SUBPIXEL DISPARITY FOR ACCURATE STEREOSCOPIC DEPTH & VELOCITY

A.A.H. Ab-Rahman[1], U.U. Sheikh[2], S.A.R. Abu-Bakar[3]
Computer Vision, Video, and Image Processing Lab (CVVIP)
Faculty of Electrical Engineering
Universiti Teknologi Malaysia
Skudai, 81310, Johor, MALAYSIA
{hadi@fke.utm.my[1], uus99@hotmail.com[2], syed@fke.utm.my[3]}

Keywords: *Computer Vision, Stereoscopic, Image Processing, Image Disparity.*

ABSTRACT

This paper presents a subpixel disparity technique for obtaining accurate depth & velocity estimation of a moving object using stereoscopic vision. In such system, depth information is obtained from calculating the disparity between two images. This value is often precise within)1 pixel. The limitation in pixel precision leads to inaccuracy in getting 3D information of an object. We propose a technique to increase the precision of disparity calculation by interpolating the values of Sum of Absolute Difference (SAD) within the matching histogram, so as to provide a more accurate depth information, which subsequently gives a more accurate value for estimating velocity of a moving object. The techniques of Least Square Method (LSM) and Doolittle factoring is utilized to perform interpolation, and solve the resulting system of linear equations. Experiments show a significant increase in the accuracy of approximating depth and velocity of a moving object.

1. INTRODUCTION

Stereoscopic vision system is mainly used to obtain depth information, i.e. the distance from the cameras to a desired object [4], [5], [6], [7]. This is achieved by applying the *stereo triangulation* technique, which starts with performing pattern matching on the images. From the depth value, 3D information of an object can also be obtained, by extending the existing equations to include the width and the height values. If 3D information for a particular object is captured at two different time frames and positions, the speed and direction of the object can be estimated.

However, the accuracy of the resulting 3D information and velocity estimation largely depends on the validity of depth value. Better accuracy in depth recovery would in turn give a higher accuracy in velocity and 3D readings. Thus it is crucial to acquire an accurate value as possible for depth, in order to achieve the correct corresponding velocity of a moving object. By integrating subpixel disparity interpolation technique, depth, 3D, and velocity estimation results can be improved.

Szeliski and Scharstein (2002) utilized subpixel interpolation in pattern matching, which is one of the central issues in stereoscopic vision. By considering the subpixel values for intensity, they were able to improve the matching operation. Rather than using the commonly used SAD, they employed the Sum of Squared Difference (SSD), in the continous Disparity Space Image (DSI). Further detailed analysis concludes that the initial subpixel computation cost is restored and compensated for the new and improved disparity value.

The technique of subpixel disparity in image processing is also proposed by Umeda & Takahasi (2000), in the application of robotic vision, with a purpose to avoid correspondence (matching) problem with small camera separation and low resolution images. Kim et al (2001) employed this particular technique in their system to acquire accurate depth and 3D information, by using Relative Stereo Disparity (RSD) in the calculation. Among other application in subpixel interpolation includes the eye gaze tracking, proposed by Zhu and Yang (2002). Again, better results were achieved by employing this technique. In the imaging device and signal level, Baba et al (2000) interpolates an analog signal through the Charge Coupled Device (CCD) to produce a higher resolution for position measurement.

The authors propose a subpixel disparity technique based on numerical methods, and proves its feasibility in the applications of 3D vector movements, which includes the accuracy effects on depth and velocity estimation.

2. SUBPIXEL INTERPOLATION

Stereoscopic *triangulation* technique [6] is used to obtain depth information from an image pair calibrated in parallel optical axes, with a general equation

$$Z = B*f / d$$

Where Z is the depth information, B is the baseline (camera separation), f is the focal length of lens and d is the image disparity. As the equation reveals, the baseline and the focal length are constant terms; depth information is inversely proportional to image disparity.

Consider the case of a stereoscopic vision system with a baseline of 1 meter, and focal length of lens of 25 millimeters. The table below shows depth values for image disparity range of 10 to 14 pixels.

Disparity (pixels)	Depth (meters)
10	200.00
11	181.82
12	166.67
13	153.85
14	142.86

Figure 1: Disparity-depth analysis of a stereoscopic vision system.

From the table above, it can be concluded that a certain disparity value computes to a one depth value. For example, a disparity of 11 pixels would give depth of exactly 181.82 meters, while a pixel disparity of 12 would work out to a depth value of 166.67 meters. Without considering the subpixels of between 11 and 12, it is not possible to obtain depth values of between 181.82 meters and 166.67 meters.

Image disparity is obtained through stereoscopic (pattern) matching. The most common technique for this is the Sum of Absolute Difference (SAD) [4]. SAD matching takes a certain window size from a reference image, and finds the corresponding window in the other image, by comparing the reference window with a searched window which has the least sum of absolute difference. This position with the least value is thus the matched pixel location, and the disparity is determined by calculating the difference between the reference image pixel location and the second image pixel location.

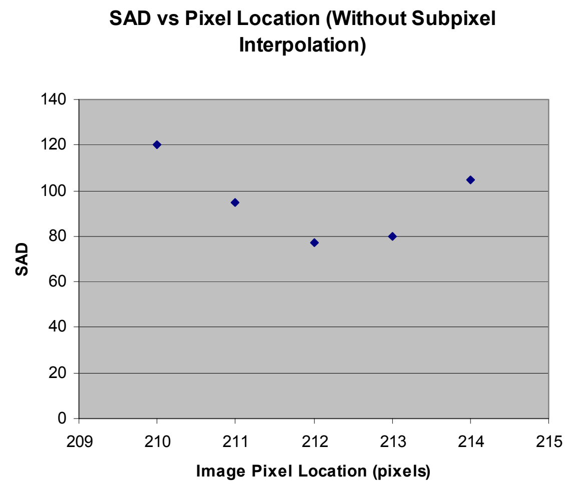

Figure 2: SAD graph without subpixel interpolation. Matched position is at the lowest SAD value, which is at 212.

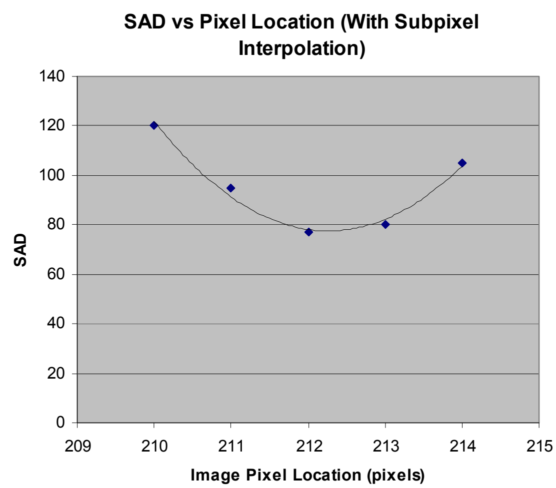

Figure 3: SAD graph with subpixel interpolation. Matched position is now at position 212.3.

Refer to figures 2 and 3. Assume reference image location of 223. Without subpixel interpolation, image disparity is 11 (223 – 212) pixels, which translates to a depth value of 181.82 meters (figure 1). However, with subpixel interpolation, the matched position is now at 212.3, which corresponds to image disparity of 10.7 pixels, and a depth value of 186.92 meters. If the actual depth of a desired object is at 190 meters, the technique without subpixel interpolation results in accuracy of 96.7%, while utilizing the subpixel interpolation technique produces an accuracy of 98.4%.

3. TECHNIQUE INTEGRATION

Integrating the subpixel disparity technique in a stereoscopic vision system starts with capturing image pairs from cameras calibrated in parallel optical axes into a workstation. SAD intensity matching is then performed on the images to get the initial disparity, i.e. without subpixel interpolation. Based on these SAD values, the pixels are interpolated to get the subpixel values. New disparity is calculated and sent to the depth and velocity estimation algorithm for further processing.

Figure 4 shows the flow chart of integrating the subpixel disparity technique. For practical purposes, the Least Square Method (LSM) is used to interpolate 5 set of points (pixel position, SAD value) to a second degree polynomial [3]. The general case for n set of coordinate points that interpolates to an m^{th} degree polynomial,

$$\begin{bmatrix} s_0 & s_1 & \dots & s_m \\ s_1 & s_2 & \dots & s_{m+1} \\ \dots & \dots & \dots & \dots \\ s_m & s_{m+1} & \dots & s_{2m} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \dots \\ a_m \end{bmatrix} = \begin{bmatrix} v_0 \\ v_1 \\ \dots \\ v_m \end{bmatrix}$$

Where:

$$s_j = \sum_{k=0}^{n} X_k^j \text{, j = 0, 1, 2 …, 2m}$$

$$v_l = \sum_{k=0}^{n} X_k^l * f_k \text{, l = 0, 1, 2, …, m}$$

X_k and f_k are the coordinate points, with k = 0, 1, 2, .., n

Solving the system of linear equations above produces an m^{th} degree polynomial function:

$$F(x) = a_0 + a_1x + a_2x^2 + \dots + a_mx^m$$

Taking the derivative of the above equation, and equates it to zero, produces the root, or in our case, the subpixel value. The system of linear equation above is easily solved using factoring techniques.

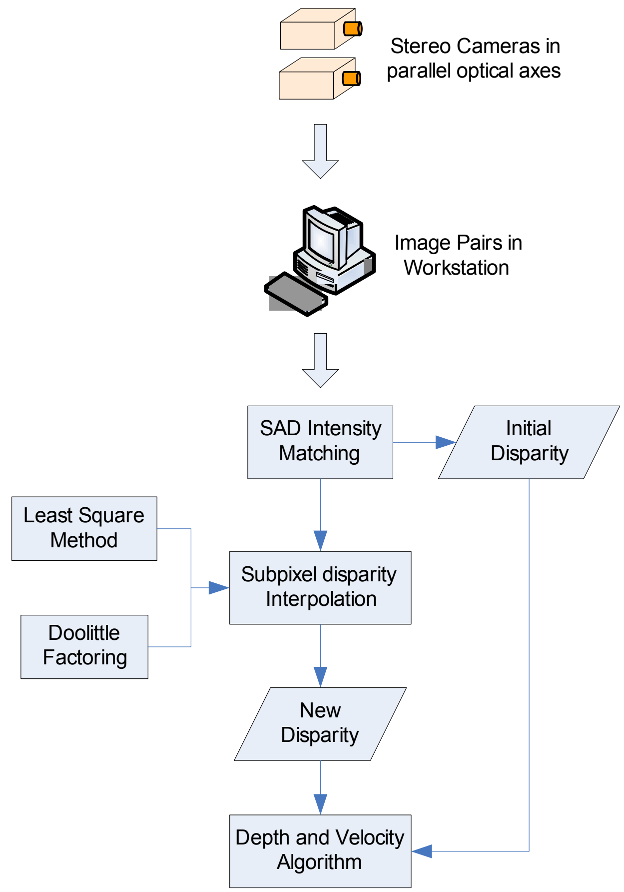

Figure 4: Flow chart of subpixel disparity technique in a stereoscopic vision system.

4. EXPERIMENTAL RESULTS

The effectiveness of subpixel disparity technique is verified through a developed stereoscopic vision system that is capable of estimating depth, and velocity of a moving object.

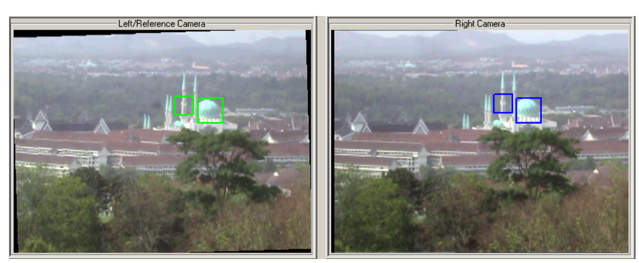

Figure 5: Finding distance of the dome from cameras. Using subpixel disparity, the accuracy of the distance measured increases from 76.3% to 90.3%.

Figure 5 shows part of the system to estimate depth up to 2 kilometers. It uses 2 cameras with 1/3” CCD format, 950 millimeter baseline, and 25 millimeter lens. The actual distance of the dome of the mosque is 1.450 kilometers from the cameras (determined using Global Positioning System technique). Without subpixel disparity technique, distance is measured at 1.901 kilometers, while integrating this technique, distance is now evaluated at 1.309 kilometers. This corresponds to an accuracy increase of 14%.

Figure 6: Finding the distance of the small tree on the lower right corner of the images, from the cameras. The left image, which has undergone software calibration, becomes the reference image to find the distance.

The next test involved a medium range distance measurement, of about 400 meters from the cameras. Figure 6 is another screenshot showing the stereoscopic vision program being used to find depth. The small tree in the figure is at a distance of 410 meters. By using subpixel disparity, we achieved a distance of 422.38 meters, while the result without subpixel interpolation is 461.64 meters. This proves another significant increase in the accuracy of depth estimation.

Figure 7: Finding the speed of the remote controlled car. The speed is initially determined using a stopwatch-measuring tape, and compared to systems with and without subpixel disparity.

Figure 7 is another screenshot of the system, this time used to estimate velocity. To verify this technique, only the speed component will be considered. From experiments and observations, the remote controlled car has an actual speed of 0.463 m/s. Without integrating this technique, the system estimates the speed at 0.454 m/s, while subpixel disparity estimates the speed at 0.459 m/s. It is concluded that at short range, this method produces little effect on accuracy. This is also most likely due to the minimized error with the choice of lens and baseline for short distance. Nevertheless, it is increased by 1.7%, from 98.06% to 99.13%.

Figure 8: Finding the speed of the car. The car's speedometer is clocked at about 20 km/h. Two separate tests were conducted, with and without subpixel disparity technique.

The final set of test is conducted outdoors in an open field, at a range of about 100 meters from the cameras (Figure 8). The car is driven at a constant speed of 20 km/h, and the result produced by the stereoscopic vision system is recorded and compared. As expected, results were improved by using subpixel disparity technique, as oppose to considering only the discrete pixels, by more than 10%.

5. SUMMARY

In this paper, we have presented a technique that could increase the accuracy of estimating depth and velocity using stereoscopic vision. It is performed by considering the subpixel values based on the sum of absolute difference technique for stereoscopic matching. The interpolated points provide a new and more accurate disparity for depth estimation. The theory is proved by integrating the technique into an existing stereoscopic vision system.

Current work involves estimating depth for long range (up to 2 kilometers), and estimating velocity for short and medium range (up to 100 meters). However, the goal is to accurately determine the speed and direction of moving objects at long range. For that, apart from subpixel disparity technique, other enhancement methods such as camera calibration techniques, histogram projection, and linear prediction are also being reviewed and implemented in order to have a more robust and reliable system.

6. REFERENCES

[1] K. Umeda, Takahashi, "Subpixel Stereo Method: A New Methodology of Stereo Vision," *IEEE International Conference on Robotics & Automation*, pp 3215-3220, April 2000.

[2] J.K. Kim, M.C. Lee, C.K. Ahn, C.Y. Kim, "Acquisition of 3D Information Using Stereo Vision System by Relative Stereo Disparity with Subpixel resolution," *ISIE 2001*, pp 1971-1676.

[3] A.R. Ahmad, “Introduction to Numerical Methods,” Universiti Teknologi Malaysia, Addison-Wesley, 2004.

[4] T. Kanade and M.Okutomi, “A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment,” *Pattern Analysis and Machine Intelligence*, pp. 920-932, 1994.

[5] T. Kanade and M.Okutomi, “A Multiple Baseline Stereo Method,” *Proceedings of DARPA Image Understanding Workshop,* pp. 409-426, Jan. 27-29 1992.

[6] K.D. Toennies, “3D Computer Vision, Stereo Vision (introduction),” Universitat Magdeburg, URL: http://isgwww.cs.unimagdeburg.de/bv/skript/cv/VLcv01.pdf

[7] L. Iocchi, “Stereo Vision: Triangulation,” Universita Di Roma, URL:www.dis.uniroma1.it/~iocchi/stereo/triang.html

[8] G.B. Thomas, “Calculus and Analytic Geometry,” Massachusetts Institute of Technology, Addison-Wesley, pp 442-480, 1983.

[9] D. Demirdjian and T. Darrel, “Motion Estimation from Disparity Images,” *Eighth IEEE International Conference on Computer Vision,* pp. 213-218, 7-14 July 2001.

[10] D. Demirdjian and T. Darrel, “Using multiple-hypothesis disparity maps and image velocity for 3D motion estimation,” *IEEE Workshop on Stereo and Multi-baseline Vision,* pp. 121-128, 9-10 Dec. 2001.

[11] M. Baba, K. Ohtani, and T. Konishi, “A Novel Position Measurement with Subpixel Resolution Using a Charged Coupled Device,” *Measurement Science & Technology,* pp. 266-272, March 2000.

[12] R. Szeliski and D. Scharstein, “Symmetric Subpixel Stereo Matching,” *Seventh European Conference on Computer Vision,* pp. 525-540, May 2002.

[13] J. Tumblin and P. Choudhury, “Bixels: Picture Samples with Sharp Embedded Boundaries,” *Eurographics Symposium on Rendering,* 2004.

[14] K. Siddiqi, B. Kimia, and C.W. Shu, “Geometry Shock Capturing ENO Schemes for Subpixel Interpolation, Computation, and Curve Evolution,” *ISCV ‘95,* 1995.