SICE Annual Conference 2007
Sept. 17-20, 2007, Kagawa University, Japan

# Parallel Processing of Figure-Ground Separation in Random Dot Kinematogram with Optical Flow Discretization

Satoshi ITO[1] and Minoru Sasaki[2]

[1]Faculty of Engineering, Gifu University, Gifu, Japan / BMC research Center RIKEN, Nagoya, Japan
(Tel : +81-58-293-2540; E-mail: satoshi@gifu-u.ac.jp)
[2]Department of Human and Information Systems, Faculty of Engineering, Gifu University, Gifu, Japan
(Tel : +81-58-293-2541; E-mail: sasaki@gifu-u.ac.jp)

**Abstract:** Parallel processing is an important factor for fast image processing in the computer vision as well as human visual systems. From this point of view, this paper treats the figure-ground separation problem in the random dot kinematogram. To achieve this separation, the optical flow is discretized in the five vectors, each of which the flag is assigned to indicate its correct direction in each pixel. The recognition process is described as the flag dynamics, which is given as the gradient system of the global potential functional. The performance is examined with computer simulations.

**Keywords:** image processing, figure-ground separation, random dot animation, parallel operation, optical flow discretization

## 1. INTRODUCTION

A human visual system is constructed with a parallel operating structure that enables us to process a large amount of image information in the real-time. Some important processes such as edge detection, an optical flow calculation, shape from shading and so on are called an early vision, and mathematically explained using the framework of regularization theories [1].

Yuasa et al. has described a visual processing algorithm of auto-associative memories[2], binocular stereograms[3] and figure-ground separations in animation [4] using a special class of differential equations, reaction-diffusion equations. In this paper, we also consider a method for the figure-ground separation in random dot animations using the reaction diffusion equations, where we aim at distinguishing multiple objects (figures) in the image.

## 2. RELATED WORKS

Animations are constructed from many slightly-different images. Fast serial switching of these images allows us to recognize the movement in the animation. Each image in the animation is called here frame.

In the normal animation, each frame possesses some features such as edges, moment, color and so on. Between two consecutive frames, the deviation of these features is detected. Based on this deviation, human could recognize the motion in the animation. However, human can recognize some moving parts (i.e., figures) from the background (ground) even in the random dot kinematogram (RDK), which is an animation consisting of the random dot images. Obviously, the frame in the RDK is a random dot image. Thus, there are no features in the random dot frame. This fact indicates that human never utilizes some features found in the still image for figure-ground separation in animation. In other words, human visual system possesses a recognition mechanism of figure-ground separation that does not need any image features. Our interest here exists in explaining such a human visual processing by the computational approach.

Regarding visual process of RDK, Ueyama et al. [4], [5] have succeeded a figure-ground separation by describing the recognition process with the bistable Ginzburg-Landau equation. Further, Okura et al. [6] proposed a method for restoring the 3D shape of objects that are separated from RDK as figures.

In these works, however, the figure-ground separations are executed in two steps. Firstly, the optical flow is calculated in each pixels. Next, the clustering of figures is performed based on the information of the optical flow. Then, each of two steps includes the minimization process. This double minimization process makes an amount of the computation large. From this point of view, we here consider an algorithm and a system structure of the RDK processing, and aim at describe the figure-ground separation as a single minimization process. We report a simple simulation to evaluate the recognition performance of this system.

## 3. SYSTEM CONSTRUCTION

### 3.1 Assumptions

To make a problem simple, we assume the followings on the RDK.

- The image consists of the random dots in black and white.
- The figure and ground neither rotate nor be scaled.
- The translation of the figure and ground are limited in four directions: up, down, right and left.
- The translation distance during one frame interval is at most 1 dot.

Owing to these assumptions, we can discretize the optical flow in only five vectors. The figures are clustered based on this restricted optical flows.

### 3.2 Basic idea

The flags are prepared to each pixel of the image. This flag represents the moving direction of its pixel which

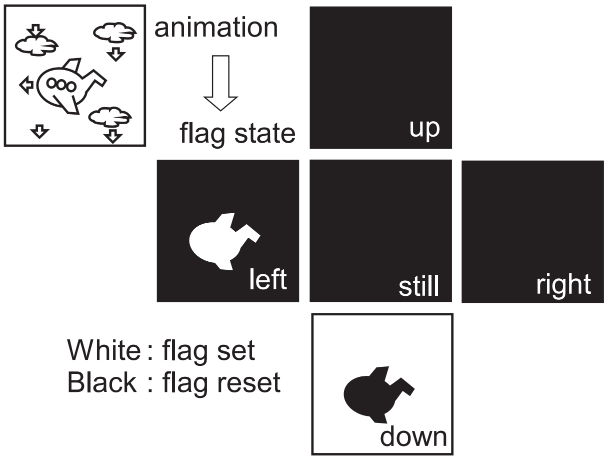


Fig. 1 Example of figure-ground separation.

is represented as the discretized optical flow. Thus, 5 flags are provided: up, left, still, right, and down flags. Between two consecutive frames, the moving direction should be uniquely determined in each pixel except the occlusion area. Hence, among five, only one flag must be set at each pixel.

Now, the figure-ground separation is represented as clustering the flags in each flag plane. The flag plane is a plane that is constructed by picking up a flag representing same optical flow from each pixel and aligning it in the same order as the pixels from which the flag is picked up. In the case of the Fig. 1, for example, the central part is moving to the left, while the background is moving to the bottom. Then the up, still and right flags should be reset in all the pixels and the left and down flags should be partly set in the corresponding pixels.

### 3.3 Recognition dynamics

We construct the dynamics from viewpoint of the local parallel processing. When we attempt to determine the flag value from the local image information, we will sometimes obtain incorrect results due to the locality of the information. These results should be corrected for the correct recognition. Here, we achieve this correction by introducing the dynamical process during which the local information spreads to the neighboring area. The spread of the local information is expected to compensate the shortness of the information on the whole images.

In order to define the dynamical process, the flag is assumed to take continuous value between 0 and 1. The dynamical process is defined to the flag values so that they are determined appropriately, i.e., the steady state should be given as illustrated in Fig. 1.

Here, we define some notations. ${}^np(x,y)$ denotes the pixel value at the coordinate $(x,y)$ in the $n$-th frame, and ${}^nf_k(x,y,t)$ is flag value there. $t$ is a time in the recognition process and $k$ distinguished the direction. $k=1,2,3,4,5$ respectively represents up, left, still, right and down direction.

Generally speaking, the optical flow possesses the following features.

- The optical flow is uniquely determined in each pixel except the occlusion area.
- The optical flow is almost continuous in the image except the boundary of the figures.

The first feature indicates that only one flag should be set among five flags of each pixels, as mentioned the above. The second feature means that the neighboring flags in the flag plane usually takes the same values. This results from the fact that the boundary between figure and ground, which causes the discontinuity of the optical flow, is relatively small in one image. Based on these features, the flag dynamics are constructed below.

To achieve the first feature, the competitive interaction is introduced among 5 flags in each pixel. As such a competitive interaction, the winner-take-all interaction which is proposed in the Synergetics[7] is adopted. This interaction is described as the gradient system with the following potential functional

$$V_r = \iint_\Omega \left[ -\frac{1}{2}\sum_{k=1}^{5} f_k^2 + \frac{1}{4}\sum_{k=1}^{5}\sum_{k'=k} f_k^2 f_{k'}^2 + \frac{1}{4}\left(\sum_{k=1}^{5} f_k^2\right)^2 \right] dxdy \tag{1}$$

Here, we omit the frame number $n$. This interaction makes the sole flag converge to 1 while all the other converge to 0.

Regarding the second feature, the diffusion process is introduced. The diffusion process is one of the local interactions that possessing the averaging action. This process should be defined in each flag plane so that the flag values are corrected by the effect of the neighboring optical flow information. The diffusion process require boundary conditions. Then, the total sum of the flag value contains global information that indicates the ratio of the optical flow direction in the entire image. Thus, boundary conditions such that keep the total sum of the flag values in the flag plane should be imposed.

The total dynamics are construct by combining the above two interactions. As a result, the flag dynamics are described as a reaction-diffusion equation:

$$\tau\frac{df_k}{dt} = -\frac{\partial V_r}{\partial f_k} + K_D\Delta f_k \tag{2}$$

Here, $\Delta$ is Laplacian operator, $\tau$ is a time constant and $K_D$ is a diffusion coefficient.

The above dynamics (2) is also described as the gradient system

$$\tau\frac{df_k}{dt} = -\frac{\partial V}{\partial f_k} \tag{3}$$

whose potential functional is given as follows.

$$V = \iint_\Omega \left[ -\frac{1}{2}\sum_{k=1}^{5} f_k^2 + \frac{1}{4}\sum_{k=1}^{5}\sum_{k'=k} f_k^2 f_{k'}^2 + \frac{1}{4}\left(\sum_{k=1}^{5} f_k^2\right)^2 + K_D(\nabla f_k)^2 \right] dxdy \tag{4}$$

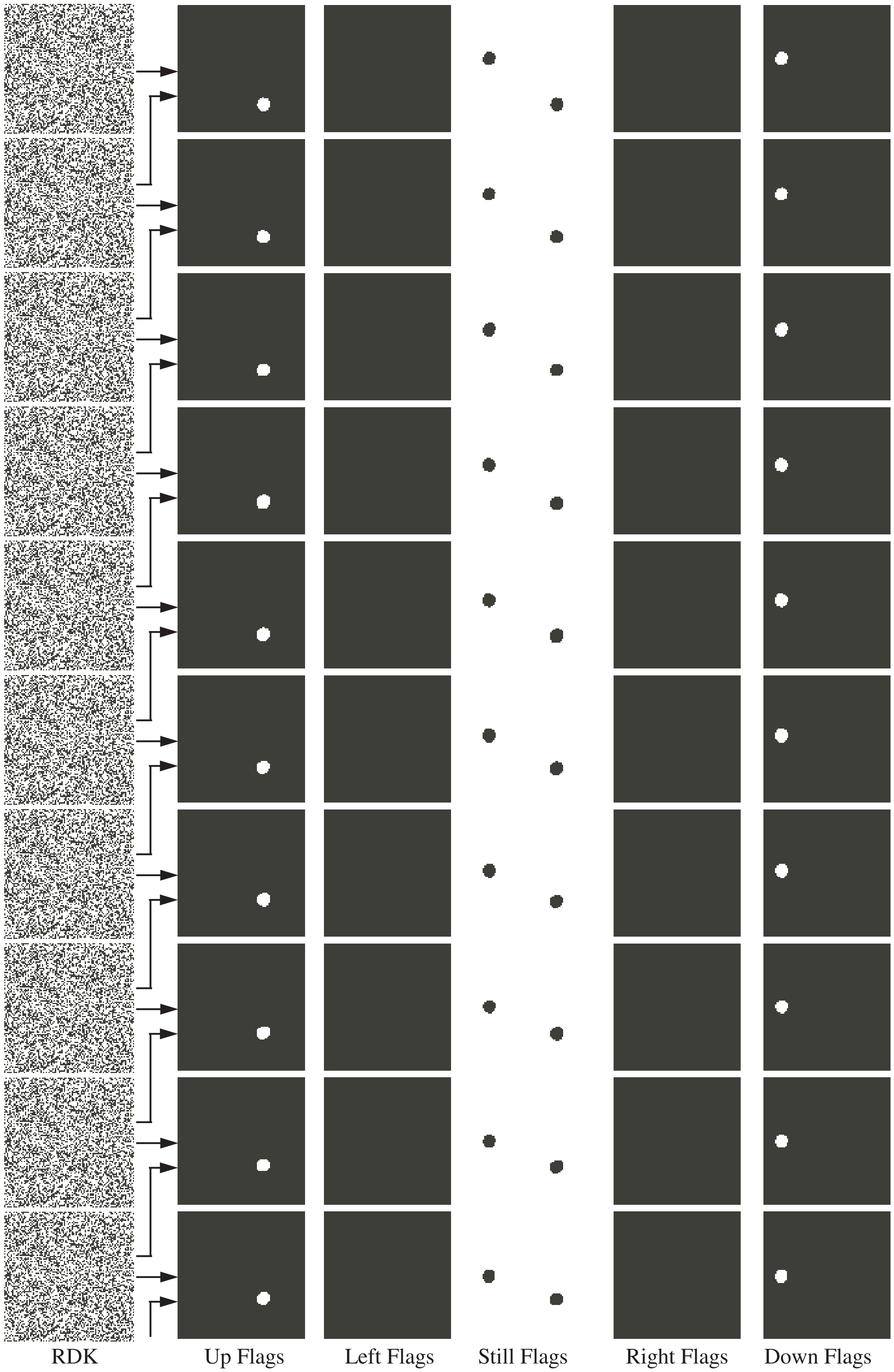


Fig. 2 Simulation result

Here, $\nabla$ is the spatial differential operator. The steady state of the dynamics (2) corresponds to the minimum of this potential function. In other words, the figure-ground separation is a minimization process of the functional (4) in this system.

### 3.4 Initialization

The initial value of the above dynamics is determined in such a way that the pixel values in two consecutive frame images are compared. Note that the initialization process should be also executable by the local parallel operation. Thus, the flag values are initialized based on only the local information of the neighboring pixels as follows.

$$^nf_1(x,y,0) = (^{n-1}p(x,y) \cdot^n p(x,y-1))_{D(x,y)} \quad (5)$$

$$^nf_2(x,y,0) = (^{n-1}p(x,y) \cdot^n p(x-1,y))_{D(x,y)} \quad (6)$$

$$^nf_3(x,y,0) = (^{n-1}p(x,y) \cdot^n p(x,y))_{D(x,y)} \quad (7)$$

$$^nf_4(x,y,0) = (^{n-1}p(x,y) \cdot^n p(x+1,y))_{D(x,y)} \quad (8)$$

$$^nf_5(x,y,0) = (^{n-1}p(x,y) \cdot^n p(x,y+1))_{D(x,y)} \quad (9)$$

$$(f \cdot g)_{D(x,y)} = \frac{\langle f \cdot g \rangle_{D(x,y)}}{\langle f \cdot f \rangle_{D(x,y)} \langle g \cdot g \rangle_{D(x,y)}} \quad (10)$$

$$\langle f \cdot g \rangle_{D(x,y)} = \iint_{D(x,y)} f(\xi,\eta) g(\xi,\eta) d\xi d\eta \quad (11)$$

where $D(x,y)$ denotes the neighbor of the coordinate $(x,y)$. The initialization is equivalent to the correlation between the current local image and the 1-dot-shifted local image at the next frame. This initialization provides the information of the moving direction, i.e., optical flow, of each pixel in a parallel operating manner.

## 4. SIMULATION

Using the RDK as shown in the left side of Fig. 2, we examine the system performance for the figure-ground separation that is proposed in the above section. This RDK consists of 20 frames, each of which has 100x100 dot in size. Although not be observed in the still frame image, two square areas are moving. One moves from the bottom to top in the right side, while the second one moves downward in the left side. This RDK satisfies the assumptions in section 3.1

In the simulation, the 4th order Runge-Kutta algorithm is used by 300 iterations with 0.001 step size. $\tau$ is set to 100, and $K_D$ is set to 100 for the first 0.1 [s] and afterwards kept to 0. The area $D(x,y)$ for initializing the flag values is defined as $3 \times 3$ area that contains the coordinate (x,y) in the center. To solve the partial differential equations, periodic boundary conditions are adopted to keep the sum of flag values.

The dynamics is computed in each frame, so 19 output is obtained for each flag plane. The first 10 outputs are shown in Fig. 2 from the second to fifth row using a flag plane representation. Because the large background does not move in this RDK, the still flags are almost set (white) except two square areas. The square area in the left side moves upward, so the up flags are set only in the corresponding area. The one in the right side moving downwards makes down flags set in this area. No areas move to the left and right sides, thus the all the flags in the left and right flag plan are reset (black). As shown in Fig. 2, almost correct separation are achieved.

## 5. CONCLUSION

In this paper, we consider a figure-ground separation in RDK from the view point of the local parallel computations. For the realization, the optical flows are restricted to the five vectors. The recognition process is defined as a system described by the reaction diffusion equation. This dynamics is represented as the gradient system of the potential functional, implying that this dynamical process is a minimization of the sole evaluation function. As future works, we will try to remove some assumptions such as no rotation or scaling, and extend this algorithm to more general conditions.


## ACKNOWLEDGEMENT

We thank Mr. Takashi Arakawa for helping us to make computer simulations.


## REFERENCES


[1] Y. Hirai, *Information processing in vision and memories*, Baifukan, 1995

[2] Hideo Yuasa, Satoshi Ito, Koji Ito, Masami Ito : "Associative memory with the reaction - diffusion equation", *Biological Cybernetics*, Vol. 76, No. 2, pp. 129–137, 1997

[3] Hideo Yuasa: "Autonomous Decentralized Image processing for Binocular Stereogram – inferring 3D structure from stereograms", *Mathematical Science*, Vol. 414, pp. 77– 83, 1997

[4] Eizo Ueyama, Hideo Yuasa, Shigeyuki Hosoe, Masami Ito: "Figure-Ground Separation for Dynamic Image with Ginzburg-Landau Equation by Use of Local Clustering", *Trans. of SICE*, Vol. 32. No. 11, pp. 1544-1552, 1996

[5] Eizo Ueyama: "Figure-Ground Separation for Time-Varying Images by Reaction Diffusion Advection Models", *Doctoral thesis*, Nagoya University, 2000

[6] Akihito Okura, Hideo Yuasa, Tamio Arai: "Autonomous Decentralized 3D Recognition System for Moving Object", *Proc. of the 19th Annual Conference of RSJ*, pp. 853-854, 2001

[7] H. Haken: *Synagetics, An introduction*, Springer-Verlag, 1978