Efficient Tracking of a Moving Object Using Representative Blocks Algorithm

Sung-Yug Choi * , Hwa-Ra Hur** and Jang-Myung Lee***

* Research Inst. of Computer Information & Communication, Pusan National University, Busan, Korea
(Tel : +82-51-510-2598; E-mail: sychoi3@korea.com)
**Department of Electrical Engineering, Songho College, Hoengseong, Korea
(Tel : +82-33-340-1030; E-mail: hrhur@songho.ac.kr)
*** Department of Electronics Engineering, Pusan National University, Busan, Korea
(Tel : +82-51-510-2598; E-mail: sychoi3@korea.com)

Abstract: In this paper, efficient tracking of a moving object using optimal representative blocks is implemented by a mobile robot with a pan-tilt camera. The key idea comes from the fact that when the image size of moving object is shrunk in an image frame according to the distance between the camera of mobile robot and the moving object, the tracking performance of a moving object can be improved by changing the size of representative blocks according to the object image size. Motion estimation using Edge Detection(ED) and Block-Matching Algorithm(BMA) is often used in the case of moving object tracking by vision sensors. However these methods often miss the real-time vision data since these schemes suffer from the heavy computational load. In this paper, the optimal representative block that can reduce a lot of data to be computed, is defined and optimized by changing the size of representative block according to the size of object in the image frame to improve the tracking performance. The proposed algorithm is verified experimentally by using a two degree-of-freedom active camera mounted on a mobile robot.

Keywords: Mobile robot , Moving Object**,** Representative Block

1. INTRODUCTION

In this paper, we suggest a new tracking method for moving objects in the part of mobile robots. In the case of real-time object tracking, ED (Edge Detection) and BMA (Block Matching Algorithm) have been used[1]. In the case of the ED method, it spends much time on recognizing the object because of computation complexity of every pixel. Specifically, in a convolution operation step, the system's memory spends too much time and then there are several cases that some real-time vision data are lost because of slow image processing. In the BMA method, motion estimation is calculated by using the block by block matching method[2-4]. This includes motion estimation by investigating the maximum correlation between a former frame (n-1) block data and a present frame (n) block data. MAD (Mean Absolute Difference) and MSD (Mean Squared Difference) are often used in the BMA[5][6]. The BMA method may have a better operation speed than the ED, but it has one drawback that local minima happen unexpectedly[1][7]. Moreover, in the BMA, a fixed camera is mainly used and block size is also fixed. Therefore, these systematic constraints cause low object-tracking performance for mobile robots.

The purpose of this paper is to improve the tracking performance in the above-mentioned applications using EA and BMA. The proposed method has two advantages in object tracking. One is the improvement of speed in the image processing and the other is improvement of the object recognition ability. These improvements can be acquired by RB (Representative Block) that is proposed in this paper, especially the variable size RB is unique in this paper. Using a 2-DOF (Degree Of Freedom) active camera mounted on a mobile robot, we verified the high tracking performance when using the variable size RB. Then, the advantages of the variable size RB are remarkably shown in the case that the various size of objects exist in an image plane.

This paper is composed of the following: Section 2 explains the BMA method whose performance is compared with the optimal RB, and section 3 includes the definition of the RB. The size decision of the RB is mentioned in section 4, which defines the optimal RB. Section 5 deals with experiments using the RB and shows the results. Finally, we summarize our ideas in section 6.

2. BLOCK MATCHING ALGORITHM (BMA)

Frames $n+1$ and n are often referred to as the present frame and previous frame, respectively. In the BMA method, the present frame of the sequence is divided into rectangular or square blocks of pixels. For each block in the current frame, we look for the block of pixels that is the closest to the block in the previous frame, according to the predetermined criterion. The relation between the likely corresponding blocks describes a motion vector.

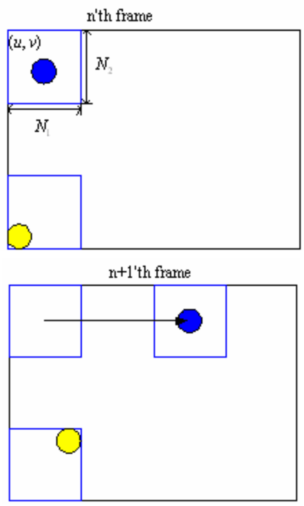


Fig. 1. Block matching algorithm.

The best matching can be found by minimizing a distortion measurement through criterion, such as the MAD (Mean Absolute Difference) and MSD (Mean Square Difference), or by maximizing a correlation function of the two blocks. The matching of the blocks can be quantified according to various criteria including the maximum cross-correlation, the minimum MSD, the minimum MAD and the maximum MPC (Matching Pixel Count). The following equations define the MAD and MSD mathematically.

$$MAD(u,v) = \frac{1}{N_1 \times N_2} \sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} \left| L^{n+1}(i+u, j+v) - L^{n}(i,j) \right| \qquad (1)$$

$$MSD(u,v) = \frac{1}{N_1 \times N_2} \sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} \left| L^{n+1}(i+u, j+v) - L^{n}(i,j) \right|^2 \qquad (2)$$

where L^n and L^{n+1} represent values of the pixel in n th and ($n+1$)th frame, respectively and (u,v) is a search point in the search area. MAD has been used more widely than MSD because MAD requires less computations and also is easier to implement the hardware than MSD.

In this paper, we assumed that the information of a moving object in a uniform surroundings was known *a prior*. So the computation of MAD to find motion vector was done more rapidly than the computation of MSD or cross-correlation because MSD is the squared form of MAD as shown in (1) and (2). In Fig. 1, the motion vector of the blue object is easily computed because the displacement of the object is bigger than the size of block. However the motion vector of the yellow object may have been lost because the object displacement is smaller than the block size. If we mainly focus on the object region, we can use cross-correlation rather than MAD or MSD. However, it was hard for cross-correlation criteria to find the motion vector because of heavy computational burden in real-time. So, in this paper, we have defined RB (Representative Block) and applied an optimal RB to not lose the motion vector. The optimal RB has been verified in this paper that it provides a high performance for real-time moving object tracking.

3. REPRESENTATIVE BLOCK MODEL

The other algorithms had a low performance due to many operations in achieving the accurate object recognition and they have many problems in real-time object tracking. In this paper, our main interest is in moving vector elements.

The moving vector elements of an object are represented in Fig. 2(a) , which are converted to a block and shown as Fig. 2(b). Here, detection of the movement vector through the simplification is focused. Then, the object on Fig. 2(a) can be distinguished from the background through the gray block sampling as shown in Fig. 2(c). Fig. 2(d) represents the simplified object region represented as the representative blocks. On the whole, Fig. 2(a) can be represented as Fig 2(b). Therefore, the gray blocks are defined as RB (Representative Block) and the block (2x2) is representative of the bigger block in Fig. 2(d).

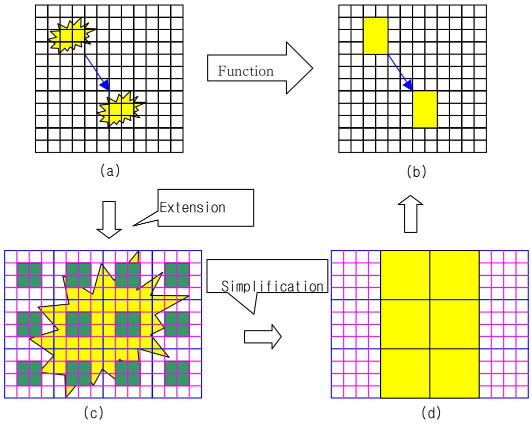


Fig. 2. Representative Block (RB).

As shown in Fig. 2, when the image data are processed in each pixel unit, the dimension of image processing is increased by 16 times. Therefore, for both to decrease the dimension of data processing and to improve the image data processing speed, the RB (Representative Block) is defined. Since RB covers its neighboring pixels, the region of RB and its neighbor is defined as S_{OB} which is generally represented as,

4. REPRESENTATIVE BLOCK

4.1 Perspective Transformation

A perspective transformation(also called an imaging transformation) projects 3-D points onto a plane. Perspective transformation plays a central role in image processing because it provides an approximation to the manner in which an image is formed by viewing a 3-D world[15].

Fig. 4.shows a model of the image formation process. The camera coordinate system (x, y, z) has the image plane coincident with the xy plane and the optical axis(established by the center of the lens) along the z axis. Thus, the center of the image plane is at the origin, and the center of the lens is at coordinates $(0,0,\lambda)$. If the camera is in focus for a distant object, λ is the focal length of the lens. Here the assumption is that the camera coordinate system is aligned with the world coordinate system (X,Y,Z).

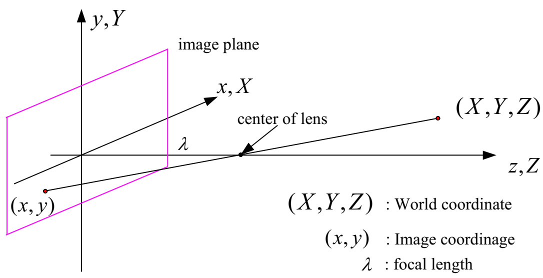


Fig. 3. Basic model of the imaging process.

Let (X, Y, Z) be the world coordinates of any point in a 3-D scene, as shown in Fig. 4. We assume throughout the following discussion $Z > \lambda$; that is, all points of interest lie in front of the lens. The first step is to obtain a relationship that gives the coordinates (x, y) of the projection of the point (X, Y, Z) onto the image plane. This is easily accomplished by the use of similar triangles properties. With reference to Fig.3,

4.2 Ball Size

The active camera on the mobile robot is configured so that the ball's information exists in the middle of the ball's input image at all times. Therefore, as the mobile robot moves closer to the target, the adjustment of the pan and tilt angle of the active camera in the robot system enables the information of the target to be in the middle of the input image at all times. Hence, both the ball's center and the center of camera's lens are always positioned on the Z -axis as shown Fig. 5. (X_0, Y_0) and (X_1, Y_1), which determine the ball's diameter as shown in Fig. 5, don't change. The ball size on the image frame can be changed by only the Z_0, which is the distance between the ball and the camera. First of all, we can obtain Z_{ccd} (camera height) and θ_p (tilt angle) by using the kinematics of the 2-DOF active camera mounted on the mobile robot to obtain R which is the radius of the ball. The corresponding points on the image plane, (x_0, y_0) and (x_1, y_1), can be computed as follows:

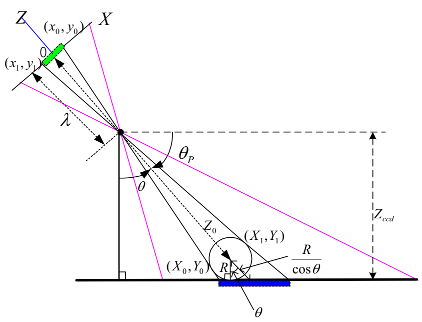


Fig. 4. Distance between robot's camera and ball.

5. EXPERIMENT AND RESULT

5.1 Equipment of experiment

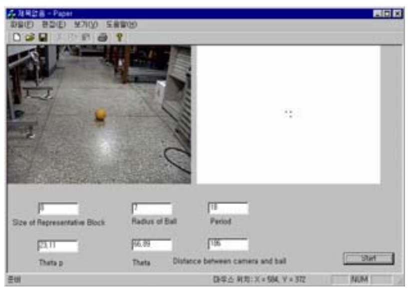
Fig. 5. Experimental mobile robot.

Fig. 5 is the photograph of the mobile robot for the real experiment. In the mobile robot, there are six controllers to control two wheels and the pan/tilt camera, to gather the gyro sensor data, and to coordinate all these controllers by one-high controller which is a Pentium 3 PC board. There is also a frame grabber card to obtain the information of the image, which is interfaced through PCI bus to the PC. The CAN-bus is used for the controllers, and there is an interface card for the active camera to control through ISA bus.

5.2 Result

We made two experiments. In the first experiment, the distance between the ball and camera is 83cm and the ball moves by 17 cm along the corridor, which is shown in the difference between Fig. 6(a) and Fig. 6(b) as an upward shift.. The comparison of the optimal RB and the BMA method is performed. The result of the first experiment is shown in Fig. 10 (a) is the previous image in (n) frame and (b) is the present image in (n+1) frame. (c) and (e) are the resultant images using the optimal RB which has 20 by 20 pixels after the object tracking. (d) and (f) are the resultant images using the BMA method which has 20 by 20 pixels after the object tracking. Intentionally, the block size of BMA is kept identical with the sub-block size of the optimal RB. However, within a block of BMA all pixels (20x20x3=1200) are to be computed, while, within a sub-block , the optimal RB (10x10x3=300) needs to be processed. Thus the computation load is reduced to a quarter. In other words, although both methods are satisfactory in the sense of tracking and recognition, the optimal RB is superior to the BMA in the meaning of operation speed. In the BMA, it elapsed about 29~31 msec, while, in the optimal RB, about 18~20 msec lapse exists for the object recognition.

(a) Fixed the size of RB



(b) Optimal RB's size(8)

Fig. 6. Comparison between a Fixed and an optimal RB.

6. CONCLUSION

This paper proposed an optimal RB method for the real time object tracking, and dealt with the comparison between optimal RB method and BMA method to show the superiority of RB method. In the case of BMA, image preprocessing is required to recognize the correct image, which results in the heavy processing load and the loss of tracking object. The missing-object happens when the displacement of the object is less than the block size in the image if BMA with MSD and MAD is applied for the moving object tracking algorithm in the real-time. When the fixed RB is utilized for the tracking, sometimes the algorithm misses the moving object in the image frame, although it reduces the computational time. However, when the optimal RB is utilized for the algorithm, the effects of both high tracking performance and less computation time are obtained concurrently. For the decision of the RB size, the size of the tracking object in the image frame is the major factor. Therefore, this RB method with the size optimization scheme can be widely used for the applications that include the tracking of fast moving objects in the real-time.

REFERENCES

[1] Bede Liu and Andre Zaccarin, "New Fast Algorithms for the Estimation of Block Motion Vectors" IEEE Trans. Circuits, vol. 3, no.2, pp.148-157, APRIL 1993.

[2] Joon-Seek Kim and Rae-Hong Park, "A Fast Feature-Based Block Matching Algorithm Using Integral Projections" Selected Areas in Communications, IEEE Journal on, vol.10, no. 5, pp968-971, June 1992 [3] Yih-Chuan Lin and Shen-Chuan Tai, "Fast Full-Search Block-Matching Algorithm for Motion-Compensated Video Compression" IEEE Transs. Communications, vol. 45, no.5, MAY 1997

[4] Wen-Jyi Hwang, Yeong-Cherng Lu and Yi-Chong Zeng, "Fast Block-Matching Algorithm for Video Coding", Electronics Letters, vol. 33, no. 10, 8th MAY 1997.

[5] Raphaël Canals, Anthony Roussel, jean-Luc Famechon, and Sylvie Treuillet,"A Biprocessor-Oriented Vision-Based Tracking System" IEEE Trans. Industrial Electronics, vol. 49, no. 2. APRIL 2002.

[6] Ahmad, M.B., Dong Yoon Kim, Kyoung Sig Roh and Tae Sun Choi, "Motion Vector Estimation Using Edge Oriented Block Matching Algorithm for Video Sequences", Image Processing, 2000. Proceedings. 2000 International Conference on, vol. 1 pp806-863, 2000
[7] Th.Zahariadis and D. Kalivas, "Fast algorithms for the estimation of block motion vectors" ICECS, pp.716-719, 1996

[8] Yilin. Zhao and Spencer L. BeMent, "Kinematics, Dynamics and Control of Wheeled Mobile Robots", IEEE Conf. Robotics and Automation, pp.91-96, 5. 1992

[9] J. David, Kreigman et al. "Stereo vision and navigation in building for mobile robots", IEEE Trans. on Robotics and Automation, vol. Ra-2, no. pp14-23, 1986

[10] Mark W. Spong, M. Vidyasagar, Robot Dynamics and Control, WILEY, 1989

[11] F. L. Lewis, C. T. Abdallah, D. M. Dawson, Control of Robot Manipulators, Machmillan, 1993

[12] Il-Myung Kim, Wan-Cheol Kim, Kyung-Sik Yun, Jang-Myung Lee, "Navigation of a Mobile Robot Using Hand Gesture Recognition", Journal of Control, Automation and Systems Engineering, vol. 8, no. 7, pp. 599-606, July, 2002

[13] Grosso, E., Tistarelli, M. "Active/Dynamic stereo vision," IEEE Trans. On Pattern Analysis and Machine Intelligence, vol. 7, pp. 868-879, 1995

[14] K. Daniilidis, C. Krauss, "Real-Time Tracking of Moving Objects with an Active Camera," Real-time imaging, Academic Press Limited, 1998

[15] Rafael C. Gonzalez, Richard E. woods, Digital Image Processing, Addison Wesley,1993