ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

AJNA – A Smart Assistance for the Visually and Cognitively Impaired

Elaiyaraja K[1], Aishwarya K[2], Akshaya G[3], Monicka P[4], Monikha R[5]

Assistant Professor (Senior grade), Dept. of IT, Valliammai Engineering College, Kattankulathur, Chennai, India [1]

UG Student, Dept. of IT, Valliammai Engineering College, Kattankulathur, Chennai, India [2]

UG Student, Dept. of IT, Valliammai Engineering College, Kattankulathur, Chennai, India [3]

UG Student, Dept. of IT, Valliammai Engineering College, Kattankulathur, Chennai, India [4]

UG Student, Dept. of IT, Valliammai Engineering College, Kattankulathur, Chennai, India [5]

ABSTRACT: Recognizing who a person is, looking at their face, is a trivial task for most humans. However, this ordinary day-to-day activity is something that is very challenging for those who are visually or cognitively impaired. In recent days, machines have become as good as or better than humans in tasks like face recognition, speech recognition, etc., owing to the increase in the amount of data and processing power. This project hopes to leverage these recent developments and develop a system which automatically recognizes faces and helps a visually or cognitively impaired person identify whom they are talking to or what they are looking at. Existing systems require expensive smart glasses which have limited processing capability and are not energy efficient. Users have reported that their glasses get heated up which might have health impacts due to smart glasses proximity to the face. The plan is to integrate a head-mounted camera in an optical glass along with microphones and speakers. ; the camera is turned on to capture the image only upon receiving a voice command from the user to recognize faces and offloads the computation to nearby devices like smart phones or laptops. This methodology does not heat up the glasses and are inexpensive.This approach is also energy efficient. Initially the system will be designed as a prototype in pc/mobile phone and later will be deployed with optical glasses for efficient use.

KEYWORDS: Face recognition, Speech recognition, Head-mounted glasses, and voice command.

I. INTRODUCTION

Of the 37 million people around the world who are visually impaired, about 15.0 million are found in India. India has now the largest number of visually impaired persons. Identifying objects and persons is a very difficult task for visually impaired. So, the aim is to develop a smart system which aids the visually impaired. It helps them in the basic and regular activities of their life. The smart system uses the first-person vision enabled by smart glasses which is a wearable device. The system detects and recognizes known faces and translates the information into speech and conveys to the visually impaired. The system is trained with known faces and objects to enable such detections.

II. PROPOSED WORK

The prototype consists of an optical glass with head-mounted camera that captures images and offloads the computation to the database. The database is built using Microsoft Azure API and consists of details of all known persons to the visually impaired. Based on the computations, the database performs facial recognition and identification. If it is a face which already exists in the database, then the person's name and face features will be returned. If it is an unknown face, then it returns that the person is unknown to the visually impaired. With the help of identification information obtained from the unknown person, he/she can be added to the database. The Text to Speech API converts the input text into human-sounding custom speech and returned as an audio file.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

III. KEY FEATURES OF THE SYSTEM

1. To provide a wearable device as an assistance for the visually impaired.
2. To help the visually impaired easily recognise known faces.
3. To help them navigate through obstacles.
4. To give knowledge to the visually impaired about the surroundings.
5. To bridge the gap between visually impaired persons and normal persons.

IV. SYSTEM SETUP

The visually impaired person wears an optical glass mounted with wireless camera and microphones. When a person comes in front of the visually impaired and speaks, the user gives voice command to the system and it turns on the camera. Once the camera is turned on, it captures the image of the person. The captured image is sent for facial recognition by checking it in the offline and the trained database. If the captured image is similar to any of the images in the offline database, it performs facial recognition and identification. It then converts the text format details of the person into speech format through Text to Speech conversion API and returns a human soundable audio file to the user through speakers.

A. BLOCK DIAGRAM

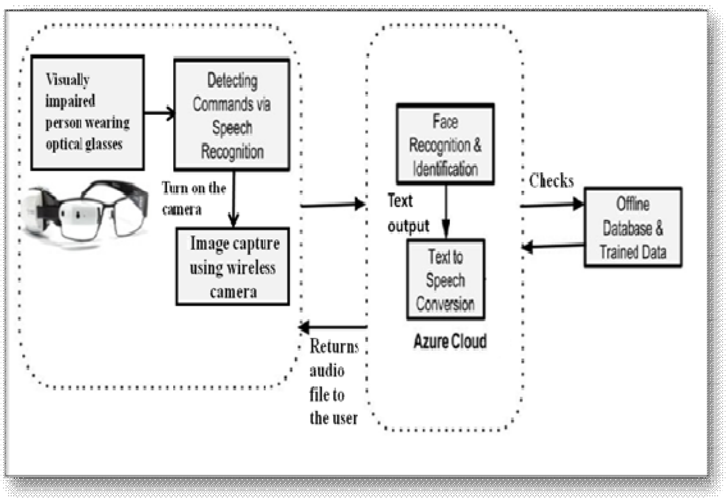

Fig.1 Architecture Diagram

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

B. METHODOLOGY

The methodologies that are adopted for AJNA – A Smart assistance for the visually and cognitively impaired are listed below.

1. CREATING DATABASE (PERSONGROUP):

The first step is to build the database which consists of all the persons details the visually impaired person is familiar to. We make use of the Microsoft Azure Cognitive Services for this purpose. A new person group with person GroupId, name, and optional user-provided user Data is created. The person group consists of the data, including face images and face recognition features of the uploaded persons. The person's face, image, and userData will be stored on server until Person Group Person - Delete or PersonGroup - Delete is called. 1,000-person groups can be created where each holds up to 1,000 persons.

2. ADDING FACES TO PERSONS IN THE DATABASE:

Initially, a face image is added to a person belonging to a person group for face identification or verification. A persistedFaceId is returned which represents the added face. Instead of the actual face image, the extracted face features will be stored on server until the person group is deleted.

1) A person in a person group can hold up to 248 faces.

2) The supported formats are JPEG, PNG, GIF and BMP and allowed image file size varies from 1KB to 6MB.

3) Sequential processing is carried out for adding or deleting faces to or from a same person and parallel processing for adding or deleting faces to or from different person.

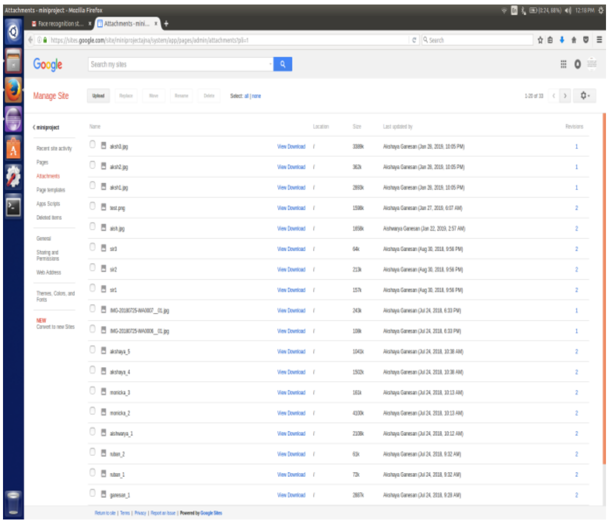

Fig.2 Adding faces to persons in the database.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

3. TRAIN THE DATABASE:

A person group is submitted for training. To perform Face-Identify training the person group is significant which is an asynchronous task.The number of person entries and their faces in a person group will determine the time taken to complete the training process which could be several seconds to minutes. The training status can also be checked.

4. FACE DETECTION AND IDENTIFICATION:

In each image the human face is detected, and face rectangles are returned along with faceIds, landmarks, and attributes. faceId, landmarks, and attributes including age, gender, emotion, are returned which are optional. faceId will be used in Face - Identify. The supported formats are JPEG, PNG, GIF, and BMP and the allowed image file size varies from 1KB to 6MB.64 faces can be returned for an image.

5. TEXT TO SPEECH CONVERSION:

The Text to Speech (TTS) API converts the text into a speech that is understandable by the user (also called *speech synthesis*). HTTP POST requests are sent from the application to the Speech service to generate speech. The input text is converted into human- sounding custom speech and returned as an audio file. Many voices and languages are supported by the service. There are more than 75 voices which supports 45 and more languages. To use these standard "voice fonts", you only need to specify the voice name with a few other parameters when you call the service's REST API.

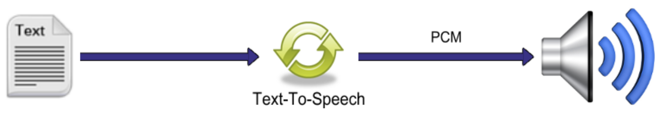

Fig.3 Text to speech conversion

6. SMART GLASSES AND THE USER INTERFACE:

The first-person vision enabled by smart glasses to capture scenes in front of a visually impaired person. However, keeping the camera on all the time will be energy consuming. Thus, we propose to turn on the camera on prompt from the user. For example, if the user asks a question to the smart system: Is there anyone in front of me? Another way is to capture the scene periodically such as once in two seconds. It is required to emulate a smart glass by connecting a wireless camera mounted on an optical glass to either a smart phone or a nearby laptop. The captured image from the smart glass is processed to recognize objects or faces. The result is converted to speech and sent to the headphones connected wirelessly which is then conveyed to the visually impaired user.

V. MODULES WITH WORKING PRINCIPLES

The modules of the system are as follows:

A. SPEECH RECOGNITION MODULE:

When a person appears in front of the visually impaired and speaks , the microphone detects the voice . On recognition of the voice it commands the system to turn on the camera.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

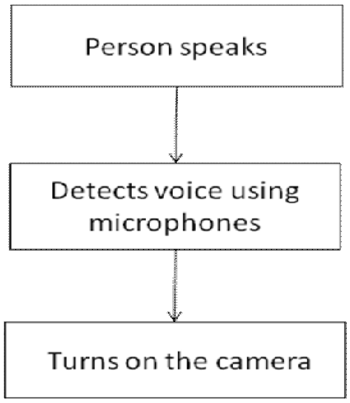

Fig.4 Speech recognition

B. IMAGE CAPTURING MODULE:

Once the camera gets turned on, it captures the image of the person standing in front of the visually impaired. Then the captured image is offloaded to the database for performing facial recognition.

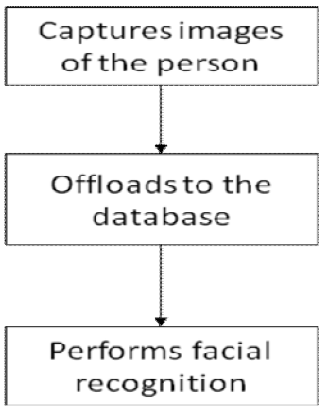

Fig.5 Image Capturing.

C. FACIAL RECOGNITION MODULE:

If the captured image is similar to the image in the offline database(known persons to the visually impaired) , then it returns the identity of that person. Otherwise, it returns the facial features of the unknown person and asks if the user wants to add that person to the existing database.

 DOI:10.15662/IJAREEIE.2019.0802017

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

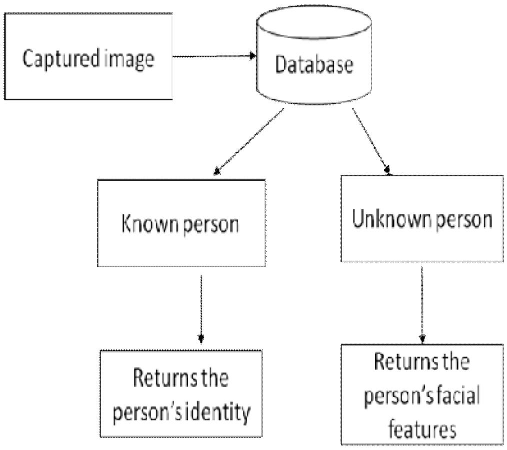

Fig.6 Facial Recognition.

D. TEXT TO SPEECH CONVERSION MODULE:

The output (text) which is returned is converted to a human soundable custom speech. The output in speech format is returned to the visually impaired as an audio file through speakers.

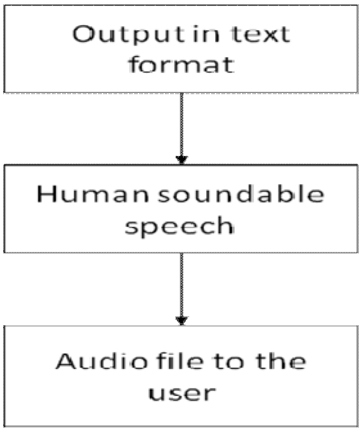

Fig.7 Text to speech conversion

E. OFFLINE DATABASE:

The image of the persons already known to the visually impaired is stored in an offline database separately. The public URL of the database is specified in the code for face recognition.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Website: www.ijareeie.com

Vol. 8, Issue 2, February 2019

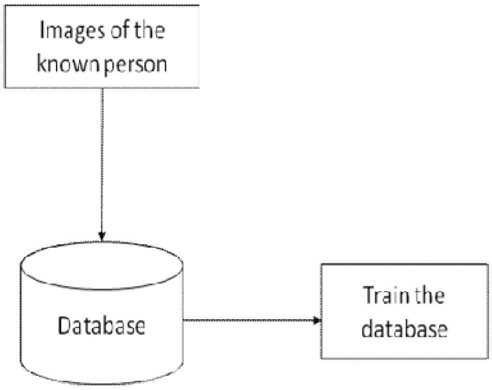

Fig.8 Offline Database

VI. RESULT AND CONCLUSION

Thus, this system provides smart assistance to the visually impaired in an efficient manner. The wearable device which is designed as a prototype is inexpensive and energy efficient which satisfies the primary aim of the system. It also helps the visually impaired navigate through obstacles and gives knowledge about the surroundings. Thus, this prototype bridges the gap between the visually impaired persons and normal persons by making the interaction more easier.

REFERENCES

[1] J.Bai, S.Lian, Z.Liu, K.Wang and D.Liu, "Smart guiding glasses for visually impaired people in indoor environment," in *IEEE Transactions on Consumer Electronics*, vol. 63, no. 3, pp. 258-266, August 2017.
[2] B. Söveny, G. Kovács and Z. T. Kardkovács, "Blind guide - A virtual eye for guiding indoor and outdoor movement," in *2014 5th IEEE Conf.Cognitive Infocommunications (CogInfoCom)*, Vietri sul Mare, 2014, pp.343-347.
[3] C. Álvarez Casado, M. Bordallo López, J. Holappa and M. Pietikaïnen, "Face detection and recognition for smart glasses," *2015 International Symposium on Consumer Electronics (ISCE)*, Madrid, 2015, pp. 1-2.
[4] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: A survey," *IEEE Trans. Systems, Man, Cybern.*, vol. 40, no. 1, pp. 25-35, Jan. 2010.
[5] W. Xu, Y. Shen, N. Bergmann, and W. Hu, "Sensor-assisted face recognition system on smart glass via multi-view sparse representation classification," in IPSN.
[6] L. Tian, Y. Tian and C. Yi, "Detecting good quality frames in videos captured by a wearable camera for blind navigation," in *2013 IEEE Int. Conf. Bioinformatics and Biomedicine*, Shanghai, 2013, pp. 334-337.
[7] M. Moreno, S. Shahrabadi, J. José, J. M .H. du Buf and J. M. F. Rodrigues, "Realtime local navigation for the blind: Detection of lateral doors and sound interface," in *Proc. 4th Int. Conf. Software Development for Enhancing Accessibility and Fighting Info-exclusion*, 2012, pp. 74-82.
[8] D. Dakopoulos and N. G. Bourbakis, "Wearable obstacle avoidance electronic travel aids for blind: A survey," *IEEE Trans. Systems, Man, Cybern.*, vol. 40, no. 1, pp. 25-35, Jan. 2010.
[9] A. Aladrén, G. López-Nicolás, L. Puig and J. J. Guerrero, "Navigation assistance for the visually impaired using RGB-D sensor with range expansion," *IEEE Systems J.*, vol. 10, no. 3, pp. 922-932, Sept. 2016.
[10] H. Fernandes, P. Costa, V. Filipe, L. Hadjileontiadis and J. Barroso, "Stereo vision in blind navigation assistance," *2010 World Automation Congr.*, Kobe, 2010, pp. 1-6.
[11] W. Heuten, N. Henze, S. Boll and M. Pielot, "Tactile wayfinder: a non-visual support system for wayfinding," *Nordic Conf. Human Computer Interaction*, Lund, 2008, pp. 172-181.
[12] Mankar, Vijay & G Bhele, Sujata. (2012). "A Review Paper on Face Recognition Technique"s. International Journal of Advanced Research in Computer Engineering & Technology. 1. 339-346.
[13] Ishwar S. Jadhav, V. T. Gaikwad, Gajanan U. Patil,"Human Identification Using Face and Voice Recognition", International Journal of Computer Science and Information Technologies, Vol. 2 (3), 2011.