Low-cost Stereo Vision System for Supporting the Visually Impaired's Walk

Thanathip Limna, Pichaya Tandayya and Nikom Suvanvorn
Department of Computer Engineering, Faculty of Engineering,
Prince of Songkla University

Hat Yai, Songkhla 90112, Thailand
Tel. +66 74287352 Fax. +66 74212895

boatkrap@gmail.com, pichaya@coe.psu.ac.th, kom@coe.psu.ac.th

ABSTRACT
This paper presents an obstacle detection system for the visually impaired to use with a walking stick. The system in this work can find obstacles applying the depth discontinuities Pixel-to-Pixel (P2P) stereo algorithm which is one of the Intensity-based Stereo Matching (ISM) techniques in stereo vision that can find objects with featureless surface such as whiteboard, door, etc. However, the ISM technique is time consuming and not fast enough for real-time usage. Applying parallel computing using Message Passing Interface (MPI) helps reduce the computing time and enables real-time usage. In this work, we estimate the distance between objects and the visually impaired by applying the V-disparity. Our system can accurately detect objects within the range of 5 meters using 12-centimeter based-line low-cost webcams.

Categories and Subject Descriptors
I.2.10 **[Vision and Scene Understanding]:** video analysis

General Terms
Experimentation

Keywords
Stereo Vision, Parallel computing, the visually impaired, Pixel-to-Pixel stereo algorithm, Intensity-based Stereo Matching technique, Message Passing Interface, Calibration.

1. INTRODUCTION
Normally, a visually impaired person walks to places using a walking stick for checking obstacles, especially which lie on the floor. However, visually impaired people still often get accidents because of hung signs, floating objects, a protection extending from the eaves against rain or sun usually found in Asia, tree branches sticking out into the walk way, etc. A walking stick with an ultrasonic sensor detecting obstacles was made to warn the visually impaired [1]. However, it can only detect whether there are obstacles ahead but cannot inform the distance, size and position of the objects.

There are several inventions using ultrasonic or infrared devices attached to walking sticks and glasses, in order to find the object distance. Although they can detect whether there are obstacles in the current surrounding environments and find the distances, they do not provide sufficient information to enable the visually impaired to continue their journey. The important information lacked is the size.

Stereo vision technology can help solving the problem by finding the distance between the object to the stereo cameras, and the object's size and position. The accuracy of information depends on the light condition and stereo vision technique.

Electro Neural Vision System (ENVS) [5] is an application of stereo vision for the visually impaired that presents the obstacles and distances via different signals alerting at their ten figures. If the object is close, a high frequency signal will be used. Likewise, if the object is far away, a lower frequency signal will be sent. By this way, the visually impaired can walk to places and avoid obstacles by themselves. However, the Feature-based stereo matching (FSM) techniques [5] used in the ENVS, although can be quickly processed, cannot detect objects with featureless surface which cause dangers to the visually impaired.

On the other hand, the Intensity-based Stereo Matching (ISM) [10] techniques can detect featureless objects. However, there are also problems in applying the techniques in real time due to complex and time consuming image processing techniques and processes, especially when more accuracy is required. However, no reports about using parallel computing with the ISM techniques to reduce computing time have been found.

Parallel computing technology can help reducing the processing time at some parts of stereo vision, i.e., finding the image edges and scan line matching. Message Passing Interface (MPI) [12] offers a parallel programming interface specification that supports several platforms and can specify the number of processes at run time.

Our work applies parallel computing using MPI and of-the-shelf multi-core computers with the Intensity-based Stereo Matching (ISM) techniques [10] in order to reduce computing time. We also investigate for suitable calibration of our low-cost stereo cameras using webcams for supporting walking journeys of the visually impaired.

This paper presents an obstacle detection system for the visually impaired to use with a walking stick. The system finds obstacles applying the depth discontinuities Pixel-to-Pixel (P2P) stereo algorithm [7][8] of the ISM techniques that can detect featureless surface objects and estimate the object distance using V-disparity [9].

The next section presents the background knowledge of stereo vision. The third section describes system design including system overview, low-cost stereo cameras, object distance estimation, enhancing P2P using MPI, and limitations. The final section will be conclusions and future work.

2. STEREO VISION

The stereo vision system imitates the visual perception of human being by processing images from a pair of cameras. The human being's vision is 3-Dimentional (3D). Human beings can perceive the object's size and position, and the distance between an object and themselves. Likewise, in the stereo vision system, the 3D images can be processed in order to find an object's size, position and distance.

In this work, we apply stereo disparity to find the depth of the images. The method to find the images' disparity can be classified into two categories [11].

- Intensity-based Stereo Matching techniques (ISM)
- Feature-based Stereo Matching techniques (FSM)

In ISM techniques, the matching process is directly applied to the intensity profiles of the two images. In FSM techniques, features are first extracted from the images and the matching process is applied to the features. In this research, we are interested in the ISM techniques as they match corresponding pixels using the image intensity that can find the disparity even in the case of featureless surface objects. For our work, the most appropriate and recognized algorithm for processing the image disparity in the ISM techniques is the Depth Discontinuities by Pixel-to-Pixel Stereo (P2P) by Birchfield and Tomasi [7][8] as it appears as one of the OpenCV functions [14] and the source code can be down loaded freely.

In the P2P algorithm, there are two processes, scan line matching and post-process. Scan line matching is a process to match corresponding pixels on left and right images on the same scan lines. In each scan line, processes match corresponding pixels independently from one another. The post-process is a process to exchange data between the scan lines in order to select the best disparity image. This post-process concerns processing data across rows and columns and is not suitable for applying parallel programming. When running the P2P algorithm sequentially on a computer with the following specification, Intel® Core™ 2 Dual 6320 1.86 GHz 1010.7 MB RAM, running on the Linux operation system kernel 2.6.26, at the maximum disparity of 100 and the image size 320x240 pixels, the average computing times are 1.168 seconds for the scan line matching process and 0.165 second for the post-process. As a result, the scan line matching process took most computing time or about 70%. Therefore, we focus on scan line matching in this work.

3. SYSTEM DESIGN

This section describes our stereo vision system for the visually impaired as follows.

3.1 System Overview

For the safety of the visually impaired, this object detection system is required to give warnings to the visually impaired when there are obstacles in the distance of 10 meters ahead as shown in Figure 1. However, the most critical distance is between 1 meter and 6 meters that is out of the walking stick's range and far enough to cover dangerous areas while walking so that the visually impaired can avoid obstacles. This system is not to replace the walking stick but aims to be used with the walking stick in order to provide more confidence with the tool they are used to.

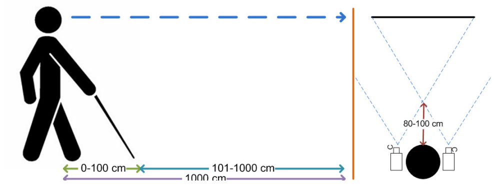

Figure 1. System usage overview.

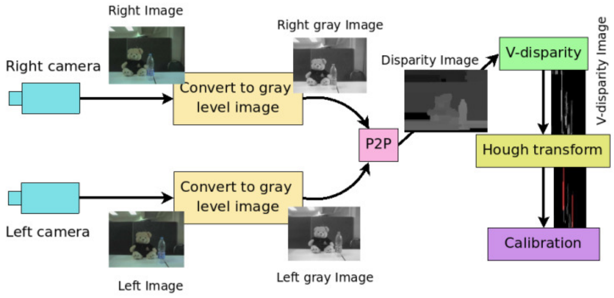

Figure 2. Stereo vision processes

Figure 2 shows the stereo vision processes. First, the color images from left and right cameras will be transformed into gray level image for faster processing and reducing unnecessary information. The disparity image created by the P2P algorithm is a binary image containing pixels that inform the relationship between left and right images. The disparity image is used to find the V-disparity image which is the summary of the disparity values in each scan lines. After that, the next process is finding depth lines from the V-disparity image using the Hough transform algorithm. Then, comparing the lines with the V-disparity can inform the object distances.

3.2 Low-cost Stereo Cameras

As normal standard stereo cameras for general stereo vision systems are rather expensive, we applied a pair of much cheaper web cams to be stereo cameras. These low-cost stereo cameras, however, are less accurate than the standard ones and may produce more noises.

Our stereo cameras are Logitech QuickCam webcams for Notebooks Pro [15]. As we want to attach the stereo cameras at the head of the visually impaired like wearing glasses or a helmet, the distance of the base line should not be more than the diameter of the head. In our work, we specify the maximum diameter to be not more than 12 centimeters which is the average head size. The design of the stereo cameras is as shown in Figure 3 and Figure 4 shows how they look like.

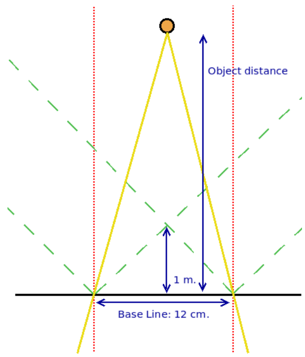

Figure 3. Design of our stereo cameras

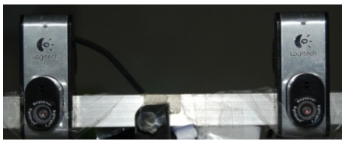

Figure 4. Our stereo cameras using webcams

3.3 Object Distance Estimation with V-disparity

The disparity image created by the P2P algorithm is a binary image that contains corresponding pixels of left and right images. The V-disparity image is calculated from the summation of the disparities in each scan line. Each vertical straight line in V-disparity image presents a depth distance. The Hough transform algorithm is then used for finding depth lines from the V-disparity image. Finally, the distance information in the depth line will be compared with the V-disparity in order to find the object distance.

4. RESULT AND DISCUSSION

4.1 Calibration and Distance Estimation

4.1.1 Stereo camera calibration

The V-disparity image transformed from the disparity image contains several straight lines. Each straight line has its length equals to the height of the associated object from the disparity image and is far from the left edge with an inverse variation to the distance from the stereo cameras to the object. We can find the distance to the object from the straight line from the V-disparity by comparing with the distance of the object measured for comparison in prior experiments.

Table 1 shows the relationship between the object distances and V-disparity line pixel positions (minimum, medium and maximum). Each distance range refers to the prior testing dataset. The data is processed to find the medium for distance estimation. The matches between object distances and pixel positions in the V-disparity can be estimated by the 6^{th} order polynomial equation shown in Equation 1. Figure 5 shows a graph plotting to fit Equation 1 using the medium dataset. In Equation 1, d is the distance to the obstacle and x is the value of the vertical line from the x axis in the V-disparity.

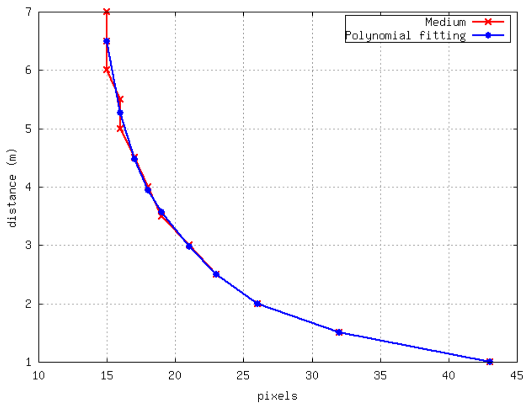

Figure 5. Graph shows the medium dataset fit for a 6^{th} order polynomial equation.

Table 1. Matches between real object distances and pixel positions from the V-disparity by finding the medium

Object distance (meters)	Pixel Position		
	Minimum	Medium	Maximum
1	42	43	44
1.5	31	32	34
2	26	26	27
2.5	21	23	24
3	20	21	21
3.5	19	19	20
4	18	18	18
4.5	17	17	17
5	16	16	16
5.5	16	16	16
6	15	15	15
6.5	15	15	15
7	15	15	15

$$d = 1.7938 \times 10^{-6}x^6 - 2.9101 \times 10^{-4}x^5 + 0.019294x^4 - 0.67064x^3 + 12.925x^2 - 131.57x + 559.10 \qquad \ldots(1)$$

4.1.2 Object distance estimation

In order to estimate the object distance, we use the V-disparity from the P2P algorithm to find straight lines using the openCV function *cvHoughLine* that gives 2 sequence pairs which are the beginning and ending sequence pairs of the straight lines. The value in the x axis represents the object depth and is used for replacing the variable x in Equation 1 in order to find the object distance. Currently, our system can find object distances within the range of 5 meters from the stereo cameras.

4.2 Enhancing P2P using MPI

Nowadays, multi-core computers are produced and sold widely in the market, i.e., Intel Core 2 Duo, Intel CORE i7, AMD Phenom, AMD Athlon64 X2, etc. In order to fully utilize the CPU, parallel computing technique is needed. In this work, we applied the parallel computing using MPI (MPICH implementation) to help reduce the computing time. In our work, we use the *ch_shmem* device of the MPICH [13] that is suitable for the symmetric multiprocessors (SMPs) architecture used in the 2-core PC notebook.

The workflow of scan line matching process is suitable for parallelism because it independently computes each scan line. Figure 6 shows the example of the data distribution in the scan line matching process of the P2P algorithm. For each image frame in the scan line matching process, we divide left and right image data into two segments and then distribute them to be separately processed in two processes. The first process computes the top left and right image data, and the second process computes the bottom left and right. After we combine the outputs and create a disparity image.

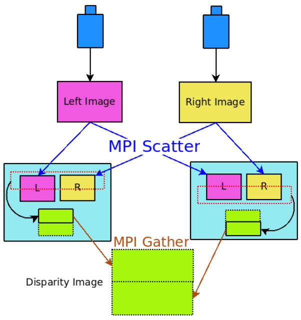

Figure 6. Overview diagram of data distribution

The parallelized P2P algorithm can reduce the total response time of the program to 0.79 second at the maximum disparity of 100, 320x240 pixels when applying two processes on a 2-core computer. The total computing time of the sequential algorithm was 1.365 second per frame. The parallelized version can reduce the computing time to about a half. Using more processes than the number of CPU cores is not recommended since it is not faster.

4.3 Limitations

Using the P2P algorithm which is an algorithm using ISM techniques requires the most parallel images so that the algorithm can accurately detect obstacles. Our prototype of stereo cameras is made of webcams that have the maximum based line diameter not more than 12 centimeters in order to enable the attachment to the head of the visually impaired. Consequently, the prototype can accurately detect obstacles only within the range of 5 meters. For further distances than 5 meters, the accuracy of the system is variable. It can detect an object of 65x65 centimeters at the distance of 10 meters but the object appeared so small that may be mixed up with noises. However, using better quality cameras will increase the detecting distance and accuracy.

The speed of the program is 0.790 second for the image of 320x240 pixels at the maximum disparity of 100. It can be used to detect objects with slow movement, i.e., human being's walk. It cannot properly detect objects with fast movement, i.e., car, in real time.

5. CONCLUSION

Our obstacle detection for the visually impaired using stereo vision and parallel computing reduces the response time of the depth discontinuities P2P stereo algorithm in the part of scan line matching using MPI running on a 2-core computer. The object distances are found by applying the V-disparity.

Our system can accurately detect objects with slow movement within the range of 5 meters using 12-centimeter based-line low-cost webcams. However, more development and detail experiments are needed to provide more information to the users.

The possible future work can be building a suitable interface for the visually impaired and a warning system in case of getting close to dangerous objects. Applying stereo vision with pattern recognition can also help informing more details of the environment to the users. In the future, when prices allow, using more processes up to the number of cores on a higher performance multi-core computer, says 4-core or 8-core, will certainly reduce much of the computing time.

6. ACKNOWLEDGMENTS

The authors are grateful for the support of the PSU Grid Center of Prince of Songkla University and the Intel Company.

7. REFERENCES

[1] Ando, B. Electronic Sensory Systems for the Visually Impaired. *Instrumentation & Measurement Magazine , IEEE Volume 6*, Issue 2, pp. 62 - 67, June 2003.

[2] Cristina Nicolescu, Pieter Jonker Lecture. Parallel low-level image processing on a distributed-memory system. *Proceedings of the 15 IPDPS 2000 Workshops on Parallel and Distributed Processing*. pp. 226 - 233, 2000.

[3] Hashimoto, H., Magatani, K., Yanashima, K. The development of the navigation system for visually impaired persons. *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE Volume 2*, pp. 1481 - 1483 vol.2, 25-28 Oct. 2001.

[4] Helal, A., Moore, S.E., Ramachandran, B. Drishti: an integrated navigation system for visually impaired and disabled. *Wearable Computers, 2001. Proceedings. Fifth International Symposium*, pp. 149 - 156, 8-9 Oct. 2001.

[5] Meers S., Ward, K. A vision system for providing 3D perception of the environment via transcutaneous electro-neural stimulation. *Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference*, pp. 546 - 552, 14-16 July 2004.

[6] Thomas Bräunl. Tutorial in Data Parallel Image Processing. *Australian Journal of Intelligent Information Processing Systems (AJIIPS)*, vol. 6, no. 3, pp. 164–174 . 2001

[7] Birchfield, S., Tomasi, C. Depth discontinuities by pixel-to-pixel stereo. *Computer Vision, 1998. Sixth International Conference*. pp. 1073-1080. 4-7 Jan 1998.

[8] Birchfield, S., Tomasi, C. A pixel dissimilarity measure that is insensitive to image sampling. *Pattern Analysis and Machine Intelligence*, IEEE Transactions. pp. 401-406, Volume: 20, 4, Apr 1998.

[9] Labayrade, R. Aubert, D. Tarel, J.-P. Real time obstacle detection in stereovision on non flat road geometry through "v-disparity" representation. *This paper appears in: Intelligent Vehicle Symposium*, IEEE. pp. 646- 651 vol.2, 17-21 June 2002.

[10] Robyn Owens. Stereo. http://homepages.inf.ed.ac.uk/rbf/ CVonline/ LOCAL_COPIES/OWENS/LECT11/lect11.html. Available on 4 Jan 2009.

[11] Stereo vision starts with two views. http://www.vision3d.com/ stereo.html. Available on 4 Jan 2009.

[12] The Message Passing Interface (MPI) standard. http://www-unix.mcs.anl.gov/mpi/. Available on 4 Jan 2009.

[13] MPICH-A Portable Implementation of MPI. http://www-unix.mcs.anl.gov/mpi/mpich1/. Available on 4 Jan 2009

[14] CvAux. http://opencv.willowgarage.com/wiki/CvAux# FindStereoCorrespondence. Available on 4 Jan 2009.

[15] Scott Clark. Logitech QuickCam for Notebooks Pro. http://www.everythingusb.com/logitech_quickcam_for_note books_pro.html. Available on 4 Jan 2009.