

Context-aware Assistive Indoor Navigation of Visually Impaired Persons

Chathurika S. Silva^{1*} and Prasad Wimalaratne^{2**}

¹Faculty of Technology, University of Colombo, Kaduwela Rd., Sri Jayawardenepura Kotte, Sri Lanka

²University of Colombo School of Computing, 35 Reid Ave, Colombo 7, Sri Lanka

(Received October 3, 2019; accepted March 17, 2020)

Keywords: sensor fusion, wearable sensors, context awareness, assistive technology, human–computer interaction, mobile–cloud computing

This paper presents an approach for context awareness in navigation for visually impaired persons via sensor-based obstacle detection, obstacle recognition, sensor fusion, and walking context analysis. Sonar and vision sensor data are fused using a complementary sensor fusion approach. A wearable belt has sonar and vision sensors that detect and recognize obstacles, respectively. A fuzzy logic model is used for safety aspect handling during visually impaired navigation. Walking context analysis handles decisions on the current walking status by using clues acquired from the smartphone application and obstacle detection process. Feedback is provided via audio and tactile cues. The usability evaluation experiment using the proof-of-concept reveals positive results and other areas of investigation have been identified.

1. Introduction

Various navigation systems have been developed to enhance the mobility of visually impaired persons. Despite the fact that assistive devices for outdoor navigation systems have progressed relatively well, indoor navigation aids remain more challenging. Outdoor navigation systems usually use the global positioning system (GPS); however, owing to the unavailability of the line of sight with satellites, the signal of GPS in an indoor environment is poor. Most existing indoor navigation systems are based on supplementing physical infrastructure with tags such as beacons and radio frequency identifiers (RFIDs).⁽¹⁾ This paper presents an approach for indoor visually impaired navigation that does not require establishing indoor setups and depends on inexpensive, lightweight, active, and passive sensors.

A single type of sensor, such as a range, vision, or inertial sensors, fails to provide adequate information on the surrounding. Hence, an approach based on the fusion of homogeneous and heterogeneous sensors to harness the capabilities of different sensors while minimizing limitations has been considered viable. Complementary sensor types are used to improve the accuracy of the proposed work.^(2,3) One of the key challenges is to determine their optimal use in terms of type and number of complementary sensory channels to aid visually impaired

*Corresponding author: e-mail: chathurika@iat.cmb.ac.lk

**Corresponding author: e-mail: spw@ucsc.cmb.ac.lk

<https://doi.org/10.18494/SAM.2020.2646>

persons over a wide range of environmental conditions. The identification of appropriate complementary sensory channels and their strengths partially contribute to the novelty of proposed work. The findings of the evaluation experiments provides insights into further research making contributions to the field.

Work has been reported in the field of indoor navigation for visually impaired persons with the use of a multitude of sensors.⁽²⁻⁷⁾ Most studies are based on sonars as range sensors.⁽³⁻⁵⁾ However, significant errors can persist owing to the broad beam width of sonar sensors and other factors due to weather conditions.

Computer vision has been used for both obstacle detection and recognition.^(6,7) Preserving obstacle detection accuracy and maintaining real-time feedback to the user are challenging owing to the high computational demands of image processing. Hence, sensor fusion can be used to overcome the limitations of individual sensors by complementing the strengths of sonar and vision sensors.⁽⁸⁻¹²⁾

Sensor fusion methods can be classified on the basis of the level of fusion as low, medium, or high,⁽¹³⁾ the type of sensor as homogeneous or heterogeneous,⁽¹⁴⁾ and the type of data as competitive, complementary, or cooperative.⁽¹⁵⁾

A fusion of vision and sonar sensors is in the proposed work. The objective of sensor fusion is to combine sensor information into a single representational format. However, converting sonar and vision sensor measurements to a single common representation format is expected to require considerable effort since the sonar sensor provides three-dimensional information and the vision sensor provides two-dimensional information.

Therefore, either high-level fusion or integration seems to be more practical than low-level sensor fusion. Thus, both rule-based sensor fusion and integration are applied in this research accordingly.

There have been a few attempts to integrate vision and sonar sensors in the field of visually impaired navigation.^(8,9) The amount of research in the area of combining multiple sensors for visually impaired navigation is minimal. Among the reported studies, most of them targeted sensor integration but not sensor fusion. Thus, the challenge of sensor fusion is that it has to blend at pieces of sensory information coming from different sensors in order to place them in one representation format.

Visually impaired persons can gain certain benefits by using context space awareness in their navigation process. Only a few studies on the context awareness of walking for blind navigation⁽¹⁶⁻¹⁸⁾ targeted the assessment of the context on the basis of the changes in the existing environment⁽¹⁹⁻²¹⁾ considering only the person when assessing the walking context. No work has been reported on the integration of both environmental changes and individual factors when determining the walking context for the next moment. Therefore, we take into account both current environmental aspects (obstacle density and distance to nearby obstacles) and personalization factors when accessing the walking context of visually impaired persons. Therefore, a hybrid walking context estimation method has been proposed on the basis of environmental adaptation and personalization in this research.

In summary, there is a significant gap in the literature regarding a single hybrid approach that can perform object detection, recognition, sensor fusion, and current context estimation.

Therefore, we propose a novel approach that can perform object detection, recognition, and context estimation using sensor fusion.

2. Methodology

A constructive research methodology has been followed, and a proof-of-concept system has been developed as an experimental testbed to develop a navigation aid that makes prime use of multiple sensors for smooth and continuous navigation.

Design Assumptions

The prototype is evaluated within selected in-house areas with a small number of barriers to detect stationary obstacles only. Only micromavigation, where there is no travel between a source and a destination is considered. The smartphone is placed close to the sensor belt as it is connected to the feedback systems through a Bluetooth connection. The user shall keep the tactile feedback device with vibration motors in contact with the body. The user wears an earphone on one side correctly to hear the voice feedback.

The architecture of the proof of the concept prototype is shown in Fig. 1 consisting of key components shown below.

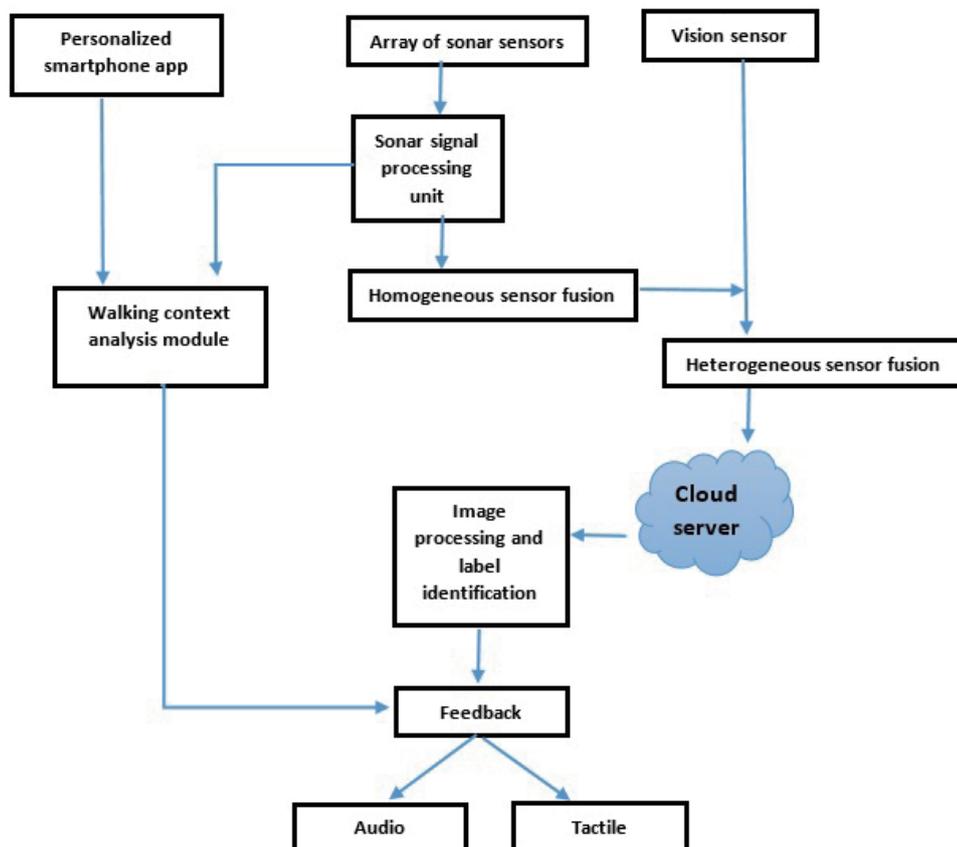


Fig. 1. (Color online) High-level architecture of the system.

A set of sonar sensors: Obstacles appearing from different directions are detected by sonar sensors.

Processing of sonar signals: The time-of-flight method is used to calculate the distances to the obstacles.

Vision sensor: This identifies the obstacles detected via sonar sensors.

Homogeneous sensor fusion: The fusion of a few sonar sensors is carried out to overcome limitations that arise owing to the broad beam width of the sonar.

Heterogeneous sensor fusion: Whenever obstacle detection occurs, the vision sensor that captures a snapshot, which in turn provides additional information about the obstacle.

Cloud server: The snapshot taken by the smartphone camera is sent to the cloud for image processing, which returns a label identifying the objects of the image.

Personalization smartphone app: This app, which is shown in Fig. 2(b), is used to input facts about the user such as age, gender, height, and visual status, which will be later used to compute the current walking context of the user.

Walking context analysis: This determines whether it is safe or dangerous to walk in a particular direction in the current context on the basis of the outputs of the obstacle detection module and the personalization app. A detailed description of the approach to walking context analysis can be found in our previous publication.⁽²²⁾

Audio feedback: Feedback corresponding to the vision sensor is prompted via the audio feedback method.

Haptic feedback: The detection of obstacles by sonar sensors is performed via tactile units.

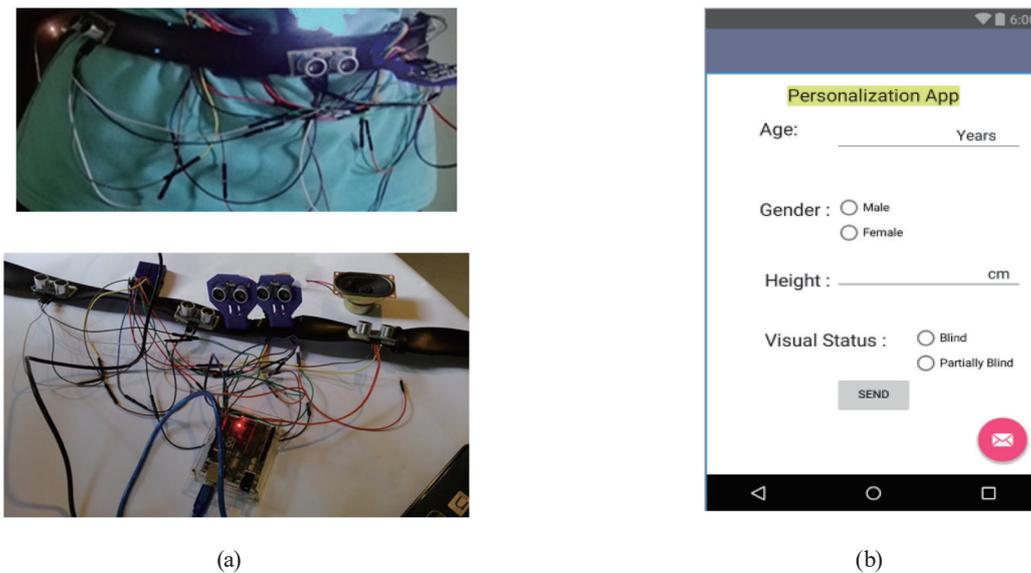


Fig. 2. (Color online) (a) Experimental waist strap and (b) personalized Android application.

3. Proof of Concept Prototype

3.1 Obstacle detection

A waist belt, with five ultrasonic sensors, as shown in Fig. 2(a), is used to detect obstacles. Two sonar sensors positioned on the left and right of the belt distinguish left-, and right-side obstacles, respectively. There are three sensors positioned in the middle of the belt: one to detect front obstacles and the other two are tilted at a certain angle to the floor to detect ground-level obstacles. A microcontroller processing module called Arduino is used to process measurements acquired from sonar sensors.

The system uses coin vibration motors to generate tactile feedback. The wearable sensor belt consists of five coin vibration motors. Sensors are placed outside of the belt, where they face the surrounding environment, and tactile units are attached to the belt such that they are in contact with the body (around the waist). The vibration motors change the intensity of the vibration concerning the gap between the waist strap and the obstacle. Ultrasonic sensor and vibration motor specifications are shown in Tables 1(a) and 1(b), respectively.

3.2 Obstacle recognition

The obstacle recognizer uses a built-in camera of the mobile phone with an Android operating system and an image recognition algorithm running on Google Cloud. The use of cloud computing-based processing overcomes the limitations in the computational power of mobile and embedded devices. A smartphone is used for image capture and streaming to cloud servers for image recognition.

The camera is triggered only when an object is detected, which will, in turn, optimize computational resources without continuously triggering the camera, leading to redundant image processing operations. Google Cloud-based image processing uses a label detection algorithm.

This audio is used as the feedback mechanism in the image identification. The output of obstacle recognition is received as a text message to the cloud vision application running on a smartphone. This text message is converted to a voice using text to speech and is sent to the earphone worn by the user.

Table 1
Specifications of (a) ultrasonic sensors and (b) coin vibration motors.

(a)		(b)	
Frequency	40 Hz	Frequency	10–55 Hz
Range (Max)	4 m	Vibration speed	15000 rpm
Range (Min)	2 cm	Motor body diameter	12 mm
Input signal (Trigger)	10 μ S TTL pulse	Motor body length	2.7 mm
Output signal (Echo)	Input TTL lever signal and the range in proportion		
Dimensions	45 \times 20 \times 15 mm ³		

3.3 Sensor fusion

3.3.1 Fusion of sonar sensors

A recursive Bayesian filter called the extended Kalman filter (EKF) is used to combine two ultrasonic sensors tilted to the floor to enhance ground level obstacle detection.

3.3.2 Vision and sonar sensor fusion

The data from sonar and vision sensors are two complementary types of sensor that provide information on different aspects of the environment, i.e., sonar gives the distances to obstacles and the vision sensors provide more descriptive information on the detected obstacles. Therefore, rule-based fusion fuses data arriving from vision and sonar sensors. During this fusion process, the vision sensor activates only when the sonar indicates a detection of an obstacle. Therefore, only the snapped frame of the scene is sent for further processing to identify the object rather than send all the frames to the image processor, which would overload the task of processing. In this way, the use of sonar sensors accelerates the process of acquiring visual data and reduces the computational power required for video image processing. More importantly, this method enables real-time data processing.

3.4 Walking context analysis module

Hybrid walking context analysis, which is used to improve the safety of the present walking situation, is based on environmental adaptation and individual personal preferences. Therefore, the hybrid walking context analysis module consists of two parts: an adaptation module and a personalization module.

The inputs to the adaptation module are acquired through the distance measured by sonar sensors attached to the waist belt. The outputs of the adaptation module are the density of obstacles and the distance to the nearest obstruction. The smartphone application provides the inputs to the personalization module. A smartphone is used to calibrate the walking aid for the user's personal data (age, gender, height, and visual status). The output of the personalization module is the walking speed of the visually impaired navigator.

The hybrid walking context analysis module is built as a fuzzy inference system (FIS) in MATLAB (Fig. 3). The Mamdani approach-based fuzzy logic controller is used, and the centroid method is used in defuzzification. Triangular and trapezoidal linear functions are used to define the membership functions of the inputs since their domains cover a wide range of values. The output of the hybrid walking context module, which is whether it is safe or not to travel the next few steps under the current walking conditions, is converted into voice messages and rendered to the user.

4. Evaluation of Approach

The subjects of the evaluation were 10 users of both genders (four female users and six male users) with two different age ranges (eight young users with ages of around 22–37 years and two

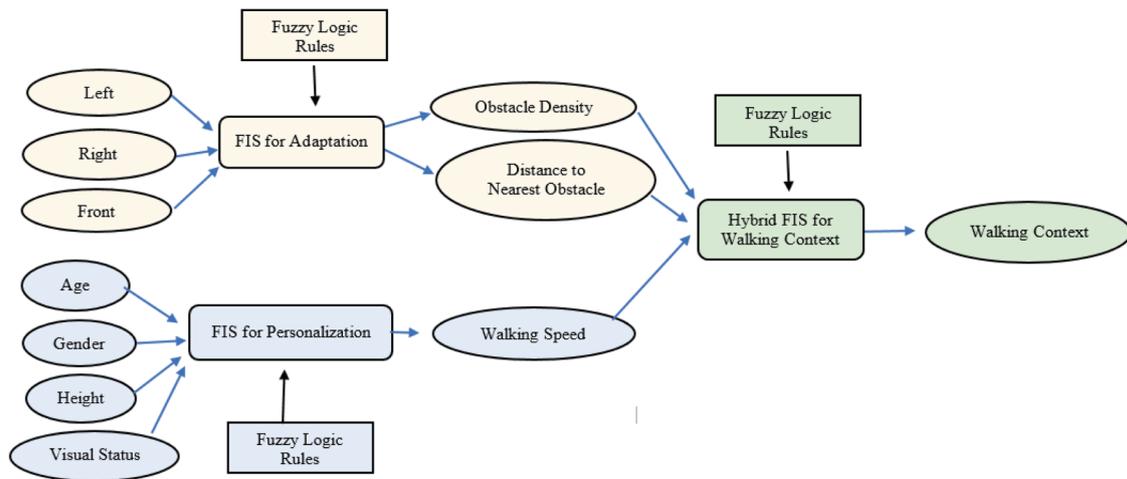


Fig. 3. (Color online) Hybrid FIS for walking context.

older users with ages around 70). Seven of the subjects were blindfolded, and their visual status was assumed to be blind. The other three subjects had age-related vision loss. All participants confirmed that they did not have disabilities related to mobility or hearing.

The evaluation environment was a previously adjusted free space consisting of different types of static object scattered in different directions, including a staircase, a corridor, and drop-offs.

The ethical consideration was taken into account during the evaluation experiment. During the experiment, both physical well-being and privacy were well protected. Safety measures were taken to minimize risks such as falls during the experiment. Extensive training was given to all the participants of the study to maintain the consistency of the results between them. Close attention was paid to the comfort of the subjects allowing them to pause the experiment any time they felt tired.

4.1 Pilot study

A pilot study was conducted using the prototype system. It was assessed whether the prototype system functioned correctly and posed any problems, and the subjects were trained on different configurations of the experimental test.

4.2 Observations and modifications based on the pilot study

Several improvements shall carry out prior to the enhanced final user evaluation experiment, on the basis of the results obtained from the pilot study. Certain areas for the improvement of the prototype were revealed during the pilot study. It was necessary to increase the frequency of vibrations generated by the motors in the tactile belt and to increase the feedback duration from 500 to 1000 ms.

4.3 Componentwise evaluation

It was identified that obstacle detection, obstacle recognition, sensor fusion, and current context analysis can be further enhanced for better usability via several directions of research.

4.4 Obstacle detection

Obstacles are categorized as left, right, front, constrained spaces, stairs, and drop-offs. Three chances are given to each user to complete one local navigation. When the user was successful in avoiding an obstacle, it was classified as a “hit”. Otherwise, it was classified as a “miss”.

4.4.1 Results of the evaluation of obstacle detection

Figure 4 shows the obstacle detection rate for each type of obstacle as a percentage. According to the chart in Fig. 4, 93% of the left obstacles, 88% of right obstacles, 100% of frontal obstacles detection rate, and 81% of stairs and walls were identified with an accuracy of 100%.

4.5 Obstacle recognition

The two most important aspects of obstacle recognition are the speed and accuracy of the response. Evaluation experiments focused on measuring the response time since it is necessary to have a rapid response of less than 1 s to preserve the real-time performance of the system. Images taken at different indoor locations were used as the test data. The camera embedded in the smartphone was used to take pictures of the scenes, which were sent to the cloud server for image classification. The mobile phone was connected to the internet through a wireless network in the indoor premises.

4.5.1 Results of obstacle recognition

The average feedback period between sending and receiving the captured image was determined for pictures with different resolutions. The feedback times corresponding to different resolutions are shown in Fig. 5.

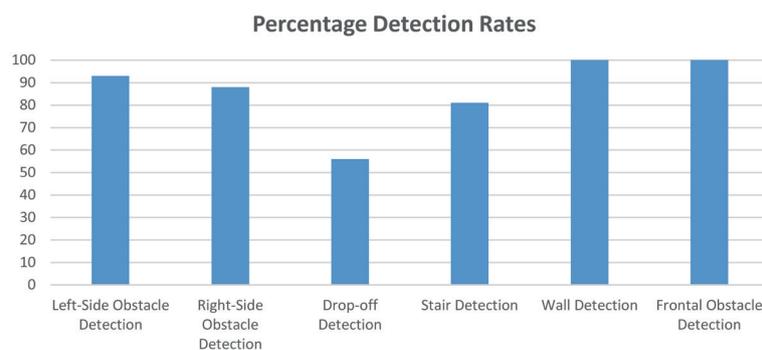


Fig. 4. (Color online) Obstacle detection rates as percentages.

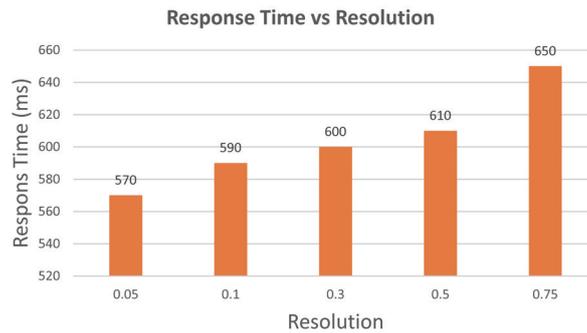


Fig. 5. (Color online) Change in response time with the resolution.

This response time indicates real-time feedback since it is less than 1 s. The lower resolutions obtained by compressing the same set of frames had shorter response times than the default resolution.

4.6 Sensor fusion

Homogeneous fusion between two ultrasonic sensors that were tilted to detect the ground-level obstacles was carried out. An extended Kalman filter was used to carry out the fusion of ultrasonic sensors.

The blue line in Fig. 6 illustrates the outcome of the fusion, whereas the red and green lines represent the data from the two ultrasonic sensors.

4.7 Evaluation of walking context analysis

4.7.1 FIS model used for adaptation

As shown in Fig. 7, the membership functions of the object density (small, medium, and large) are based on left, right, and front ultrasonic sensor readings. Figure 8 shows the membership functions of the “nearest obstacle” variable, which consists of three fuzzy sets called near, medium, and far. The range of the distance to the nearest obstacle is from 0 to 4 m.

4.7.2 FIS model for personalization

The customization of the assistive travel based on individual personal factors was performed through a user review. The assistive prototype was tested with and without calibration according to individual factors such as age, gender, height, and visual status. After the evaluations, subjects were asked whether calibration concerning individual factors improved their mobility.

Figure 9 illustrates user responses on whether the customization of the prototype according to the above four factors affected their navigation.

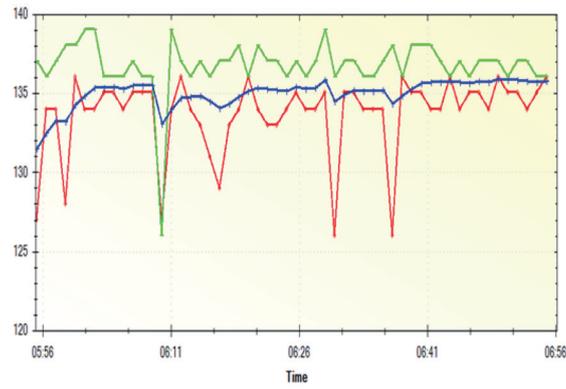


Fig. 6. (Color online) Homogeneous fusion of two ultrasonic sensors.

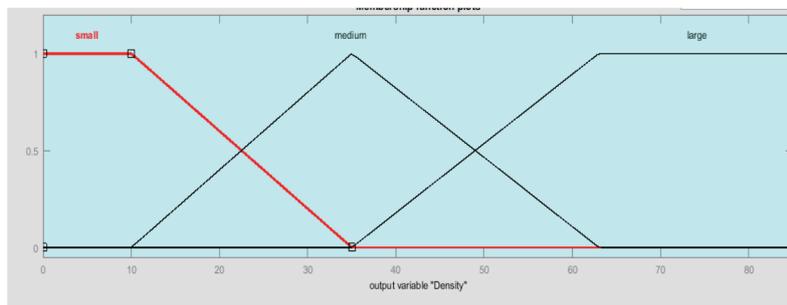


Fig. 7. (Color online) Membership functions of the obstacle density.

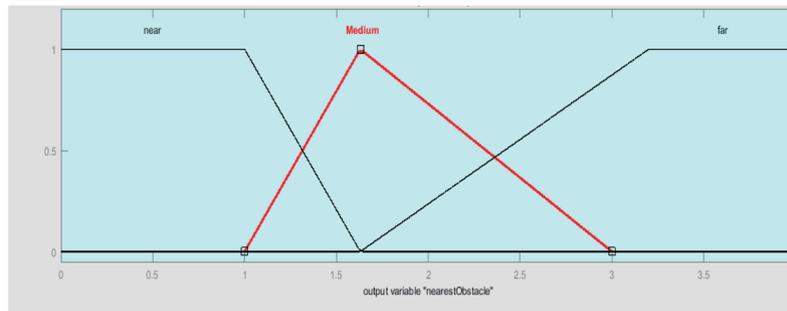


Fig. 8. (Color online) Membership functions of the output variable of the nearest obstacle.

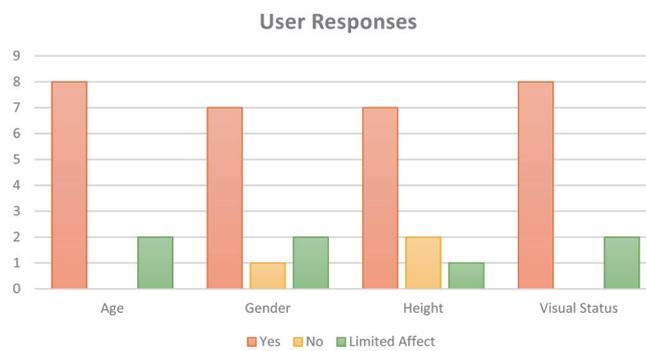


Fig. 9. (Color online) Effect of personalization.

According to Fig. 9, the personalization of all four factors was considered by all or nearly all the users to improve their navigation. Figure 10 shows the output of the personalization model based on the inputs of age, gender, height, and visual status.

4.7.3 Hybrid FIS model

The hybrid fuzzy inference model for walking context analysis was developed by combining adaptation and personalization models. The output of this hybrid FIS is the walking context, which is used to investigate whether it is safe to proceed with walking under the given conditions. The inputs to this hybrid model are the outputs of the adaptation model (obstacle density and distance to the nearest obstacle) and the personalization model (walking speed). Walking context is output as the final result when the obstacle density, distance to the nearest obstruction, and walking speed are inputted to the hybrid fuzzy system as shown in Fig. 11.

4.8 Evaluation of feedback

When giving feedback to visually impaired persons, alternative sensory skills, such as auditory and tactile sensing, were taken into consideration. Users were provided feedback during user training to evaluate only the feedback identifying ability. Each command was given 10 times in a random order to the users as shown in Fig. 12. Feedback was presented in periods of 5 and 10 s.

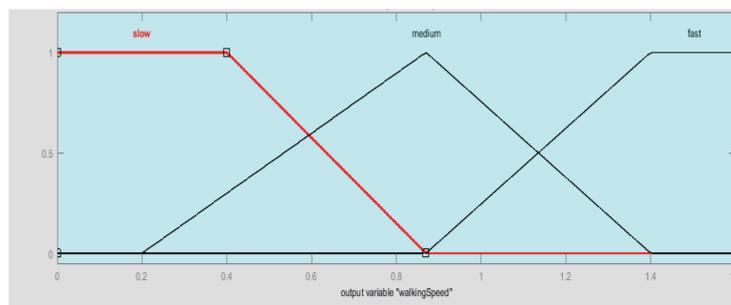


Fig. 10. (Color online) Membership functions of outputs of walking speed factor.

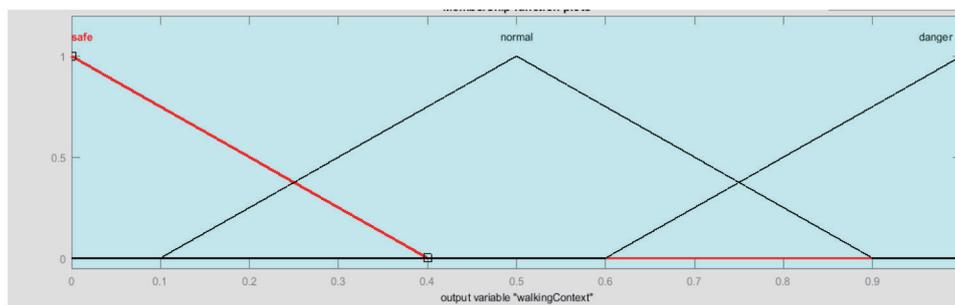


Fig. 11. (Color online) Membership functions of the walking context.

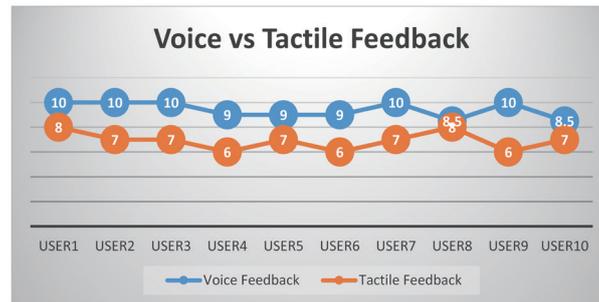


Fig. 12. (Color online) Voice feedback vs tactile feedback.

4.9 Analysis and interpretation of the feedback evaluation

In the user evaluation, higher scores were given for voice input evaluation as pictorially depicted in Fig. 12. Therefore, the voice input module has earned higher user satisfaction and was considered more convenient.

The audio and haptic cues showed comparable scores. Therefore, the fusion of the audio and tactile techniques improved the feedback of the proof-of-concept system.

4.10 User evaluations

Contextual inquiry was also part of each experiments. Subjects were asked questions regarding their experiences during the experiment. When they faced any inconveniences, it was hoped that they would express them in an interview and they were also encouraged to think of modifications or improvements of the prototype.

4.10.1 Analysis and interpretation of user evaluations

Subjects of the experiment provided positive feedback on the ability to personalize based on their age, height, gender, and visual status. It was observed that the wearable system is more accurate and user-friendly than a white cane. However since there is a need to make surrounding persons aware that the person is visually impaired, it is necessary to have the white cane as the symbol. In addition users will feel more comfortable to have augmented support by using dual methods.

5. Conclusion

Developing navigation aids for visually impaired persons increases their independence in their day-to-day life activities. The results indicate that the prototype system allows visually impaired persons to navigate indoors without requiring any infrastructure to be set up in the environment. The work provided an insight into the functionality of homogeneous and heterogeneous sensors and their fusion with wearable microprocessors.

In this research, vision and ultrasonic sensors were used in synergy for efficient navigation in an indoor environment. The usage of computer vision allows users to determine the objects around them, which was not possible when using an ultrasonic sensor alone.

Context awareness affects the assistance of visually impaired persons in traversing local pathways safely and efficiently. A hybrid fuzzy inference module was developed by integrating personalization and adaptation fuzzy inference modules. The acoustic method was used to alert the user of the current context, i.e., whether it is safe or not to continue navigation. Tactile feedback was generated to inform users about the closest objects on the path. The results of the hybrid FIS proved the safety by understanding the potential dangers during the navigation.

There is a considerable scope for future improvements to the proposed approach. Further work is underway in both indoor and outdoor environments. The current focus investigates whether feedback mechanisms can be improved to adjust their walking speed according to the current context.

References

- 1 Infsoft GmbH: (2019). https://www.infsoft.com/Portals/0/Images/solutions/basics/whitepaper/infsoft-Whitepaper-EN-Indoor-Positioning_download.pdf
- 2 F. Shahdib, W. U. Bhuiyan, K. Hasan, and H. Mahmud: *IJCA* **66** (2013) 6. <https://doi.org/10.5120/11114-6074>
- 3 Y. Priyadarshana and G. Wimalaratne: *IJSCE* **3** (2014) 1. <http://www.ijscce.org/wp-content/uploads/papers/v3i6/F1992013614.pdf>
- 4 C. Gearhart, A. Herold, B. Self, and L. Slivovsky: *Proc. Sensors Applications Symposium (IEEE, 2009)*. <https://doi.org/10.1109/SAS.2009.4801815>
- 5 A. Rodríguez, J. Yebes, P. Alcantarilla, L. Bergasa, J. Almazán, and A. Cela: *Int. J. Sens.* **12** (2012) 17476. <https://doi.org/10.3390/s121217476>
- 6 D. Kim, K. Kim, and S. Lee: *Sensors* **14** (2014) 10412. <https://doi.org/10.3390/s140610412>
- 7 Y. Lee and G. Medioni: *Advanced Reasoning with Depth Camera Workshop (2016)*. <https://doi.org/10.1016/j.cviu.2016.03.019>
- 8 G. Koshmak, M. Lindén, and A. Loutfi: *J. Sens.* **10** (2016) 1. <https://doi.org/10.1155/2016/6931789>
- 9 S. W. Yoon, S. B. Park, and J. S. Kim: *J. Sens.* **2015** (2015) 1. <https://doi.org/10.1155/2015/347379>
- 10 B. Mocanu, R. Tapu, and T. Zaharia: *Sensors* **16** (2016) 1807. <https://doi.org/10.3390/s16111807>
- 11 S. Y. Hwang, J. T. Park, and J. B. Song: *Proc. of IEEE Workshop on Advanced Robotics and Its Social Impacts (IEEE, 2010)*. <https://doi.org/10.1109/ARSO.2010.5679632>
- 12 C. Fernández: <https://www.diva-portal.org/smash/get/diva2:852457/FULLTEXT01.pdf>
- 13 F. Castanedo: *Sci. World J.* **2013** (2013) 1. <http://doi.org/10.1155/2013/704504>
- 14 W. Elmenreich: *An Introduction to Sensor Fusion* **47** (2001) 1. https://www.academia.edu/649942/An_introduction_to_sensor_fusion
- 15 H. F. Durrant-Whyte: *IJRR* **7** (1988) 97. <https://doi.org/10.1177/027836498800700608>
- 16 Q. Lin and Y. Han: *Sensors* **14** (2014) 18670. <https://doi.org/10.3390/s141018670>
- 17 Q. Lin and Y. Han: *Sensors* **16** (2016) 667. <https://doi.org/10.3390/s16050667>
- 18 V. Yerubandi, Y. Reddy, and M. Kumar: *IJSRP* **5** (2015) 2. <https://pdfs.semanticscholar.org/73ab/07ce475bb3bd8b0361469391d2c07dcc01cd.pdf>
- 19 N. Rahman, I. Abustan, S. Talib, M. Abustan, M. Ali, and H. Gotoh: *ICCEE* **34** (2018) 1. <https://doi.org/10.1051/e3sconf/20183401023>
- 20 Daily Nation: <https://www.nation.co.ke/lifestyle/How-we-walk-depends-on-who-we-walk-with/1190-4774810-awkjfd/index.html> (accessed September 2019).
- 21 M. Fenton: *The Complete Guide to Walking for Health, Weight Loss, and Fitness* (Lyons Press, 2008).
- 22 C. S. Silva and P. Wimalaratne: *Sens. Mater.* **31** (2019) 4. <https://doi.org/10.18494/SAM.2019.2232>

