

A Review on Smart Assistant for Visually Impaired

B.L.Srinivasa Prasad

Associate Professor – T. John College, Bangalore

Abstract- The most common means of communication is language and the basic medium is speech. The interaction among humans and computer that is their communication is called human computer interface (HCI). Physical movement is one of the biggest challenges for the visually impaired. People with complete blindness or low vision often have a difficult time in self-navigating unfamiliar environments. Eventually people with blindness or low vision pose great difficulty in crossing lanes or move in huge crowd, So many people with blindness or low vision tend to bring a friend or family member for assistance. A personal digital assistant (PDA) was chosen because it combines small-size, computational resources and low cost price. Key technologies necessary are: Voice processing, Image processing, object recognition and speech synthesis. So that the problems faced by blind people can be reduced.

Index Terms- HCI (Human computer interface), Object recognition, Image processing, Voice processing.

INTRODUCTION

India is now the world's largest number of blind visually impaired, India is home to 1.2 crore blind people of which three million suffer from corneal disorder, present India has 115 eye donation center. Many of the people cannot afford eyes treatments. To be categorized as blind, there is a complete loss of vision. Blindness cannot be improved by simple visual aids. So this paper present a system to aid the blind people. The Assistor Device is a device which is a passive type intelligent glasses that focuses on aiding the blind or people with low vision capabilities to move around from one place to another place without having to worry about anything around. This navigation can mainly be categorized as vision enhancement, vision replacement and vision substitution. Vision replacement systems provide the visual cortex of the human brain with the information either directly or via the optic nerve. Vision substitution and Vision enhancement systems have almost same working principles with respect to

environment detection process, however, each provides the environmental information differently. Vision enhancement puts the information in a visual manner, whereas vision substitution typically uses tactual or auditory perception or both. We propose a Smart system to help such visually impaired people in their basic activities of life.

Our System mainly consists of two components: 1) Wireless Camera 2) Android Device. The main objective is to simplify the system for the users and make interaction between the two entities as easy as possible. The entire project is dependent on the Smartphone App and its reliability. A separate database is designed, where the definition of the objects are found. In the system level we could say that the originality lies in the real-time application working on the Smartphone.

LITERATURE SURVEY

There are a multiple devices which assist the visually challenged for navigation indoor and outdoor. All these devices depends mainly on Global Positioning System (GPS) alone, to navigate around.

A Stereo Image Processing System for Visually Impaired: The main task is to identify and assign importance to the objects based on its distance. In this work, the objects in both the images are identified by locating their boundaries. The above paper proposes a system utilizes stereo vision, image processing methodology to support blind navigation. The system consists of wearable system, Stereo camera, Helmet and Stereo earphone.

Drawbacks:

- Devising a real time image processing algorithm is relatively complex than conventional offline image processing.
- Cost of hardware of this is expensive and occupies more volume.

Visually Impaired Assistive System using VIAS user tracking information: This paper presents visually impaired assistive system (VIAS) which focuses on independent mobility of visually impaired or blind people who suffer in an unknown environment without any manual assistance. The VIAS use the RC8660FP TTS module to convey received text information in the form of speech to the user via headphone.

Drawbacks:

- The size is still large and giving service to multiple persons is difficult.
- This system uses low data rate as it transfers text data instead of audio file.

Abandoned object detection via temporal consistency modeling and back-tracing verification for visual surveillance: This paper presents effective approach for detecting abandoned luggage in surveillance video. Here the camera is fixed at a position making image captured to be in static if there are any irregularity is observed in the surveillance the objects are detected.

Drawbacks:

- Image captured is static.
- System background is needed to be known before use.

METHODS

Text-to-Speech:

Text to speech synthesis is a rapidly growing aspect of computer technology and is increasingly playing a more important role in the way we interact with the system and interfaces across a variety of platforms. We have identified the various operations and processes involved in text to speech synthesis. This module comprises of image and speech processing. The main aim of this module is to acquire a 3D world real image of any text constraints area and to convert this image into text followed by providing audio output using speech processing. We are implementing a dynamic system that makes use of Google API's for conversion of Text to Speech dynamically provided that good internet connectivity is present. For instance, if a book or any write-up is held in front of the camera frame capture the image of the text to be read by the user by using the camera.

Send the data to the android device via Bluetooth. Match each and every letter and provide the output to the user.

Object recognition:

Object recognition is a computer vision technique for identifying objects in images or videos. Object recognition is a key output of deep learning and machine learning algorithms. When humans look at a photograph or watch a video, we can readily spot people, objects, scenes, and visual details. The goal is to teach a computer to do what comes naturally to humans: to gain a level of understanding of what an image contains. . It becomes a tedious task for visually impaired person to locate these objects. This system gives sort of visual aid that dynamically identifies objects and locates them. For Example, a visually impaired person is sitting on his study table. He has multiple objects in front of him such as Water Bottle, Walking Stick, Fruits, etc. So this System will help him locate all his objects.

Face recognition:

Facial recognition is a category of biometric software that maps an individual's facial features mathematically and stores the data as a face print. The software uses deep learning algorithms to compare a live capture or digital image to the stored face print in order to verify an individual's identity. For example, an algorithm may analyze the relative position, size, and/or shape of the eyes, nose, cheekbones, and jaw.

ALGORITHMS

Text-to-Speech: There are a lot of things that go into making a TTS system from scratch and you need to have some low level knowledge to make the speech portion work. I've broken the steps down into a simple little list that covers the basics of what you would need to do.

Step 1: Text to words.

This may seem like something very easy to do, but remember; a computer is extremely dumb and does exactly what we tell it to. Written language is very ambiguous. We have words that sound different even when they have the same letters, for example read and read. (I read this post, you will read my response.)

Step 2: Words to phonemes

Phonemes - any of the perceptually distinct units of sound in a specified language that distinguish one word from another, for example p, b, d, and t in the English words pad, pat, bad, and bat. Compare with allophone.

There are only 26 letters in the English alphabet, but over 40 phonemes. Which makes all the different sounds used in actual speech?

Step 3: Phonemes to Sound

There are three different approaches. One is to use recordings of humans saying the phonemes, another is for the computer to generate the phonemes itself by generating basic sound frequencies (a bit like a music synthesizer), and a third approach is to mimic the mechanism of the human voice.

First option is what Dragon Speak uses, second option is used by Microsoft's TTS engine, and third option is used by Microsoft's Cortana.

Step 4: Putting it all together

Face Recognition:

Step1: prepare a data of known or recognized faces.

Step2: prepare the known data to be read.

Step3: Calculate the Eigen faces from the dataset (Training set), keeping only m images that correspond to highest eigenvalues. These m images define the face space.

Step 4: Loading any image format (.bmp, .jpg, .png) from given source. Detecting the faces present in the images using the Cascade Classifier in opencv.

Step5: We will use Face Recognizer for face recognition.

Step 6: Calculate set of weights based on input image and the m Eigen faces by projecting the input image onto each Eigen face

Step 7: Determine if image is known or unknown by checking to see if the image is sufficiently close to face space.

CONCLUSION

I've learned many things through my research of the visually impaired. Of all the things, the most important thing I've read is the obstacles and experiences people who are visually impaired go through. The shift in their lives and how they must adapt to their lack of eyesight is astounding. This

paper puts forward a system which allows blind people or people with low vision to detect and avoid hurdles/obstacles were implemented as an android app. n. The proposed system could be also applied to smart phones that have android OS and infrared emitter & sensor, thus enabling us to measure a depth map of the environment. This system could be advanced with GPS data, an important feature to most of the android smart phones.

REFERENCES

- [1] Autonomous Walking Stick for the Blind Using Echolocation and Image Processing.
- [2] A Stereo Image Processing System for Visually Impaired.
- [3] Kevin Lin, Shen-Chi Chen, Chu-Song Chen, Daw- Tung Lin, Senior Member, IEEE, and Yi-Ping Hung.
- [4] Smart vision for the blind people
- [5] Semantic scholar.org
- [6] acadpubl.eu