International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 08 | Aug 2020 www.irjet.net p-ISSN: 2395-0072

DETECTION OF HAND GESTURE, CONVERSION TO SPEECH AND TEXT

Rohith H S[1], Chethan kumar H M[2], Chetan shivapura shetru[3], Girish T[4], K Nidish[5]

[1]Assistant Professor, Department of Electronics and Communication Engineering, East West Institute of Technology, Karnataka.
[2-5]BE. Students, Department of Electronics and Communication Engineering, East West Institute of Technology, Karnataka.

ABSTRACT Generally deaf-dumb people use sign language for communication, but they find difficulty in communicating with others who don't understand sign language. in our country around 2.78% of peoples are not able to speak (dumb). Their communications with others are only using the motion of their hands and expressions. We proposed a new technique called artificial speaking mouth for dumb people. It will be very helpful to them for conveying their thoughts to others. And a Text-to-speech synthesizer is an application that converts text into spoken word, by analyzing and processing the text using Natural Language Processing (NLP) and then using Digital Signal Processing (DSP) technology to convert this processed text into synthesized speech representation of the text. Here, we developed a useful text-to-speech synthesizer in the form of a simple application that converts inputted text into synthesized speech and reads out to the user which can then be saved as an mp3.file. The development of a text to speech synthesizer will be of great help to people with visual impairment and make making through large volume of text easier.

KEYWORDS Raspberry Pi, Hand Gestures, Image processing.

I.INTRODUCTION

Text-to-speech synthesis -TTS - is the automatic conversion of a text into speech that resembles, as closely as possible, a native speaker of the language reading that text. Text-tospeech synthesizer (TTS) is the technology which lets computer speak to you crisis picture of public funds going to waste, our Microcontroller commands the stepper motor to stop, at the same time with a delay the DC motor connected to the load cell with a plastic container attached rotates by a measure of 90- degree hence emptying the contents in a bowl which belong to the user. The TTS system gets the text as the input and then a computer algorithm which called TTS engine analyses the text, pre-processes the text and synthesizes the speech with some mathematical models. The TTS engine usually generates sound data in an audio format as the output. The text-to-speech (TTS) synthesis procedure consists of two main phases. The first is text analysis, where the input text is transcribed into a phonetic or some other linguistic representation, and the second one is the generation of speech waveforms, where the output is produced from this phonetic and prosodic information. These two phases are usually called high and low-level synthesis.

II. BLOCK DIAGRAM

The generalized block diagram of hand gesture recognition system shown in the figure Total flow of the implementation is divided in to the gesture acquisition, image preprocessing, feature extraction and classification of gesture. Gesture recognition is contour based recognition therefore there is no use of color information of image. Thus color images are converted in to the gray image which further reduce size of image and reduces image processing time. Sometimes gesture images may affected due to blur caused by movement of camera human, unequal light distribution and shadow.

To remove this histogram equalization is used. For the texture feature extraction LBP is used and for classification KNN classifier is used.The concept of local binary pattern (LBP) was introduced for texture.

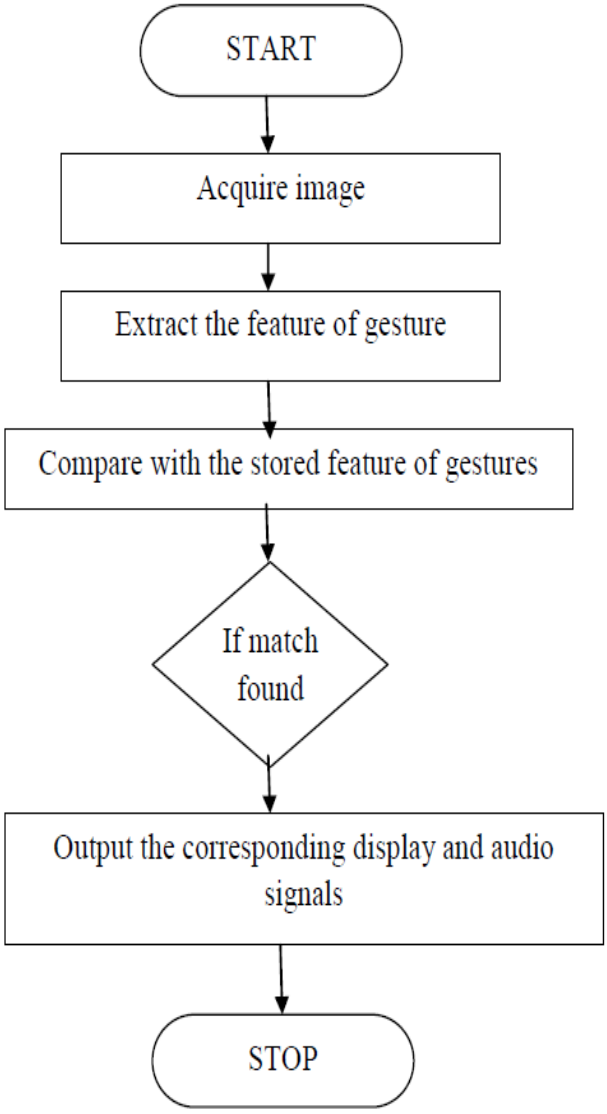

Figure 1: Block diagram of proposed system.

III. DESIGN AND IMPLEMENTATION

A. RASPBERRY PI

The Raspberry Pi is a series of small single-board computers developed in the United Kingdom by the Raspberry Pi Foundation to promote teaching of basic computer science in schools and in developing countries. The original model became far more popular than anticipated, selling outside its target market for uses such as robotics. It is now widely used even in research projects, such as for weather monitoring because of its low cost and portability.

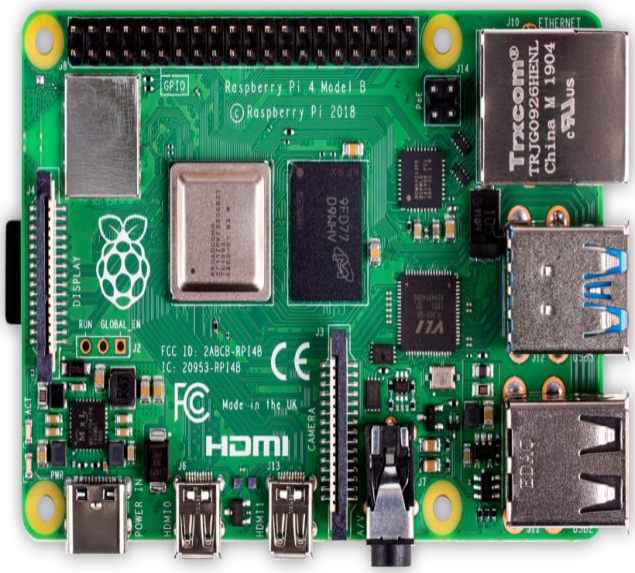

Figure 2: Raspberry pi

B. Tesseract

In geometry, the **tesseract** is the four-dimensional analogue of the cube; the tesseract is to the cube as the cube is to the square. Just as the surface of the cube consists of six square faces, the hypersurface of the tesseract consists of eight cubical cells. Since each vertex of a tesseract is adjacent to four edges, the vertex figure of the tesseract is a regular tetrahedron. The dual polytope of the tesseract is called the regular hexadecachoron , or 16-cell, with Schläfli symbol {3,3,4}, with which it can be combined to form the compound of tesseract and 16-cell. Projections on the 2D-plane become more instructive by rearranging the positions of the projected vertices. In this fashion, one can obtain pictures that no longer reflect the spatial relationships within the tesseract, but which illustrate the connection structure of the vertices.

C. e-Speak

eSpeak s a compact open source software speech synthesizer for English and other languages, for Linux and Windows. eSpeak uses a "formant synthesis" method. This allows many languages to be provided in a small size. The speech is clear, and can be used at high speeds, but is not as natural or smooth as larger synthesizers which are based on human speech recordings. Includes different Voices, whose characteristics can be altered. Can produce speech output as a WAV file., SSML (Speech Synthesis Markup Language) is supported (not complete), and also HTML. Compact size. The program and its data, including many languages, totals about 2 Mbytes

Figure 4: eSpeak

D. Implementation

- Step 1: Start.
- Step 2: Choose option to convert image to speech.
- Step 3: Capture and read the gesture using the camera.
- Step 4: Find contours and object of the image.
- Step 5: Remove the defect edges.
- Step 6: Convert RGB image to Grey scale
- Step 7: Compare the image with the stored data.
- Step 8: If the data is not found again recapture image.
- Step 9: perform Grey scale conversion.
- Step 10: Stop.

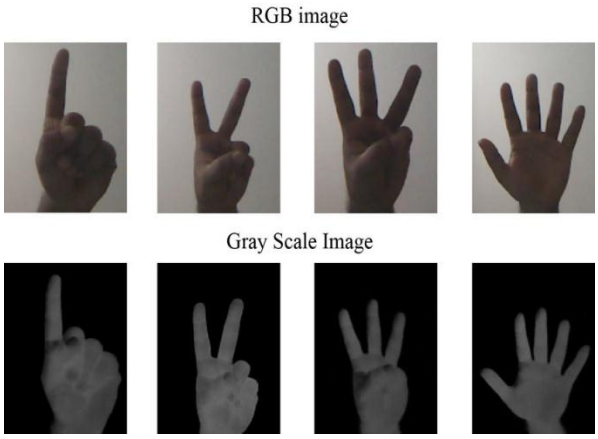

Figure 5: Image transformation.

Conclusion

The practical adaption of the interface solution for visually impaired and blind people is limited by simplicity and usability in practical scenarios. As an easy and practical way to achieve human-computer- interaction, in this solution hand gesture to speech and text conversion has been used to facilitate the reduction of hardware components. On the whole, the solution aims to provide aid to those in need thus ensuring social relevance. The people can easily communicate with each other. The user-friendly nature of the system ensure that people can use it without any difficulty and complexity. The application is cost efficient and eliminates the usage of expensive technology.

References

- D. K. Sarji, "HandTalk: assistive technology for the deaf,"Computer, vol. 41, pp. 84-86, 2008
- Ibrahim, P., and Srinivasa, R. “Automated speech recognition approach to continuous cue symbols generation”, International journal of power control signal and computation, Vol. 18, No. 8,pp. 434-520,2007.
- Emil M. Petriu, Qing Chen, Nicolas D. Georganas, Real-time vision-based hand gesture recognition using haar-like features, 2017.
- T.H. Speeder, "Transformation human hand motion for Telemanipulation," Presence, Vol.1, no. 1, pp.63-79,1992.
- A.W. Fitzgibbon, and J. Lockton, “Hand Gesture Recognition Using Computer Vision”, BSc.Graduation Project,
- D. Xu, "A neural network approach for hand gesture recognition in virtual reality driving systern

of SPG," presented at the 18 Inc.Conf. Pattern recognition, 2006.

- Mehdi S.A., Khan Y. N., "Sign language recognition using sensor gloves", Proceedings of the 9th International Conference, Volume:5, IEEE Conference Publications, 2002.
- K.C. Shriharipriya and K. Arthy, "Flex sensor based hand gesture recognition system", Proceedings of International Journal Of Innovative Research And Studies (IJIRS), Vellore, India, May 2013.
- Kunal Kadam, Rucha Ganu, Ankita Bhosekar, Prof. S. D. Joshi, "American Sign Language Interpreter", Proceedings of the IEEE Fourth International Conference on Technology for Education, 2012.
- Mehdi S.A., Khan Y. N., "Sign language recognition using sensor gloves", Proceedings of the 9th International Conference, Volume:5, IEEE Conference Publications, 2002.