2018 International Conference on Signal Processing and Information Security (ICSPIS)

X-EYE: A Bio-smart Secure Navigation Framework for Visually Impaired People

[1]Rabia Afzal Minhas, [2]Ali Javed, Member IEEE
afzalminhasrabia@gmail.com, ali.javed@uettaxila.edu.pk
[1]Computer Science Department, City University of Science and Information Technology, Peshawar, Pakistan.
[2]Software Engineering Department, University of Engineering and Technology, Taxila, Pakistan.

Abstract- This paper presents an effective method of providing day-to-day mobility aid to visually impaired people. An android application named X-EYE using LOOXCIE wearable camera is designed for blind people to navigate safely. Existing navigation aid systems use various hardware components such as sensors that are expensive and cause health hazards. The proposed system presents an economical solution using a wearable camera and a smart phone to provide safe navigation facility to the visually impaired user. X-EYE provides the features of obstacle detection, person recognition, location tracking and sharing, SMS reader, and language translation. Audio messages are specifically generated to provide better usability to the blind/visually impaired user. The proposed system is robust to egocentric video limitations i.e. partial appearance of objects, sudden background change, jitter effects, and illumination conditions. Performance of the proposed method is evaluated on ten real-time egocentric videos. Experimental results indicate the effectiveness of our method in terms of providing safe mobility service to the visually impaired people.

***Index Terms*—Face Recognition, Location Sharing, Location Tracking, Obstacle Detection, Visually Impaired, Wearable Camera.**

I. INTRODUCTION

According to a survey [1], almost 285 million people in the world are visually impaired and 39 million are blind. It is difficult for the visually impaired people to be dependent on their cognitive sense most of the time; and needs assistance of their caretaker to perform daily life activities. Therefore, there exists a need to develop automated systems that can assist the visually impaired people to perform their routine tasks independently.

In the last decade we have witnessed a variety of portable and wearable devices that were designed to provide safe navigation services to visually impaired people in their environment. White cane and guide dogs were popularly used as navigation aid tools; however, these tools are limited in functionality in terms of inability to predict the speed, size and distance of the obstacles exist in surroundings. The rapid evolution of technology has replaced these manual tools with various automated systems that use various devices i.e. wearable cameras, mobile sensors, etc. to propose navigation aid systems for visually impaired persons. The proposed system could be deployed easily in the industry which could be economical solution in comparison to the existing systems.

Existing navigational aid systems [2-4] largely provide the obstacle detection and avoidance facility to assist the blind persons in their orientation by detecting obstacles on their way to avoid any injury. Wang et al. [2] used a wearable system consisting of camera, embedded chip and a haptic belt to provide information about the surrounding environment to the blind person. Motion and depth information was used to identify the distance between obstacles (i.e. chair) and the camera wearer. Finally, the haptic belt sensor was used to inform the blind person regarding these obstacles through vibrations. This system [2] provides useful information to the visually impaired person but requires expensive equipment. Zheng et al. [3] used mobile sensors to collect data and applied acoustic localization approach for obstacle detection. Smart phone was used to generate a system alert to the blind person in case of obstacle detection within 2 m range. For experimentation purpose walls were considered as an obstacle in this system [3]. Li et al. [4] proposed an indoor navigational system for visually impaired persons using the smartcane and cell phone. A hybrid approach comprises of Kalman filtering and connected component labeling was used for obstacle detection. A built-in text to speech API was used to provide alerts to the user. In addition, a smartcane device consisting of vibration motor and keypad was used to guide the user for navigational purpose.

Existing literature [5-7] reported for visually impaired facilitation applications also provide the face recognition feature to inform the blind user about the people around him. Neto et al. [5] proposed a facial recognition system using a smart watch to facilitate the individuals with low-vision or blind. Samsung gear watch camera worn by the user was used to capture the nearby individuals. Histograms of Oriented Gradient (HoG) features were extracted and trained on the K-NN algorithm to classify between the known or un-known person. An audio message was also generated to inform the blind person regarding the identity of the nearby person. Neto et al. [6] also used a Kinect wearable device to develop a real-time facial recognition system for visually impaired users. A 3780-dimensional HoG descriptor was employed for feature extraction. PCA algorithm was applied for dimensionality reduction. Finally, KNN classifier was used in combination of Euclidean distance metric to perform the face recognition. Similarly, Chaudhry et al. [7] proposed a blind assistance system based on facial recognition. Viola jones algorithm [8] was used for face detection. OpenCV face recognition library based on local binary patterns histogram was used for face recognition.

978-1-7281-0257-3/18/$31.00 ©2018 IEEE

II. PROPOSED SYSTEM

The proposed system presents an effective solution to aid the visually impaired persons to mobilize independently in performing their routine tasks. An android application is specifically designed for blind people that provides useful features comprising of obstacle detection, person identification, location tracking & sharing, SMS reader and language translator. Egocentric video is captured at real-time through a wearable camera worn by the user on the ear. The proposed system processes the egocentric video to detect the obstacles (pedestrians and vehicles) in-front of the blind person. In addition, face recognition is used to inform the blind user about the identity of nearby people. Auxiliary features like SMS reader, location tracking and sharing are also developed to provide assistance for independent mobility to the visually impaired persons. In addition, translator utility (English to Urdu and vice versa) is also provided to deliver the native support to the user. The proposed system is robust to illumination conditions, partial appearance of objects, rapidly changing background and jitter effects due to movement of camera wearer. Shown in Fig. 1 is the block diagram of the proposed system.

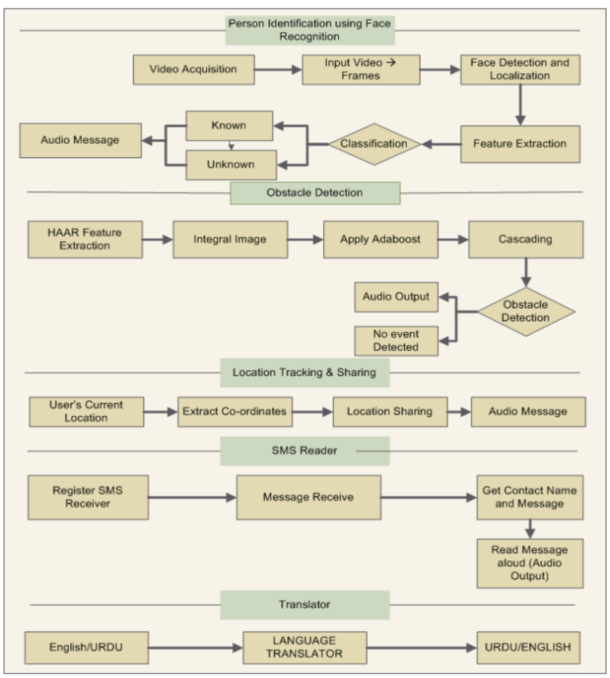

Fig. 1 Block Diagram of Proposed System.

A. Person Identification through Face Recognition

Person identification is performed on the real-time egocentric videos through face recognition. Face detection is commonly employed initially to localize the face, thereby reducing the computational cost for frame processing. Haar features are used to train the Adaboost algorithm for face detection in the proposed work. After face localization we represented the extracted face images through local tetra patterns (LTrPs) descriptor. Finally, extreme learning machine (ELM) classifier was used for facial recognition.

1) Feature Extraction: For LTrPs computation, we computed the first-order derivative in vertical and horizontal directions. For image I, the first-order derivative along 0 and 90 degrees are computed as:

$$I_0^1(P_c) = I(P_h) - I(P_c) \quad (1)$$

$$I_0^{90}(P_c) = I(P_v) - I(P_c) \quad (2)$$

Where P_c represents the center pixel in the given neigborhood of I, P_h and P_v represents the horizontal and vertical neighbors of P_c. The direction of center pixel P_c is computed as:

$$I_d^1(P_c) = \begin{cases} 1, & I_0^1(P_c) \geq 0 \text{ and } I_{90}^1(P_c) \geq 0 \\ 2, & I_0^1(P_c) < 0 \text{ and } I_{90}^1(P_c) \geq 0 \\ 3, & I_0^1(P_c) < 0 \text{ and } I_{90}^1(P_c) < 0 \\ 4, & I_0^1(P_c) \geq 0 \text{ and } I_{90}^1(P_c) < 0 \end{cases} \quad (3)$$

As we can observe from Eq. (3), four different values are computed that are the direction values. Similarly, we define the second order $LTrP^2(p_c)$ as follows:

$$LTrP^2(P_c) = \{f_1(I_1^d(p_c), I_1^d(p_1)),\ f_1(I_1^d(p_c), I_1^d(p_2)), \\ .., f_1(I_1^d(p_c), I_1^d(p_n))\} \quad (4)$$

$$f_1\left(I_1^d(p_c),\ I_1^d(p_n)\right) = \begin{Bmatrix} 0, & I_1^d(p_c) = I_1^d(p_n) \\ I_1^d(p_n), & Otherwise \end{Bmatrix} \quad (5)$$

We obtained the 8-bit tetra pattern against each center pixel from Eqs. (4) and (5). After generating the LTrPs we represented each frame through the histogram as:

$$H = \frac{1}{M \times N} \sum_{k=1}^{M} \sum_{l=1}^{N} f_2(LTrP(k,l), w) \quad (6)$$

$$f_3(x,y) = \begin{Bmatrix} 1, & if\ x = y \\ 0, & Otherwise \end{Bmatrix} \quad (7)$$

Where $M \times N$ represents the size of image.

2) ELM Classification: For classification we use the LTrPs feature vectors of face images to train the ELM classifier [9] for face recognition. In the proposed work, ELM is used as a multi-class classifier to recognize 11 different persons. In multi-class classification, the number of output neurons will be consequently set equivalent to the number of classes. In a single ELM classifier, the multi-class issue is executed with a design of multi-output nodes which is equivalent to the number of pattern classes. We split the training data into one against all (OAA) and pairwise one against one (OAO) techniques. The reason of using these schemes is that ELM-OAA and ELM OAO requires less hidden nodes than the single ELM classifier. In addition, ELM-OAO is computationally efficient as compared to the single ELM classifier in case of more than ten classes.

B. Obstacle Detection

Obstacle detection technique is implemented using the adaboost algorithm [2-4]. In the first stage, integral image representation is created that allows the features to be computed quickly at different scales. It can be calculated from the image by applying few operations on the pixels. Integral image facilitates the calculation of Haar-like features at multiple scale. Secondly, AdaBoost learning algorithm is employed to design a classifier by picking few key features. Each stage of the boosting procedure picks a new weak classifier for feature selection. In the next stage complex classifiers are integrated in a cascade. The computation speed of the detector is increased significantly by integrating complex classifiers consecutively in a cascade structure.

If the classifier successfully detects the pedestrians or vehicles, then an audio message alert is generated to aid the blind person to avoid occlusions with the obstacles.

C. Location Tracking and Sharing

Location tracking feature is implemented to facilitate the visually impaired people to get the information regarding their current location. Geo-fencing technique [5] based on acquiring the global positioning system coordinates (GPS) is used for location tracking. The location of visually impaired person is being traced followed by generating an audio message to inform the user regarding his current location. Moreover, our system automatically broadcasts an SMS to his/her caretaker regarding the current location of the blind person that is very useful in case of emergency. In addition, geo-fencing approach is used to provide useful information regarding the nearby places to the blind person.

D. SMS Reader

We developed a key feature of SMS reader that generates an audio message by converting the text message into speech using text to speech converter [6]. This feature is added to facilitate the blind person to receive the message of his/her caretaker without any support. The proposed system can also be used by the elderly people to navigate safely during traveling. URDU translator feature is specifically added to the application to provide the native support to the user.

III. PERFORMANCE EVALUATION

Performance of the proposed system is evaluated on a custom video dataset captured from LOOXCIE wearable camera. We employed objective evaluation metrics such as accuracy, recall, precision, and F-1 score for performance evaluation.

A. Dataset

A custom egocentric video dataset consisting of ten videos of 18 hours of length is created for performance evaluation. Each video in the dataset has a frame resolution of 640 x 480 and a frame rate of 30 fps. The dataset videos are diverse in terms of length, environment (i.e. indoor/outdoor), object appearance, rapidly varying backgrounds, and illumination conditions. For face recognition method, we used 11000 face images of 11 different test subjects. We collected 1000 face images of each person (test subject). We used 60% images for training and remaining 40% for testing.

B. Experimental Results

For performance evaluation of the proposed facial recognition method, we computed the precision, recall, accuracy and F-1 score by calculating true positive, true negative, false positive and false negative values. The results obtained on each test subject is shown in Table 1. The proposed method achieves an average precision of 0.96, recall of 0.95, accuracy of 0.97, and F-1 score of 0.95. It can be observed from Table I that the proposed method is very effective in terms of recognizing faces for person identification.

TABLE I
Recognition Performance of the Proposed Method.

Users	Accuracy	Precision	Recall	F1-score
User-1	0.98	0.92	0.9	0.91
User-2	0.99	0.94	0.94	0.94
User-3	0.97	0.93	0.97	0.95
User-4	0.95	0.98	0.99	0.99
User-5	0.96	0.97	0.97	0.97
User-6	0.98	0.95	0.94	0.95
User-7	0.99	0.98	0.97	0.98
User-8	0.97	0.95	0.97	0.96
User-9	0.98	0.99	0.97	0.98
User-10	0.94	0.97	0.93	0.95
User-11	0.96	0.95	0.93	0.94
Average	**0.97**	**0.96**	**0.95**	**0.95**

In the second experiment, we evaluated the performance of our obstacle detection method for each real-time video. In the proposed work we considered the people and vehicles as obstacles. The results obtained on each video for obstacle detection are shown in Table II.

TABLE II
Obstacle Detection Result for Pedestrian and Vehicles.

Videos	Precision	Recall	Accuracy	F-1 Score
1	1	0.96	0.98	0.98
2	1	1	1	1
3	1	1	1	1
4	0.85	0.87	0.9	0.86
5	1	0.95	0.98	0.98
6	1	0.9	0.94	0.95
7	1	1	1	1
8	0.9	0.9	0.95	0.9
9	1	0.91	0.96	0.95
10	0.8	1	0.95	0.9
Average	**0.96**	**0.95**	**0.97**	**0.95**

A confusion matrix is designed in our next experiment to show the classification accuracy of vehicle and pedestrian detection (Table III). Our obstacle detection method performs better for pedestrian detection as compared to vehicle detection (Table IV). The marginal drop reported in vehicle detection results is observed due to the brisk movement of vehicles. Moreover, it is observed that the videos captured in sun light reports more false positives as compared to videos captured under cloudy conditions. This is attributed to the fact that the shadows observed under sunlight are mistakenly detected as persons.

TABLE III
Confusion Matrix for Pedestrian and Vehicle Detection.

	Vehicle Detection	**Pedestrian Detection**
Vehicle Detection	0.95	0.05
Pedestrian Detection	0.01	0.99

TABLE IV
Obstacle Detection Result with Existing State of the Art

Framework	**Recall**	**F1-Score**
Lee et. al [7]	0.93	0.93
Yankun et. al [8]	0.94	0.94
Proposed System	**0.95**	**0.95**

Moreover, location tracking, SMS reader, and translator features are evaluated by taking 500 test samples (See Fig. 2). We evaluated location tracking and sharing features at different places and on multiple platforms. SMS-reader feature is tested by sending different messages at the same time to different mobile devices.

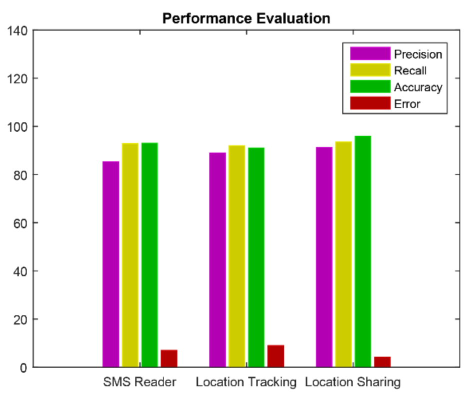

Fig. 2. Performance Evaluation of SMS Reader, Location Tracking, and Location Sharing.

We also designed an experiment to compare the performance of our system with existing state-of-the-art blind assistance systems [2, 3, 5, 6,7]. Accuracy and error rate metrics are used for performance comparison (See Table V). From the results in Table V we can observe that the proposed system provides superior detection performance as compared to the existing systems.

IV. CONCLUSION AND FUTURE WORK

We proposed an effective low-cost blind assistance system for visually impaired people. Our application allows the users to perform their routine tasks by avoiding any obstacles, information about the nearby persons, location tracking & sharing, SMS reader and language translation. The proposed application is specifically designed to provide better usability to the blind person. Our system is robust to illumination conditions and the aforementioned egocentric video limitations. Currently, we are extending the proposed system by working on sign board text recognition. In future, we will be working on audio input commands in order to make the system more user friendly to the blind.

TABLE V
Performance Comparison.

Blind Assistance Systems	**Accuracy**	**Error**
Wang et al. [2]	0.68	0.32
Zheng et al. [3]	0.96	0.04
Neto et al. [5]	0.84	0.16
Neto et al. [6]	0.72	0.28
Chaudhary et al. [7]	0.61	0.39
Proposed System	0.97	0.03

REFERENCES

1. Chauhan, D., N.M. Patel, and M. Joshi. Automatic summarization of basketball sport video. in Next Generation Computing Technologies (NGCT), 2016 2nd International Conference on. 2016. IEEE.
2. Viola, P. and M. Jones, Robust real-time object detection. International Journal of Computer Vision, 2001. **4**.
3. Viola, P., M.J. Jones, and D. Snow, Detecting pedestrians using patterns of motion and appearance. International Journal of Computer Vision, 2005. **63**(2): p. 153-161.
4. Jones, M.J. and D. Snow. Pedestrian detection using boosted features over many frames. in Pattern Recognition, 2008. ICPR 2008. 19th International Conference on. 2008. IEEE.
5. Petscharnig, S. and K. Schöffmann, Learning laparoscopic video shot classification for gynecological surgery. Multimedia Tools and Applications, 2018. **77**(7): p. 8061-8079.
6. Wei, W.-L., et al. Deep-net fusion to classify shots in concert videos. in Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on. 2017. IEEE.
7. Lee, C.-H., Y.-C. Su, and L.-G. Chen. An intelligent depth-based obstacle detection system for visually-impaired aid applications. In Image Analysis for Multimedia Interactive Services (WIAMIS), 13th International Workshop on, 2012.
8. Yankun, Z., C. Hong, and N. Weyrich. A single camera based rear obstacle detection system. in Intelligent Vehicles Symposium (IV), 2011 IEEE. 2011. IEEE.
9. Bertozzi, M. and A. Broggi, GOLD: A parallel real-time stereo vision system for generic obstacle and lane detection. IEEE transactions on image processing, 1998. **7**(1): p. 62-81.
10. Chen, L., B.-l. Guo, and W. Sun. Obstacle detection system for visually impaired people based on stereo vision. in Genetic and Evolutionary Computing (ICGEC), 2010 Fourth International Conference on. 2010. IEEE.
11. Lee, C.-H., Y.-C. Su, and L.-G. Chen. An intelligent depth-based obstacle detection system for visually-impaired aid applications. in 2012 13th International Workshop on Image Analysis for Multimedia Interactive Services. 2012. IEEE.