

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)



---

# Stereo Algorithm with Anisotropic Reaction-Diffusion Systems

---

Atsushi Nomura, Koichi Okada, Hidetoshi Miike, Yoshiki Mizukami,  
Makoto Ichikawa and Tatsunari Sakurai

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/46025>

---

## 1. Introduction

The computer vision research aims at a better understanding of the human visual system and building artificial visual systems. Vision researchers in psychology and physiology have explored biological visual systems including the human vision, and obtained much knowledge on nature and architecture of their visual information processing. For example, some of previous results in experimental psychology suggested integration of several visual cues [1–3], and others of them showed evidence of anisotropy in the stereo depth perception [4, 5]. Mathematical models and computer algorithms developed according to previous experimental results help us to understand the human visual system and to build artificial visual systems.

The human visual system has two eyes aligned on a horizontal line. When the system captures an object located in a three-dimensional space, it perceives depth of the object. The visual system projects the object onto both of left and right retinæ and the projected intensity distributions are referred to as retinal images. Since the eyes see the object from slightly different three-dimensional positions, the object is projected at slightly different positions of both retinal images. The positional difference which is referred to as stereo disparity gives the depth of the object according to the concept of triangulation [6]. Detection of the stereo disparity requires a task of finding a reliable one-to-one correspondence between the left and right images. It is difficult to deal with this task, because there are ambiguities such as repeated texture and uniform color. The human visual system seems to have some robust mechanism for finding stereo disparity.

Motivated by the human visual system, Marr and Poggio presented a computer algorithm of stereo disparity detection [7–9]. Their algorithm named “cooperative algorithm” solves the stereo correspondence problem with a biologically inspired grid system, in which they placed cells at grid points and connected neighboring cells. In order to solve the stereo

correspondence problem and to obtain a dense stereo disparity map, they proposed imposing two constraints: uniqueness and continuity on the disparity map. The uniqueness constraint states that a point on the stereo disparity map has a unique disparity level except for transparent surfaces and object boundaries having multiple disparity levels. The continuity constraint states that neighboring grid points share the same or similar disparity levels except object boundaries. Marr and Poggio designed the grid system so as to satisfy the two constraints, by connecting neighboring cells cooperatively for the continuity constraint, and by connecting multi-layered grid systems exclusively for the uniqueness constraint.

Psychological and biophysical research results have affected the computer vision research including the cooperative algorithm. According to the Gestalt psychology [10, 11], when the human visual system captures an image consisting of small figures such as short lines and small crosses, it perceives a group of neighboring elements sharing the same or similar visual properties [12]. The Gestalt psychologists originally found this phenomenon and clearly stated the laws of closure, similarity, proximity, symmetry, continuity and common fate. In addition, previous biophysical research results showed that biological cells respond to external stimuli and exhibit a nonlinear excitation-inhibition process. Therefore, we understand that these previous results have affected the cooperative algorithm by Marr and Poggio. We can find similarities between the continuity constraint and the perceptual grouping exhibiting the laws of similarity, proximity and continuity, and between behavior of artificial cells in the cooperative algorithm and the cell responses in the excitation-inhibition process.

We previously presented several reaction-diffusion algorithms for segmentation and stereo disparity detection under the concept of reaction-diffusion systems [13, 14]. A reaction-diffusion system refers to the system of diffusively coupled elements exhibiting an excitation-inhibition process [15]. The reaction-diffusion system is mathematically described with a set of time-evolving partial differential equations consisting of diffusion terms and reaction ones. By numerically computing the reaction-diffusion system, we can simulate spatio-temporal phenomena such as pulse propagation observed in natural and biological systems, for example, in biological information transmission processes. We can expect that the pulse propagation phenomenon in a reaction-diffusion system serves as the continuity constraint in the stereo correspondence problem, and thus we proposed a stereo algorithm consisting of exclusively connected multi-layered reaction-diffusion systems [14]. In addition, inspired by the strong inhibitory diffusion causing Turing patterns [16], and suggested by a lateral inhibition mechanism in a biological visual system [17], we have imposed the strong inhibitory diffusion on the reaction-diffusion stereo algorithm.

This chapter presents recent advances in the reaction-diffusion stereo algorithm introducing anisotropy in diffusion processes. After quickly reviewing previous related work achieved in several different areas of psychology, stereo algorithms, reaction-diffusion and physiology in Section 2, we describe elementary stereo geometry, the cooperative algorithm and a reaction-diffusion system as preliminaries in Section 3. Then, we proceed to the original reaction-diffusion stereo algorithm with isotropic diffusion processes and its recent advances introducing anisotropic diffusion processes in Section 4. The section also presents the cooperative algorithm revised with a reaction-diffusion system. Then, we demonstrate comparison among the reaction-diffusion stereo algorithms including the original and

anisotropic ones for several stereo image pairs provided on the Middlebury stereo vision page [18, 19]. We discuss the experimental results and consider future research topics for the reaction-diffusion stereo algorithms. Finally, we conclude this chapter by summarizing the reaction-diffusion stereo algorithms, the experimental results and the future research topics.

## 2. Background

### 2.1. Human stereo depth perception

When the human visual system captures a three-dimensional scene with two eyes, it perceives depth structure of the scene. Even if the system is exposed to a pair of random-dot stereo images having only randomly dotted pattern (random-dot stereogram), it can perceive depth structure of objects embedded into the stereo images. Julesz generated random-dot stereograms by utilizing computers [20, 21]. Thimbleby et al. later proposed a computer algorithm generating a single image named “autostereogram”, in which they embedded a pair of stereo images; the autostereograms successfully caused stable depth perception for the human visual system [22]. These previous findings suggested that structural image pattern or higher knowledge on objects is unnecessary for the human stereo depth perception. In addition, the findings suggested that the human visual system has a function or a module of detecting disparity in its early vision. With the technique artificially generating random-dot stereograms many researchers have explored the nature of the human visual system. For example, it was shown that there exists the analogy between stereo depth perception and brightness perception [23, 24]. It was strongly suggested that a common mechanism underlies both the stereo depth perception and the brightness one.

In the depth perception, the human visual system integrates several visual cues such as stereo disparity, motion parallax and monocular configuration, each of which brings depth information. Landy et al. presented a weak fusion model that linearly integrates the visual cues with weighted averaging [1]. In contrast to the weak fusion model, Bradshaw and Rogers presented a strong fusion model that integrates the cues of stereo disparity and motion parallax with a nonlinear manner [2]. The strong fusion model states that each of the two modules processing stereo disparity and motion parallax takes an output of the other module as feedback, and thus the two modules are depending each other. Ichikawa et al. examined the depth perception and presented an integration model of the three visual cues: stereo disparity, motion parallax and monocular configuration [3]. Their model states that the cues of disparity and motion parallax are integrated with the strong fusion model at particular spatial frequency-tuned channels, and then outputs obtained at all the channels are furthermore integrated linearly with the cue of monocular configuration with the weak fusion model.

There is anisotropy in the human stereo depth perception; the human visual system perceives differently a horizontally slanted surface and a vertically slanted one. Rogers and Graham measured depth effect for an object having a one-dimensionally curved surface with the Cornsweet type depth profile [4]. When the surface of the object slants horizontally, the human visual system perceives a part of the surface to be nearer than the true depth. However, when the surface slants vertically, the system perceives depth of the surface correctly. From these experimental results, they presented a hypothesis. There exist two different processes for perceiving vertical and horizontal slant surfaces; this brings anisotropy in the human

stereo depth perception. Ichikawa more carefully examined the anisotropy with respect to latency and adaptation of the stereo depth perception for three different depth profiles and for a wide range of orientation [5]. His results also showed the similar anisotropy and presented evidence to support the hypothesis presented by Rogers and Graham.

These previous experimental results inspired us to develop the reaction-diffusion stereo algorithm. The former evidence showing the integration of several visual cues does not straightforwardly indicate the integration of image edge information into stereo disparity detection. However, it motivated us to study the integration of the image edge information, which is obtained with another algorithm having a mechanism similar to reaction-diffusion [25]. The latter evidence showing anisotropy encouraged us to divide a diffusion process into two diffusion processes, that is, horizontally and vertically oriented ones, with different diffusion coefficients.

## 2.2. Previous stereo algorithms

Dev proposed a feature segmentation model and its application to the stereo disparity detection [26]. The segmentation model employed a multi-layered network of which each grid point is described with a neural process having excited and resting states. When a point on the network enters an excited state, it becomes a member of the group of the associated feature. For achieving the segmentation, she imposed two interactions on the multi-layered network; in one of the interactions a point on a network layer should have excitatory links to neighboring points, and in the other one of the interactions the point in an excited state should inhibit excitation of neighboring points on other network layers. She applied the segmentation model to stereo disparity detection, in which each layer of the network is associated with each layer of a possible disparity level. The former interaction is similar to the continuity constraint, and the latter one is similar to the uniqueness constraint imposed on the cooperative algorithm [7–9]. Although the model by Dev [26] and the cooperative algorithm impose the similar interactions or constraints, the algorithm by Dev has an additional inhibitor network, which controls the inhibitory interaction; this is the main difference between the algorithm by Dev and the cooperative algorithm by Marr and Poggio.

Following the earlier work of the cooperative algorithm, many stereo algorithms have been proposed [27]. There are three main categories of the cooperative algorithm, the matching algorithm [28] and the regularization algorithm [29, 30]. The matching algorithm deals with the stereo correspondence problem by utilizing a similarity measure such as a cross-correlation coefficient computed between left and right images; in the case of the block matching algorithm, a rectangular area is utilized for the similarity measure. If images are rectified, an object in the left image should exist at the same vertical position in the right image; thus, a search area for the matching algorithm is usually restricted on a horizontal line (the epipolar constraint). Since a larger value of the cross-correlation coefficient indicates a higher probability of a correspondence between left and right images, the disparity with the maximum coefficient among possible disparities is employed as the detected disparity. The regularization algorithm formulates a functional consisting of a data term and a smoothness term. The data term denotes difference between left and right images at a disparity level, and the smoothness term denotes the continuity constraint. By minimizing the functional, the regularization algorithm provides a disparity map. Thus, the regularization algorithm avoids



explicitly dealing with the stereo correspondence problem by replacing the problem with a kind of the optimization problem.

In order to obtain a reliable stereo disparity map, we need to solve several problems such as the aperture problem, the occlusion problem and the problem arising from transparent surfaces. The following describes what the three problems refer to, and how classical stereo algorithms have approached the problems.

The matching algorithm needs to estimate an optimal window size for measuring the similarity. Real images usually contain untextured or featureless areas, which do not provide information to find a stereo correspondence. If we extend a window size so as to cover neighbor textured or feature-rich areas, we may obtain the drawback of a highly blurred disparity map, in which we can not expect detailed depth structure around object boundaries, resulting in lack of information necessary for later stages of the visual system. On the other hands, a smaller window size results in an unreliable disparity map. The aperture problem refers to the trade-off problem in estimating the window size. Kanade and Okutomi proposed a block matching stereo algorithm with an adaptive correlation window [28]; the algorithm controls the size and the shape of the window area for reducing uncertainty and simultaneously for keeping detailed depth structure.

Let us consider a situation in which there are two objects in a three-dimensional scene and one of the two objects partly occludes the other one on captured stereo images. Since two eyes capture the scene from slightly different positions, a part of the occluded object appears in one of the images and remains to be occluded in the other one. There is no corresponding points for the part of the occluded object in the other image, and thus stereo algorithms tend to find false correspondences for the part. This is the occlusion problem. A bi-directional matching technique provides a cue for detecting occluded areas [31–33], as follows. In the case where an interest point  $p_l$  in the left image is not occluded in the right image, the corresponding point in the right image,  $p_r$ , can be detected by searching for a point with the maximum similarity, and then a point  $p'_l$  in the left image, which corresponds to the detected point  $p_r$ , will be searched. Now the coordinate of  $p'_l$  is expected to be same with that of  $p_l$ . On the other hand, in the case where an interest point  $q_l$  in the left image does not have its corresponding point in the right image due to the occlusion, the point  $q_r$  with maximum similarity in the right image have the corresponding point  $q'_l$  on the left image, whose coordinate is different from that of  $q_l$ . It means that by detecting a two-step corresponding point and comparing it with the interest point, it is possible to judge if the interest point is occluded in the other image or not. As another approach, Zitnick and Kanade proposed a modern cooperative algorithm that can detect the occlusion areas [34]. As similar to the classical cooperative algorithm, their algorithm iteratively updates states of network. In contrast to the classical algorithm, the modern algorithm multiplies similarity distributions and states of the network at every iteration. This simple modification for the classical cooperative algorithm effectively detected occlusion areas.

Most stereo algorithms assume that objects are opaque and a point on a disparity map has only one disparity level; that is, they impose the uniqueness constraint on a disparity map. However, when we see an object through a window of glass material or wire fences, we

perceive two different surfaces [35]. The human visual system can perceive multiple disparity levels from a single pair of stereo images. According to the result of Tsirlin et al. [35], the human visual system can perceive up to six overlaid surfaces; Akerstrom and Todd examined the performance of the human visual system for random-dot stereograms including transparent surfaces [36]. Later, Shizawa presented a mathematical model describing the transparency upon the principle of superposition [37]. Szeliski and Golland proposed a stereo algorithm for transparent objects [38]. The occlusion problem can be also considered as the stereo transparency problem. Since an object occludes another object at an occluding boundary, there exist two disparity levels at the boundary. Assumption of two disparity levels brought another approach to solve the occlusion problem [39].

Since there have been proposed many stereo algorithms, Scharstein and Szeliski built a website named “Middlebury Stereo Vision Page” [18, 19] for quantitative evaluations of the algorithms. The website provides several stereo images, an evaluation system for stereo disparity maps and tables ranking stereo algorithms submitted to the website. There are two state-of-the-art algorithms in addition to the above mentioned three algorithms. One of the two state-of-the-art algorithms is the belief propagation algorithm and the other one is the graph-cuts algorithm. The belief propagation algorithm, which was originally proposed by Sun et al. [40], has grid points on a disparity map, and propagates the belief into their neighboring points. The belief denotes a kind of probability showing existence or non-existence of its associated disparity level at a grid point. By updating the state of a grid point with messages of belief received from its neighboring points, the algorithm builds a disparity map iteratively. The graph cuts algorithm was originally proposed by Kolmogorov and Zabih [41]. In the algorithm, we consider a graph network so as to express a functional shown in the regularization algorithm. By minimizing the number of points cutting the graph network, the algorithm provides a disparity map.

### 2.3. Reaction-diffusion systems related to image processing and vision research

Reaction-diffusion systems are common in nature, in particular, in chemical and biological systems [15]. Let us focus on a photo-sensitive chemical reaction-diffusion system. Busse and Hess firstly reported that the two-dimensional chemical reaction system senses illumination and generates a circular pattern of a chemical concentration wave at an illuminated point [42]. Kuhnert et al. reported that the photo-sensitive system has the functions of image memory, edge enhancement and segment detection for image pattern projected onto a surface of the two-dimensionally extended chemical solution [43, 44]. They also mentioned that the reaction-diffusion system is applicable to image processing and becomes a candidate of a new computer architecture with parallel processing. After their findings, many researchers examined experiments and proposed mathematical models on the photo-sensitive system. In particular, Sakurai et al. experimentally demonstrated a traveling path of a chemical reaction wave guided by a feedback control system having an illumination light [45]. These preceding research results suggested that reaction-diffusion systems provide new research topics in image processing and vision research.

An example of a biological reaction-diffusion system exists in an active pulse transmission process along a nerve axon. A mathematical model of the system is formulated with a set

of two reaction-diffusion equations consisting of diffusion terms and the FitzHugh-Nagumo type reaction terms [46, 47]. If we stimulate a point on the reaction-diffusion system, we can observe pulses traveling from the point on the system. Stereo algorithms presented in this chapter utilize the FitzHugh-Nagumo type reaction-diffusion system; the nature of traveling pulses helps to realize the continuity constraint. Later, Section 3.3 describes how the reaction-diffusion system works for the stereo correspondence problem.

With a reaction-diffusion system Turing proposed a mechanism for explaining a stationary pattern formation process observed in a biological system [16]. He considered a set of two reaction-diffusion equations having two variables: activator and inhibitor. In general, a diffusion process brings a spatial uniform distribution. For example, when a chemical species distributes non-uniformly in a space, it diffuses according to a gradient of the distribution and finally distributes uniformly. However, by considering two diffusion processes on activator and inhibitor and by assuming that the inhibitor diffuses more rapidly than the activator does, Turing found that those diffusion processes bring a non-uniform stationary pattern. According to the mechanism proposed by Turing [16], Gierer and Meinhardt [48] proposed more biologically plausible models explaining how biological systems self-organize spatial patterns. More recently, several researchers reported evidence showing that the Turing mechanism causes pattern formation observed on a fish skin [49], and other researchers identified two proteins as an activator and its inhibitor for a hair follicle spacing pattern of a mouse [50]. As the results of these researches, biologists have accepted the Turing mechanism as a possible mechanism explaining biological pattern formation.

Mach found an edge enhancement phenomenon for a step-wise illumination change in the human brightness perception; it is now known as the Mach-bands pattern [51]. Hartline and Ratliff found that there are excitatory and inhibitory interactions among outputs of individual eye units in compound eyes of *Limulus*, and proposed a linear model of equations describing the interactions [52]. Later, Barlow and Quarles proposed a nonlinear model for explaining the Mach-bands pattern observed in the visual system of *Limulus* [17]; they derived the nonlinear model by modifying the original model proposed by Hartline and Ratliff [52]. By comparing laboratory experiments with numerical results of the two models of equations, they indicated the importance of the long-range inhibition and the nonlinearity in the modified model for explaining the Mach-bands pattern. Gierer and Meinhardt pointed out that the long-range inhibition is analogous to the rapid inhibitory diffusion of the Turing condition imposed in modeling biological pattern formation processes [48]. These results suggested that the long range inhibition or the rapid inhibitory diffusion may play an important role in biological pattern formation and vision.

As pointed out by Gierer and Meinhardt [48], we have been interested in some common mechanism organizing visual functions, or underlying biological visual systems and pattern formation processes. We believe that the common mechanism is reaction-diffusion with the rapid inhibitory diffusion proposed by Turing [16], or the long-range inhibition found in eyes of *Limulus*. In order to show this, we have built the visual functions of edge detection [53–55], segmentation (grouping) [13] and stereo disparity detection [14], explored how the mechanism works, and confirmed how much the mechanism is effective in the visual functions.



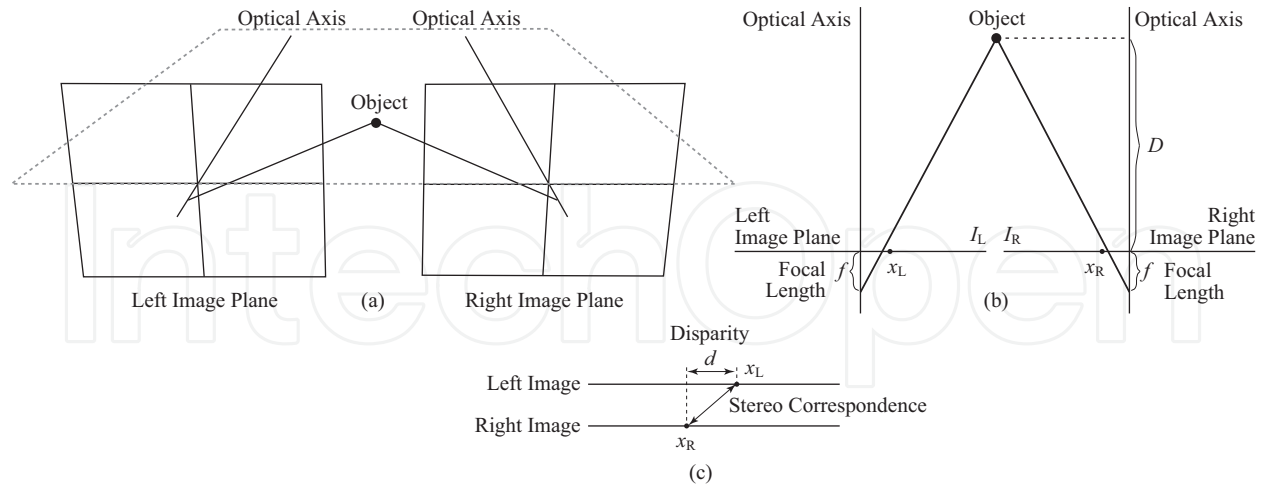
### 3. Preliminaries

#### 3.1. Stereo geometry

A stereo vision system consists of two cameras, which independently project an object located in a three-dimensional space onto the image planes of the cameras. Let us consider a simple stereo vision system, in which optical axes of the two cameras are parallel and a horizontal line passing through the origin of the left image plane also passes through the origin of the right image plane, as shown in Fig. 1. In this situation, an epipolar line refers to each of horizontal lines shared by the two image planes. If we utilize a pin-hole camera model, we can simply obtain depth of the object from stereo disparity  $d$  (pixel) [6]. Let  $f$  be a focal length of the cameras and  $\ell$  be distance between the two optical axes. The stereo vision system projects the object at the position  $x_L$  on the left image  $I_L$  and at the position  $x_R$  on the right image  $I_R$ . Then, the depth  $D$  becomes  $D = f(\ell - d)/d$  with  $d = x_L - x_R$ . Thus, if we can solve the stereo correspondence problem at particular points from the stereo images, we can obtain a full disparity map  $M(x, y)$  and reconstruct its depth structure.

#### 3.2. Cooperative algorithm for stereo disparity detection

A simple way of solving the stereo correspondence problem is to utilize a similarity measure such as a cross-correlation coefficient computed for local areas between stereo images. Let  $\mathcal{B}$  be a local area for computation of the similarity measure; for example, the area is defined as a  $3 \times 3$  pixels square area consisting of a target discrete position  $(i, j)$  and its relative positions  $\mathcal{B}_{3 \times 3} = \{(i', j') | -1 \leq i' \leq 1, -1 \leq j' \leq 1\}$  in a discretized coordinate system, or a cross-like local area consisting of the target position and its nearest neighboring four relative positions  $\mathcal{B}_5 = \{(0, 0), (-1, 0), (1, 0), (0, -1), (0, 1)\}$ . If we utilize a cross-correlation



**Figure 1.** Stereo vision system. Figure (a) shows the stereo cameras having left and right image planes; their optical axes are parallel and perpendicular to the image planes. An object in a three-dimensional space is projected onto the image planes. Figure (b) shows a top view of the vision system. The cameras have the same focal length  $f$ . The object is located at the distance  $D$  from the image planes, and is projected to the position  $x_L$  on the left image plane  $I_L$  and at the position  $x_R$  on the right one  $I_R$ . Figure (c) shows that the stereo disparity  $d$  (pixel) refers to the difference between the two corresponding positions  $x_L$  and  $x_R$ .

coefficient as the similarity measure and the local area  $\mathcal{B}$  as its local correlation area, we can compute a similarity measure  $C_{d,i,j}$  between  $I_{L,i,j}$  and  $I_{R,i-d,j}$  as follows:

$$C_{d,i,j} = \frac{1}{s_{L,i,j} \times s_{R,i-d,j}} \sum_{(i',j') \in \mathcal{B}} \left[ I_{L,i+i',j+j'} - \overline{I_{L,i,j}} \right] \times \left[ I_{R,i+i'-d,j+j'} - \overline{I_{R,i-d,j}} \right], \quad (1)$$

in which  $d$  (pixel) is a discrete disparity level;  $s_{L,i,j}$  is the standard deviation of  $I_{L,i+i',j+j'}$  and  $s_{R,i-d,j}$  is that of  $I_{R,i+i'-d,j+j'}$  for  $(i',j') \in \mathcal{B}$ ;  $\overline{I_{L,i,j}}$  is the average of  $I_{L,i+i',j+j'}$  and  $\overline{I_{R,i-d,j}}$  is that of  $I_{R,i+i'-d,j+j'}$  for  $(i',j') \in \mathcal{B}$ . The similarity measure of Eq. (1) provides 1.0 for a pair of completely matched local areas of the left and right images. If the similarity measure provides a high value such as  $C_{d,i,j} \simeq 1.0$  for the disparity level  $d$  and provides low values for other disparity levels, the position  $(i-d, j)$  on the right image  $I_R$  becomes a candidate for the corresponding point of the position  $(i, j)$  on the reference left image  $I_L$ . Let  $\mathcal{D}$  be a set of possible discrete disparity levels;  $\mathcal{D} = \{d_0, d_1, \dots, d_{(N-1)}\}$ . Thus, detection of the maximum of  $C_{d,i,j}$  for  $d \in \mathcal{D}$  at a position  $(i, j)$  provides a stereo disparity map  $M_{i,j}$ , as follows:

$$M_{i,j} = \operatorname{argmax}_{d \in \mathcal{D}} C_{d,i,j}. \quad (2)$$

As stated in Section 2.2, there are several typical problems in the stereo correspondence problem. The aperture problem results from situations in which the local area defined by  $\mathcal{B}$  does not have enough intensity information such as distinguishable patterns from other areas and thus the area is matchable to each of the most areas on its epipolar line. In the occlusion problem, since a position on one of the stereo images is occluded on the other image, searching the maximum of the similarity measure may bring a false position as its corresponding point. For stereo images containing transparent objects, there are multiple maximum values at a position; it is difficult to detect multiple disparity levels without information of its multiplicity or under noisy situations.

Marr and Poggio proposed a classical cooperative algorithm by focusing on a random-dot stereogram [7–9]. They imposed two constraints: continuity and uniqueness on a stereo disparity map  $M_{i,j}$ . They considered a cell network in which cells enter an excited state or a resting state and a state of a cell propagates into its neighboring cells, and utilized a multi-layered cell network, in which each network layer is associated with a disparity level  $d \in \mathcal{D}$ . The cooperative algorithm iteratively computes states of cells on the network. Let  $S_{d,i,j}^k$  be a state of the cell located at a position  $(i, j)$  on a network layer associated with a disparity level  $d$ ; a large value of  $S_{d,i,j}^k \simeq 1.0$  denotes that the cell at  $(i, j)$  on the network layer is in an excited state, and a small value of  $S_{d,i,j}^k \simeq 0.0$  denotes that the cell is in a resting state. Then, the state  $S_{d,i,j}^{k+1}$  at the  $(k+1)$ -th iteration is updated, as follows:

$$S_{d,i,j}^{k+1} = \sigma \left( \sum_{(i',j',d') \in \Omega(i,j,d)} S_{d',i',j'}^k - \xi \sum_{(i',j',d') \in \Theta(i,j,d)} S_{d',i',j'}^k + C_{d,i,j}, T \right), \quad (3)$$

in which  $\sigma(S, T)$  is a step function which returns 1 for  $S > T$  and returns 0 for  $S \leq T$  with the threshold level  $T$ ;  $\xi$  is a constant for inhibition;  $\Omega$  denotes an excitatory domain for the

continuity constraint and  $\Theta$  denotes an inhibitory domain with  $\mathcal{D} \setminus \{d\}$  for the uniqueness constraint (for more detail, see Fig. 1 in Ref. [34]). Finally, as similar to Eq. (2), the cooperative algorithm provides a disparity map  $M_{i,j}^k$  with

$$M_{i,j}^k = \operatorname{argmax}_{d \in \mathcal{D}} S_{d,i,j}^k. \quad (4)$$

In Eq. (3) the term  $\sum_{\Omega} S_{d',i',j'}^k$  works for the continuity constraint and the term  $\sum_{\Theta} S_{d',i',j'}^k$  works for the uniqueness constraint, as follows. If most of neighboring cells in  $\Omega$  are in excited states with  $S_{d',i',j'}^k \simeq 1.0$ ,  $\sum_{\Omega} S_{d',i',j'}^k$  also becomes large. Then, according to Eq. (3), if  $\sum_{\Omega} S_{d',i',j'}^k$  becomes larger than the threshold level  $T$ , the next state of  $S_{d,i,j}^{k+1}$  also enters an excited state. Thus, all cells in the local area denoted by  $\mathcal{B}$  are in excited states, and the area becomes to be in the disparity level  $d$ . This is what the continuity constraint indicates. If there is a situation in which  $\sum_{\Theta} S_{d',i',j'}^k$  is large, that is, if cells on other network layers associated with other disparity levels are already in excited states, the threshold level  $T$  becomes relatively large. Thus, the cell on the network layer of  $d$  tends to remain in a resting state, even if its surrounding cells on the same network layer are in excited states. This is what the uniqueness constraint indicates.

### 3.3. Reaction-diffusion system and coupled nonlinear elements

A reaction-diffusion system is described with a set of reaction-diffusion equations consisting of diffusion terms and reaction ones. Let us consider a two dimensional space  $\mathcal{L}_x \times \mathcal{L}_y$ , in which a process governed by the reaction-diffusion system proceeds with time  $t$  and the process has two distributions of activator  $u(x, y, t)$  and inhibitor  $v(x, y, t)$  defined at a position  $(x, y) \in \mathcal{L}_x \times \mathcal{L}_y$  and  $t \geq 0$ . The reaction-diffusion system has a general form of two reaction-diffusion equations with their reaction terms  $f(u, v)$  and  $g(u, v)$ , as follows:

$$\partial_t u = D_u \nabla^2 u + f(u, v), \quad (5)$$

$$\partial_t v = D_v \nabla^2 v + g(u, v), \quad (6)$$

in which  $\partial_t = \partial/\partial t$  and  $\nabla$  is a two-dimensional gradient operator;  $D_u$  and  $D_v$  are diffusion coefficients. For example, the FitzHugh-Nagumo type reaction-diffusion equations have the following reaction terms [46, 47]:

$$f(u, v) = [u(u - a)(1 - u) - v]/\varepsilon, \quad (7)$$

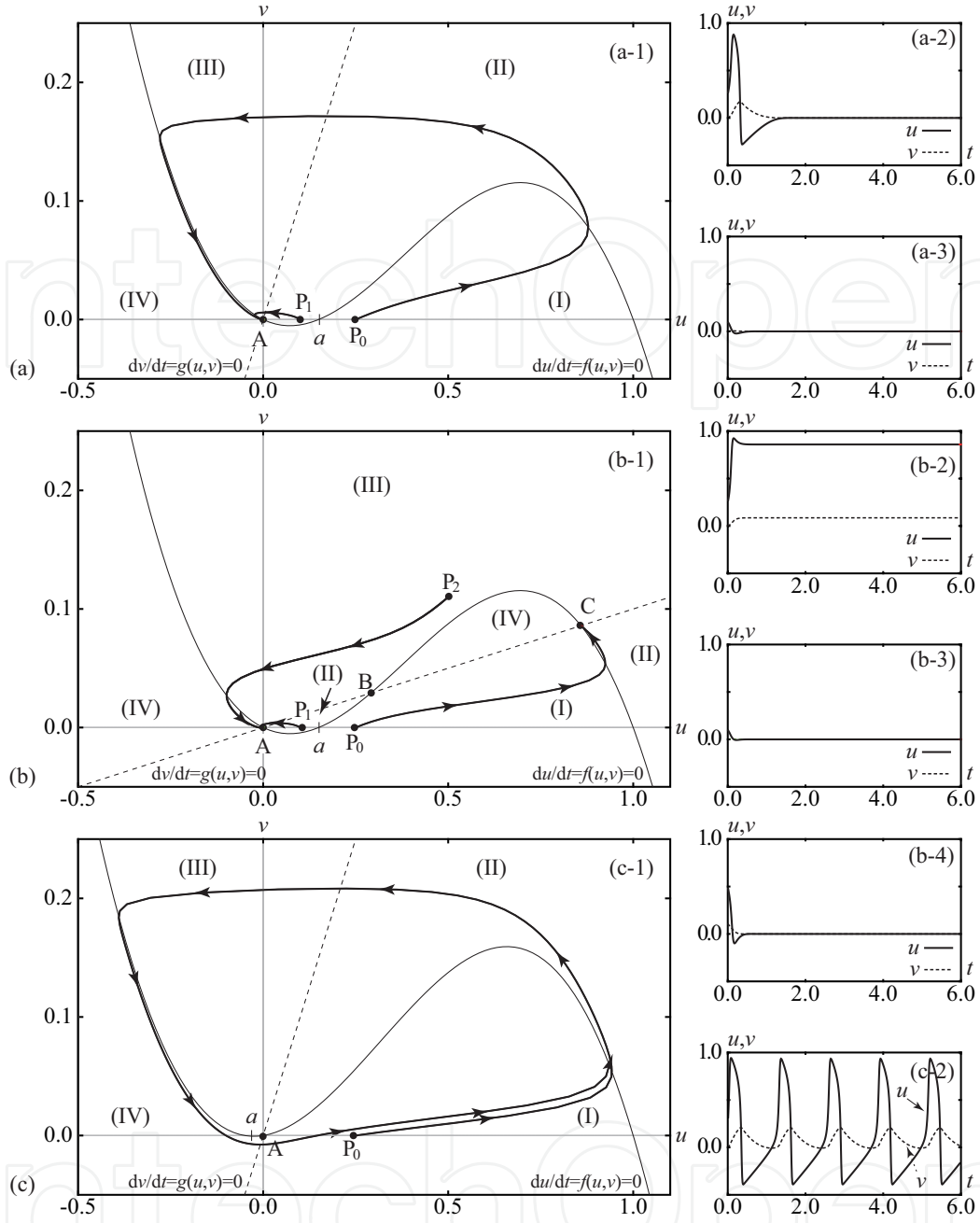
$$g(u, v) = u - bv, \quad (8)$$

in which  $a$  and  $b$  are constants and  $\varepsilon$  is a small constant ( $0 < \varepsilon \ll 1$ ).

In order to understand temporal behavior of the FitzHugh-Nagumo type reaction terms, let us consider a single element governed by the following ordinary differential equations:

$$\frac{du}{dt} = f(u, v) = [u(u - a)(1 - u) - v]/\varepsilon, \quad (9)$$

$$\frac{dv}{dt} = g(u, v) = u - bv. \quad (10)$$



**Figure 2.** Phase-portraits, solution trajectories and temporal changes of solutions in the FitzHugh-Nagumo type ordinary differential equations. Figure (a) shows those of a mono-stable element with the parameter settings of  $a = 0.15, b = 1.0$ ; Fig. (b) shows those of a bi-stable element with  $a = 0.15, b = 10$ ; Fig. (c) shows those of an oscillatory element with  $a = -0.05, b = 1.0$ ; the parameter  $\varepsilon$  was fixed at  $\varepsilon = 1.0 \times 10^{-2}$ . Figures (a-1), (b-1) and (c-1) show the phase-portraits and the solution trajectories; Figs. (a-2), (b-2) and (c-2) show the temporal changes of solutions starting from a point  $P_0$ ; Figs. (a-3) and (b-3) show those from a point  $P_1$ ; Fig. (b-4) shows that from a point  $P_2$ . The null-clines  $du/dt = f(u, v) = 0$  and  $dv/dt = g(u, v) = 0$  divide each of the phase-portraits into four areas (I), (II), (III) and (IV). A set of solutions  $(u, v)$  traces a trajectory, depending on the signs of  $du/dt$  and  $dv/dt$  in each area denoted by (I), (II), (III) and (IV). For example, the combination of  $du/dt > 0$  and  $dv/dt > 0$  in the area (I) increases the both solutions  $u$  and  $v$  as time proceeds.

Equations (9) and (10) describe temporal changes of  $u(t)$  and  $v(t)$ . Figure 2 shows the temporal changes as well as phase-portraits of Eqs. (9) and (10). Depending on the parameter settings of  $a$  and  $b$ , the element of Eqs. (9) and (10) exhibits three different types of behavior, such as a mono-stable element shown in Fig. 2(a), a bi-stable element shown in Fig. 2(b), and an oscillatory element shown in Fig. 2(c). The mono-stable element has one stable steady state, to which all solutions converge; the bi-stable element has two stable steady states and solutions converge to either of the two states. When a solution of an element is in areas having large values of  $u > 0.5$ , we denote that the element is in an excited state; when a solution is in areas having small values of  $|u| \simeq 0$ , we denote that the element is in a resting state. In contrast to the mono-stable and bi-stable elements having stable steady states, an oscillatory element does not have any stable steady state, and autonomously alternates between an excited state and a resting one as time proceeds.

In addition to the reaction-diffusion system of diffusively coupled elements such as the FitzHugh-Nagumo type, there is also an interesting system consisting of discretely coupled elements; the former system is a continuous system and the latter one is a discrete system. For example, the following equations describe a discrete system with reaction terms of  $f(\cdot, \cdot)$  and  $g(\cdot, \cdot)$  in a two-dimensional grid system  $(i, j)$  on  $\mathcal{L}_x \times \mathcal{L}_y$ , as follows:

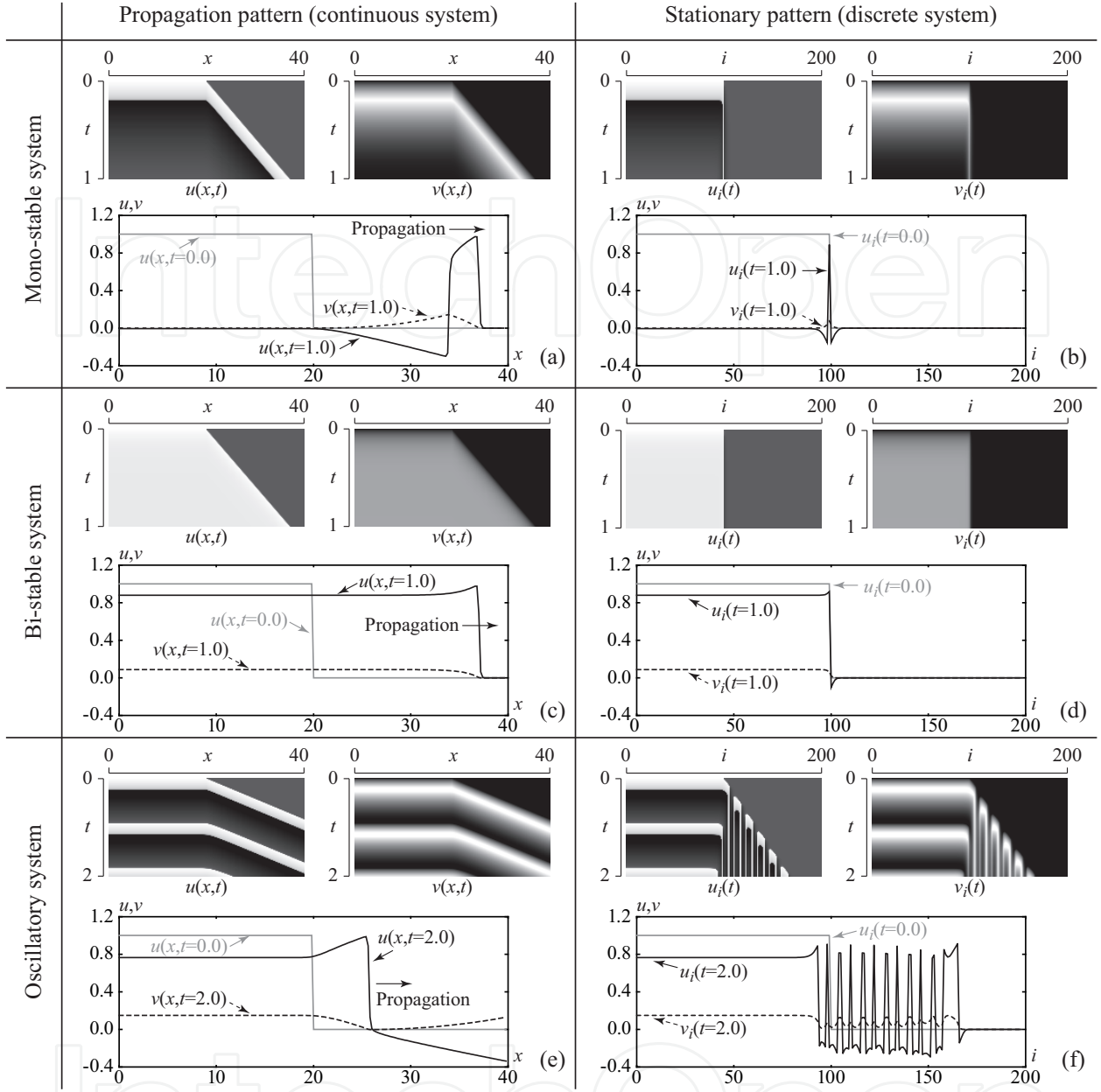
$$\frac{du_{i,j}(t)}{dt} = C_u \left[ \sum_{(i',j') \in \mathcal{B}_5 \setminus \{(0,0)\}} u_{i+i',j+j'} - 4u_{i,j} \right] + f(u_{i,j}, v_{i,j}), \quad (11)$$

$$\frac{dv_{i,j}(t)}{dt} = C_v \left[ \sum_{(i',j') \in \mathcal{B}_5 \setminus \{(0,0)\}} v_{i+i',j+j'} - 4v_{i,j} \right] + g(u_{i,j}, v_{i,j}), \quad (12)$$

in which  $C_u$  and  $C_v$  are positive coupling strength;  $\mathcal{B}_5$  defines a local area consisting of a target position and its nearest neighboring four positions, as introduced in Section 3.2. Equations (11) and (12) are similar to a spatially discretized version of the reaction-diffusion system of Eqs. (5) and (6); thus, both of the reaction-diffusion system and the system of discretely coupled elements have common mechanisms such as the nonlinear reaction denoted by  $f(\cdot, \cdot)$  and  $g(\cdot, \cdot)$  and spatial coupling of the nonlinear elements.

Let us confirm spatio-temporal patterns generated by the continuous system and the discrete one; both the systems consist of the FitzHugh-Nagumo type reaction terms of Eqs. (7) and (8). Figure 3 shows spatio-temporal representations of  $u$  and  $v$  and their one-dimensional snapshots obtained in one-dimensional continuous or discrete space. The continuous mono-stable system generates a traveling pulse as shown in Fig. 3(a), and the continuous bi-stable system generates a propagating wave as shown in Fig. 3(c); velocity of the pulse and the wave is almost constant, depending on the parameter settings of the system. The continuous oscillatory system generates multiple pulses that also travel in the space as shown in Fig. 3(e). In contrast to these, the discrete system with strong-inhibitory coupling ( $C_u \ll C_v$ ) generates stationary pulse(s) and a wave, as shown in Figs. 3(b), 3(d) and 3(f). For example, the discrete mono-stable system generates a pulse fixed at the edge position of a step-wise initial distribution  $u_i(t = 0)$ . This result indicates that the discrete system is applicable to edge detection for its initial condition.





**Figure 3.** Spatio-temporal representations of solutions and their one-dimensional snapshots obtained by numerical computations for a reaction-diffusion system (continuous system) and a system of discretely coupled elements (discrete system). See Eqs. (5) and (6) for the continuous system, and Eqs. (11) and (12) for the discrete system; both the systems have the FitzHugh-Nagumo type reaction terms of Eqs. (7) and (8). Figures (a) and (b) show the results of mono-stable systems with the parameter settings of  $a = 0.05, b = 1.0$ , Figs. (c) and (d) show the results of bi-stable systems with the parameter settings of  $a = 0.05, b = 10$ , and Figs. (e) and (f) show the results of oscillatory systems with the parameter settings of  $a = -0.05, b = 1.0$ ; the parameter setting of  $\varepsilon$  was fixed at  $\varepsilon = 1.0 \times 10^{-3}$  across (a)~(f). Figures (a), (c) and (e) show the results of propagation pattern obtained by the continuous system with the parameter settings of  $D_u = 1.0, D_v = 0.0$ , and Figs. (b), (d) and (f) show the results of stationary pattern obtained by the discrete system with the parameter settings of  $C_u = 4.0, C_v = 12$ . The spatio-temporal representations of  $u(x, t)$  and  $u_i(t)$  visualize the range of  $-0.3 \leq u \leq 1.0$  and those of  $v(x, t)$  and  $v_i(t)$  visualize the range of  $0.0 \leq v \leq 0.15$ . The numerical computations were done with a spatial finite difference  $\delta h = 0.2$  and a temporal one  $\delta t = 1.0 \times 10^{-5}$ .

In the neural network approach, Wang and his coworkers proposed a segmentation algorithm (eg [56]). The algorithm consists of FitzHugh-Nagumo type oscillatory elements with complex spatial coupling rules and a global inhibitor that controls the oscillatory elements. In the algorithm, an oscillation period is automatically divided into the number of segments to be detected. During one oscillation period, one segment emerges in a part of the period and another segment emerges in another part of the period. The idea of utilizing the global inhibitor is similar to the inhibitory array of the algorithm proposed by Dev [26]. However, in contrast to the inhibitory array in the algorithm of Dev, the algorithm by Wang and his coworkers has the global inhibitor represented by one variable. The approach by Wang and his coworkers is also similar to our approach of reaction-diffusion algorithms. Although both approaches utilize FitzHugh-Nagumo type elements, our approach imposes the strong inhibitory coupling on the algorithms without any global inhibitor nor any complex spatial coupling rules. We would like to emphasize this point as the main difference between the two approaches.

## 4. Reaction-diffusion stereo algorithm

### 4.1. Original reaction-diffusion stereo algorithm

This section presents a stereo algorithm solving the stereo correspondence problem, by utilizing bi-stable reaction-diffusion systems. Let us recall that the bi-stable system with the FitzHugh-Nagumo type reaction terms generates a traveling wave as shown in Fig. 3(c). Areas in excited states extend outward, as their interface facing an excited state and a resting state propagates into areas in resting states. We apply the nature of the traveling wave to the continuity constraint imposed on the stereo algorithm. According to the original cooperative algorithm mentioned in Section 3.2, we consider multiple reaction-diffusion systems and associate each of the systems with a possible disparity level. Each system governs areas having the associated disparity level; if a point of the system enters an excited state and all of the other systems are in resting states at that point, the algorithm judges the point to have the associated disparity level. We consider a state of each reaction-diffusion system as a kind of possibility and thus the traveling wave works as the continuity constraint. If the reaction-diffusion systems are independent, the traveling wave fills in everywhere of the space  $\mathcal{L}_x \times \mathcal{L}_y$ . In order to partition a disparity map into segments having correct disparity levels, we need to link the multiple reaction-diffusion systems exclusively for the uniqueness constraint. This can be done via the parameter  $a$ ; however  $a$  is a constant in the original FitzHugh-Nagumo type reaction term of Eq. (7), we consider  $a$  as a threshold level depending on states of other reaction-diffusion systems; the role of  $a$  is similar to that of the threshold level  $T$  of the cooperative algorithm described with Eq. (3).

The above consideration brings a set of reaction-diffusion equations associated with a disparity level  $d \in \mathcal{D}$ , as follows:

$$\partial_t u_d = D_u \nabla^2 u_d + [u_d(u_d - a_d)(1 - u_d) - v_d] / \varepsilon + \mu C_d(x, y), \quad (13)$$

$$\partial_t v_d = D_v \nabla^2 v_d + (u_d - b v_d), \quad (14)$$

$$a_d = \alpha + [1 + \tanh(d_a - \beta)] \times \max_{d' \in \Theta} u_{d'} / 2, \quad d_a = \left| d - \operatorname{argmax}_{d' \in \Theta} u_{d'} \right|, \quad (15)$$

**Algorithm 1** Original reaction-diffusion stereo algorithm.

---

```

1: for all  $d \in \mathcal{D}$  do
2:   for all  $(i, j) \in \mathcal{L}_x \times \mathcal{L}_y$  do
3:     Compute  $C_{d,i,j}$  from  $I_L$  and  $I_R$ . ▷ with Eq. (1).
4:     Set initial conditions of  $u_{d,i,j}^{k=0} = v_{d,i,j}^{k=0} = 0$ .
5:   end for
6: end for
7:  $k \leftarrow 0$ 
8: while  $k < L_t / \delta t$  do
9:   for all  $d \in \mathcal{D}$  do
10:    for all  $(i, j) \in \mathcal{L}_x \times \mathcal{L}_y$  do
11:      Compute  $d_a$  and  $a_n$  ▷ with Eq. (15)
12:      Compute  $u_{d,i,j}^{k+1}$  and  $v_{d,i,j}^{k+1}$  ▷ with Eqs. (17) and (18)
13:    end for
14:  end for
15:  for all  $(i, j) \in \mathcal{L}_x \times \mathcal{L}_y$  do
16:    Compute  $M_{i,j}^k$ . ▷ with Eq. (16)
17:  end for
18:   $k \leftarrow k + 1$ 
19: end while

```

---

in which  $C_d(x, y)$  denotes an external stimulus and  $\mu$  is its coefficient; the reaction-diffusion stereo algorithm utilizes a cross-correlation coefficient of Eq. (1) as  $C_d(x, y)$ ;  $\alpha$  and  $\beta$  are constants.

After sufficient duration of time  $L_t$ , a disparity map is obtained by

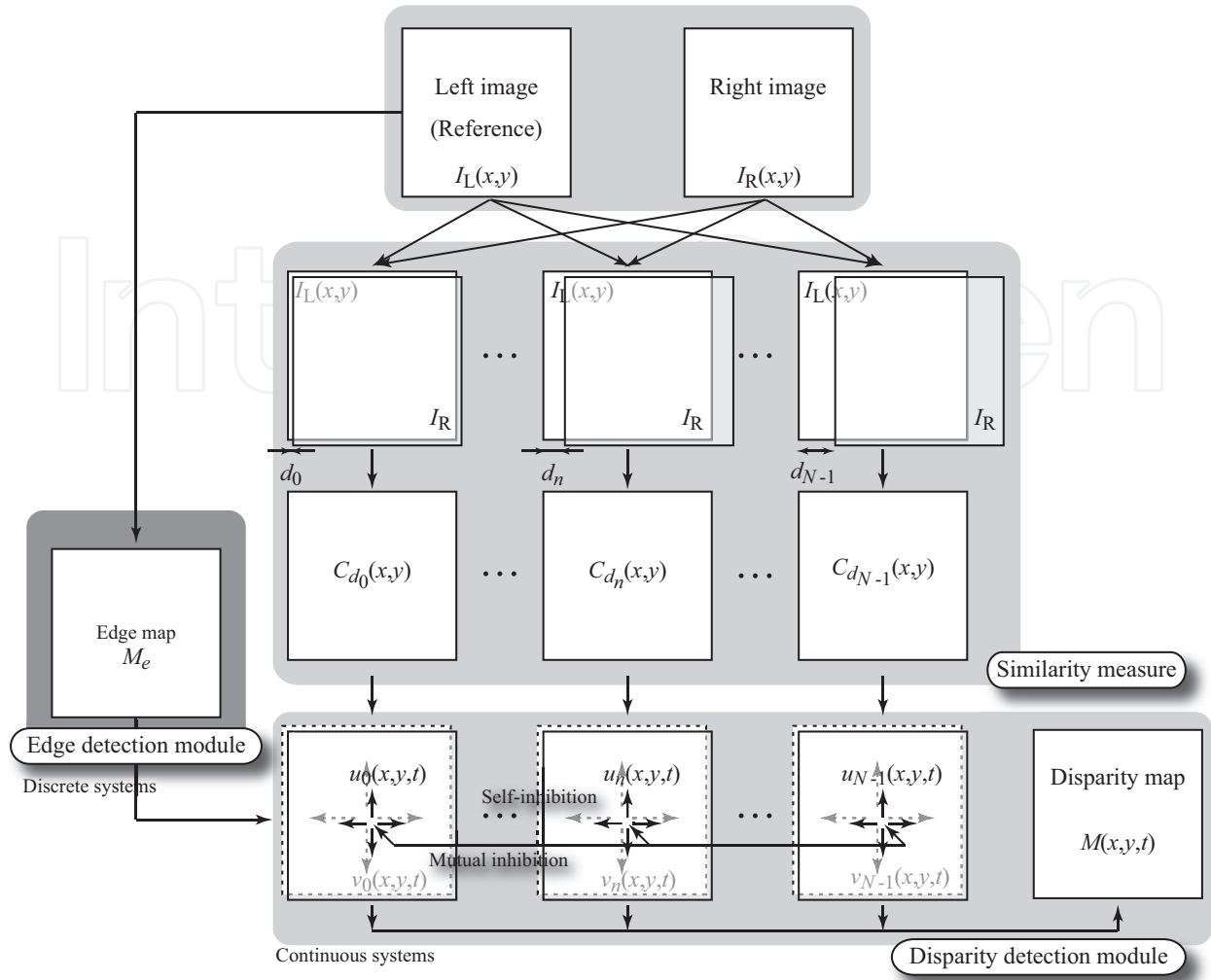
$$M(x, y, L_t) = \operatorname{argmax}_{d \in \mathcal{D}} u_d(x, y, L_t). \quad (16)$$

Numerical implementation of the reaction-diffusion stereo algorithm requires discretization of Eqs. (13) and (14). Spatially and temporally discretized coordinate systems of  $i = \lfloor x / \delta h \rfloor, j = \lfloor y / \delta h \rfloor$  and  $k = \lfloor t / \delta t \rfloor$ , in which  $\lfloor \cdot \rfloor$  is the floor function, derive the following set of equations approximately describing Eqs. (13) and (14), as follows:

$$u_{d,i,j}^{k+1} - C_u \left[ \sum_{(i',j') \in \mathcal{B}_5 \setminus \{(0,0)\}} u_{d,i+i',j+j'}^{k+1} - 4u_{d,i,j}^{k+1} \right] = u_{d,i,j}^k + \delta t f(u_{d,i,j}^k, v_{d,i,j}^k) + \delta t \mu C_{d,i,j}, \quad (17)$$

$$v_{d,i,j}^{k+1} - C_v \left[ \sum_{(i',j') \in \mathcal{B}_5 \setminus \{(0,0)\}} v_{d,i+i',j+j'}^{k+1} - 4v_{d,i,j}^{k+1} \right] = v_{d,i,j}^k + \delta t g(u_{d,i,j}^k, v_{d,i,j}^k), \quad (18)$$

in which  $C_u = \delta t D_u / \delta h^2$  and  $C_v = \delta t D_v / \delta h^2$ . For example, the Gauss-Seidel method provides solution for a set of linear equations such as each of Eqs. (17) and (18). Algorithm 1 describes a pseudo-code designed for the reaction-diffusion stereo algorithm originally proposed in [14]. Later, Section 5.1 shows a simple demonstration of how reaction-diffusion systems work for the continuity constraint and the uniqueness one.



**Figure 4.** Processing diagram of the reaction-diffusion stereo algorithm integrating image intensity edge information. Firstly, the algorithm computes similarity measures  $C_d$  between left and right images denoted by  $I_L$  and  $I_R$  for possible disparity levels  $d \in \mathcal{D}$ . Secondly, the similarity measures are provided for reaction-diffusion systems having  $u_d$  and  $v_d$  as their external stimuli. Each of the reaction-diffusion systems has the self-inhibition mechanism and has exclusive links to the other systems via the mutual inhibition mechanism. A reaction-diffusion stereo algorithm presented in Section 4.2 integrates an edge map obtained from the left image into inhibitory diffusion coefficients; another edge detection algorithm [25] provides the edge map  $M_e$ . Finally, the algorithm provides a stereo disparity map  $M(x, y, t)$  by gathering the results of the reaction-diffusion systems. The diagram without integrating the edge map becomes that of the original reaction-diffusion stereo algorithm [14].

#### 4.2. Integration of intensity edge information into the reaction-diffusion stereo algorithm

As mentioned above, the continuity constraint states that disparity distribution varies smoothly or neighboring points share the same disparity level on a disparity map. However, object boundaries causing depth discontinuity violate the continuity constraint. Thus, it is important to identify depth discontinuity areas from other areas satisfying the continuity constraint. If the depth discontinuity areas are preliminary known, the stereo correspondence problem becomes much easier. In general, the object boundaries or the depth discontinuity

areas are unknown in advance of solving the stereo correspondence problem. Although intensity edge areas do not always correspond to depth discontinuity areas, some of them coincide with object boundaries. A practical idea to overcome the problem caused by the depth discontinuity is to utilize other visual cues such as image intensity. Previous psychological results show that the human visual system integrates several other cues into the depth perception [1–3]. Sun et al. also showed that the integration of image edge information into their belief propagation algorithm improves its performance [40]. Martin et al. proposed a contour detection algorithm [57], that can be applied to the stereo correspondence problem.

In order to prevent the continuity constraint from working across object boundaries, we consider integrating another cue of image intensity edge information into the reaction-diffusion stereo algorithm. The diffusion terms  $D_u \nabla^2 u_d$  and  $D_v \nabla^2 v_d$  in the stereo algorithm mainly control the continuity constraint. There are several choices of integrating the edge information into the algorithm; one of them is to weaken the excitatory diffusion coefficient  $D_u$  in image intensity edge areas, and another one is to strengthen the inhibitory diffusion coefficient  $D_v$  in the areas. This is because a larger value of  $D_u$  drives wave propagation and increases its velocity, and a larger value of  $D_v$  makes strong diffusion of an inhibitor distribution, and thus inhibits the propagation of the wave. Between the two choices, we take the latter one of modulating the inhibitory diffusion coefficient  $D_v$  with intensity edge information. Thus, we obtain another algorithm by replacing Eq. (14) of the original reaction-diffusion algorithm with the following equation

$$\partial_t v_d = \nabla \cdot [D_v(x, y) \nabla v_d] + (u_d - b v_d), \quad (19)$$

in which the spatial distribution of  $D_v(x, y)$  is obtained from combination of the intensity edge map denoted by  $\mathcal{M}_e$  and an initial guess of a stereo disparity map  $M(x, y, 0)$ . If an intensity edge exists at the point  $(x, y)$ , and simultaneously if the initial guess is almost uniform in a neighboring area around the point  $(x, y)$ ,  $D_{v_{\max}}$  is given for  $D_v(x, y)$ , otherwise  $D_{v_{\min}}$  is given for  $D_v(x, y)$ , as follows:

$$D_v(x, y) = \begin{cases} D_{v_{\max}} & \text{if } (x, y) \in \mathcal{M}_e \text{ and } |\nabla M(x, y, 0)| < 2 \\ D_{v_{\min}} & \text{otherwise} \end{cases}, \quad (20)$$

in which  $D_{v_{\max}} > D_{v_{\min}}$ . In addition,  $D_v(x, y)$  is diffused during a short period of time  $L_{d_i}$  for its smoothness. The discrete system of spatially coupled FitzHugh-Nagumo elements provides the intensity edge map  $\mathcal{M}_e$  [25] (recall that the discrete system can detect edges as shown in Section 3.3). Figure 4 shows a processing diagram of the stereo algorithm.

### 4.3. Anisotropic inhibitory diffusion processes depending on pre-specified orientation

Inspired by the anisotropy observed in the human stereo depth perception, we try to introduce anisotropy into an inhibitory diffusion coefficient of the reaction-diffusion stereo algorithm. We pre-specify horizontal or vertical orientation ( $\phi = 0, \pi/2$  radian) as the anisotropy. According to the formulation proposed by Shoji et al. for modeling strip patterns self-organized on fish skins [58], the coefficient is spatially modulated with difference between  $\phi$  and  $\theta = \arctan(\partial_y v_d / \partial_x v_d)$ , in which  $\partial_x = \partial/\partial x$  and  $\partial_y = \partial/\partial y$ ;  $\theta$  denotes a gradient



direction of the inhibitor distribution. If  $\theta$  coincides with  $\phi$ , the diffusion is strengthened at the direction; as the difference between  $\theta$  and  $\phi$  becomes larger, the diffusion is weakened. Thus, we obtain the anisotropic reaction-diffusion stereo algorithm by replacing Eq. (14) of the original algorithm with the following equation:

$$\partial_t v_d = D_v \nabla \cdot [A(\theta) \nabla v_d] + (u_d - b v_d), \quad (21)$$

$$A(\theta) = 1 / \sqrt{1 - \rho \cos(2\theta - 2\phi)}, \quad (22)$$

in which  $\rho$  indicates strength of the anisotropy ( $0 \leq \rho < 1$ ); if  $\rho = 0$ , Eq. (22) becomes  $A(\theta) = 1.0$ , and thus Eq. (21) returns to the original isotropic Eq. (14). Expansion of the first term in the right side of Eq. (21) derives

$$\nabla \cdot [A(\theta) \nabla v_d] = \underbrace{\partial_x A(\theta) \cdot \partial_x v_d}_{\text{Horizontal}} + \underbrace{\partial_y A(\theta) \cdot \partial_y v_d}_{\text{Vertical}} + \underbrace{A(\theta) \nabla^2 v_d}_{\text{Isotropic}}, \quad (23)$$

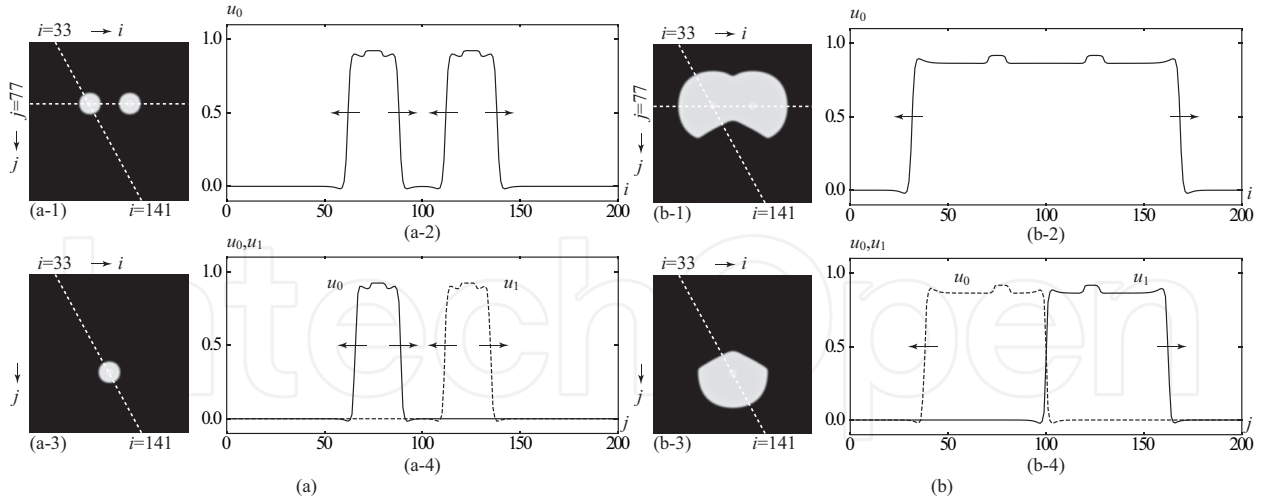
which consists of the horizontal and vertical advection terms, of which the velocity is  $-\nabla A(\theta)$ , in addition to the isotropic diffusion term. Thus, in accordance with the psychological hypothesis proposed by Rogers and Graham [4] and its evidence shown by Ichikawa [5], the anisotropic reaction-diffusion stereo algorithm has two processes of horizontal and vertical advection terms that propagate information of existence of its associated disparity level.

#### 4.4. Cooperative algorithm revised with a reaction-diffusion equation

As described in Section 4.1, the original reaction-diffusion stereo algorithm satisfies the continuity constraint and the uniqueness one presented in the original cooperative algorithm. Let us focus on the FitzHugh-Nagumo type ordinary differential equations of Eqs. (7) and (8). Under the condition of a fixed  $v = 0$ , the ordinary differential equation  $du/dt = f(u, 0) = [u(u - a)(1 - u)]/\varepsilon$  behaves as a switching element having a threshold level  $a$ . This is because the solution  $u(t)$  increases as time proceeds and finally converges to  $u = 1$  for the initial condition  $u(0) > a$ , and decreases and finally converges to  $u = 0$  for  $u(0) < a$ . Thus, the reaction term  $f(u, 0) = [u(u - a)(1 - v)]/\varepsilon$  can qualitatively replace the step function  $\sigma(S, T)$  of Eq. (3), in which  $S$  corresponds to  $u$  and  $T$  does to  $a$ . These correspondences between the step function  $\sigma(S, T)$  and the behavior of the FitzHugh-Nagumo equations with the fixed  $v = 0$  bring a revised version of the cooperative algorithm proposed by Marr and Poggio, as follows:

$$\partial_t u_d = D_u \nabla^2 u_d + u_d(u_d - a_d)(1 - u_d)/\varepsilon + \mu C_d(x, y), \quad (24)$$

in which  $a_d$  is obtained by Eq. (15), other parameters  $D_u, \varepsilon, \mu$  are constants, and  $C_d(x, y)$  denotes a similarity measure. The diffusive coupling  $D_u \nabla^2$  in Eq. (24) works for the continuity constraint described with the term  $\sum_{\Omega} s_{d', i', j'}^k$  in Eq. (3); the threshold level  $a_d$  in Eq. (24) works for the uniqueness constraint described with  $\xi \sum_{\Theta} s_{d', i', j'}^k$  and the fixed threshold level  $T$  in Eq. (3). By comparing the revised cooperative algorithm of Eq. (24) and the reaction-diffusion stereo algorithm of Eqs. (13) and (14), we can also understand that the original reaction-diffusion stereo algorithm is an extended version of the cooperative



**Figure 5.** Simple situation having two disparity levels  $d = 0, 1$  pixel ( $N = 2$ ) for the original reaction-diffusion stereo algorithm. Figures (a) and (b) show snap shots of  $u_0$  and  $u_1$  obtained at  $t = 1.0$  and  $t = 4.0$ . (a-1)  $u_{0,i,j}(t = 1.0)$ ; (a-2)  $u_{0,i,j=77}(t = 1.0)$  observed along the white broken line at  $j = 77$  indicated in (a-1); (a-3)  $u_{1,i,j}(t = 1.0)$ ; (a-4)  $u_{0,i,j}(t = 1.0)$  and  $u_{1,i,j}(t = 1.0)$  observed along the white broken lines indicated in (a-1) and (a-3). (b-1)  $u_{0,i,j}(t = 4.0)$ ; (b-2)  $u_{0,i,j=77}(t = 4.0)$  observed along the white broken line at  $j = 77$  indicated in (b-1); (b-3)  $u_{1,i,j}(t = 4.0)$ ; (b-4)  $u_{0,i,j}(t = 4.0)$  and  $u_{1,i,j}(t = 4.0)$  observed along the white broken lines indicated in (b-1) and (b-3). See Table 1 for the parameter settings of the algorithm RDSA(org).

algorithm, and the main difference between the two algorithms is the existence of the inhibitory distribution  $v_d$  described with Eq. (14). Section 5.2 demonstrates the effect of the inhibitory distribution.

## 5. Experimental results and discussion

### 5.1. Demonstration for a simple stereo image pair

This section presents how the original reaction-diffusion stereo algorithm works for a simple situation, in which there are two possible disparity levels  $d = 0, 1$  pixel. Let us suppose that a distribution  $C_{d,i,j}$  obtained by a similarity measure has high values at three points and zero at other points; the two of the three points have the high values of  $C_d$  at the disparity level  $d = 0$  pixel, and the reminder has the high value in the disparity level  $d = 1$  pixel; the three points located at different positions. Figure 5 shows snap shots of  $u_0$  and  $u_1$  obtained by the stereo algorithm at two different time instances  $t = 1.0, 4.0$ . During the early period of the algorithm, for example, at the time instance  $t = 1.0$ , waves originated from the three points began to extend outward as shown in Figs. 5(a-1)~5(a-4). After that, on the distribution of  $u_0$  the two circular waves collided and became one, which occupied a large continuous area as shown in Figs. 5(b-1) and 5(b-2). In contrast to this, on the distribution of  $u_1$  the wave originated from the point in  $u_1$  exclusively collided with the waves originated from the two points on  $u_0$ . As the result, the wave on  $u_1$  did not extend beyond the collision boundary. The former result shown in Figs. 5(b-1) and 5(b-2) denotes that the continuity constraint indeed works, and the latter result shown in Figs. 5(b-3) and 5(b-4) denotes that the uniqueness constraint does.

Algorithm	Equations	Parameter settings
COR5	Eqs. (1), (2)	–
M&P	Eqs. (24), (15)	–
RDSA(org)	Eqs. (13), (14), (15), (16)	$D_v = 3.0$
RDSA(edge)	Eqs. (13), (19), (20), (15), (16)	$D_{v_{\max}} = 15.0, D_{v_{\min}} = 0.5, L_{d_t} = 10.0$
RDSA(aniso-H)	Eqs. (13), (21), (22), (15), (16)	$D_v = 2.0, \phi = 0, \rho = 0.9$
RDSA(aniso-V)		$D_v = 2.0, \phi = \pi/2, \rho = 0.9$

**Table 1.** Summary of the stereo algorithms presented in this chapter and their parameter settings utilized in experiments. RDSA(org) denotes the original reaction-diffusion stereo algorithm, RDSA(edge) denotes the algorithm integrating intensity edge information, RDSA(aniso-H) denotes the algorithm with anisotropic diffusion process depending on pre-specified horizontal orientation, RDSA(aniso-V) denotes that depending on pre-specified vertical orientation, and M&P denotes the cooperative algorithm originally proposed by Marr and Poggio and revised with a reaction-diffusion equation. These algorithms shared the same parameter settings of  $D_u = 1.0, \varepsilon = 1.0 \times 10^{-2}, \alpha = 0.13, \beta = 1.5, b = 10.0, \mu = 3.0, L_t = 100$ , and the same finite differences of  $\delta h = 1/5, \delta t = 1/100$  for numerical computation, and utilized the same similarity measure of Eq. (1) with the correlation window  $\mathcal{B}_5$ . The algorithm COR5 provided a stereo disparity map from the only similarity measure.

## 5.2. Results for Middlebury stereo image pairs

The Middlebury stereo vision page [18, 19] provides four stereo image pairs named TSUKUBA, VENUS, TEDDY and CONES as well as their ground truth data of stereo disparity maps, and definitions of performance evaluation areas such as nonocclusion areas (nonocc.), all areas (all) and depth discontinuity areas (disc.). The page also provides ranking tables of stereo algorithms with respect to bad-match-percentage error measure, in addition to an evaluation system for submitted stereo disparity maps. The bad-match-percentage error measure  $\text{BMP}_T$  evaluates an obtained stereo disparity map  $M_{i,j}$  with the ground truth disparity map  $G_{i,j}$ , as follows:

$$\text{BMP}_T = \frac{1}{|\mathcal{E}|} \sum_{(i,j) \in \mathcal{E}} \sigma(|M_{i,j} - G_{i,j}|, T) \times 100 (\%), \quad (25)$$

in which  $\mathcal{E} \in \{\text{nonocc.}, \text{all}, \text{disc.}\}$  and  $|\mathcal{E}|$  denotes the number of pixels in the area  $\mathcal{E}$ ;  $\sigma(\cdot, T)$  denotes the step function and  $T$  (pixel) denotes a threshold level for judgment of a bad-match pixel or a correct one; we chose  $T = 0.5, 1.0$  pixel in this section. In addition to the error measure, this section also utilized the root-mean-square error measure denoted by RMS and defined by

$$\text{RMS} = \sqrt{\frac{1}{|\mathcal{E}|} \sum_{(i,j) \in \mathcal{E}} (M_{i,j} - G_{i,j})^2} \quad (\text{pixel}). \quad (26)$$

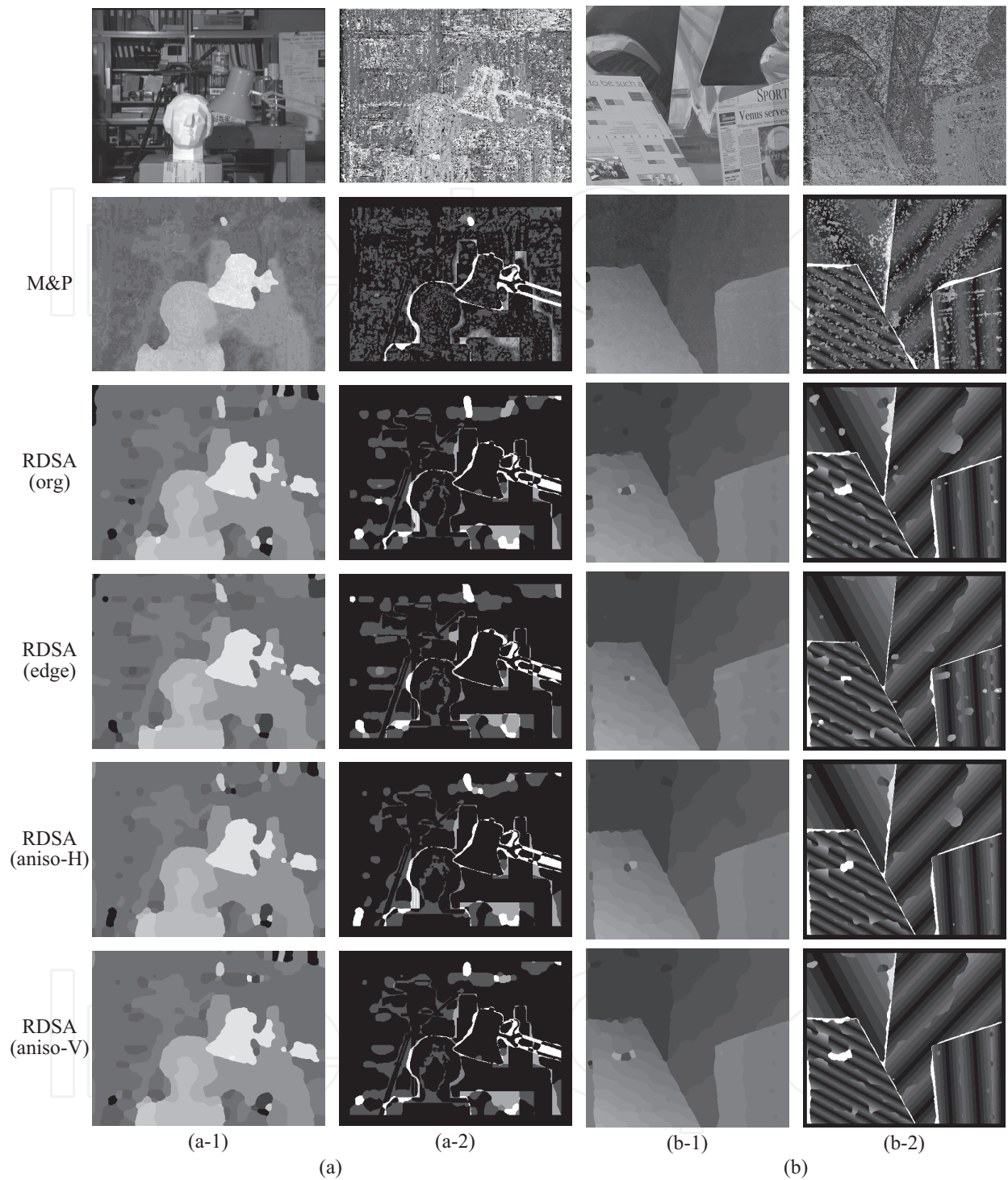
We applied the algorithms presented in this chapter to the four stereo image pairs and evaluated obtained disparity maps with the error measures of  $\text{BMP}_{1.0}$ ,  $\text{BMP}_{0.5}$  and RMS. Table 1 summarizes the algorithms utilized here; COR5 provides a stereo disparity map with the only similarity measure, M&P is the cooperative algorithm revised with a reaction-diffusion equation, RDSAs are the reaction-diffusion stereo algorithms, in which RDSA(org) is the original algorithm, RDSA(edge) is the algorithm integrating intensity edge

Algorithm	Measure	TSUKUBA			VENUS		
		nonocc.	all	disc.	nonocc.	all	disc.
COR5	BMP <sub>1.0</sub> (%)	51.31	52.22	47.02	60.65	61.24	55.62
	BMP <sub>0.5</sub> (%)	68.74	69.41	64.52	69.21	69.68	63.80
	RMS (pixel)	4.36	4.39	4.13	6.14	6.17	5.46
M&P	BMP <sub>1.0</sub> (%)	5.50	7.10	21.94	5.45	6.58	27.89
	BMP <sub>0.5</sub> (%)	39.84	40.99	45.10	29.14	30.01	40.93
	RMS (pixel)	1.29	1.46	2.53	0.84	0.97	2.37
RDSA (org)	BMP <sub>1.0</sub> (%)	7.01	8.81	19.82	2.81	3.97	21.64
	BMP <sub>0.5</sub> (%)	22.82	24.24	27.58	10.93	12.04	25.70
	RMS (pixel)	1.43	1.62	2.50	0.75	0.92	2.01
RDSA (edge)	BMP <sub>1.0</sub> (%)	6.46	7.99	18.54	1.14	2.36	7.48
	BMP <sub>0.5</sub> (%)	23.71	24.92	27.11	10.70	11.87	16.53
	RMS (pixel)	1.43	1.61	2.60	0.57	0.79	1.48
RDSA (aniso-H)	BMP <sub>1.0</sub> (%)	6.78	8.57	20.47	1.99	3.44	18.69
	BMP <sub>0.5</sub> (%)	19.84	21.34	28.38	9.61	10.97	23.57
	RMS (pixel)	1.41	1.61	2.54	0.71	0.91	1.95
RDSA (aniso-V)	BMP <sub>1.0</sub> (%)	6.33	8.12	20.20	2.46	3.90	19.69
	BMP <sub>0.5</sub> (%)	21.02	22.48	27.52	9.40	10.76	24.82
	RMS (pixel)	1.36	1.56	2.52	0.75	0.95	1.94

Algorithm	Measure	TEDDY			CONES			Average Rank
		nonocc.	all	disc.	nonocc.	all	disc.	
COR5	BMP <sub>1.0</sub> (%)	70.93	73.83	69.72	58.55	63.17	62.44	122.0
	BMP <sub>0.5</sub> (%)	76.16	78.55	76.35	64.31	68.34	68.60	122.0
	RMS (pixel)	17.03	18.09	15.25	15.39	17.24	15.22	–
M&P	BMP <sub>1.0</sub> (%)	12.99	20.48	29.68	6.98	13.98	17.83	108.0
	BMP <sub>0.5</sub> (%)	28.08	35.02	43.80	20.00	26.95	30.28	115.3
	RMS (pixel)	2.30	4.70	3.48	2.20	3.15	3.76	–
RDSA (org)	BMP <sub>1.0</sub> (%)	14.00	20.03	29.42	5.03	12.12	14.09	102.7
	BMP <sub>0.5</sub> (%)	22.48	29.29	39.03	10.29	17.45	22.15	88.4
	RMS (pixel)	2.16	3.21	3.35	1.94	3.08	3.35	–
RDSA (edge)	BMP <sub>1.0</sub> (%)	14.53	20.59	27.85	5.41	13.59	14.56	98.3
	BMP <sub>0.5</sub> (%)	23.93	30.54	38.60	12.07	19.96	23.11	90.3
	RMS (pixel)	2.29	3.32	3.45	1.90	3.15	3.20	–
RDSA (aniso-H)	BMP <sub>1.0</sub> (%)	13.52	19.56	29.36	5.21	13.68	14.38	103.3
	BMP <sub>0.5</sub> (%)	22.58	29.28	39.30	10.80	19.13	22.60	86.6
	RMS (pixel)	2.14	3.33	3.32	1.97	3.18	3.39	–
RDSA (aniso-V)	BMP <sub>1.0</sub> (%)	13.87	19.83	29.19	5.64	13.77	15.76	104.8
	BMP <sub>0.5</sub> (%)	22.81	29.52	39.46	11.03	19.07	23.94	87.6
	RMS (pixel)	2.09	3.25	3.33	2.07	3.05	3.56	–

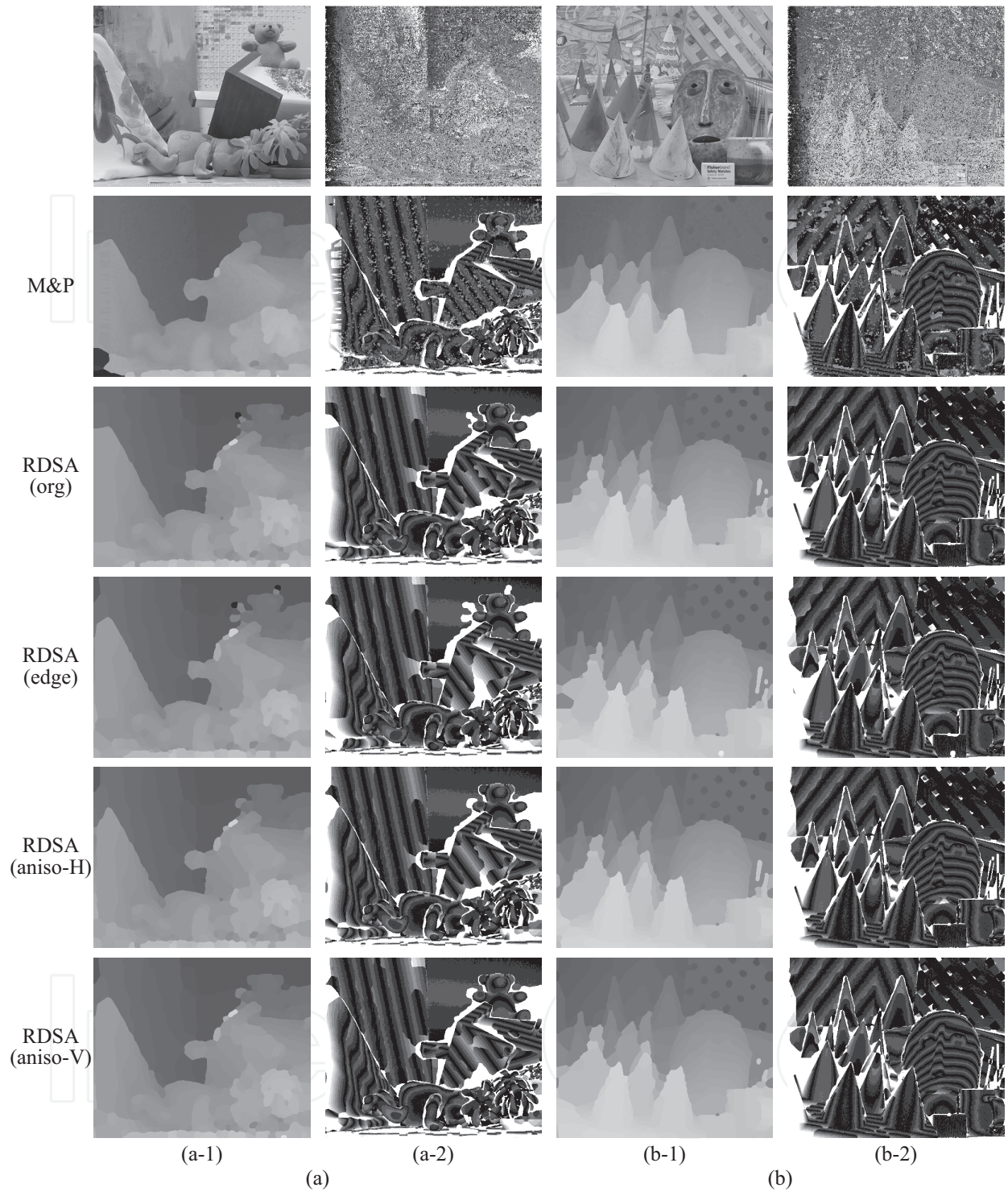
**Table 2.** Evaluation of the stereo algorithms summarized in Table 1 with BMP<sub>1.0</sub>, BMP<sub>0.5</sub> and RMS [see Eqs. (25) and (26)]. Figures 6 and 7 show obtained stereo disparity maps; the Middlebury stereo vision page [18] provides the stereo image pairs of TSUKUBA, VENUS, TEDDY and CONES, their ground truth disparity maps and definition of evaluation areas: non occlusion areas (nonocc.), all areas (all) and depth discontinuity areas (disc.). We computed average ranks for the algorithms according to the ranking tables of the page on March 15, 2012.



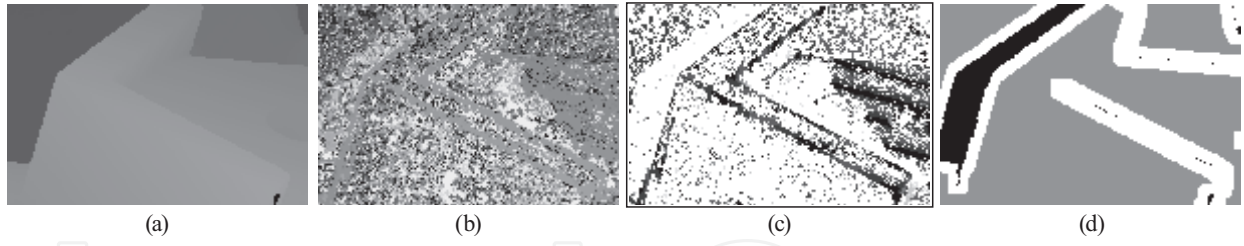


**Figure 6.** Results of stereo disparity maps obtained for (a) TSUKUBA and (b) VENUS. From top to bottom in Figs. (a-1) and (b-1), the left reference image and the disparity maps obtained by M&P, RDSA(org), RDSA(edge), RDSA(aniso-H) and RDSA(aniso-V) are shown. From top to bottom in Figs. (a-2) and (b-2), the disparity map obtained by COR5 and absolute error distribution maps evaluated for the disparity maps shown in Figs. (a-1) and (b-1) are shown; gray levels indicate absolute error and a brighter level indicates larger error. See Table 1 for the algorithms and their parameter settings and Table 2 for their quantitative performance evaluation results.





**Figure 7.** Results of stereo disparity maps obtained for (a) TEDDY and (b) CONES. From top to bottom in Figs. (a-1) and (b-1), the left reference image and the disparity maps obtained by M&P, RDSA(org), RDSA(edge), RDSA(aniso-H) and RDSA(aniso-V) are shown. From top to bottom in Figs. (a-2) and (b-2), the disparity map obtained by COR5 and absolute error distribution maps evaluated for the disparity maps shown in Figs. (a-1) and (b-1) are shown; gray levels indicate absolute error and a brighter level indicates larger error. See Table 1 for the algorithms and their parameter settings and Table 2 for their quantitative performance evaluation results.



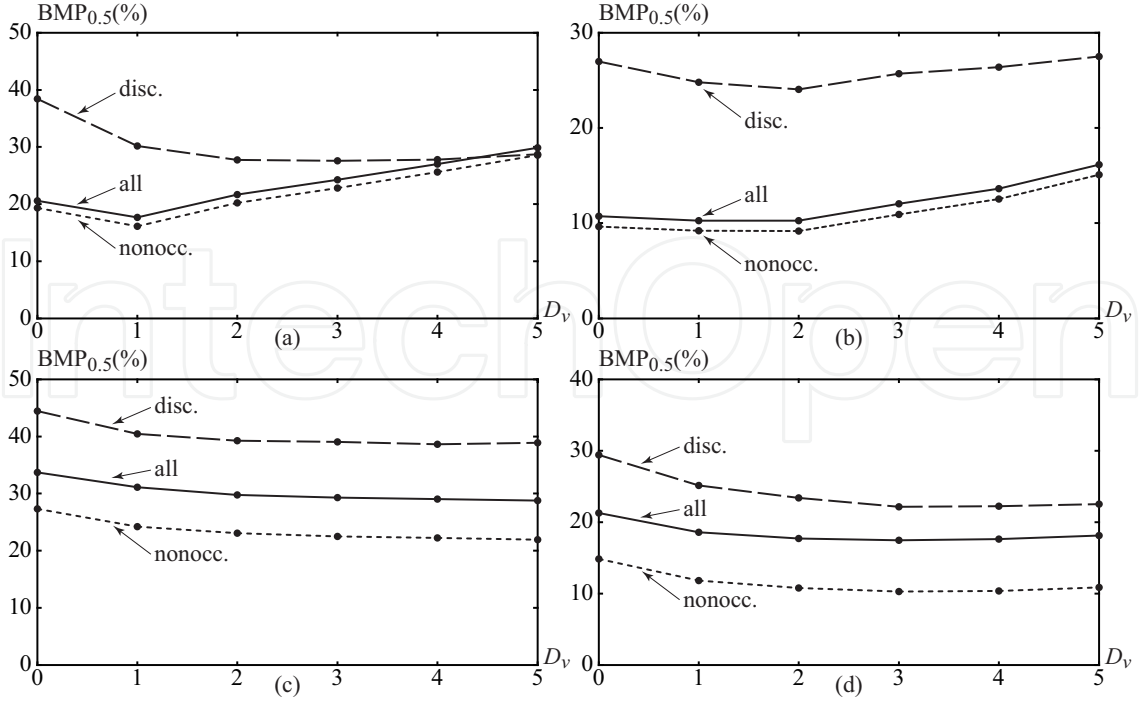
**Figure 8.** Enlarged stereo disparity map obtained by COR5 for TEDDY; (a) the groundtruth disparity map  $G_{i,j}$ , (b) the obtained disparity map  $M_{i,j}$ , (c) the absolute error distribution  $|M_{i,j} - G_{i,j}|$  and (d) the definition of depth discontinuity areas (disc.). See Fig. 7(a) for the obtained full disparity map and the full absolute error distribution.

information, and RDSA(aniso-H) and RDSA(aniso-V) are the algorithms with anisotropic diffusion processes depending on the pre-specified orientation. We fixed their parameter settings across the four stereo image pairs except for the image size of  $\mathcal{L}_x \times \mathcal{L}_y$  and the possible disparity levels  $\mathcal{D}$ , which were provided as the Middlebury stereo vision page indicates.

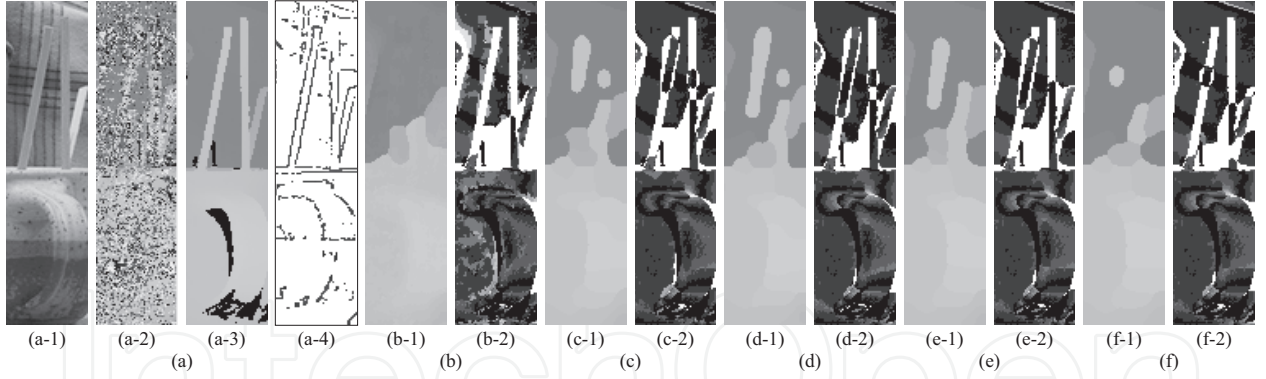
Table 2 shows evaluation results and average ranks of the algorithms, and Figs. 6 and 7 show stereo disparity maps and their absolute error distributions. From these results, we can state the following. On each of the algorithms: RDSAs, average ranks evaluated with  $BMP_{0.5}$  were better than those with  $BMP_{1.0}$ . The algorithms RDSAs were better than M&P in most cases excluding for TSUKUBA and including average ranks. Among RDSAs, average ranks were almost similar; although there was no distinguishable difference on performance among RDSAs, RDSA(edge) was slightly better than RDSA(org) and RDSA(aniso-V), and RDSA(edge) indicated similar error measures with RDSA(aniso-H) according to the sum of average ranks obtained with  $BMP_{1.0}$  and  $BMP_{0.5}$ . When focusing on the results for VENUS, in particular, the results for the depth discontinuity areas (disc.), RDSA(edge) was quite effective in comparison with other RDSAs. This can be also easily confirmed with the disparity maps and their error distributions shown in Fig. 6(b). In most cases, the results with COR5 denoted that the error measures evaluated in the areas of disc. were less than those evaluated in the areas of nonocc. and all, except for the results of CONES with  $BMP_{1.0}$  and  $BMP_{0.5}$ . This implies that the depth discontinuity areas bring rather reliable disparity information than other nonocclusion areas do in COR5; Fig. 8 supports this implication. We consider that this is because there is rather rich information of image intensity distribution around depth discontinuity areas. In contrast to this, the results for M&P and RDSAs indicated the worst error measures for the depth discontinuity areas.

Figure 9 shows dependence of error measures on the inhibitory diffusion coefficient  $D_v$  in RDSA(org). Let us focus on the dependence confirmed for the depth discontinuity areas (disc.). For all of the four image pairs TSUKUBA, VENUS, TEDDY and CONES, the error measures achieved the least values at certain positive values  $D_v > 0$ ; for example, for the image pair TSUKUBA, the error measure decreased as  $D_v$  increased and achieved the least value at  $D_v = 3.0$ ; for other image pairs, they indicated similar trends.

Figure 10 shows a representative example in which RDSA(edge) works in comparison to other algorithms. The example shows a small rectangular area capturing a coffee cup and chopsticks in the left image of CONES, and stereo disparity maps obtained in the area. An edge detection result shown in Fig. 10(a-4) has almost completely detected object boundaries



**Figure 9.** Dependence of error measures  $BMP_{0.5}$  on  $D_v$  in the original reaction-diffusion stereo algorithm RDSA(org) applied to the stereo image pairs (a) TSUKUBA, (b) VENUS, (c) TEDDY and (d) CONES. See Table 1 for the other parameter settings; see Figs. 6 and 7 for disparity maps obtained with  $D_v = 3.0$ .



**Figure 10.** Enlarged results of disparity maps obtained for the rectangular area capturing a coffee cup and chopsticks in CONES. Figure (a) shows the left image in (a-1), the initial disparity map  $M_{i,j}^0$  in (a-2), the ground truth disparity map  $G_{i,j}$  in (a-3) and an edge map  $\mathcal{M}_e$  obtained by another edge detection algorithm [25] in (a-4). Figure (b) shows results by M&P. Figure (c) shows results by RDSA(org). Figure (d) shows results by RDSA(edge) integrating the edge information (a-4). Figure (e) shows results by RDSA(aniso-H). Figure (f) shows results by RDSA(aniso-V). In each figure of (b)~(f), (b-1), (c-1), (d-1), (e-1) and (f-1) are disparity maps  $M_{i,j}$  in the rectangular area, and (b-2), (c-2), (d-2), (e-2) and (f-2) are their absolute error distributions  $|M_{i,j} - G_{i,j}|$ . See Fig. 7(b) for their full stereo disparity maps.

of the chopsticks. Although the result by RDSA(edge) was not perfect as shown in Fig. 10(d), it was better than the results by the other algorithms of RDSA(org) and RDSA(aniso-V). Since there exists depth discontinuity along the chopsticks standing almost vertically, the algorithm



RDSA(aniso-H) causing strong inhibitory diffusion in horizontal orientation worked better than RDSA(aniso-V) and provided a result similar to that of RDSA(edge).

Figure 11 shows temporal changes of error measures evaluated during processes of stereo disparity detection in the stereo algorithms RDSA(org), RDSA(aniso-H) and RDSA(aniso-V). In the early period less than  $t = 10$ , the error measures dynamically changed; after that, they changed slowly and achieved almost constant at  $t = 100$ . From these results, we can roughly state that the algorithms converged.

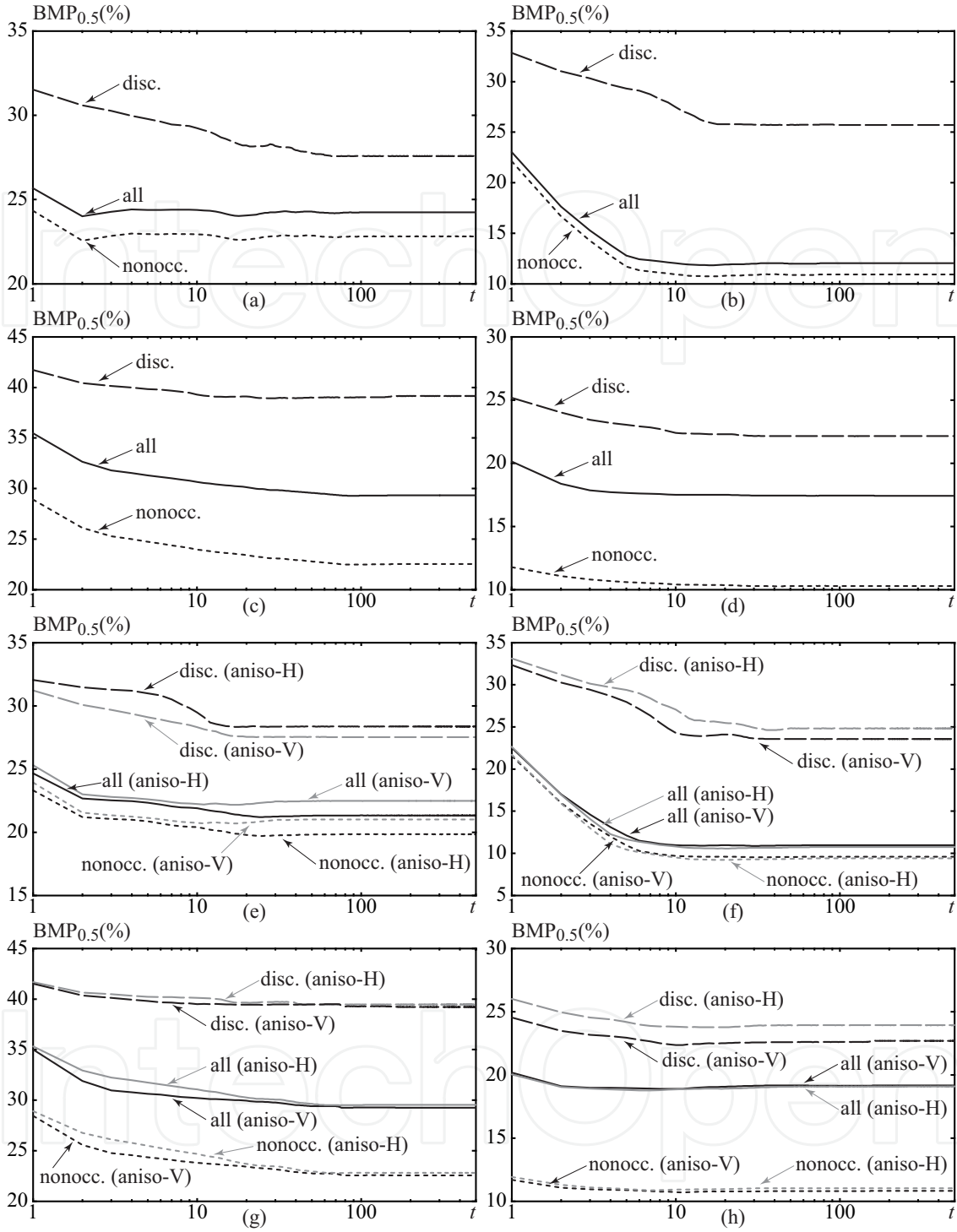
### 5.3. Discussion

Each of the reaction-diffusion systems utilized in the original reaction-diffusion stereo algorithm has diffusion terms  $D_u \nabla^2 u_d$  and  $D_v \nabla^2 v_d$  on excitation and inhibition. The dependence of the inhibitory diffusion coefficient  $D_v$  on performance shows that the strong inhibitory diffusion compared with the weak excitatory one improves performance in depth discontinuity areas. In addition, the reaction-diffusion stereo algorithm having the inhibitory distribution works better than the cooperative one revised with a single reaction-diffusion equation having only excitatory distribution. These results indicate the importance of the strong inhibitory diffusion, that is, the long-range inhibition. If turning our attention to an edge detection algorithm proposed by Marr and Hildreth [59], we can find that the algorithm has a long-range inhibition and a short-range excitation. Thus, we hypothesize that the long-range inhibition is an important factor and underlies visual functions including edge detection and stereo disparity detection.

In low curvature areas found along straight lines of depth discontinuity areas (see Fig. 10), we can not expect the effect of the strong-inhibitory diffusion. In addition, previous psychological results implied that the human depth perception integrates several visual cues [1–3]. From these two reasons, we designed the reaction-diffusion stereo algorithm integrating intensity edge information so as to work in the depth discontinuity areas, if the intensity edge areas coincide with the depth discontinuity areas. Figure 10 showed an example of a successful situation, in which the integration of intensity edge information compensated for difficulties in the depth discontinuity areas. For more performance improvement, it is necessary to integrate not an edge map, but a contour map defining object boundaries into the algorithm. The study by Martin et al. [57] should be helpful for obtaining a contour map.

Image processing and computer vision algorithms utilizing diffusion equations require estimating stopping time [60]; this is called the termination problem [56]. Reaction-diffusion algorithms with a nonlinear reaction such as the FitzHugh-Nagumo type do not require solving the termination problem, if they can spend sufficient duration of time for their processing, as shown in Figs. 11(a)~(d).

Between the two algorithms RDSA(aniso-H) and RDSA(aniso-V) of which diffusion coefficients depend on pre-specified orientation, we could not confirm distinguishable difference on their convergence trends, as shown in Figs. 11(e)~(h). According to previous results of psychological experiments [4, 5], the human stereo depth perception indicates anisotropy on latency. In order to confirm the correctness of the idea introducing anisotropic diffusion processes, we need to apply the algorithms to random-dot stereograms utilized



**Figure 11.** Temporal changes of the error measures  $BMP_{0.5}$  on the algorithms: RDSA(org), RDSA(aniso-H) and RDSA(aniso-V). Figures (a)~(d) show the results on RDSA(org) and Figs. (e)~(h) show those on RDSA(aniso-H) and RDSA(aniso-V); Figs. (a) and (e) show those for the image pair TSUKUBA, Figs. (b) and (f) show those for VENUS, Figs. (c) and (g) show those for TEDDY, and Figs. (d) and (h) show those for CONES; each of Figs. (a)~(h) include the results in the areas: nonocc., all and disc. See Figs. 6 and 7 for disparity maps at  $t = 100$ , Table 1 for their parameter settings, and Table 2 for quantitative error measures at  $t = 100$ .



in the psychological experiments and confirm their convergence. At this moment, our experimental results are insufficient to mention the correctness of the idea.

There are two main future research topics. One of them is intended for the occlusion problem. As shown in Fig. 7, the results obtained for TEDDY and CONES indicate that occluded areas have much error or large deviation from the true disparity level. In order to solve the occlusion problem, we believe that the idea of bi-directional matching is helpful also in the reaction-diffusion stereo algorithm. That is, by feeding the result of the bi-directional matching to the iteration process of the reaction-diffusion systems, we build a dynamical system with feedback and try to solve the occlusion problem. This feedback system may furthermore bring a hint for building a strong fusion model proposed as the human depth perception integrating several visual cues. The other future research topic is accuracy improvement for slanted surfaces. The reaction-diffusion stereo algorithms presented here imposed the continuity constraint on each reaction-diffusion system associated with a disparity level; the continuity constraint does not work across more than two disparity levels. Thus, for slanted surfaces, we need to impose the continuity constraint on the surfaces. We believe that the imposition of the constraint on the slanted surfaces is possible. In addition to the accuracy improvement, the imposition may bring a hint for our understanding the anisotropy of the human stereo depth perception.

## 6. Conclusion

This chapter presented a reaction-diffusion stereo algorithm [14] solving the stereo correspondence problem with multiple reaction-diffusion systems. The algorithm converts the stereo correspondence problem into a segmentation problem with respect to stereo disparity levels. Each of the reaction-diffusion systems is associated with a particular disparity level. A set of FitzHugh-Nagumo type reaction-diffusion equations describes each of the reaction-diffusion systems. Depending on the parameter settings of the set, it behaves as a mono-stable system, a bi-stable system and an oscillatory system. In order to govern existence or non-existence of a stereo disparity level associated with a reaction-diffusion system, the algorithm utilizes bi-stable reaction-diffusion systems. The bi-stable systems have two stable steady states, of which the excited state denotes the existence, and of which the resting state denotes the non-existence. Since the bi-stable systems have the nature of driving wave propagation, they extend their associated areas with the nature and fill-in undefined areas of stereo disparity. This filling-in process works as the continuity constraint presented by Marr and Poggio. For the uniqueness constraint assuming only one disparity level at a particular point, the algorithm exclusively linked the multiple reaction-diffusion systems with mutual inhibition mechanism. In addition to these constraints, the reaction-diffusion algorithm has the self-inhibition mechanism caused by rapid inhibitory diffusion. The mechanism contributes to preserving detailed depth structure or high curvature areas such as areas of corner points in a disparity map.

Besides the original reaction-diffusion stereo algorithm, this chapter presented two additional stereo algorithms with anisotropic diffusion processes. One of the algorithms integrated the intensity edge information, which is provided by another edge detection algorithm designed with discretely coupled nonlinear elements [25]; the stereo algorithm was inspired by the human depth perception integrating several visual cues. The other algorithm introduced

anisotropic diffusion processes depending on pre-specified horizontal or vertical orientation. The human visual system differently perceives a horizontally slanted surface and a vertically slanted one; this anisotropy inspired the stereo algorithm with anisotropic diffusion processes depending on pre-specified orientation.

The experimental section demonstrated performance of the reaction-diffusion stereo algorithms as well as a cooperative algorithm revised with a reaction-diffusion equation. Experimental results showed that the reaction-diffusion stereo algorithms perform better than the cooperative algorithm, and the original reaction-diffusion stereo algorithm works better with larger inhibitory diffusion coefficients especially in depth discontinuity areas. Thus, we reached the conclusion that the strong inhibitory diffusion or the strong inhibitory coupling is an important condition for the algorithms, and the strong inhibitory coupling underlies visual functions; a previous edge detection algorithm [59] and biological evidence in vision [52] also support the conclusion. Integrating image intensity edge information is effective, if intensity edge areas coincide with depth discontinuity areas. This also implies that the algorithm achieves better, if object contour information is available; a future research topic exists in how to obtain the reliable contour information. For the algorithms with anisotropic diffusion processes depending on horizontal or vertical orientation, experimental results demonstrated convergence of the algorithms for four stereo image pairs. However, there was no distinguishable difference on the convergence between the two algorithms. We need to confirm the convergence by applying the algorithms to random-dot stereograms utilized in psychological experiments. In addition to these future topics, development of the depth perception model integrating several visual cues is also an interesting topic, for which the previous strong and weak fusion models may provide hints.

## Acknowledgment

The present study was supported in part by the Grant-in-Aid for Scientific Research from the Japan Society for the Promotion of Science (No. 20500206, No. 23500278).

## Author details

Atsushi Nomura, Koichi Okada, Hidetoshi Miike and Yoshiki Mizukami  
*Yamaguchi University, Japan*

Makoto Ichikawa and Tatsunari Sakurai  
*Chiba University, Japan*

## 7. References

- [1] Landy M S, Maloney L T, Johnston E B, Young M (1995) Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research* 35:389–412.
- [2] Bradshaw M F, Rogers B J (1996) The interaction of binocular disparity and motion parallax in the computation of depth. *Vision Research* 36:3457–3468.
- [3] Ichikawa M, Saida S, Osa A, Munechika K (2003) Integration of binocular disparity and monocular cues at near threshold level. *Vision Research* 43:2439–2449.

- [4] Rogers B J, Graham M E (1983) Anisotropies in the perception of three-dimensional surfaces. *Science* 221:1409–1411.
- [5] Ichikawa M (1992) Effects of the slant orientation in depth on binocular stereopsis. *The Japanese Journal of Psychonomic Science* 11:9–17 [in Japanese].
- [6] Gonzalez R C, Woods R E (1992) *Digital Image Processing*. New York: Addison-Wesley.
- [7] Marr D, Poggio T (1976) Cooperative computation of stereo disparity. *Science* 194:283–287.
- [8] Marr D, Palm G, Poggio T (1978) Analysis of a cooperative stereo algorithm. *Biological Cybernetics* 28:223–239.
- [9] Marr D, Poggio T (1979) A computational theory of human stereo vision. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 204:301–328.
- [10] Koffka K (1955) *Principles of Gestalt Psychology*. London: Routledge and Kegan Paul Ltd.
- [11] Köhler W (1969) *The Task of Gestalt Psychology*. New Jersey: Princeton University Press.
- [12] Beck J (1966) Effect of orientation and of shape similarity on perceptual grouping. *Perception & Psychophysics* 1:300–302.
- [13] Nomura A, Ichikawa M, Miike H (2004) Realizing the grouping process with the reaction-diffusion model. *IPSI Transactions on Computer Vision and Image Media* 45(SIG8/CVIM-9):26–39 [in Japanese].
- [14] Nomura A, Ichikawa M, Miike H (2009) Reaction-diffusion algorithm for stereo disparity detection. *Machine Vision and Applications* 20:175–187.
- [15] Murray J D (1989) *Mathematical Biology*. Berlin: Springer-Verlag.
- [16] Turing A M (1952) The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 237:37–72.
- [17] Barlow Jr R B, Quarles Jr D A (1975) Mach bands in the lateral eye of *Limulus*. *Journal of General Physiology* 65:709–730.
- [18] Scharstein D, Szeliski R. The Middlebury stereo vision page. Available: <http://vision.middlebury.edu/stereo/>. Accessed 2012 March 15.
- [19] Scharstein D, Szeliski R (2002) A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision* 47:7–42.
- [20] Julesz B (1960) Binocular depth perception of computer-generated patterns. *Bell System Technical Journal* 39:1125–1162.
- [21] Julesz B (1964) Binocular depth perception without familiarity cues. *Science* 145:356–362.
- [22] Thimbleby H W, Inglis S, Witten I H (1994) Displaying 3d images: algorithms for single-image random-dot stereograms. *IEEE Computer* 27:38–48.
- [23] Brookes A, Stevens K A (1989) The analogy between stereo depth and brightness. *Perception* 18:601–614.
- [24] Lunn P D, Morgan M J (1995) The analogy between stereo depth and brightness: a reexamination. *Perception* 24:901–904.
- [25] Nomura A, Ichikawa M, Sianipar R H, Miike H (2008) Edge detection with reaction-diffusion equations having a local average threshold. *Pattern Recognition and Image Analysis* 18:289–299.
- [26] Dev P (1975) Perception of depth surfaces in random-dot stereograms: a neural model. *International Journal of Man-Machine Studies* 7:511–528.

- [27] Brown M Z, Burschka D, Hager G D (2003) Advances in Computational Stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25:993–1008.
- [28] Kanade T, Okutomi M (1994) A stereo matching algorithm with an adaptive window: theory and experiment. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16:920–932.
- [29] March R (1988) Computation of stereo disparity using regularization. *Pattern Recognition Letters* 8:181–187.
- [30] Marroquin J, Mitter S, Poggio T (1987) Probabilistic solution of ill-posed problems in computational vision. *Journal of the American Statistical Association* 82:76–89.
- [31] Fua P (1993) A parallel stereo algorithm that produces dense depth maps and preserves image features. *Machine Vision and Applications* 6:35–49.
- [32] Luo A, Burkhardt H (1995) An intensity-based cooperative bidirectional stereo matching with simultaneous detection of discontinuities and occlusions. *International Journal of Computer Vision* 15:171–188.
- [33] Pan H, Magarey J (1998) Multiresolution phase-based bidirectional stereo matching with provision for discontinuity and occlusion. *Digital Signal Processing* 8:255–266.
- [34] Zitnick C L, Kanade T (2000) A cooperative algorithm for stereo matching and occlusion detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22:675–684.
- [35] Tsirlin I, Allison R S, Wilcox L M (2008) Stereoscopic transparency: constraints on the perception of multiple surfaces. *Journal of Vision* 8:1–10.
- [36] Akerstrom R A, Todd J T (1988) The perception of stereoscopic transparency. *Perception & Psychophysics* 44:421–432.
- [37] Shizawa M (1992) On visual ambiguities due to transparency in motion and stereo. *Proceedings of European Conference on Computer Vision* 411–419.
- [38] Szeliski R, Golland P (1999) Stereo matching with transparency and matting. *International Journal of Computer Vision* 32:45–61.
- [39] Shizawa M (1994) Computational theory of binocular stereo transparency and a single-shot model for computing two-fold disparities. *IEICE Transactions on Information and Systems* J77-D-II:1245–1254 [in Japanese].
- [40] Sun J, Zheng N-N, Shum H-Y (2003) Stereo matching using belief propagation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 25:787–800.
- [41] Kolmogorov V, Zabih R (2004) What energy functions can be minimized via graph cuts? *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26:147–159.
- [42] Busse H, Hess B (1973) Information transmission in a diffusion-coupled oscillatory chemical system. *Nature* 244:203–205.
- [43] Kuhnert L (1986) A new optical photochemical memory device in a light sensitive chemical active medium. *Nature* 319:393–394.
- [44] Kuhnert L, Agladze K I, Krinsky V I (1989) Image processing using light-sensitive chemical waves. *Nature* 337:244–247.
- [45] Sakurai T, Mihaliuk E, Chirila F, Showalter K (2002) Design and control of wave propagation patterns in excitable media. *Science* 296:2009–2012.
- [46] FitzHugh R (1961) Impulses and physiological states in theoretical models of nerve membrane. *Biophysical Journal* 1:445–466.
- [47] Nagumo J, Arimoto S, Yoshizawa S (1962) An active pulse transmission line simulating nerve axon. *Proceedings of the IRE* 50:2061–2070.

- [48] Gierer A, Meinhardt H (1972) A theory of biological pattern formation. *Kybernetik* 12:30–39.
- [49] Kondo S (2002) The reaction-diffusion system: a mechanism for autonomous pattern formation in the animal skin. *Genes to Cells* 7:535–541.
- [50] Sick S, Reinker S, Timmer J, Schlake T (2006) WNT and DKK determine hair follicle spacing through a reaction-diffusion mechanism. *Science*, 314:1447–1450.
- [51] Wade N J (2007) Image, eye, and retina. *Journal of Optical Society of America, A* 24:1229–1249.
- [52] Hartline H K, Ratliff F (1958) Spatial summation of inhibitory influences in the eye of *Limulus*, and the mutual interaction of receptor units. *Journal of General Physiology* 41:1049–1066.
- [53] Nomura A, Ichikawa M, Miike H, Ebihara M, Mahara H, Sakurai T (2003) Realizing visual functions with the reaction-diffusion mechanism. *Journal of the Physical Society of Japan* 72:2385–2395.
- [54] Ebihara M, Mahara H, Sakurai T, Nomura A, Osa A, Miike H (2003) Segmentation and edge detection of noisy image and low contrast image based on a reaction-diffusion model. *The Journal of the Institute of Image Electronics Engineers of Japan* 32:378–385 [in Japanese].
- [55] Kurata N, Kitahata H, Mahara H, Nomura A, Miike H, Sakurai T (2009) Stationary pattern formation in a discrete excitable system with strong inhibitory coupling. *Physical Review E* 79:056203.
- [56] Chen K, Wang D (2002) A dynamically coupled neural oscillator network for image segmentation. *Neural Networks* 15:423–439.
- [57] Martin D R, Fowlkes C C, Malik J (2004) Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26:530–549.
- [58] Shoji H, Iwasa Y, Mochizuki A, Kondo S (2002) Directionality of stripes formed by anisotropic reaction-diffusion models. *Journal of Theoretical Biology* 214:549–561.
- [59] Marr D, Hildreth E (1980) Theory of edge detection. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 207:187–217.
- [60] Mrázek P, Navara M (2003) Selection of optimal stopping time for nonlinear diffusion filtering. *International Journal of Computer Vision* 52:189–203.