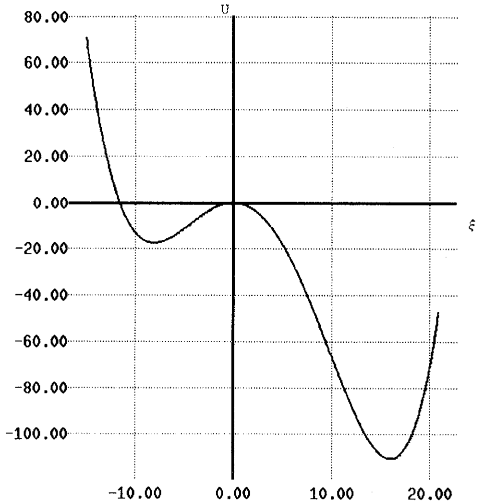


Fig. 1 Example of a potential function.

Since λ_{ab}^{uv} is a deviation of the similarity measure from its mean value in the DSA, α depends not only on the initial value of $\xi_{ab}^{uv}(t)$ itself, but also on those of binocular neurons in the DSA. Moreover, α is time-dependent through the values of $\xi_{a'b'}^{uv}(t)$ and $\xi_{ab}^{u'v'}(t)$ of binocular neurons in the DSA and CA, respectively.

First, we discuss the equilibrium of ξ_{ab}^{uv}, based on the assumption that the global minimum point of the potential function can give the unique equilibrium point of ξ_{ab}^{uv}. When $\xi_{ab}^{uv}(t) = 0$, we can obtain the following equations from Eqs. (4), (7), and (8):

$$\alpha_0\xi_{ab}^{uv} - E\xi_{ab}^{uv\,2} + C\xi_{ab}^{uv\,3} = 0, \qquad (9)$$

where

$$\begin{aligned} \alpha_0 = & -\lambda_{ab}^{uv} + (B+C)\sideset{}{'}\sum_{a'b'}\xi_{a'b'}^{uv\;2} \\ & - D\sideset{}{'}\sum_{u'v'}\xi_{ab}^{u'v'\,2}. \qquad (10) \end{aligned}$$

We then get three solutions for the above equations (9):

$$\xi_{ab}^{uv} = 0, \qquad (11)$$

$$\xi_{ab}^{uv} = \xi_{\pm} = \frac{1}{2C}(E \pm \sqrt{E^2 - 4C\alpha_0}), \quad (12)$$

where the order of sign ± is the same in the two places. For binocular neurons, there are several cases relating to the equilibrium, depending on the values of $E^2 - 4C\alpha_0$, $E^2 - 4.5C\alpha_0$, and α_0, as shown below (**Fig. 2**).

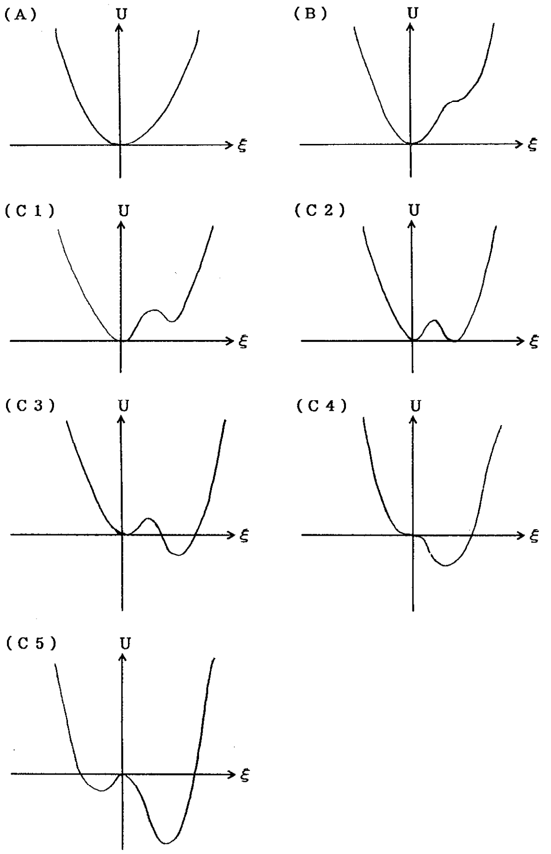


Fig. 2 Schematic diagram of potential functions.

A. $E^2 - 4C\alpha_0$ is negative.

The potential function has a global minimum at $\xi_{ab}^{uv} = 0$. In this case, the binocular neuron becomes inactivated at the point of equilibrium.

B. $E^2 - 4C\alpha_0$ is zero.

The potential function has a global minimum at $\xi_{ab}^{uv} = 0$, and a saddle point at $\xi_{ab}^{uv} = \xi_+ = \xi_- = \frac{E}{2C}$ (positive). Therefore, in this case, the binocular neuron becomes inactivated at the point of equilibrium.

C. $E^2 - 4C\alpha_0$ is positive.

(1) α_0 is positive. $E^2 - 4.5C\alpha_0$ is negative.

The potential function has a global minimum at $\xi_{ab}^{uv} = 0$, where $U(0) = 0$, a local maximum at $\xi_{ab}^{uv} = \xi_-$ (positive), and a local minimum at $\xi_{ab}^{uv} = \xi_+$ (positive), where $U(\xi_+)$ is positive. Therefore, in this case, at the point of equilibrium, the binocular neuron becomes inactivated.

(2) α_0 is positive. $E^2 - 4.5C\alpha_0$ is zero.

The potential function has a global minimum at $\xi_{ab}^{uv} = 0$, where $U(0) = 0$, a local maximum at $\xi_{ab}^{uv} = \xi_-$ (positive), and a global minimum at $\xi_{ab}^{uv} = \xi_+$ (positive), where $U(\xi_+) = 0$. Therefore, in this case, at the point of equilibrium, the binocular neuron becomes inactivated or becomes excited at $\xi_{ab}^{uv} = \xi_+$. The solution of equilibrium is not unique.

(3) α_0 is positive. $E^2 - 4.5C\alpha_0$ is positive.

The potential function has a local minimum at $\xi_{ab}^{uv} = 0$, where $U(0) = 0$, a local maximum at $\xi_{ab}^{uv} = \xi_-$ (positive) and a global minimum at $\xi_{ab}^{uv} = \xi_+$ (positive), where $U(\xi_+)$ is negative. Therefore, in this case, at the point of equilibrium, the binocular neuron becomes excited at $\xi_{ab}^{uv} = \xi_+$.

(4) α_0 is zero.

The potential function has a saddle point at $\xi_{ab}^{uv} = \xi_- = 0$ and a global minimum at $\xi_{ab}^{uv} = \xi_+ = \frac{E}{C}$ (positive). Therefore, in this case, at the point of equilibrium, the binocular neuron becomes excited at $\xi_{ab}^{uv} = \xi_+ = \frac{E}{C}$.

(5) α_0 is negative.

The potential function has a local minimum at $\xi_{ab}^{uv} = \xi_-$ (negative), a local maximum at $\xi_{ab}^{uv} = 0$, and a global minimum at $\xi_{ab}^{uv} = \xi_+$ (positive). Therefore, in this case, at the point of equilibrium, the binocular neuron becomes excited at $\xi_{ab}^{uv} = \xi_+$.

As mentioned above, it is indispensable for the binocular neuron to be excited at a point of equilibrium where $E^2 - 4C\alpha_0$ is positive, that is, where α_0 is smaller than $\frac{E^2}{4C}$. From Eq. (8), it is clear that large λ_{ab}^{uv}, small $\sum'_{a'b'} {\xi_{a'b'}^{uv}}^2$, and large $\sum'_{u'v'} {\xi_{ab}^{u'v'}}^2$ can help the binocular neuron to become excited at the point of equilibrium. In other words, a high similarity function in the DSA, little competition in the DSA, and strong cooperative coupling in the CA help the binocular neuron to become excited at the point of equilibrium.

Next, let us consider the nonequilibrium of $\xi_{ab}^{uv}(t)$. In general, even if α is negative, $\xi_{ab}^{uv}(0)$ with a negative value may not be able to realize excited equilibrium, because of the local minimum of the potential function at a negative value of $\xi_{ab}^{uv}(t) = \frac{1}{2C}(E - \sqrt{E^2 - 4C\alpha})$ and its local maximum at $\xi_{ab}^{uv}(t) = 0$. Even if α is smaller than $\frac{E^2}{4.5C}$ and α is positive, $\xi_{ab}^{uv}(0)$ with a value smaller than $\frac{1}{2C}(E - \sqrt{E^2 - 4C\alpha})$ may not be able to realize excited equilibrium, because of the local minimum of the potential function at $\xi_{ab}^{uv}(t) = 0$ and its local maximum at $\xi_{ab}^{uv}(t) = \frac{1}{2C}(E - \sqrt{E^2 - 4C\alpha})$. Moreover, since α is time-dependent, it may happen that the type of shape of the potential function for each binocular neuron changes drastically in the course of the CPR process, as described in Section 3, Numerical Results and Discussion.

As suggested in the discussion on equilibrium and nonequilibrium, the potential function is time-dependent and can have local minimum, local maximum, and saddle points. In addition, a nonequilibrium phase transition [10)] may occur in the CPR processes. Therefore, equilibrium may not be realized if the initial condition or the kinetics of $\xi_{ab}^{uv}(t)$ is inappropriate. However, the initial stage or the state kinetics of the binocular neurons in our brains is considered to have some variety. Therefore, the model for stereoscopic vision should have some degree of robustness with respect to both the initial condition and the kinetics. In this sense, noise is introduced into the present model, using the so-called simulated annealing method described by the following equations:

$$\Delta\xi_{ab}^{uv}(t) = -\Delta t \frac{\partial U}{\partial \xi_{ab}^{uv}(t)} + n(T) \qquad (13)$$

$$T(s) = \frac{\beta^2}{\log(1+s)} \qquad (14)$$

$t = \Delta ts$, and $n(T)$ denotes random noise with values from $-T$ to T. s is a step in the iteration scheme. The control parameter T corresponds to the temperature of a physical system. β is given appropriately, and is independent of s. As T decreases, $\xi_{ab}^{uv}(t)$ is forced to approach the value that gives the global minimum of the potential function.

Since a variable, $\xi_{ab}^{uv}(t)$, in the present modified CPR equations is treated as the wide-ranging activity of binocular neurons, it should have negative, zero, and positive values. Therefore, a kind of simulated annealing method is indispensable for simulating the CPR processes of stereoscopic vision with the present model.

3. Numerical Results and Discussion

After describing a simulation of one-dimensional stereoscopic vision, we will give the results of a two-dimensional case.

In the one-dimensional case, $\xi_{ab}^{uv}(t)$ was simplified as $\xi_a^u(t)$; the DSA and the CA were also one-dimensional.

The following random 50 dots were used in the present study:

Left (0th to 49th)
10001011011000010010010011101010111001100000011101

Right (0th to 49th)
10001011011000010010010100111010111001100000011101

The displaced right dots were produced by placing the 20th to 29th dots from the left sequence in the 22nd to 31st positions of the right sequence.

In the present study, a was chosen to be from -3 to 3, which corresponded to the DSA. The window function was defined in the range of a from -3 to 3. The value l of the CA was chosen to be 2. The values of the constants B, C, D, and E used in the potential function (6) were 0.01, 0.01, 0.001, and 0.08, respectively. These constant values were selected in such a way that B and C were the same as the values that gave successful simulation results with the previous CPR equations (3), and D and E were selected by trial and error in order to obtain successful simulation results with the modified CPR equations (4) and (6). The first selected value of E in Eq. (1) was 4, which gave successful simulation results with the previous CPR equations (3). Then, the second selected value of E in Eq. (1) was -4, which gave poor simulation results with the previous CPR equations (3). The poor simulation results indicated that there was no winner among the binocular neurons in the CPR processes. The first case presented an initial condition in which every activity of binocular neurons was positive. On the other hand, the second case presented an initial condition in which every activity of binocular neurons was negative.

Figure 3 shows the potential function in the first case of E. The arrow indicates the values of $\xi_2^{20}(t)$ and the potential function at the time. As can be seen in the figure, the CPR process to the point of equilibrium for the winner was successfully simulated. **Figures 4** and **5** show competing processes in the cases where $u = 10$ and $u = 20$, respectively. In the case where $u = 10$, the binocular neuron with $a = 0$ was the winner. On the other hand, in the case where $u = 20$, the binocular neuron with $a = 2$ was the winner, with depth perception achieved.

Next, **Figs. 6** and **7** show competing processes in the second case of E, in which no successful results were given by the previous CPR equations (3), at least in our simulation. As shown in the figures, the CPR processes to

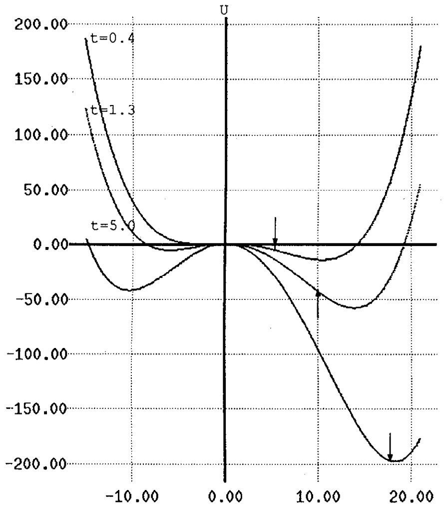


Fig. 3 Potential in the case where $E = 4$, $u = 20$, $a = 2$.

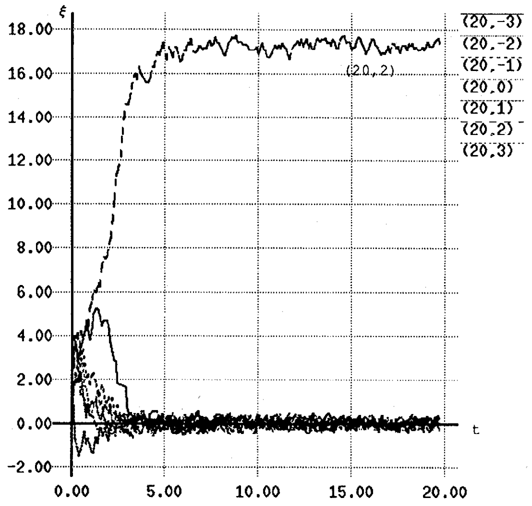


Fig. 5 Competing processes in the case where $E = 4$, $u = 20$.

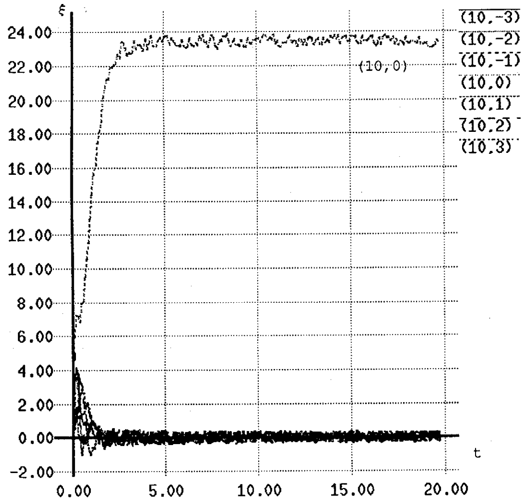


Fig. 4 Competing processes in the case where $E = 4$, $u = 10$.

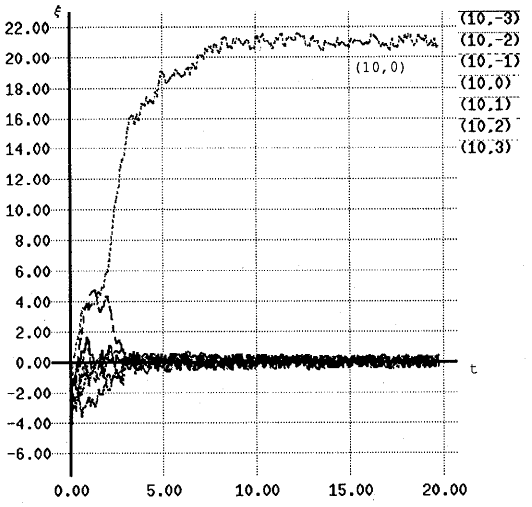


Fig. 6 Competing processes in the case where $E = -4$, $u = 10$.

the point of equilibrium were successfully simulated in our model. In the cases where $u = 10$ and $u = 20$, the binocular neurons with $a = 0$ and $a = 2$ were the winners, respectively. In the case where $u = 20$, depth perception was achieved. The modified CPR equations (4) and (6) gave the successful CPR processes to the point of equilibrium for both values of E, where the final winner among the binocular neurons did not depend on the value of E. However, the kinetics of the CPR processes with the initially negative activity were considerably slower than those with the initially positive activity. Examples of the winner's potential and the loser's potential related to Fig. 7 are shown in **Figs. 8** and **9**, respectively. In Fig. 8, the added numbers indicate the time and the arrows show the value of $\xi_2^{20}(t)$ at each time. In Fig. 8, the potential function is actually changing over time from 0.3 to 9.1. However, since the change is negligibly small, the potential at a time of 0.8 is given as a representative case. Figure 8 shows the process escaping from the local minimum of the potential. The potential function changed drastically in a time-dependent manner after the escape from the local minimum of the potential, which is similar to the time-dependency of the potential function shown in

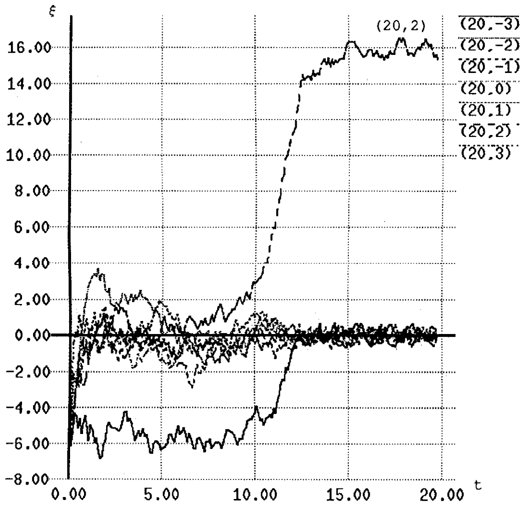


Fig. 7 Competing processes in the case where $E = -4$, $u = 20$.

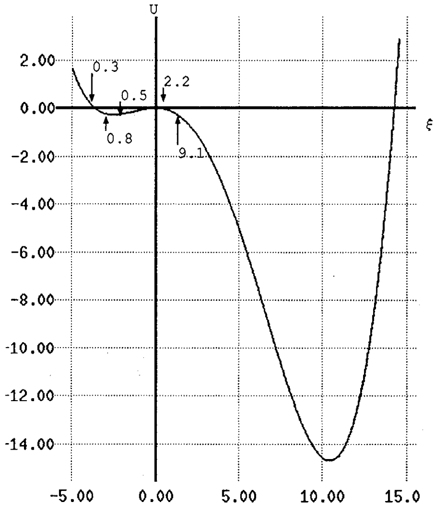


Fig. 8 Potential in the case where $E = -4$, $u = 20$, $a = 2$.

Fig. 3. On the other hand, in the case of the loser in the CPR process, there is always one minimum of the potential function at $\xi_a^u(t) = 0$ in the CPR process, as shown in Fig. 9. The arrows in Fig. 9 also indicate the values of $\xi_1^{20}(t)$ and the potential function at the time.

There can be some changes in the simulation results even if the simulation is carried out under the same conditions, because of random noise in the simulated annealing method. In some cases, the activity of the binocular neuron $\xi_a^u(t)$ may be trapped in the local minimum of the potential, owing to inappropriate annealing.

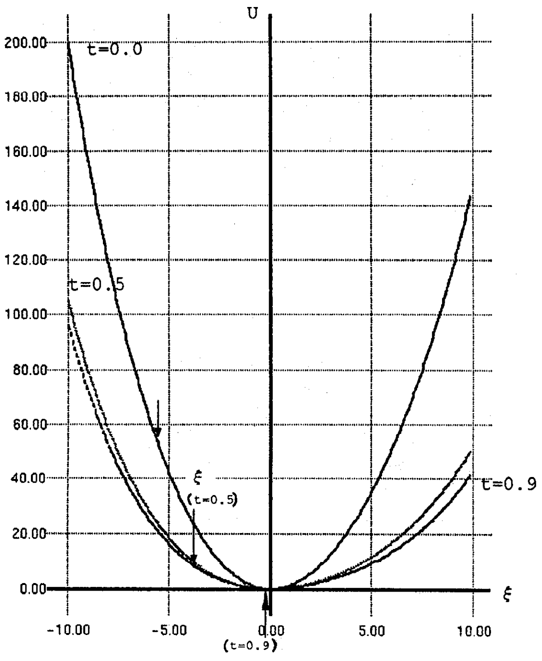


Fig. 9 Potential in the case where $E = -4$, $u = 20$, $a = 1$.

Next, 50 × 50 two-dimensional random dots were used. In the region of $y = 10$ to 39, the displaced right dots were produced by placing the dots of $x = 10$ to 39 from the left sequence in the positions of $x = 12$ to 41 of the right sequence.

a and b were respectively chosen to be from $a = -3$ to 3, $b = -1$ to 1, which corresponded to the DSA. The window function was defined in the range of a from -3 to 3 and b from -1 to 1. The value l of the CA was chosen to be 2. The values of the constants B, C, D, and E used in the potential function (6) were -0.001, 0.01, 0.00025, and 0.08, respectively. These constant values were selected by trial and error in order to obtain successful simulation results with the modified CPR equations (4) and (6). The values of E in Eq. (1) were selected to be the same as those in the case of 1-dimensional simulation. The first selected value of E in Eq. (1) was 4, which gave the successful simulation results with the previous CPR equations (3). The second selected value of E in Eq. (1) was -4, which was not permitted when the previous CPR equations (3) were used. The first case presented an initial condition in which every activity of binocular neurons was positive. On the other hand, the second case presented an initial condition in which every activity of

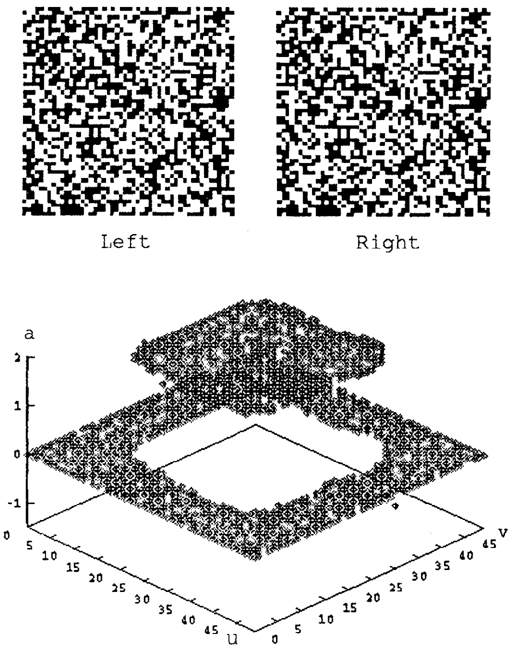


Fig. 10 (Top) Pair of random-dot stereograms presented to the left and right eyes.
(Bottom) 3-dimensional image of the random-dot stereograms viewed by the present neural network model.

binocular neurons was negative.

In both the first and second cases, the CPR process to the point of equilibrium was simulated by using our equations, with the result that depth perception was successfully achieved. **Figure 10** shows that in the second case depth perception was achieved with the present model. Except in cases where the prospective winner was trapped at the local minimum of the potential, the CPR processes to the point of equilibrium were successfully simulated. The data corresponding to these exceptions (about 15% of all cases) are shown as blanks in the lower part of Fig. 10. We obtained the similar results in the first case.

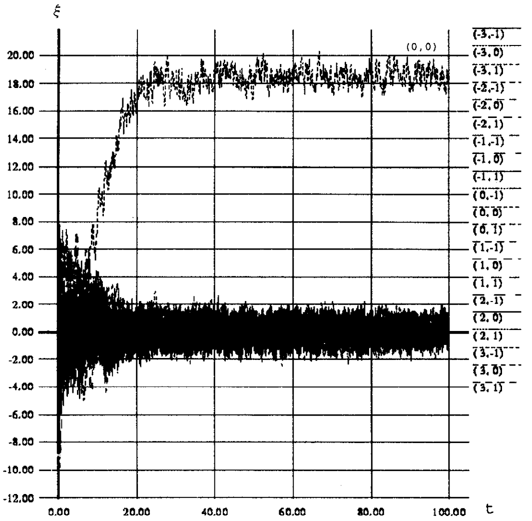


Fig. 11 Competing processes in the case where $E = -4$, $u = 5$, $v = 5$.

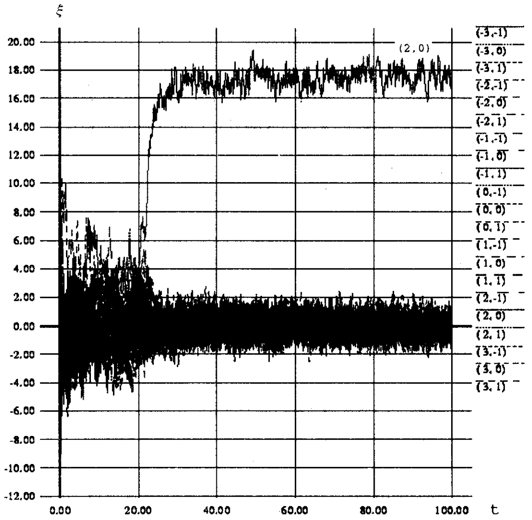


Fig. 12 Competing processes in the case where $E = -4$, $u = 25$, $v = 25$.

Figures 11 and **12** show competing processes in the second case, in which no successful results would be given with the previous CPR equations (3), at least in our investigation. As can be seen in the figures, the CPR processes to the point of equilibrium were successfully simulated in our model. In the cases where $u = 5$, $v = 5$ and $u = 25$, $v = 25$, the binocular neurons with $a = 0$, $b = 0$ and $a = 2$, $b = 0$ were the winners, respectively. In the case where $u = 25$, $v = 25$, depth perception was achieved against the background landscape ($u = 5$, $v = 5$, included), as shown in Fig. 10. The modified CPR equations (4) and (6) presented the successful CPR processes to the point of equilibrium for both values of E, where the final winner among the binocular neurons did not depend on the value of E. However, the kinetics of the CPR processes with the initially negative activity were considerably slower than those with the initially positive activity. Examples of the winner's potential and the loser's potential related to Fig. 12 are shown in **Figs. 13** and **14**,

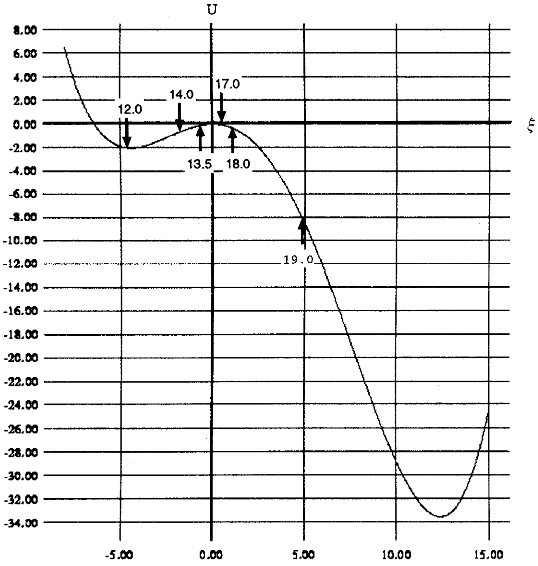


Fig. 13 Potential in the case where $E = -4$, $u = 25$, $v = 25$, $a = 2$, $b = 0$.

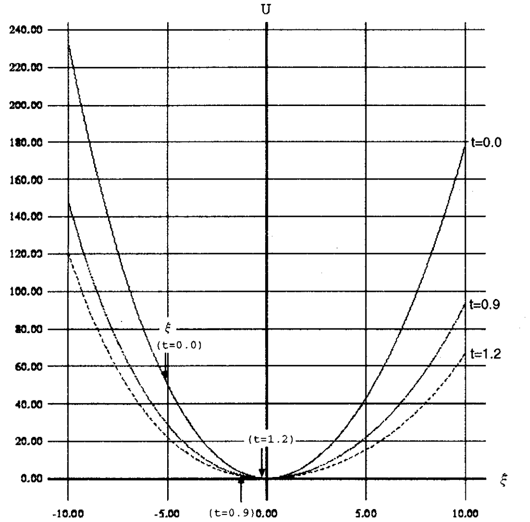


Fig. 14 Potential in the case where $E = -4$, $u = 25$, $v = 25$, $a = 1$, $b = 0$.

respectively. In Fig. 13, the added figures indicate the time and the arrows show the value of $\xi_{2,0}^{25,25}(t)$ at each time. In Fig. 13, the potential function actually changes over time from 12.0 to 19.0. However, since the change is small, the potential at a time of 14.0 is given as a representative case. Figure 13 shows the process escaping from the local minimum of the potential, with the help of annealing. The potential function changed dramatically in a time-dependent manner after the escape from the local minimum of the potential, which is similar to the time-dependency of the potential function shown in Fig. 3. On the other hand, in the case of the loser in the CPR process, there is always one minimum of the potential function at $\xi_{ab}^{uv}(t) = 0$ in the CPR process, as shown in Fig. 14. The arrows in Fig. 14 also show the value of $\xi_{1,0}^{25,25}(t)$ at each time.

According to the present study, the special selection of the constant E for the similarity function in the previous study[7)] is considered to have the role of avoiding poor competition in which there is no winner. The special selection of the initial condition of $\xi_{ab}^{uv}(t)$ seems to have essentially the same effect as escaping from a local minimum in the potential function. However, in the present model, special selection of the initial condition of $\xi_{ab}^{uv}(t)$ is unnecessary because of the simulated annealing method. Therefore, in a sense, the existence of the constant E in the similarity function is not indispensable.

4. Conclusion

An algorithm for stereoscopic vision was developed on the basis of recent pattern recognition equations describing a dynamic self-organizing process with competition and cooperation. The effects of exponential growth, competition, and cooperation in the CPR processes are expressed by a single parameter. A potential model is introduced, allowing the dynamic situation to be seen clearly in the potential function. A simulated annealing process is also introduced, as a result of which the algorithm can be used to simulate stereoscopic depth perception for any initial situation of binocular neurons. It has been shown that the simulated annealing is very useful in allowing a disparity map to develop by overcoming the local minimum in the potential. This might suggest that some kind of noise in our brains is indispensable for stereoscopic depth perception.

References

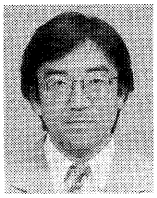
1) Julesz, B.: *Foundations of Cyclopean Perception*, University of Chicago Press, Chicago (1971).

2) Amari, S. and Arbib, M.A.: Competition and Cooperation in Neural Nets, *Systems Neuroscience*, pp.119–165, Academic Press (1977).

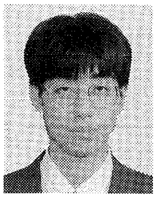
3) Hentschel, H.G.E. and Fine, A.: Statistical Mechanics of Stereoscopic Vision, *Phys. Rev. A*, Vol.40, pp.3983–3997 (1989).

4) Geman, S. and Geman, D.: Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol.PAMI-6, No.6, pp.721–741 (1984).
5) Yuille, A., Geiger, D. and Bulthoff, H.H.: Stereo Integration, Mean Field Theory and Psychophysics, *Network*, Vol.2, pp.423–442 (1991).
6) Terzopoulos, D., Witkin, A. and Kass, M.: Stereo Matching as Constrained Optimization Using Scale Continuation Methods, Optical and Digital Pattern Recognition, *SPIE*, Vol.754, pp.92–99 (1987).
7) Reimann, D. and Haken, H.: Stereo Vision by Self-Organization, *Biol. Cybern.*, Vol.71, pp.17–26 (1994).
8) Dhond, U.R. and Aggarwal, J.K.: Structure from Stereo: A Review, *IEEE Trans. Systems Man Cybern.*, Vol.19, No.6, pp.1489–1510 (1989).
9) Barnard, S.T. and Fischler, M.A.: Computational Stereo, *Comput.Surveys*, Vol.14, pp.553–572 (1982).
10) Haken, H.: *Synergetic Computers and Cognition*, Springer, Berlin, Heidelberg, New York (1991).

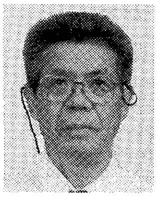
(Received August 5, 1996)
(Accepted October 1, 1997)



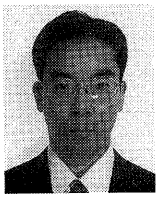
Yasunari Yoshitomi was born in 1956. He received his B.E., M.E. and Dr.Eng. degrees in Applied Mathematics and Physics from Kyoto University in 1980, 1982 and 1991, respectively. He had worked in Nippon Steel Corporation since 1982 and had been engaged in research on image analysis application and development of soft magnetic materials. Since 1995 he has been in Miyazaki University as an associate professor at the Department of Computer Science and Systems Engineering. His current research interests are image recognition, neural network, and stochastic programming problem. He is a member of IPSJ, IEICE, JSIAM, JNNS, ORSJ and PSJ.
e-mail: yoshi@cs.miyazaki-u.ac.jp



Tetu Kanda was born in 1974. He received his B.E. degree in Computer Science and Systems Engineering from Miyazaki University in 1996. He has been with Nintendo since 1996. He is engaged in producing TV game instrument and in its production control. e-mail: kanda@hamster.nintendo.co.jp



Teturou Kitazou was born in 1937. He received Dr.Sci. degree from Dept. of Physics, Osaka University in 1966. He was an assistant professor from 1966 to 1972 and an associate professor from 1972 to 1991 at Dept. of Physics, Kobe University. He moved to Faculty of Engineering, Dept. of Computer Science and Systems Enginering, Miyazaki University as a professor. His current research interests are image recognition by neural nets equation, speech recognition and quantum cosmology at the early stage of universe. He is a member of IPSJ, RSJ and PSJ.
e-mail: kitazoe@cs.miyazaki-u.ac.jp



Tomohiro Shii was born in 1974. He received his B.E. degree in Computer Science and Systems Enginering from Miyazaki University in 1997. He is a student of master course of Graduate school of Engineering in Miyazaki University. His current research interests are image recognition by neural network.
e-mail: siitomo@cs.miyazaki-u.ac.jp