The stereo vision and distance sensors fusing for the visually impaired

Teerapon Chongpipattanasiri[1], Watcharin Kaewapichai[2], Pichaya Tandayya[*]

Computer Engineering, Prince of Songkla University, Hat Yai, Songkhla 90112
E-mail: 5410120072@email.psu.ac.th[1], watcharin@coe.psu.ac.th[2], pichaya@coe.psu.ac.th[*]

Abstract

Classical methods to support the visually impaired include using a cane and a guided dog. These methods enable the visually impaired to identify only nearby obstacles. However, they fail to allow them to detect obstacles that are located relatively farther and are mostly visible to normal sighted people. In order to address this problem, we propose the Pixel-to-Pixel (P2P) stereo matching algorithm to find objects that are in front of the visually impaired. The computing time of this algorithm is reduced by the use of parallel computing that uses Message Passing Interface (MPI), but the algorithm results in noisee or inaccuracy of the object distance approximation. In this work, the distance data from the stereo vision and distance sensors are used to approximate the distance and increase the accuracy by applying the discrete Kalman filter. This paper presents an obstacle detection system for the visually impaired by enhancing P2P with MPI while Kalman filter is applied to improve accuracy of system.

Keywords: Stereo Vision, Sensor Fusion, the Visually Imparied.

1. Introduction

Traditional navigation systems for the visually impaired are white canes and guide dogs. The white cane is used to find objects on the ground sight within the range of 1 meter. Its risk is that the accident is very high, especially in case of "floating objects obstructed along the walkway" such as tree-branches, damaged wire, and others. Possible working areas of the traditional navigation systems are limited. Therefore, many support devices and electronic technologies have been developed for so many years to help the visually impairered's walk, called "Electronic Travel Aids (ETA)" [1]. ETA has been classified into two groups which are the distance sensor system and stereovision system. Two popular products from distance sensor systems are ultrasonic canes [2] and glasses [14] communicating to the user with sounds and/or vibration. Ultrasonic wave is the most useful and produces high accuracy in approximating the distance of objects along the walkway. Comparing with another distance measurement, the laser-based sensor system, the ultrasonic system gives lower resolution and magnitude and is less expensive. Advantages of the ultrasonic system are that it can detect the obstacles with high confidence and require no light sources. The ultrasonic system provides a good result of object distances within the range of 5 meters and 30 degrees of the working area. However, the working area of the ultrasonic system is limited, and only the closest objects in the working area are detected with poor angular resolution.

The stereovision system has been popularly mounted on helmets and communicates to the users with sounds [3] and vibration [4]. The distance of the object is estimated by a disparity image created by stereo matching algorithms. These algorithms have been classified into two groups: Intensity-based Stereo Matching (ISM) and Feature-based Stereo Matching (FSM) techniques [5]. The ISM technique [6, 7] is a matching algorithm of the pixel to pixel (P2P) algorithm by finding the intensity of pixels. An advantage of ISM is that this system can work very well with low texture images. However, the computational cost is the main problem of this technique. There are other techniques applied to reduce the computation time of the P2P technique: Message Passing Interface (MPI) [9] and V-disparity [10]. The accuracy of both stereo vision systems is within the range of 5 meters using low-cost web cams. However, the stereo vision system does not work if there is no external light source.

In this paper, we present a new ETA for the visually impaired, integrating the advantages of the ultrasonic sensor system and stereo vision system in order to increase the efficiency and effectiveness .

2. System overview

Our system consists of two main parts: Distance measurement and Data Fusion. Distance measurement values are extracted from the stereo vision device and ultrasonic sensors. The distance values are fed into Data Fusion. Sensor noise, system noise, and user movement are modeled by the Kalman filter which is an algorithm to increase the accuracy of the system by linear approximation model. However, the reliability areas of both systems are overlapped with each other. Data Fusion is applied to increase the reliability and accuracy of the system.

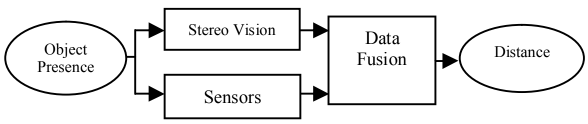

Figure 1. Flow chart of our system.

2.1 System design

There are three main scopes of the system: 5 meters working area, mobility system (light weight and low power consumption), and keeping it low cost. Two webcams and 3 ultrasonic sensors are mounted on a safety helmet. The distance between the cameras is 12 centimeters and the working area of the ultrasonic system is 15 degrees as shown in Figure 2.

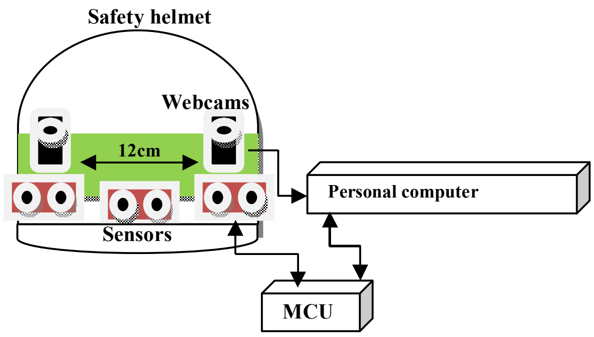

Figure 2. The prototype of the system.

Two images from webcams are fed into the embedded personal computer (EPC) for "Stereo Distance Estimation" explained in Section 2.2. Three distance values from ultrasonic sensors are calculated by the microcontroller (Atmega 168) and sent to EPC via RS-232.

2.2. Stereo vision and distance sensors fusion for the visually impaired

1 Distance Estimation of the Stereovision

We use the V-disparity [10] technique to approximate the distances of objects. The V-disparity is calculated from the disparity map produced from the stereo image processing in the Intensity-based Stereo Matching category. We use the Pixel-to-Pixel stereo matching technique [7] to create a disparity map because it is a widely used technique in intensity matching.

1. Pixel-to-Pixel stereo matching algorithm

Birchfield and Tomasi proposed Pixel-to-Pixel Stereo Matching (P2P) [7] in 1998. There are two processes: scan line matching and post-processing. Scan line matching matches equivalent pixels between left and right images on the same scan lines. Post-processing manipulates data between the scan lines, and selects the best disparity value to identify the position. As a result, the scan line matching process takes a long computing time when running the P2P algorithm sequentially. To reduce the cost of computation, Parallel programming [9] is applied because the scan line matching process independently computes in each scan line.

2. V-disparity

The disparity map created from the P2P algorithm is the depth distance equivalent pixels of left and right images. The V-disparity map is accumulative of the disparities in each scan line. Hough transform is used to find the depth line in V-disparity. It represents the depth distance information uses for approximate the object distance. The depth line from the V-disparity compared with the real distance of object is shown in Figure 3.

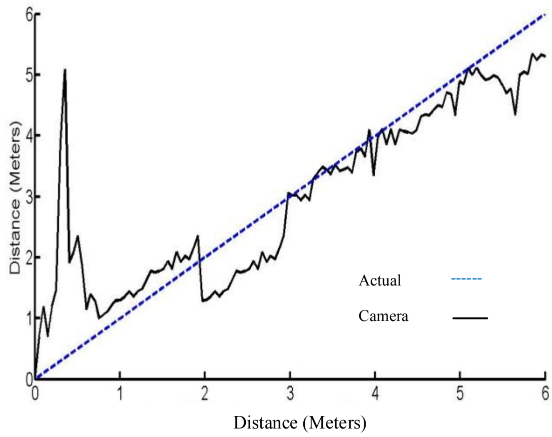

Figure 3. Comparing the V-disparity distance with the actual distance.

The dash line is the actual value. The continuous line is the distance data from V-disparity (100 measurements per step).

2.3 Distance Estimation of Distance Sensors

In this paper, we use the PING [15] ultrasonic sensor for object detection by sending a short ultrasonic burst and waiting for reflects from the echo PING signal from the object. The ultrasonic burst wave is 40 kHz. The burst travels through the air with the speed of 340 meters per second, when the burst hits an object and echo-back to the sensor. The time travel of the burst signal from the sensor to the object and back to the sensor is used to calculate the distance as shown in Equation 1 as follows:

$$s = vt/2 \quad (1)$$

When s is the distance between the sensor and the object, v is the speed of the ultrasonic wave (340 meters/second), and t is the time travel of the burst signal from the sensor to the object and back to the sensor.

In Figure 4, the actual and measurement distances are plotted to compare the accuracy. The results show that the ultrasonic sensor working area is within the range of 4 meters.

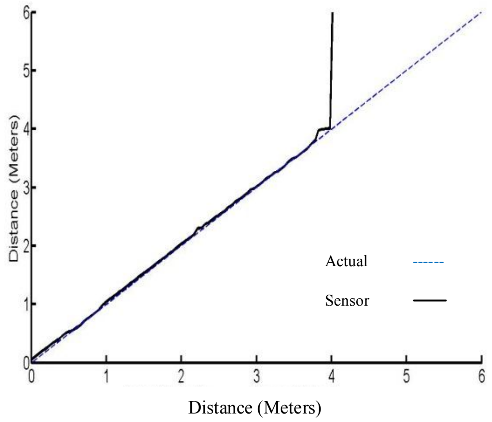

Figure 4. Distance measurement and actual distances of the ultrasonic sensor.

3. The Discrete Kalman Filter Algorithm

In 1960, R.E. Kalman published his famous paper describing a recursive solution of the discrete-data linear filtering problem. Kalman filter is a set of mathematical equations [12] which can be applied to the distance Measurement and Multi-sensor Data Fusion [13]. Kalman filter is a process that is a forms of feedback control: Kalman filter runs the process state at some time and obtains feedback in forms of noise measurements of the Kalman filter algorithm which are classified into two groups: time update equations and measurement update equations. The time update equations are responsible for analyzing the current state and error covariance estimation to obtain the priori estimates for the next step shown in Figure 5. The measurement update equations are responsible for feedback of incorporating new measurement into the priori estimate. It is a method to improve the posteriori estimate for next iteration.

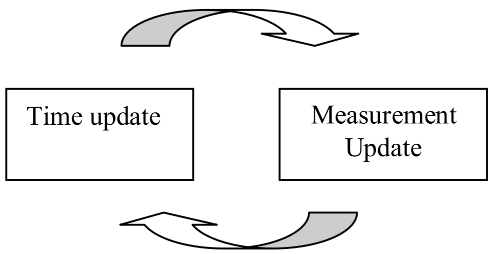

Figure 5. Discrete Kalman filter cycle.

In this paper, we apply the discrete Kalman filter (DKF) for distance measurements (Ultrasonic and Stereo vision). There are two main processes. First is estimating distances by using the initial distance values from ($\hat{X}_{k-1}$) in Equation 2. Initially, a priori (P_k^-) is an estimation error covariance by processing noise covariance from (Q_k) in Equation 3. Second, measurement update shown in Equation 3 uses the result from the first process ($\hat{X}_k^-$) and the result from a sensor to update the distance value of the system ($\hat{X}_k^+$) In Equation 4. We used the Kalman gain (K_k) In Equation 5 to update the final result ($\hat{X}_k^+$)). The Posteriori estimates error covariance (P_k^+) In Equation 6 is used to find the Kalman gain (K_k). R is a noise measurement covariance. The process of Kalman filter is used to reduce noise from systems by using a form of feedback control.

$$\hat{X}_k^- = \Phi_{k-1}\hat{X}_{k-1} \quad (2)$$

$$P_k^- = \Phi_{k-1}^2 P_{k-1} + Q_k \quad (3)$$

$$\hat{X}_k^+ = \hat{X}_k^- + K_k\left(z_k - H_k\hat{X}_k^-\right) \quad (4)$$

$$K_k = P_k^- H_k (P_k^- H_k^2 + R_k)^{-1} \quad (5)$$

$$P_k^+ = (1 - K_k H_k)^2 P_k^- + K_k^2 R_k \quad (6)$$

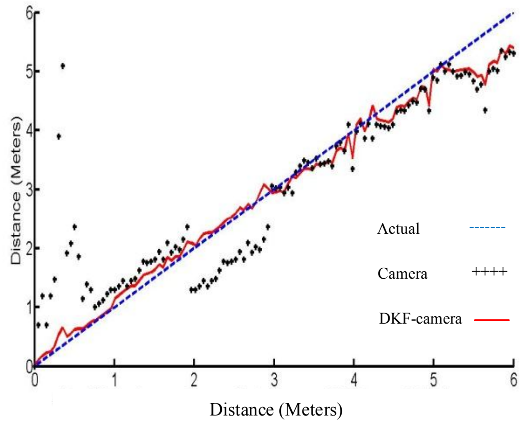

Figure 6. Result of applying Discrete Kalman filter (DKF) to the stereo vision.

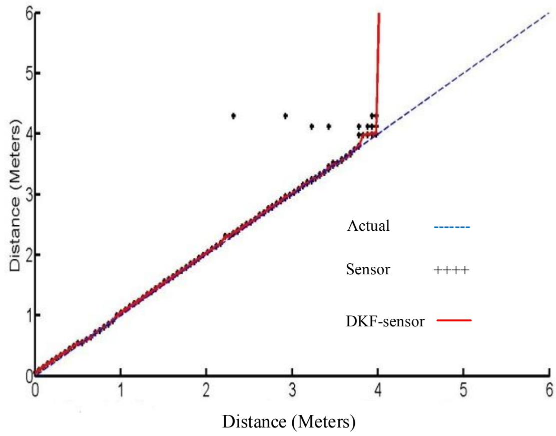

Figure 7. Result of applying Discrete Kalman filter (DKF) to the ultrasonic sensor.

4. Data Fusion

We set off connections with three front overlapping sonar sensors and the stereo vision in the viewing area. After running the Kalman filter process to improve the accuracy of distance approximation, there are 4 results from stereo vision and sensors. We can combine results from sensors and stereo vision by using the value of confidence that is the probability of sensors and stereo vision. The result shows that the output from this step is more accurate as shown in Figure 8.

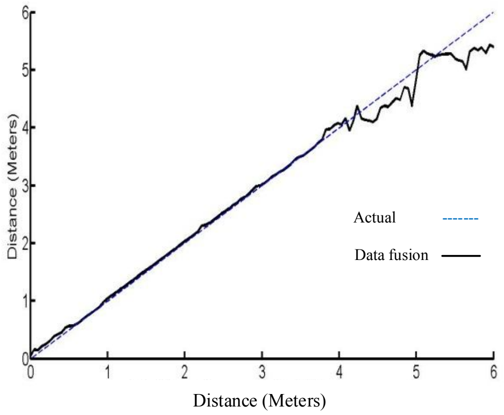

Figure 8. Comparing the result of the data fusion distance with the actual distance

A basically method to fusion measured distance from both sensors is weighting shown in Equation 7.

$$Distance = \alpha D_{stereo} + \gamma D_{sensor} \quad (7)$$

where *Distance* is fusion distance measurement

D_{stereo} is distance answer from stereo vision.

D_{sensor} is distance answer from sonar sensor.

$$\alpha = \begin{cases} 0.5 & if\ D_{stereo} < 4 \\ 1 & if\ D_{stereo} > 4 \end{cases}$$

$$\gamma = \begin{cases} 0.5 & if\ D_{sensor} < 4 \\ 0 & if\ D_{sensor} > 4 \end{cases}$$

5. Result and discussions

5.1 Stereo vision distances estimation

The result from stereo vision is non-linear and made-up with noises. Stereo vision algorithm working area is within the range of 5 meters. There are 3 positions of distances which are left, right and center positions. The primary advantage of stereo vision is that it gives good angular resolution and can be done at low cost, but the output becomes noisy due to smooth surfaces or low contrast image. The result is shown in Figure 6.

5.2 Sensors distances estimation

The result from ultrasonic sensors is linear and gives high accuracy sensors is that they do not require 1-4 meter distance. The primary advantage of sensor is not requires light source. The result is shown in Figure 7.

6. Conclusion

In our work, we have implemented the supporting system for the visually impaired using the fusing distances from the ultrasonic sensors and the stereo vision applying parallel computing in order to reduce the response time. We have used stereo vision and sensors of which the estimation of relative distances is improved by using the Kalman filter algorithm. This system can accurately detect objects within the range of 5 meters, 60 degree angular resolutions and 0.79 second for the image of 320x240 pixels. The future work can be construction of an appropriate interface for the visually impaired in 3D sound.

Acknowledgments

This work was supported by the Higher Education Research Promotion and National Research University Project of Thailand , Office of the Higher Education Commission.

References

[1] Rusha Patra, M. Manjunatha, J. Mukhopadhyay and A. K. Majumdar Amit,An Electronic Travel Aid for Navigation of Visually Impaired Persons, (COMSNETS), 2011 4-8 Jan. 2011.

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

[3] L. Dunai, G. P. Fajarnes, V. S. Praderas, B. D. Garcia, L. Lengua, Real–Time Assistance Prototype –a new Navigation Aid for blind people, IECON, 2010.

[4] D. Dakopoulos, N. Bourbakis Preserving, Visual Information in Low Resolution Images During Navigation of Visually Impaired, PETRA 08, 2008.

[5] Meers S., Ward, K. A, Vision system for providing 3D perceptionof the environment via transcutaneous electro neural stimulation. Information Visualisation, 2004. IV 2004, 14-16 July 2004.

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

[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. Volume:2 ,4 Apr 1998.

[9] Thanathip Limna, P. Tandayya and N. Suvanvorn, Low-cost Stereo Vision System for Supporting the Visually Impaired's Walk, I-CREATe , 22-26 April 2009.
[10] Labayrade, R. Aubert, D. Tarel, J.-P. Real time obstacle detection in stereovision on non flat road geometry through"v-disparity" representation, IEEE 17-21 June 2002.
[11] Roger Meier1, Terrence Fong, Charles Thorpe, and Charles Baur, A Sensor Fusion Based User Interface for Vehicle Teleoperation, IEEE , 1999.
[12] Greg Welchand Gary Bishop, An Introduction to the Kalman filter, July 24, 2006.
[13] Brown, Robert Grove and Patrick Y. C. Hwang. Introduction to random signals and applied Kalman filtering. John Wiley & Sons., 1997.
[14] Nikolaos Bourbakis, Sensing Surrounding 3-D Space for Navigation of the Blind, Engineering in Medicine and Biology Magazine IEEE, Jan.-Feb. 2008.