An Electronic Travel Aid for Navigation of Visually Impaired Persons

Amit Kumar
Checktronix India Pvt. Ltd.
Chennai, India
amitkumar@burning-glass.com

Rusha Patra, M. Manjunatha, J. Mukhopadhyay and A. K. Majumdar
Indian Institute of Technology, Kharagpur
Kharagpur, India

***Abstract*—This paper presents an electronic travel aid for blind people to navigate safely and quickly, an obstacle detection system using ultrasonic sensors and USB camera based visual navigation has been considered. The proposed system detects the obstacles up to 300 cm via sonar and sends feedback (beep sound) to inform the person about its location. In addition to this, an USB webcam is connected with eBox 2300™ Embedded System for capturing the field of view of the user, which is used for finding the properties of the obstacle in particular, in the context of this work, locating a human being. Identification of human presence is based on face detection and cloth texture analysis. The major constraints for these algorithms to run on Embedded System are small image frame (160x120) having reduced faces, limited memory and very less processing time available to achieve real time image processing requirements. The algorithms are implemented in C++ using Visual Studio 5.0 IDE, which runs on Windows CE™ environment.**

Keywords-Electronic Travel Aid; Visual Impairment; Navigation Aid; Mobility; Ultrasonic Range Sensor; Embedded System; Face Detection

I. INTRODUCTION

Mobility is one of the main problems encountered by visually impaired persons in their daily life. Over decades, these peoples were using navigational aids like white cane, guide dogs etc. Long white cane is a traditional mobility tool used to detect obstacles in the path of the blind person. The length of the white cane depends upon height of the user and extends from the ground to the user's sternum. On the other hand, guide dogs are assistance dogs, trained to lead visually impaired around obstacles. Due to the development of modern technology, many different types of navigational aids [1] are now available to assist the blinds. They are commonly known as Electronic Travel Aids.

Some of these aids are Sonic Pathfinder [2], Mowat-Sensor [3] and Guide-Cane [4], but having very narrow directivity. However, Sonic-Guide [5], NavBelt [6] and other ETA devices [7, 8] are having wide directivity and are able to search several obstacles at the same time. These devices are all based on producing beams of ultrasonic sound or laser light. In such a system, the device receives reflected waves, and produces either an audio or vibration in response to nearby objects, but its market acceptance is rather low as useful information obtainable from them are not significantly more than that from the long cane and responses received from them are not user friendly. So recent research efforts are being directed to produce new navigational system in which digital video camera is used as vision sensor. Some of these ETA devices are vOICe [9], NAVI [10], SVETA [11] and CASBLIP [12].

In vOICe, the image is captured using single video camera mounted on a headgear and the captured image is scanned from left to right direction for sound generation. The sound is generated by altering the top of the image into high frequency tones and the bottom portion into low frequency tones. The loudness of sound depends on the brightness of the pixels. Similar work has been carried out in NAVI where the captured image is resized to 32x32 and the gray scale of the image is reduced to 4 levels. The image is differentiated into foreground and background using image processing techniques. The foreground and background are assigned with high and low intensity values respectively. Then the processed image is converted into stereo sound where the amplitude of the sound is directly proportional to intensity of image pixels, and the frequency of sound is inversely proportional to vertical orientation of pixels. In SVETA, an improved area based stereo matching is performed over the transformed images to compute dense disparity image. Low texture filter and left/right consistency check are carried out to remove the noises and to highlight the obstacles. To map the disparity image to stereo musical sound, sonification procedure is used. In CASBLIP, the object is detected through sensors and stereo vision. In addition to this, orientation is computed using GPS system. This system is embedded on the Field Programmable Gate Array (FPGA).

The most important factors, which enable blind users to accept these devices readily, are portability, low cost, and above all simplicity of controls. Hence, ETA device should be small in size and lightweight for portability. Since a blind person is not able to see the display panel or control buttons, hence the device should be easily controllable. Moreover, the ETA device should be of low-cost so as to be affordable by a common man. Considering all these requirements, eBox 2300™ is used as a processing unit in the proposed system. The processing of the ETA device is performed on this embedded system. It controls all the modules which are used to navigate the user.

II. AN OVERVIEW OF THE ELECTRONIC TRAVEL AID

The system is based on an embedded system eBox 2300™ [13], a small (4.5" x 4.5") low-cost X86 processor based

978-1-4244-8953-4/11/$26.00 ©2011 IEEE

embedded computer system. The ultrasonic sensors are connected with sensor circuit. It feeds the distance data to eBox 2300™ through a RS-232 serial cable. A USB webcam is connected with eBox 2300™ for capturing the field of view of the person, which is used for locating a human being. A headphone is connected with eBox 2300™ to get the audio feedback (beep sound) of the obstacle distance and presence of human being. The eBox 2300™ is powered by 5V, 3A DC adapter and sensor circuit is powered by two 9V alkaline batteries. The algorithms are implemented in C++ using Visual Studio 5.0 IDE, which runs on WinCE environment.

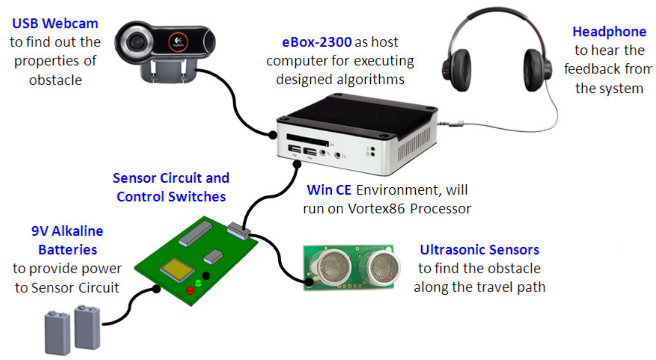


Figure 1. Proposed Electronic Travel Aid

Figure 1 depicts the proposed system for ETA. For better field of view, USB webcam is mounted on a helmet and ultrasonic sensors are placed on the user's belt. Three easy control switches are provided to control the ultrasound based distance measurement system, human detection system and motion detection system respectively. The eBox 2300™ and sensor circuit are kept in the bag which will be held on the waist of the user. The user has to operate the system manually and he/she will get the auditory feedback till the switches are pressed.

III. Ultrasound Based Distance Measurement

In this work ultrasonic sensors are used to detect the obstacles in the path of the blind person. The overall operation is divided into two parts. First part consists of a development of sensor unit and the other part deals with processing of sensor circuit data (distance data) in eBox 2300™. Sensor unit has three subunits: transmitter, receiver and processing units. The transmitter unit generates 40 kHz ultrasonic signal and the receiver unit receives the reflected echo signal. Finally, the processing unit computes the time difference between the transmitting and receiving pulses. The sensor unit provides the raw distance data in the form of pulse count as its output. These counts are used to compute actual distance through processing in eBox 2300™.

The pulse count data is obtained as raw data for computation of distance of the object from sensor unit through RS232 cable. This data is processed in eBox 2300™ embedded system. The eBox 2300™ generates a temporary file to store the serial data coming through its COM port. A C++ program running on eBox 2300™ reads this file to process the data. Finally, object distance is calculated and accordingly the feedback (beep sound) is provided to the user.

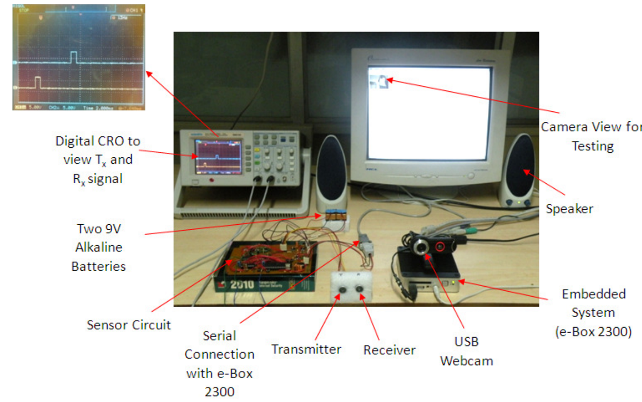


Figure 2. Experimental setup of the ETA device

The performance of the ultrasound based measurement system has been evaluated by using the experimental setup as portrayed in the Figure 2. The system is calibrated by placing an obstacle at a measured distance in front of the sensor. The calibrated distance (*d_cal*) is measured by using first order interpolation and the mapping between the pulse count and the distance. This calibrated distance is compared with distance computed from the velocity of sound (*d_v*). From the experimental data it is found that for distances 15cm to 150 cm, there is no error between the two distances (*d_cal* and *d_v*). However, after a distance of 150cm, due to one to many mapping, there is error in the calibrated distance. In order to reduce this error, average of three consecutive calibrated distances has been considered. The system detects the obstacles up to 300 cm. The system is also inert to background noises since the ultrasound frequency (40 kHz) is well beyond the audible frequency range (20 Hz – 20 kHz).

IV. Detection of Human Presence

The detection of human presence is carried out by detecting the human face. However there are situations, when the face is not present in the field of view of the camera in spite of the presence of a human in front of the visually impaired person. Such type of presence of human being may be asserted by detecting cloth and human skin. To detect the presence of human, if cloth is found in the vicinity of human skin, and face is not detected (if Side faces), then it will be considered as human.

A. Face Detection

Face detection method consists of three main processing blocks: segmentation of skin regions, computation of shape features and classification based on thresholding. In segmentation of skin region, the input *RGB* image is transformed to YC_rC_b color space and the skin pixels are detected based on their chrominance values. A connected component labeling algorithm using contour tracing technique is used to find out largest skin region in the image frame. This is followed by the computation of the shape features of the largest skin region.

ICOP-eBox 2300™ and Windows Embedded™ are respective Trademarks of Microsoft Corporation and ICOP Technology.

The computed shape features for making decisions are as follows:

a) F_{be}: Ratio between area of bounding box and area of fitted ellipse.

b) F_{ab}: Ratio between semimajor and semiminor axes of fitted ellipse.

c) F_{oe}: Ratio between total number of pixels in skin region and number of pixels inside the fitted ellipse of that skin region.

Classification of human face is performed by thresholding different shape parameters. The feature values are summarized into mean (μ), standard deviation (σ) and the 3σ-control limits are taken as threshold. To analyze the features, our in-house database is used, where 240 frontal face images of size 160x120 of twelve individuals have been taken at distances 30 cm, 60 cm, 90 cm and 120 cm. To examine the accuracy of the algorithm, 60 object images (non-face) have been taken randomly.

The final threshold values of each feature are given in Table I. Considering normal distribution of the features, the probabilities of misclassification or error (Type-I error) [14] is calculated and are given in Table I. Constrained by real time requirements of processing, proposed methodology is based on simple heuristics. Yet it is found to be more effective than other existing algorithms in the given scenario. This algorithm is implemented on Windows CE based eBox 2300™ embedded system.

TABLE I. SUMMARY STATISTICS AND CONTROL LIMITS OF THE FEATURES

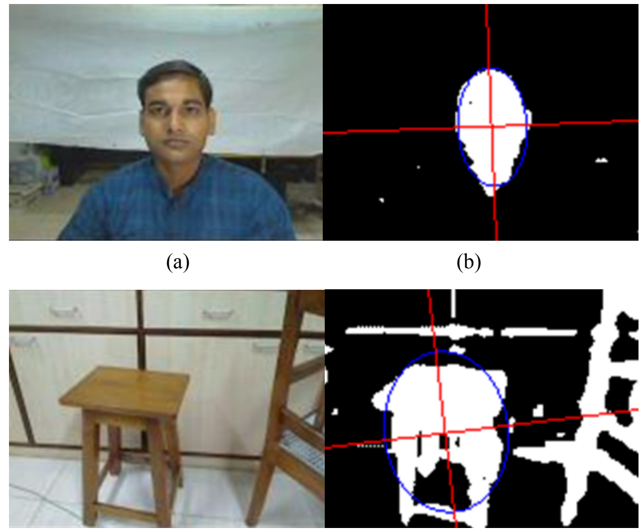
Feature	Lower threshold	Upper threshold	Type-I error
F_{be}	1	1.5	0.06
F_{ab}	1.5	2.4	0.1
F_{oe}	1.05	1.15	0.1

The major constraints for this algorithm to run on embedded system are small image frame (160x120) having reduced faces, limited memory and very less processing time available to achieve real time image processing requirements. USB webcam connected with eBox 2300™ embedded system is providing image frames at a rate of 15 fps, which restricts the execution time of algorithm within 66 ms. Taking all the above constraints into consideration, the proposed face detection algorithm targets the basic shape features of object to classify human face to reduce computational complexity. The execution time of proposed algorithm is 20 ms.

Figure 3 represents one face and one non-face object images and their results. In non-face image though false skin region has been segmented out, but its feature values are not satisfying all three conditions (F_{be} = 1.48, F_{ab} = 1.28 and F_{oe} = 1.188), hence it is classified as non-face.

B. Detection of Cloth

The image is subdivided into 48 non-overlapping subimages of size 20x20 pixels and these sub images are processed by using 'Haar' wavelet decomposition at level 1. The energy values of the approximate (eA) and detail coefficients (eH for horizontal, eV for vertical and eD for diagonal) coefficients are computed for each sub image. These energy values are considered as features to classify cloth texture.



(a) (b) (c) (d)

Figure 3. (a), (c) Input images, (b), (d) Result of ellipse fitting on skin region.

(a) (b)

Figure 4. Results of Cloth Detection algorithm for cloth at distance 120 cm (a) Input image (b) Resultant Image.

The feature values are summarized into mean (μ) and standard deviation (σ) and the 3σ-control limits are taken as threshold, as given in Table II. The reason behind taking 3σ-control limits is that, irrespective of the sampling distribution of the feature, the probability of having values in the interval (μ-3σ, μ+3σ) is very large (nearly 0.9973). Prior to classification, the significance of the features is tested using unpaired student's t-test [14] and the p-values are given in Table II. It is observed from the table that all the features are significant at 1% level of significance. Using these thresholds, the decision has been made whether the texture is cloth or not.

The result of the cloth detection algorithm is shown in Figure 4. The input image is having different cloth textures at 120 cm distance. In the resultant image white 20x20 pixel blocks (subimages) are the detected cloth regions. Along with the tee-shirt of the person, the algorithm also detects the cloth in the background of the image.

TABLE II. SUMMARY STATISTICS AND CONTROL LIMITS OF THE FEATURES

Feature	Mean (µ)	SD (σ)	µ-3σ	µ+3σ	p-value
eA	99.5233	0.8274	97.0411	102.0055	0.0011
eH	0.0480	0.0829	-0.2007	0.2967	0.0000
eV	0.4234	0.8140	-2.0186	2.8654	0.0027
eD	0.0052	0.0113	-0.0287	0.0391	4.159×10^{-11}

V. RESULTS AND DISCUSSION

In order to test the ETA device on the blind folded persons, a system prototype is developed as shown in Figure 5 where USB webcam is mounted on a helmet and ultrasonic sensors are placed on the user's belt. Three easy control switches are provided to control the ultrasound based distance measurement system, human detection system and motion detection system respectively. Sensor circuit unit and eBox 2300™ is carried by the user in the bag.

To evaluate the performance of the Electronic Travel Aid device, it is tested on trained and novice peoples in the laboratory environment. Three easy control switches are provided to manually operate the device. First switch is to find the obstacles in the path of the blind person and second one is used to find the human presence in field of view of the camera and the last switch is used to detect any movement in front of the person. The device provides auditory feedback to the user in response to the switch pressed. For example, if user presses the first button to find the obstacles on his/her path, device will produce the beep sound whose loudness will increase or decrease with respect to the obstacle distance. To easily operate the device and to understand the auditory feedback, a proper training is required.

Figure 5. Prototype of ETA device.

A total of eight tests have been carried out in laboratory environment and outside on three blind folded persons in which two are trained subjects and one subject is novice. After blindfolding the person, he/she is asked to walk through the corridor where different type of obstacles has been placed within 10 meter range. During the experiment, user's walking motion is recorded in video camera. Time taken by the users (trained and novice) for successfully walking through the obstacles is measured and travel speed for each test has been calculated as depicted in Table III.

TABLE III. PERFORMANCE ANALYSIS OF ETA DEVICE

	User Type	Obstacles	Human Presence	Cleared Obstacles	Travel Speed (m/s)
Test 1	Novice	5	No	5	0.40
Test 2	Novice	5	No	5	0.45
Test 3	Trained	5	No	5	0.90
Test 4	Novice	5	No	3	X
Test 5	Trained	5	No	5	0.83
Test 6	Trained	5	No	5	0.77
Test 7	Novice	7	Yes	7	0.66
Test 8	Trained	7	Yes	7	0.86

It is apparent from the Table III that average speed of a trained and novice users are 0.84 and 0.50 m/s respectively. In comparison with the traveling speed of the sighted people (1.3 m/s), this result is acceptable. The accuracy of the device in finding out obstacles is 95.45%. This result shows that training of the user is one of the important factors for gaining high traveling speed and also to increase user's confidence to choose optimal path.

To evaluate the detection range of the system, five tests has been performed in the laboratory environment. The response of the proposed ETA device for different type of obstacles is shown in Table IV. It is observed from the table that the range of detection of cardboard box is 210 to 285 cm whereas that of human body is 95 to 160 cm. It is apparent from this result that the refection or obstacle detection depends on the object surface, the more smooth the surface the higher will be the detection range. In addition to this, detection range of human presence by face detection is also tested which is 85 to 150 cm. The reason behind the low range for this is the small face size, which will be reduced further as the distance between the human and device increases.

TABLE IV. RESPONSE OF THE ETA DEVICE FOR DIFFERENT TYPE OF OBSTACLES

	Detection range (in cm)				
	Test 1	**Test 2**	**Test 3**	**Test 4**	**Test 5**
Reflection by concrete wall	185	210	195	205	200
Reflection by human body (static)	120	95	155	105	160
Reflection by cardboard box	230	210	285	245	225
Reflection by wooden material	150	185	160	190	210
Human presence by face detection	85	135	155	105	90

Figure 6 shows the ETA device testing, which have been performed on blind folded volunteers in different environments. Figure 6 (a) shows that a blind folded volunteer is crossing the hazards (pillars and carton) successfully in the laboratory environment and in Figure 6 (b) a blind folded volunteer is walking through the obstacles in the corridor.

(a)

(b)

Figure 6. Blind folded volunteers walking in the different environment (a) laboratory and (b) department corridor.

VI. CONCLUSION

An electronic travel aid to navigate visually impaired persons has been proposed here. The aid has been tested on blind folded volunteers in laboratory environment. Using this ETA device blind user can pass through the unknown environment independently. The major issues for the users to accept these aids are that they should be unobtrusive, easy to carry and for the convenience of the blind user, device should be small and light weight. The proposed system is designed considering all these factors. User needs to wear helmet in which camera and headphone are mounted. The user has to carry eBox 2300™ and sensor unit of nearly 500 gm. This device can detect obstacles up to 300 cm and human presence within 120 cm. A portable ETA device has been developed taking into account the blind users' requirements and it fills the gap between these requirements and the presently available aids.

ACKNOWLEDGMENT

This work was supported by the project grant 1(23)/2006-ME & TMD, Dt. 07/03/2007 sponsored by the Ministry of Communication and Information Technology, Govt. of India.

REFERENCES

[1] Yen, D H. "Currently Available Electronic Travel Aids for the Blind." September 21, 2005. http://www.noogenesis.com/eta/current.html.

[2] A. Dodds, D. Clark-Carter, and C. Howarth, “The sonic PathFinder: an evaluation,” Journal of Visual Impairment and Blindness, vol. 78, no. 5, pp. 206–207, 1984.

[3] A. Heyes, “A polaroid ultrasonic travel aid for the blind,” Journal of Visual Impairment and Blindness, vol. 76, pp. 199– 201, 1982.

[4] I. Ulrich, and J. Borenstein, “The guide cane-Applying mobile robot technologies to assist the visually impaired,” IEEE Transaction on Systems, Man, and Cybernetics-Part A: Systems and Humans, vol. 31, no. 2, pp. 131-136, 2001.

[5] J. Barth, and E. Foulhe, “Preview: A neglected variable in orientation and mobility,” Journal of Visual Impairment and Blindness, vol. 73, no. 2, pp. 41–48, 1979.

[6] S. Shoval, J. Borenstein, and Y. Koren, “The NavBelt- A computerized travel aid for the blind based on mobile robotics technology,” IEEE Transactions on Biomedical Engineering, vol. 45, no 11, pp. 1376-1386, 1998.

[7] L. Kim, S. Park, S. Lee and S. Ha, “An electronic traveler aid for the blind using multiple range sensors,” IEICE Electronics Express, vol. 6, no 11, pp. 794-799, 2009.

[8] C. Gearhart, A. Herold, B. Self, C. Birdsong, L. Slivovsky, "Use of ultrasonic sensors in the development of an Electronic Travel Aid," Sensors Applications Symposium, 2009. SAS 2009. IEEE, pp.275-280, 17-19 Feb. 2009.

[9] P. Meijer, “An Experimental System for Auditory Image Representations,” IEEE Transactions on Biomedical Engineering, vol. 39, no 2, pp. 112-121, Feb 1991.

[10] G. Sainarayanan, “On Intelligent Image Processing Methodologies Applied to Navigation Assistance for Visually Impaired”, Ph. D. Thesis, University Malaysia Sabah, 2002.

[11] G. Balakrishnan, G. Sainarayanan, R. Nagarajan and S. Yaacob, “Wearable Real-Time Stereo Vision for the Visually Impaired,” Engineering Letters, vol. 14, no. 2, 2007.

[12] G. P. Fajarnes, L. Dunai, V. S. Praderas and I. Dunai, “CASBLiP- a new cognitive object detection and orientation system for impaired people,” Proceedings of the 4th International Conference on Cognitive Systems, ETH Zurich, Switzerland, 2010.

[13] Hamblen, James O. "Using a Low-Cost SoC Computer and a Commercial RTOS in an Embedded Systems Design Course." IEEE Transactions on Education, vol. 51, no. 3, 2008.

[14] Gun, A. M., Gupta, M. K., Dasgupta, B. Fundamentals of Statistics. vol. 2, World Press. 2008.