競合・協調型神経回路によるカラー画像の立体認識に関する研究*

華 希俊[†]・横道 政裕[‡]・河野 通夫[‡]

Stereo Vision from Color Images Based on Competitive and Cooperative Neural Networks*

Xijun HUA[†], Masahiro YOKOMICHI[‡] and Michio KONO[‡]

Stereo vision is a passive method of obtaining depth information of visible surface by measuring disparity of corresponding points in left and right stereo images. Stereo matching, which belongs to ill-posed problems, is the key issue of stereo vision. In the literature, almost all matching techniques only use gray-scale images. The color information is usually neglected, but undoubtedly it may be of great use in helping robot stereo systems perform better. In this paper, we use two layered self-organization neural network model to simulate the competitive and cooperative interaction of binocular neurons. We propose a special similarity function as initial input to make full use of the color information. In RGB color space, the similarity map is established by taking logical AND calculation of red, green and blue color value similarity. In stereo matching experiment, we first consider color random-dot stereogram for stereo correspondence. Then we take real image experiments by means of different color stereo matching approaches. Experiments results have shown that the quality and convergence speed of the stereo matching can be efficiently improved by using appropriate color matching algorithm comparing with the conventional grey value algorithm.

1. Introduction

Stereo vision is a well-known technique for obtaining depth information from stereo digital images. As one of the main components of computer vision, stereo vision is a process to reconstruct 3D structure of visible surface by measuring disparity of two corresponding points in a stereo image pair. It is a direct and passive method of obtaining depth information of the visual world which makes it attractive for applications like 3D reconstruction, object recognition and navigation[1].

Stereo vision is composed of the following steps: image acquisition, camera calibration, stereo matching, triangulation and surface reconstruction. The depth of scene is determined by the baseline distance,

focal length and disparity. Since the baseline distance and focal length are geometrically determined by a vision system, the Stereo matching, or stereo correspondence problem, which is to identify corresponding points in two or more images of the same scene, has become the most important and difficult issue of stereo vision and it is also an ill-posed problem. So far, various correspondence algorithms such as feature-based, area-based, phase-based, energy based, multiple-baseline and data fusion have been studied[2].

There are some limits for existing approaches to recover precise and dense disparity map. Feature-based stereo can produce more precise matching quickly but only sparse disparity map, leaving the rest of the surface to be reconstructed by interpolation. On the other hand, area-based approaches can provide dense disparity map but less precise matching. It is also very sensitive to the factors such as noise, lighting variation.

For an ideal reconstruction of complex scene it is essential to obtain dense disparity maps defined for every pixel in the entire image. Therefore, the neural network techniques to conventional stereo vision problems have received increasing attention in recent years. Neurons that respond selectively to binocu-

* 原稿受付 2003年12月1日

[†] 宮崎大学大学院工学研究科 システム工学専攻 Graduate School of Engineering, University of Miyazaki; 1-1, Gakuen Kibana Dai Nishi, Miyazaki 889-2192, JAPAN

[‡] 宮崎大学工学部 情報システム工学科 Department of Computer Science and Systems of Engineering, Faculty of Engineering, University of Miyazaki; 1-1, Gakuen Kibana Dai Nishi, Miyazaki 889-2192, JAPAN

Key Words: stereo vision, stereo correspondence, RGB color space, neural network, self-organization.

lar disparity were first described three decades ago, and have subsequently been observed in similar numbers in many primate visual areas including V1, V2, V3, MT and MST[3]. The importance of competitive and cooperative neural network was shown through interactions mediated by excitatory and inhibitory axon collateral pathways in rat visual cortical slice[4]. By means of appropriate neural network models, it has been demonstrated that the calculation of stereo disparities can be achieved through cooperation and inhibition among disparity-selective neurons. Amari and Arbib [5] proposed a famous neural network model for the stereo matching. A neural network model for solution of the correspondence problem was proposed by Marr, which is composed of groups of neurons selectively responsive to different disparities of visual stimuli shown to the two eyes[6-7]. Poggio [8] considered the stereo matching as a kind of optimization and solved the ill-posed problem by means of regularization theory. Reimann and Haken [9] proposed a new mathematical model to solve the correspondence problem by a dynamical self-organizing process. A feature-based technique based on three-layer neural network was proposed for classifying true and false matches[10]. Kitazoe et al. [11] researched sequential stereo vision with hysteresis in competition and cooperation neural network. On the other hand, Achour and Mahiddine proposed a 2D Hopfield network based optimization method to solve correspondence problem[12].

In our past work, we have used the intensity information of the stereo images for stereo matching[13]. Usually, in stereo system, the digital cameras capture color images, which are afterwards translated into intensity (grey value) images for further processing with a drastic reduction of information. In the literature, most of the stereo matching algorithms have been limited to grey level images, only few authors have dealt with color images. Actually, in the basic color space(RGB), a color pixel is multi-valued (3 bytes, one for each band), while in the associated grey value image the information content is single-valued (1 byte). Color images can provide more useful information than grey value images for stereo matching since red pixels cannot match with blue pixels although their intensities are similar. The color information is usually neglected, but undoubtedly it may be used to make the stereo matching perform better[14]. Recently, several authors try to utilize color information (such as the sum of absolute RGB differences, the Euclid distance between two colors, color segmentation) of color stereo images to reduce the ambiguity between potential matches. The matching quality based on color stereo matching can be considerably improved when using adaptive color algorithm instead of the grey value algorithm. It has been reported that the precision of matching results can be improved by 20 to 25%[15].

The goal of this paper is to explore an efficient

color stereo matching method for obtaining a reliable and accurate dense disparity map. We utilize pixel-based matching method based on self-organization neural network to calculate dense disparity map. In order to improve the accuracy of disparity map we proposed a new color stereo matching algorithm to deal with color images straightly. The paper is organized as follow: In section 2, we introduce two layered self-organization neural network model and define a special similarity function to establish R, G, B similarity map. In section 3, Effectiveness of the proposed algorithm is examined by experiments with random color stereogram. In section 4, a segmentation method is proposed to deal with the input similarity maps of real images. We also compared the RGB color matching method with single grey value matching method and other color matching methods.

2. Neural network model by self-organization

The neural network model is a competitive, cooperative network shown as Fig. 1. a represents range of disparity, and u represents pixel position. To restrict our search, we define a search area for a possible match. Each possible disparity in this area corresponds to a binocular neuron. The initial values of the evolution of the activity of binocular neurons are given by an appropriate similarity measure. The binocular neurons compete over the disparity search area and cooperate over the cooperation area in such a manner that in the course of time just one binocular neuron can win; all others end with zero, therefore, resulting in a stereo disparity perception. Based on the neural network equations, for each image pixel we obtain a unique disparity (uniqueness constraint), the cooperative coupling of binocular neurons of the same disparity is a local formulation of the continuity constraint with the property that it allows for discontinuities to appear.

Fig. 1 Neural network with two layers (1D model)

The neural network equations are written as eqs. (1) - (7). α_{ab}^{uv} is membrane potential of neuron. ξ_{ab}^{uv} is output membrane potential of the neuron. A , B and C are positive constants which have to be chosen

appropriately.

$$\begin{aligned} \tau_1 \dot{\alpha}_{ab}^{uv} = & -\alpha_{ab}^{uv} + A\lambda_{ab}^{uv} - B \sum_{a' \neq a} g(\xi_{a'b}^{uv}(t)) \\ & + C \sum_{u' \in U_u, v' \in V_v} g(\xi_{ab}^{u'v'}(t)), \end{aligned} \quad (1)$$

where U_u and V_v denote certain cooperation area of u and v respectively.

$$\tau_2 \dot{\xi}_{ab}^{uv}(t) = -\xi_{ab}^{uv}(t) + f(\alpha_{ab}^{uv}) \quad (2)$$

$$\lambda_{ab}^{uv}(R) = 1 - |R_l(u, v) - R_r(u + a, v + b)| \quad (3)$$

$$\lambda_{ab}^{uv}(G) = 1 - |G_l(u, v) - G_r(u + a, v + b)| \quad (4)$$

$$\lambda_{ab}^{uv}(B) = 1 - |B_l(u, v) - B_r(u + a, v + b)|. \quad (5)$$

Here $\tau_1 \ll \tau_2$, $g(x)$ is a function given by

$$g(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (6)$$

and $f(x)$ is a sigmoid function given by

$$f(x) = \frac{\text{th}(x) + 1}{2}. \quad (7)$$

In eq. (1), a and b stands for the epipolar and the perpendicular disparity respectively. They run over the disparity search area. The input data λ_{ab}^{uv} are obtained as the product of R(red), G(green) and B(blue) similarity function (3), (4), (5). $R(u, v), G(u, v), B(u, v)$ represent the color intensity value of pixels respectively. l and r stand for left and right image respectively. α_{ab}^{uv} has inhibitory connection for a competition (the third term) and excitatory connection for a cooperation (the fourth term). For simplicity, here we consider the case where the two eyes are located horizontally so that we can neglect “ b ” in eqs. (1) - (5).

3. Stereo matching for random color stereogram

The image pair is composed of three color random stereograms (Red, Green and Blue). Each band of the color stereogram is generated in a manner similar to the conventional random dot stereogram. Therefore, the color of each dot in the stereogram is randomly chosen in eight kinds of color. The random stereogram is composed of 20 scan-lines. Every scan line is composed of 80 random dots, in which from 25th to 64th dots, the relative disparity to the background is given as 5. The R, G, B similarity map are shown in Figs. 2, 3, 4, respectively. Here, the coordinate U stands for pixel position and A stands for disparity. The final AND similarity map is shown as Fig. 5. In Fig. 5, the potential corresponding points with high similarity value become much fewer than that in the R, G, B similarity map (Figs. 2-4).

From Fig. 6, a satisfactory disparity map is obtained, since the advantages of multi-spectral images is taken so as to reduce false matches efficiently. Here, the coordinate V stands for different scan-lines. Thou-

Fig. 2 Similarity map (Red)

Fig. 3 Similarity map (Green)

Fig. 4 Similarity map (Blue)

gh the images are synthetic and free of noise, there are some mismatches at the left corner due to occlusion. In the numerical computation, the differential eq. (1) and (2) are discretized. The notation “Loop” in the caption of Fig. 6 denotes the iteration time of the Euler scheme in solving eqs. (1)-(2).

When we use the conventional grey value matching algorithm to deal with the grey value images, its similarity map is shown in Fig. 7. As shown in its disparity map (Fig. 8), the mismatching points are caused by both occlusions and false targets in the boundary region. In addition, the convergence

Fig. 5 Similarity map (AND)

Fig. 6 Disparity map (AND, Loop=70)

Fig. 7 Similarity map (gray value)

speed of the color matching is fast than the grey value matching.

4. Stereo matching for real color images

4.1 Experimental method

For simplicity, here we only consider the parallel cameras case, in which corresponding points in the two images lie on the same horizontal epipolar lines. Therefore, it is necessary to keep two camera axes

Fig. 8 Disparity map (gray value, Loop=95)

parallel and the baseline of stereo camera system is at a right angle to the camera axes. In this case, the situation $a > 0$ will not occur.

In RGB color space, the R, G, B color images (8bits) are taken by CCD camera with 640×480 pixels. For simplicity the image size is compressed into 320×240 pixels. The element values of similarity map vary in the range $[0, 1]$ because the intensity values of different pixels is always varying in the range $[0, 255]$. In the other hand, there are many factors such as lighting condition, camera setting, discretization and quantization effects, and signal noise, which will make precise correspondence computation more difficult than the matching of random stereogram. There are always a little intensity variation caused by system error and random error between the two corresponding pixels. Therefore we propose a segmentation method to deal with the similarity map. A threshold function, given as formula (8), for filtering the similarity map is proposed so as to get binary similarity map for the next logical AND calculation

$$\lambda = \begin{cases} 1, & \lambda_i > S \\ 0, & \lambda_i \leq S. \end{cases} \quad S: \text{Threshold gate} \quad (8)$$

4.2 Color stereo matching

The real image is shown as Fig. 9. In the visual scene, two layer books are placed on the table. We expect to perceive the disparity of different layer books with our methods.

Fig. 9 Real image

For simplicity we select the middle area (120×20) of real image to discuss. The RGB AND similarity map is shown in Fig. 10. From Fig. 11, by using

chromatic similarity map, we can obtain satisfactory disparity map in which the ambiguity between potential matches can be obviously reduced. Here, the parameters of neural networks equations are chosen as below: $A = 1.0$, $B = 4.0$, $C = 0.7$.

Fig. 10 Similarity map (AND, $S=0.9$)

Fig. 11 Disparity map (AND, $S=0.9$, Loop=90)

We also consider the affecting of parameter S in the stereo matching process. In the formula (8), the parameter S need to be chosen appropriately in order to get a satisfactory similarity map. S is expected as big as possible theoretically but it will be affected by many factors in the two real images. In this experiment, when S is chosen as 0.8, the similarity map is shown in Fig. 12, in which many false targets emerge. Its disparity map is shown in Fig. 13.

There are some points mismatched because S is too small. When S is chosen as big as 0.97, the similarity map is shown in Fig. 14 and its disparity map is shown in Fig. 15. There are many unmatched points appeared when S is too big. The experimental results show that if S is too large or too small, there will be more miss-correspondences appeared. Therefore, S has to be chosen appropriately according to the actual real images.

Fig. 12 Similarity map (AND, $S=0.8$)

Fig. 13 Disparity map (AND, $S=0.8$, Loop=200)

Fig. 14 Similarity map (AND, $S=0.97$)

4.3 Comparison with single grey value algorithm

When we used the single grey value of HSV color space for the stereo matching, the similarity map is shown in Fig. 16. Its disparity map is given in Fig. 17, in which there are more mismatches appeared than in the Fig. 11. In addition, we also found that in the matching of real images, the convergence speed of color matching method is much faster than grey value matching method.

Fig. 15 Disparity map (AND, $S=0.97$, Loop=90) Fig. 16 Similarity map (gray value, $S=0.9$) Fig. 17 Disparity map (gray value, $S=0.9$, Loop=150)

4.4 Comparison with other color matching algorithm

The real image is given in Fig. 18. There are three layer books placed on the table. The RGB AND similarity map and disparity map are shown in Fig. 19 and Fig. 20, respectively. When using grey value matching method, the similarity map and disparity map are shown in Fig. 21 and Fig. 22, respectively. In the similarity map, S is chosen to keep the disparity variation of perceived surface continuous. The experimental re-

Fig. 18 Real image

Fig. 19 Similarity map (AND, $S=0.84$) Fig. 20 Disparity map (AND, $S=0.84$, Loop=90)

sult is similar to the above example in the section 4.2 and 4.3.

When we consider the R, G, B color space, the R, G, B coordinates of a multi-spectral pixel are defined as

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B} \text{ and } b = \frac{B}{R+G+B}$$

We can obtain a AND similarity map of R, G , and B with the same approach as the former method, as shown in Fig. 23. Its disparity map is shown as Fig. 24. There are very little system error of R, G, B between left and right images due to the relativity of R, G, B . Therefore, S can be chosen as high as 0.98. However, because the R, G, B values in the most pixels are very similar to each other, there are many pixels mismatched, as shown in Fig. 24.

Fig. 21 Similarity map (gray value, $S=0.84$) Fig. 22 Disparity map (gray value, $S=0.84$, Loop=150) Fig. 23 Similarity map (AND of R, G, B , $S=0.98$)

When we consider the Ohta color space[16], the three coordinates I_1 , I_2 and I_3 are defined as

$$I_1 = \frac{R+G+B}{3}, I_2 = \frac{R-B}{2} \text{ and } I_3 = \frac{2G-R-B}{4}$$

The AND similarity map is shown as Fig. 25. Its disparity map is shown as Fig. 26. The quality of disparity map is similar to Fig. 20. Therefore, high quality of stereo matching can be efficiently obtained by appropriate color stereo matching algorithm.

Fig. 24 Disparity map (AND of R, G, B , $S=0.98$, Loop=90) Fig. 25 Similarity map (AND of I_1 , I_2 and I_3 , $S_1=0.84$, $S_{2,3}=0.95$) Fig. 26 Disparity map (AND of I_1 , I_2 and I_3 , $S_1=0.84$, $S_{2,3}=0.95$, Loop=90)

4.5 Other example

In order to test the effectiveness of our matching method further, we also take a normal natural scene to discuss. Here, an inclined color wall is perceived as shown in Fig. 27. The disparity maps are obtained by single gray value matching method and color matching method, as shown in Fig. 28 and Fig. 29 respectively. The experiment results show that the use of

Fig. 27 Real image

Fig. 28 Disparity map (gray value, $S=0.93$, Loop=250) Fig. 29 Disparity map (AND of R, G, B , $S=0.93$, Loop=100)

color information can efficiently reduce the ambiguity between potential matches, resulting in increased accuracy of dense disparity map.

5. Conclusion

In this paper, we use two layered self-organization neural network model to simulate the competitive and cooperative interaction of binocular neurons. For the neural network we define a special similarity function by taking logical AND calculation of red, green and blue color value similarity. Experiments with both synthetic random-dot stereogram and real images have showed that the disparity quality as well as the convergence speed can be efficiently improved by using appropriate color stereo matching technique comparing with the conventional grey value algorithm.

For the real images, it is necessary to use the segmentation method to deal with the similarity maps for the logical AND calculation. Because the chromatic intensity of each pixel is very sensitive to many factor, in the future work we will continue to improve our approach and investigate more reliable stereo matching approaches based on appropriate color matching features of stereo images.

References

- [1] G. N. Desouza and A. C. Kak: Vision for mobile robot navigation: a survey; *IEEE Trans Pattern Anal. Machine Intell*; Vol. 24, No. 2, pp. 237–267 (2002)
- [2] D. Scharstein and R. Szeliski: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms; *Int. Journal of Computer Vision*, Vol. 47, No. 1/2/3, pp. 7–42 (2002)
- [3] M. S. Gazzaniga: *The New Cognitive Neurosciences*, The MIT Press, pp. 263–277 (2000)
- [4] K. Toyama and M. Tanifuji: Imaging a computational process in the visual cortex; *Neural Networks*, Vol. 9, No. 8, pp. 1351–1356 (1996)
- [5] S. Amari and M. A. Arbib: Competition and cooperation in neural nets; *Systems Neuroscience*, Academic Press, pp. 119–165 (1977)
- [6] D. Marr and T. Poggio: Cooperative computation of stereo disparity; *Science*, Vol. 194, pp. 283–287 (1976)
- [7] D. Marr and T. Poggio: A computational theory of human stereo vision; *Proc. R. Soc. Lond. B.* 204, pp. 301–328 (1979)
- [8] T. Poggio, V. Torre and C. Koch: Computational vision and regularization theory; *Nature*, Vol. 317, No. 26, pp. 314–320 (1985)
- [9] D. Reiman and H. Haken: Stereo vision by self-organization; *Biological Cybernetics*, Springer-Verlag, pp. 17–26 (1994)
- [10] M. C. Jesus, P. Gonzalo and A. Joaquin: A neural network model in stereo matching; *Neural Networks*, Vol. 8, No. 5, pp. 805–813 (1995)
- [11] T. Kitazoe, J. Tomiyama and Y. Yoshitomi: Sequential stereoscopic vision and hysteresis; *Proceeding of 5th International Conference on Neural Information Processing*, pp. 391–396 (1998)
- [12] K. Achour and L. Mahiddine: Hopfield neural network based stereo matching algorithm; *Journal of Mathematical Imaging and Vision*, Vol. 16, pp. 17–29 (2002)
- [13] X. Hua, Y. Tang, M. Yokomichi et al: 3-D perception for monochromatic surface by self-organization neural network; *Proc. of the 8th International Symposium on Artificial Life and Robotics*, pp. 479–483 (2003)
- [14] S. Taraglio and A. Zanela: Improving a real-time neural-based stereo vision system; *Real-Time Imaging*, No. 7, pp. 59–76 (2001)
- [15] A. Koschan, V. Rodehorst and K. Spiller: Color stereo vision using hierarchical block matching and active color illumination; *Proc. of 13th International Conference on Pattern Recognition ICPR'96*. Vol. I, pp. 835–839 (1996)

- [16] Y. Ohta, T. Kanade and T. Sakai: Color information for region segmentation; *Computer Graphics and Image Processing*, Vol. 13, No. 2, pp. 222-241 (1980)
-