Disparity Estimation for Objects of Interest

Yen San Yong and Hock Woon Hon

Abstract—An algorithm for estimating the disparity of objects of interest is proposed. This algorithm uses image shifting and overlapping area to estimate the disparity value; thereby depth of the objects of interest can be obtained. The algorithm is able to perform at different levels of accuracy. However, as the accuracy increases the processing speed decreases. The algorithm is tested with static stereo images and sequence of stereo images. The experimental results are presented in this paper.

Keywords—stereo vision, binocular parallax

I. INTRODUCTION

HUMAN eyes are horizontally separately; thereby image captured by each eye is slightly different [1]. This binocular vision [2] enables human to perceive the depth information. To enable computer or robot to "see" depth, it requires stereo vision to find the corresponding points between the left image and the right image, then calculate the disparity and finally estimate the depth.

There are two types of depth map can be generated via different algorithms. Correlation-based algorithms [3][4] require higher processing time to produce the dense disparity map which covers the whole captured image. On the other hand, feature-based algorithms [5][6] require pre-processing to extract features as the matching criteria to produce a sparse disparity map which concentrated on the high-informative feature such as edges.

In most of the stereo vision algorithm, solving the correspondence problem is the main key point. However, searching for correspondence points or features can be time consuming. In addition, the quality of the extracted features may directly affect the accuracy and processing time required for correspondence search. For example, if two images are taken with slightly different brightness, the area of the same object covered in both images would be different and thus the algorithm used might not be able to match this object. The proposed algorithm excludes searching for correspondence to improve the speed of the algorithm and the dependency on the quality of the extracted features. Consequently, it lost the dense disparity information as it can only estimate the average disparity of an object of interest.

Yen San Yong is with MIMOS Berhad Tachnology Park Malaysia 57000 Kuala Lumpur, Malaysia. e-mail: yyen.san@mimos.my

This paper is divided into 4 sections. Section 2 demonstrates the proposed algorithm in detail whereas Section 3 illustrates and analyses the experimental results. The paper ends with conclusion and future work in Section 4.

II. DISPARITY ESTIMATION ALGORITHM

The proposed algorithm presented in this paper employs image shifting and the overlapping area to estimate the disparity value of an object of interest. Therefore, the first step is to extract the object of interest. However, this is not part of the proposed algorithm as user can employ any algorithm to extract object of interest and convert the resulting grey scale or colour image to binary image. The next step, disparity estimation, is the one covered in this paper.

L_x: object's x-position registered on left image

R_x: object's x-position registered on right image
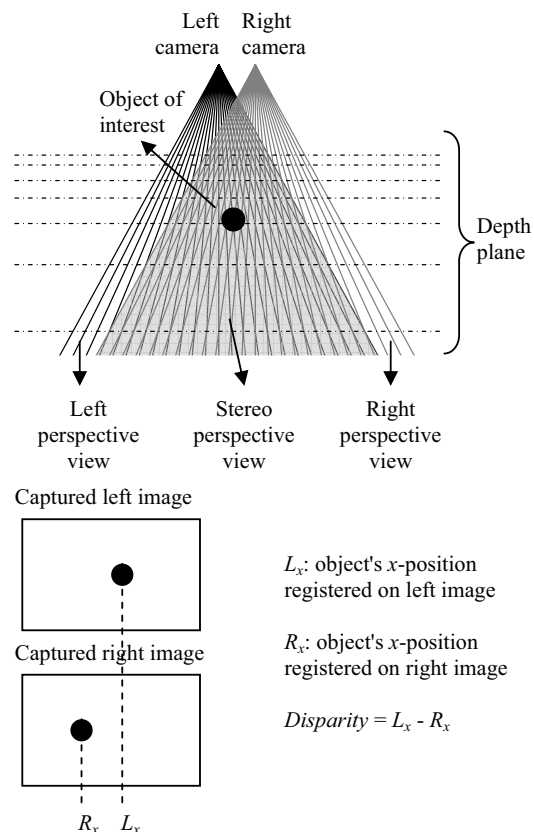
Disparity = L_x - R_x

Fig. 1 Concept of stereo capturing system.

Fig. 1 illustrates the concept of stereo capturing system. Object placed at different depth plane produces a different

separation in the stereo images. Therefore, this separation, called disparity or parallax, is able to estimate the object's position in terms of depth plane.

Fig. 2 and Fig. 3 illustrate the disparity estimation algorithm. As described in Fig. 1, the object of interest registered on different x-position in an image. However, both images share similar features. By repeating shifting one of the images in x-axis, in this case the algorithm shift the right image, at a certain point, both images will exhibit the maximum number of overlapping pixels. As a result, the shift value that produced this maximum overlapping area is used to represent the disparity value for this object of interest.
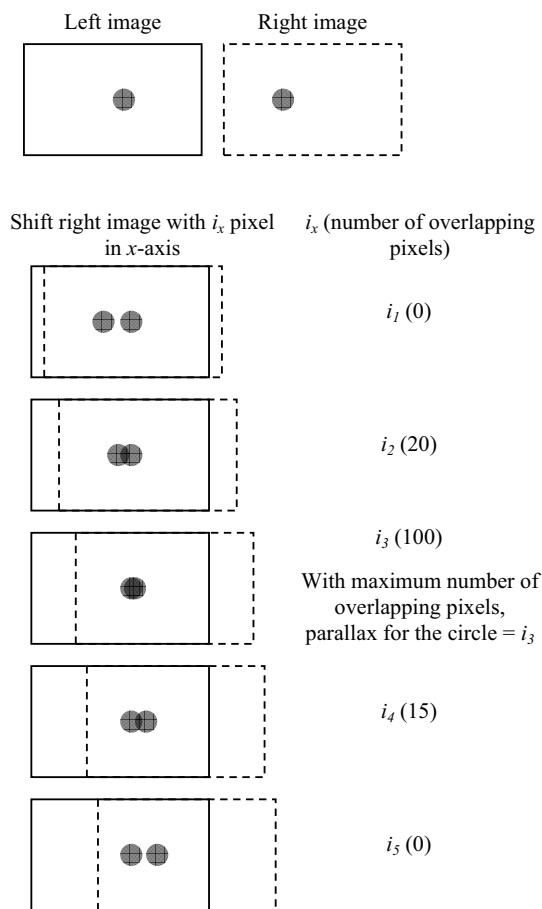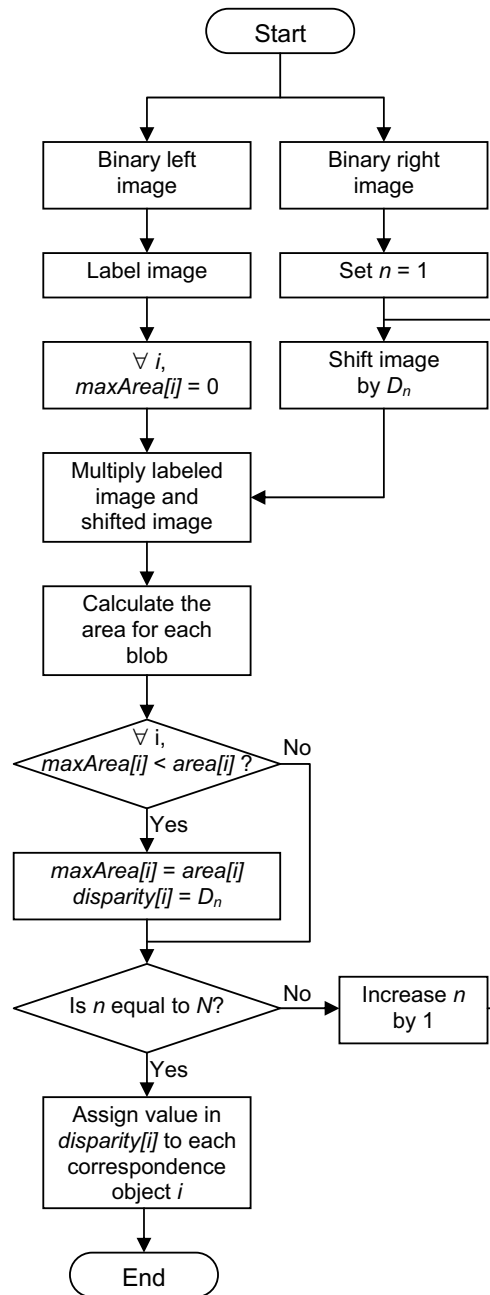

Fig. 2 Graphical representation of the disparity estimation algorithm.

The flow chart for this algorithm is shown in Figure 3. Assume that the objects of interest are extracted and binarised such that the background pixels are assigned to 0 whilst the foreground pixels are assigned to 1. Also, the disparity range, $[D_1, D_2, ..., D_N]$ is defined.

N = total number of disparity value in the range
I = total number of labeled blob
n = disparity counter
i = blob counter
D_n = nth parallax value
$maxArea[i]$ = maximum area for blob i
$disparity[i]$ = disparity value assigned to blob i
$area[i]$ = current area for blob i

Fig. 3 Flow chart for the disparity estimation algorithm.

The binary left image is labeled and kept as a reference image while the right image is shifted in x-axis to find the optimum disparity value for each blob as demonstrated in Fig. 2. The values, D_n, in the disparity range are used as x-axis translation parameter. Then the labeled left image and shifted right image are multiply together to extract the overlapping area between two images while maintaining the blob labeling number.

Next, the area of each blob in the resultant image is calculated. If the blob area is more than the stored *maxArea[]*, then the value of *maxArea* is substitute by the current area, *area*, and the current pixel shift is recorded in *disparity[]*.

After completing the shifting sequence, the values stored in *disparity[]* is the disparity value for each corresponding object of interest.

III. EXPERIMENTAL RESULT

Two set of experiments are presented in this paper. The disparity estimation algorithm is tested on static image and on image sequence.

A. Static Image

Two static images as illustrated in Fig. 4(a), which are slight difference in terms of brightness, are used as the input for the algorithm. A set of preprocessing is applied on the images to extract the markers' cap as the objects of interest. The preprocessing algorithm employed is described in the following.

Preprocessing Algorithm (applied to both images):
1. Convert the colour image to grey scale image
2. Convert the grey scale image to binary image with threshold value = 35
3. Apply opening and dilation to remove small blobs and fill up the small gaps

The resultant images are illustrated in Fig. 4(b), which then feed into the proposed algorithm to estimate the depth. Two set of disparity list, ranging from -15 to 45, are applied to test the algorithm, The dense list, D_{dense}, is the list where every disparity value within the range is included whereas the sparse list, D_{sparse}, is the list which only contains 7 disparity values for experiment.

D_{dense} = {-15, -14, -13, ..., -1, 0, 1, ..., 43, 44, 45}
D_{sparse} = {-15, -5, 5, 15, 25, 35, 45}

Fig. 4 illustrates the images involved in the process whereas Table I tabulates the experimental results. As shown in Fig. 4, the left image is brighter than the right image. With same preprocessing and threshold value, it gives two different binary images. If matching algorithm implies feature properties such as area or bounding box, some of the object may not be matched and thus no depth information can be estimated. However, the proposed algorithm, which excludes the matching process, is able to estimate the disparity with high accuracy. The time taken for the whole process

(including preprocessing) and the accuracy depend on the number of values in the disparity list. This allows user having the flexibility to choose whether the time or the accuracy should come first in particular application.

(a)   Original Left Image   Original Right Image

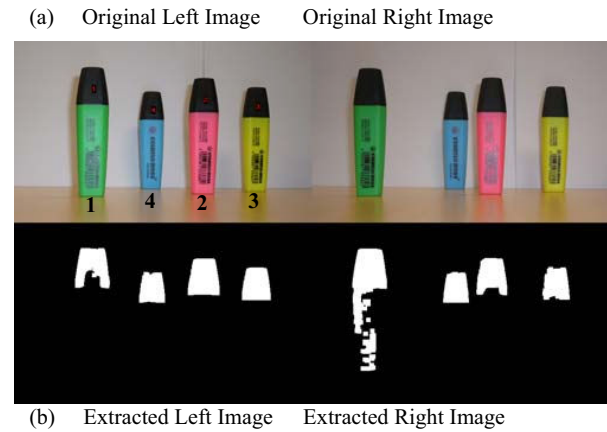(b)   Extracted Left Image   Extracted Right Image

Fig. 4 (a) Original colour images (b) binary images showing the objects of interest (markers' cap).

TABLE I
EXPERIMENTAL RESULT FOR STATIC STEREO IMAGES

Object	Manual Measurement	Measurement with D_{dense}		Measurement with D_{sparse}	
	Disparity	Disparity	|Error|	Disparity	|Error|
1	42	42	0	45	3
2	21	21	0	25	4
3	-2	-3	1	-5	3
4	-13	-12	1	-15	2
Time taken	-	0.23s		0.06s	

B. Image Sequence

Two image sequences are fed to the proposed algorithm with the following preprocessing to extract the moving object as the object of interest.

Preprocessing Algorithm (applied to both images):
1. Colour subtraction between previous frame and current frame
2. Convert the subtracted colour image to grey scale image using the formula
3. $G(x, y) = \dfrac{R(x, y) + G(x, y) + B(x, y)}{3}$

4. Convert the grey scale image to binary image with threshold value = 70
5. Apply opening and dilation to remove small blobs and fill up the small gaps

Fig. 5(a) illustrates the previous frame (frame 10) whereas Fig. 5(b) illustrates the current frame (frame 20). The algorithm does not employ the immediate previous frame for comparison so that the extracted moving blob can be more obvious to the algorithm. The parts of the moving object are extracted, as illustrates in Fig. 5(c), and feed into the proposed algorithm to estimate the disparity. This experiment is tested with two disparity lists as applied in previous experiment, D_{dense} and D_{sparse}. The experimental results are tabulated in Table II. There are more than one blob extracted from the left image but there is only one blob extracted from the right image. Therefore only a single disparity value is estimated.

(a) Left Image (frame 10)   Right Image (frame 10)

(b) Left Image (frame 20)   Right Image (frame 20)

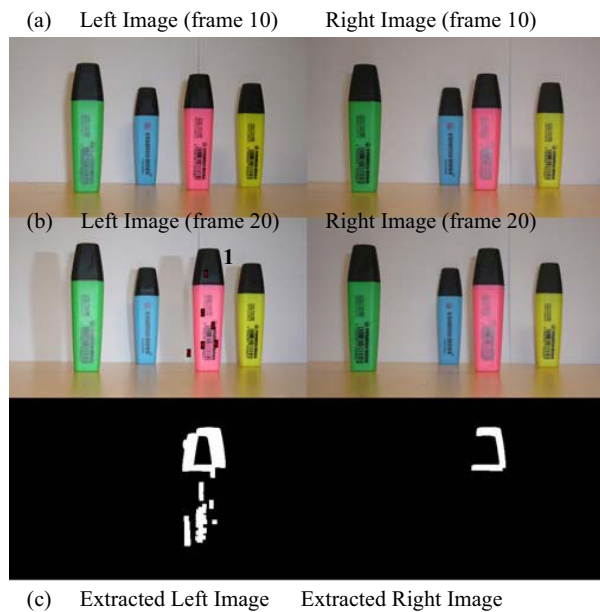(c) Extracted Left Image   Extracted Right Image

Fig. 5 (a) Original colour images (frame 10) (b) Original colour images (frame 20) (c) binary images showing the objects of interest (moving object).

IV. CONCLUSION AND FUTURE WORK

The experimental results indicate that the proposed disparity estimation algorithm is able to locate depth of objects of interest even though the images are slightly different in terms of brightness. The accuracy is high when the dense disparity list is used, but longer time is required. On the other hand, the user can used shorter disparity list to improve the speed with lower accuracy.

TABLE II
EXPERIMENTAL RESULT FOR STEREO IMAGE SEQUENCES

Object	Manual Measurement	Measurement with D_{dense}		Measurement with D_{sparse}	
	Disparity	Disparity	|Error|	Disparity	|Error|
1	36	35	1	35	1
Time taken	-	0.47s		0.19s	

This algorithm may useful for real-time application such as video surveillance [7] where at most of the time the objects of interest are moving in the scene. However, the scenario that involves two or more objects of interest overlapping each other is still unsolved in this algorithm. In future, this problem can be investigated and thus the algorithm can then be applied in wider area of application.

REFERENCES

[1] Kaufman, P.L., and Alm, A., *Adler's Physiology of the Eye*, Tenth Edition, Elsevier Health Sciences, Nov 2002.
[2] Sekuler, R., Blake, R., *Perception*, Third Edition, McGraw-Hill, 1994.
[3] S. Birchfield and C. Tomasi, "Depth Discontinuities by Pixel-to-pixel Stereo", International Journal of Computer Vision, Vol. 35 No. 3, pp. 269-293, 1999
[4] C.M. Sun, "Fast Stereo Matching using rectangular Subregioning and 3D Maximum-surface Techniques", International Journal of computer Vision, Vol 47, 2002, pp 99-117
[5] L.D. Stefano, M. Marchionni, S. Mattoccia, "A Fast Area-based Stereo Matching Algorithm", *Image and Vision Computing*, Elsevier, Vol 22, 2004, pp 983-1005.
[6] F. Tombari, S. Mattoccia, L.D. Stefano, "*Segmentation-based Adaptive Support for Accurate Stereo Correspondence*", IEEE Pacific-Rim Symposium on Image and Video Technology (PSIVT), 2007
[7] R.T. Collins, A.J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, and O. Hasegawa, "A System for Video Surveillance and Monitoring: VSAM Final Report", *Technical report CMU-RI-TR-00-12*, Robotics Institute, Carnegie Mellon University, May 2000.