Autonomous Stereo Object Tracking using BMA and JTC

Jae-Soo Lee, Jung-Hwan Ko and Eun-Soo Kim
Dept. of Electronics, Kwangwoon University, #447-1, Wolge-Dong, Nowon-Ku, Seoul, Korea

Abstract

General stereo vision system shows things in 3D, using two visions of left and right side. When the viewpoints of left/right sides are not in accord with each other, it gives fatigue to human eyes and prevents them from having the 3-D feeling. Also, it would be difficult to track mobile objects that are not in the middle of a screen. Therefore, the object tracking function of stereo vision system is to control tracking objects to always be in the middle of a screen while controlling convergence angles of mobile objects in the input image of the left/right cameras. In this paper, object-tracker in stereo vision system is presented which would track mobile objects by using block matching algorithm of preprocessing and JTC.

1. Introduction

Generally, for the method of the 2-D object tracking, the Centroid method, which tracks object by using the center of gravity on the inputting screen, and the Correlation method, which gets the degree of accordance through the continuously inputted images' correlation, are used. However, because its efficiency depends tremendously on the observing object, the degree of separation of the background and the condition of the background, there may be a lot of limitations if background exists. Thus, in this paper, a new, approaching method to track an object with the background in the stereo system was presented. The stereo object tracking system, which separates the background of a frame through Motion Estimation and controls both the camera convergence angles and the pan/tilt function at the same time by activating the JTC on two frames and by getting the inputting relative locations of the moving objects with the left/right cameras, will be proposed and presented.

2. Stereo object tracking

The dynamic stereo tracking can be very efficient if the stereo object tracker is formed by controlling the tracking object to be centralized on the frame, as controlling the convergence angles of a moving object. Fig.1 shows the process of the stereo vision system's object tracking when the moving object inputted by the left/right cameras

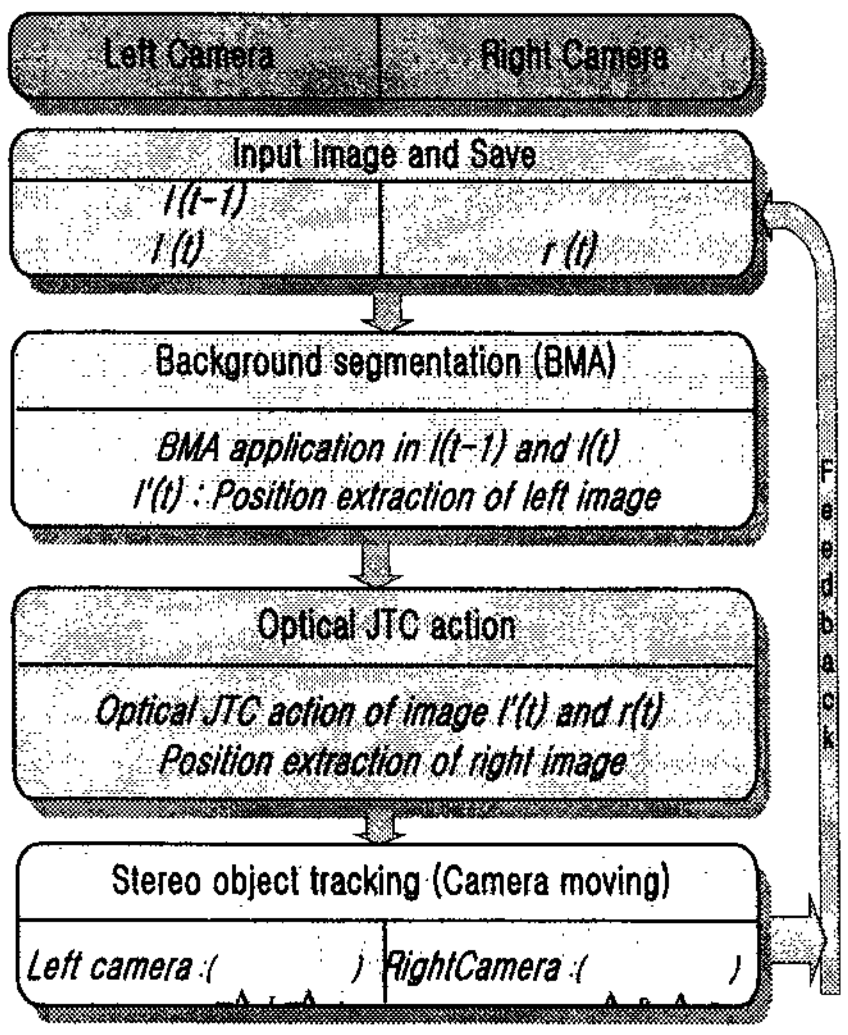
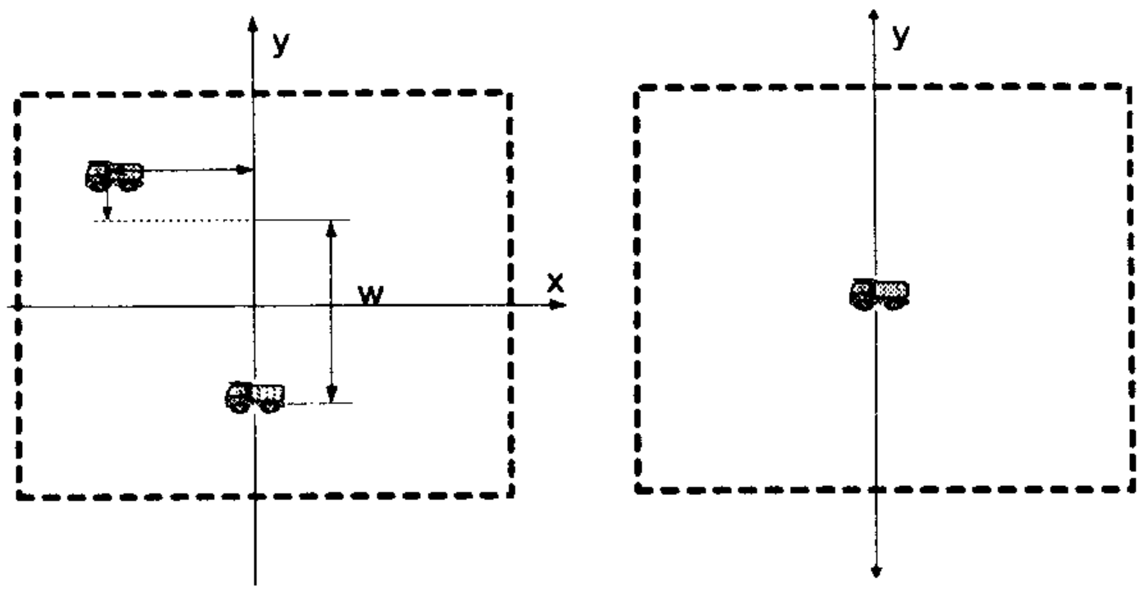


Fig.1.Flowchart of stereo object tracking

In Fig.1, l (t-1) indicates the left frame before inputted during the (t-1) time, and l (t) and r (t) indicates the left/right, current frames of the moved object during t time. Here, because l (t-1) and l (t) frames are captured without any move of the camera, they have the identical background.

In two frames, only the object, which has the moving component, can be found when non-moving block, the background, is removed. Also, the location coordination of an object can be known and using the center can move the object. Therefore, the location coordinates of the object $l_t(x-x_{Lt}, y-y_{Lt})$, (x_{Lt}, y_{Lt}) is obtained when the left frame of t time apart from the center. l' (t) is the centralized frame, of which an object of the left frame moved $(-x_{Lt}, -y_{Lt})$. of Fig.2 is the left inputted frame, l (t), and (b) is the $l'(t)$ in which the object of $l(t)$ is centralized



(a) **(b)**

Fig. 2. l(t) image and l'(t) image of left camera (a) l(t) image, (b) l'(t) image

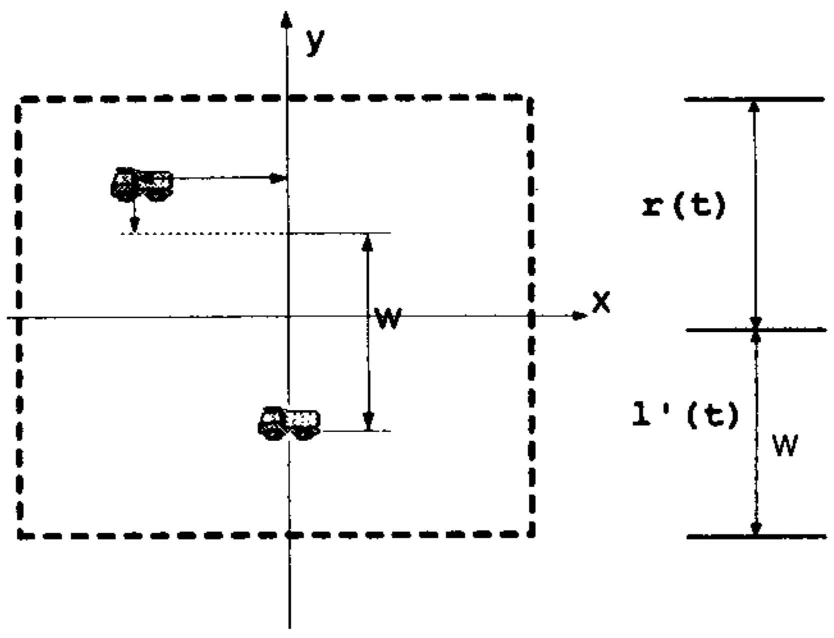


Fig. 3. Input plane of optical JTC system

In Fig.3, the upper half plane is the reference plane and the lower half plane is the comparing plane, r (t). And these two frames are placed at the same time and they are to be correlated.

By using this input plane, if the JTC was executed, the location of a moving object of the right frame, from the frame's center can be obtained in the Correlation plane. In Fig.3, inputted left frame, l' (t), and right frame, r (t), can be represented as $l_t'(x, y-w/2)$ and $r_t(x-x_{Rt}, y-y_{Rt}+w/2)$, respectively, where $2w$ is the height of the display unit (LCD), and (x_{Rt}, y_{Rt}) is the moved distance of the right image object at t time.

$$E_{JTC} = |L_t'(u,v)+R_t(u,v)|^2 \quad (1)$$
$$= |L'_t(u,v)|^2 + |R_t(u,v)|^2 + R_t^*(u,v)L'_t(u,v) + R_t(u,v)L'^*_t(u,v)$$

Eq. (1) shows the JTPS (joint transform power spectrum), the optical interference intensity distribution which is obtained by Fourier transform the reference plane frame, l' (t), and the comparing plane frame, r (t).

where, (u, v) represents the spatial frequency coordinates, and the asterisk denotes the complex conjugate. In Eq. (1), L_t' (u, v) and R_t (u, v) denotes the Fourier transform of the reference image, l_t', and the current image, r_t, respectively. The Inverse Fourier transform of Eq. (1) can be expressed as Eq. (2). In Eq. (2), the first and the second terms represent the components of self-interference, and they can be removed digitally by reconstructing the JTPS.

$$C_{JTC}(x,y) = F^{-1}\{E_{JTC}(u,y)\}$$
$$= l'_t(x,y) \otimes l'_t(x,y) + r_t(x,y) \otimes r_t(x,y)$$
$$+ l'_t(x,y) \otimes r_t(x,y) * \delta[x + \Delta x_R, y + \Delta y_R - \omega] \quad (2)$$
$$+ r_t(x,y) \otimes l'_t(x,y) * \delta[x - \Delta x_R, y - \Delta y_R + \omega]$$

The third and the fourth terms represent the cross-correlation of two frames and correlating peak points appear as the origin symmetry in the coordinate, [x_{Rt}, (y_{Rt}-w)], the location coordinate, (x_{Rt}, y_{Rt}), of an moving object can be obtained.

3. Results of digital simulation

Fig. 4 shows one frames of stereo input image from the left/right cameras. l' (t) is a centralized left image which removed the background by the motion estimation on the left camera's two images, [l(t-1), l(t)]. In this process, the moved distance (x_{Lt}, y_{Lt}) of an object moved from the center in the left camera can be obtained. r (t) is the right image that was inputted at t time.

Fig.5 has the images for executing the JTC. l' (t) and r(t) frames are used as the JTC Input Plane. As a result of executing the JTC, the correlation peak of Fig.5 shows the values, which appear on the Correlation plane. By using these values, the distance (x_{Rt}, y_{Rt}) that an object moved from the middle in the right frame, can be obtained.

Fig.6 shows the combined images, both before the tracking and after the tracking, from the Fig.11's inputted images. Like shown in Fig.6, the combined image before the tracking gives fatigue to human eyes when it is seen in 3-D, because two images are appeared overlapped. Also, since the object is not in the middle, it is difficult to track it. On the other hand, because in the combined image after the tracking, two images appear as one when they are seen in 3-D, the human eyes will not get tire. And, since the object is placed in the middle of a screen, it is possible for the stereo object tracking.

4. Conclusion

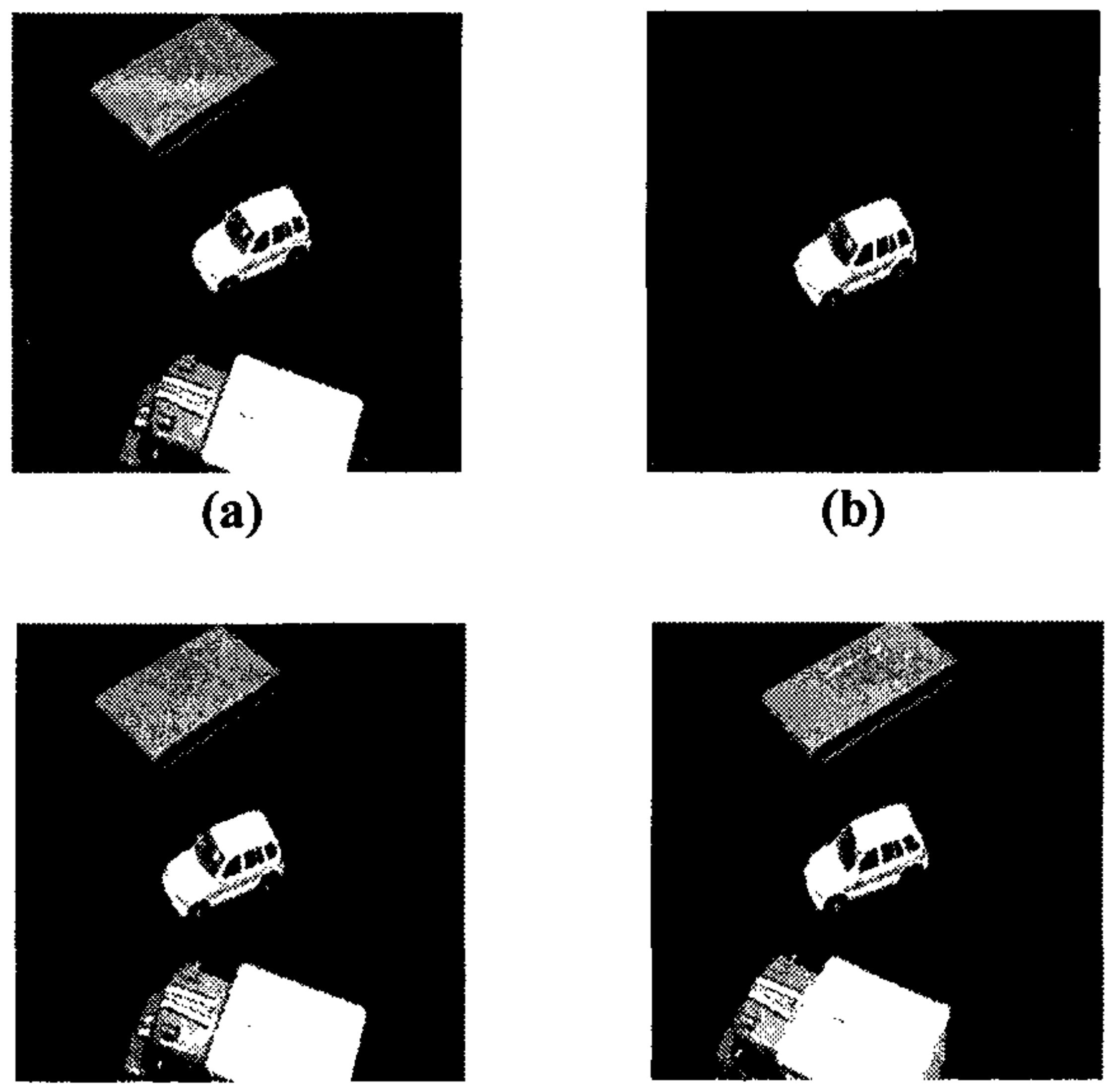
In this paper, the stereo object tracking system within stereo vision system, which uses the motion estimation and JTC as a method to track an object where the background exists, was demonstrated. As a result, by adapting, it was possible to track the exact location of a moving object with the stereo, and the background noise was not an affecting factor. Also, the formation of the stereo vision system's object traker was possible through controlling the stereo camera's convergence angles and the pan/tilt function by using the value obtained above, and this paper proposed the materialization of actual-time stereo vision's object tracking system in case of using the optical JTC.

Reference

[1]T. J. Olson and D. J. Coombs, "Real-Time Vergence Control for Binocular Robots," Intl. J. of Computer Vision, Vol.7, No.1, pp.67-89, 1991

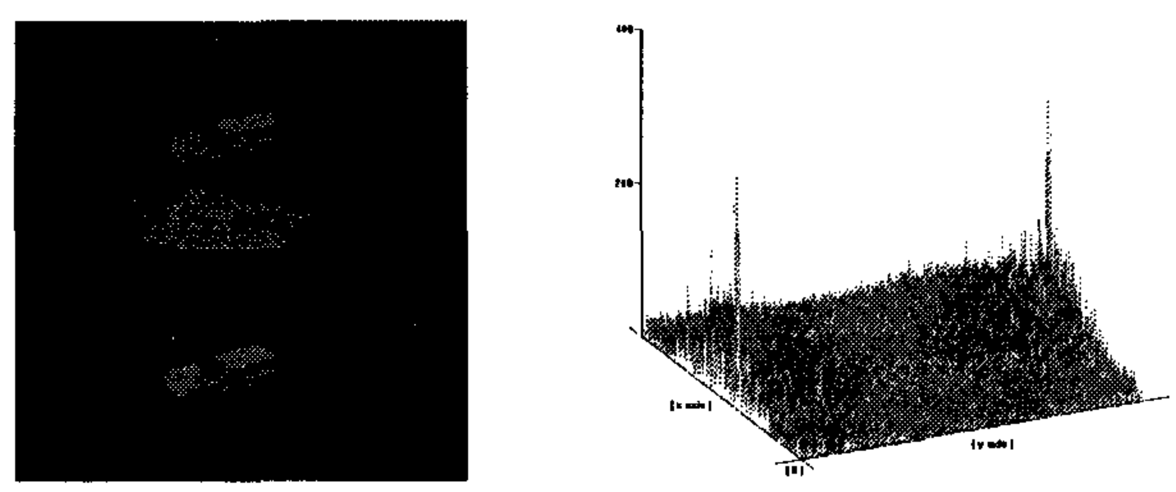
[2]David Drascic, "An Investigation of Monosopic and Stereoscopic Video for Teleopration," *MASc Thesis, Univ. of Toronto, April* 1991

[3]Stephen T. Barnard, and william B. Thompson, "Disparity Analysis of Images", IEEE Transaction on Pattern Analysis and Machine Intelligence , Vol PAMI-2, No.4, pp.333-340, July, 1980



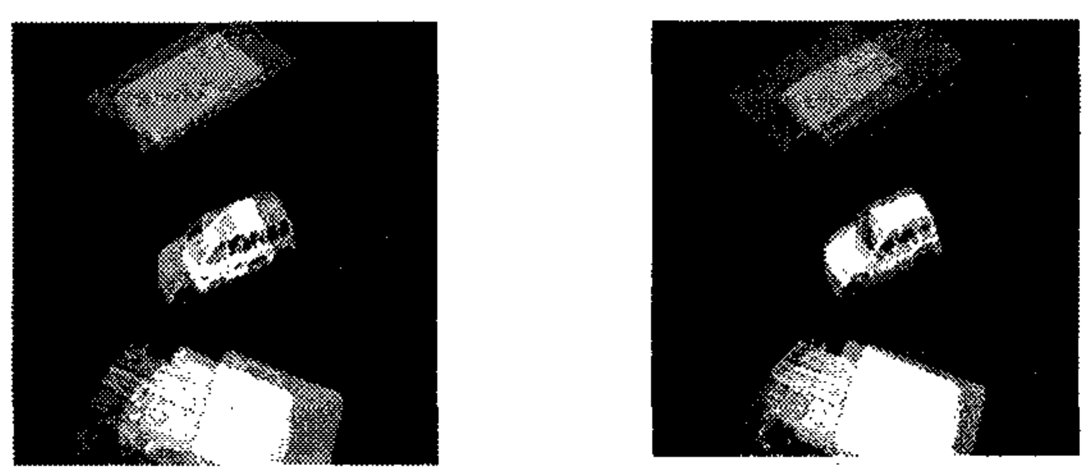
(a) (b) (c) (d)

Fig. 4. Stereo input image (1 frame)
(a) l(t-1), (b) l'(t+1), (c) l(t), (d) r(t)



(a) (b)

Fig.5. Simulation result of optical JTC
(a) Input plane of JTC, (b) Correlation Peak



(a) (b)

Fig.6. Simulation tracking results for moving object
(a) Composition image before tracking, (b) Composition image after tracking