FORM 2

THE PATENTS ACT 1970

(39 OF 1970)

&

The Patents Rules, 2003

COMPLETE SPECIFICATION

(See Section 10 and Rule 13)

A SYSTEM AND METHOD FOR COMMUNICATION BY A VOCAL AND AUDITORY IMPAIRED PERSON

Applicant:

Name: B.S ABDUR RAHMAN CRESCENT INSTITUTE OF SCIENCE AND TECHNOLOGY

Address: SEETHAKATHI ESTATE, GST ROAD, VANDALUR, CHENNAI - 600048, TAMIL NADU

Nationality: INDIAN

PREAMBLE TO THE DESCRIPTION

The following Specification particularly describes the Invention and the manner in which it is to be performed.

FIELD OF THE INVENTION

The present subject matter relates to the field of assistive technology. In particular, the present subject matter relates to communication enablement for vocal and auditory impaired person.

BACKGROUND OF THE INVENTION

Sign language is the nonverbal communication of the vocally & hearing impaired. A sign language combines hand shapes, orientation and movement of the hands, arms, or body, facial expressions, and lip patterns to convey meaning instead of using words. The Vocally & hearing impaired community seeks the assistance of sign language interpreters to help them communicate their thoughts to normal people and vice versa. The Interpreters are basically made of flex sensors and hence very expensive.

Suthagar S. et.al ("Translation of Sign Language for Deaf and Dumb People", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-5, January 2020) was to analysed and translated the sign language that is hand gestures into text and voice. For this process, Real Time Image were made by vocally and hearing impaired people is captured and it is given as input to the pre-processor. Then, the feature extraction process by using otsu's algorithm and classification by using SVM [support Vector Machine] can be done. The corresponding sign was produced after the text conversion. The obtained text was converted into voice with the use of MATLAB function. Thus hand gestures made by vocally & hearing impaired people had been analysed and translated into the text and voice for better communication. The Recognition of sign language involves pre-processing level and classification level. Pre-processing level involves gray scale conversion, noise reduction, background subtraction, brightness normalization and scaling operation. For gesture recognition, feature extraction using Otsu's algorithm and classification using SVM classifier were used.

Dipali Dhake et.al [Prof. Dipali Dhake , Manisha P. Kamble , Shrushti S. Kumbhar, Sana M. Patil , “Sign language communication with dumb and deaf people” , International Journal of Engineering Applied Sciences and Technology, 2020]. In this application, Image processing technique was been used it is known as Histogram of Gradient (HOG) with Artificial Neural Network (ANN). These two technique are used to train the system or create database and that database is stored in memory. The web camera is used to capture a continuous image of various sign gestures, which is sent to the Raspberry Pi. Raspberry pi controller recognizes the image and compare the same image with the database which had been stored in the memory of raspberry pi controller. This is how vocally & hearing impaired people's sign language is recognised. By using Voice Recognition Module(VRM), voice of normal people is converted to text with the help of Artificial Neural Network(ANN) that text is converted to various sign. In this way, two way communication between vocally & hearing impaired people and normal people takes place.

However, the processing of real time communication is computationally intensive and the devices are expensive given the high end technologies used. Further, in situations where most of the communications are along expected lines, similar amount of processing would not be necessary. For instance, if a hearing impaired person goes to railway station, they only have to have to ask standard questions like “where is the ticket counter?”, “where is the waiting room?”. Similarly, when going to a restaurant, the standard questions will be “I have a reservation”, “please give the menu card”, “please bring the bill”. Using similar processing tools to process such standard comments is a waste of computing power and also time.

Hence a system which has a simpler cheaper construct to optimise processing of standard communication can augument a system which is designed for general deciphering making it more efficient. Further, as a stand alone unit this will improve the communication in standard settings by reducing the processing time.

OBJECTIVE OF THE INVENTION

The primary objective is to develop a system to enable communication between hearing impaired and normal people.

Another objective is to develop a system which is adapted to handle standard conversation in well-defined places and situations to reduce computation power.

Yet another objective is to develop a system for faster communication in standard settings.

SUMMARY OF THE INVENTION

The present subject matter is a system for communication between a hearing impaired person and a normally able person. The system comprises a Raspberry Pi 4, Display, Raspberry Pi camera, a speaker, a mic and a location sensor. Raspberry pi 4 is configured to classify a specific hand gesture. Raspberry Pi camera and location sensor are interfaced with raspberry pi 4. Based on the inputs from the location sensor, the user history, the real time conversation, external factors, and the person addressed, a set of standard phrases are generated by an AI module and can be overruled by an user. Each standard phrase is linked to a specific hand gesture. The hand gestures when captured through pi camera, processed with CNN model and corresponding voice command is audible through speaker or earphone. When a hearing-impaired person shown the gesture to pi-camera and the output will be heard to ordinary person. When an ordinary person speaks, the speech is converted using an application to text and the output will be displayed to hearing-impaired person in LCD Display.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates system and method for communication between a hearing impaired person and a normally able person, as an embodiment of the present subject matter.

Figure 2 illustrates the Raspberry Pi 4 unit, as an embodiment of the present subject matter.

Figure 3 illustrates the LCD display unit, as an embodiment of the present subject matter.

Figure 4 illustrates the Raspberry Pi camera, as an embodiment of the present subject matter.

Figure 5 is the example implementation of the system and method for five hand gesture signs, as an embodiment of the present subject matter.

Figure 6 is the Voice to Text Conversion Experimental Setup, as an embodiment of the present subject matter.

Further, those skilled in the art will appreciate that elements in the figures are illustrated for simplicity and may not have necessarily been drawn to scale. Furthermore, the figures may show only those specific details that are pertinent to understanding the embodiments of the present invention so as not to obscure the figures with details that will be readily apparent to those skilled in the art having the benefit of the description herein.

DETAILED DESCRIPTION OF THE INVENTION

The present subject matter is a system for communication between a hearing impaired person and a normally able person. The system comprises a Raspberry Pi 4, LCD Display (16 x 2), Raspberry Pi camera, and a location sensor. CNN algorithm is implemented in the raspberry pi 4 to classify a specific hand gesture. Raspberry Pi camera and location sensor are interfaced with raspberry pi 4. Based on the inputs from the location sensor, the user history, the real time conversation, external factors, and the person addressed, a set of standard phrases are generated by a AI module and can be overruled by an user. Each standard phrase is linked to a specific hand gesture. The hand gestures when captured through pi camera, processed with CNN model and corresponding voice command is audible through speaker or earphone. When a hearing-impaired person shown the gesture to pi-camera and the output will be heard to ordinary person. When an ordinary person speaks, the speech is converted using an application to text and the output will be displayed to hearing-impaired person in LCD Display.

Figure 1 illustrates system and method for communication between a hearing impaired person and a normally able person, as an embodiment of the present subject matter. The system comprises a PI camera, a Raspberry Pi, a speaker, a display, a mic, and a location sensor. The location sensor identifies a place and generates an input to the data modelling module in the Raspberry PI. The gestures and the mapped standard phrases dataset is formulated and used for faster deciphering of signs and improving the communication speed. The hand gestures are deciphered in the CNN classifier module and used for identifying the mapped standard phrases in the text module. The text is sent to the speaker for output of the communication message to the listener. An application designed for converting speech of normal person to text. The converted text message is sent to the display for viewing by the hearing impaired person.

The method of formulating standard phrases based on a given location involves using history data of a user conversations, and a database of location based details including typical user behaviour, the real time conversation, external factors, and the person addressed.

Figure 2 illustrates the Raspberry Pi 4 unit, as an embodiment of the present subject matter. The Raspberry Pi 4 unit acts as the processor for the system operations. It comprises an input output header, USB ports, Camera port, SD card port, audio port, display port, wireless and Ethernet units. The camera port is used for interfacing with the Pi camera and receiving the inputs of hand gesture images. Similarly, the speaker and the LCD unit are connected to the respective audio and display ports. User based inputs on the standard phrases of their choice is input through any of the input ports.

Figure 3 illustrates the LCD display unit, as an embodiment of the present subject matter. The display unit displays all alphabets, numbers, and custom made symbols. There are small pins along the printed small PCB board used for connection to the micro controller and there are 16 pins marked up with the numbers. The data pins ranging from 0-7 is used for data transmisssion to the micro controller. While Vss is connected to the ground, Vcc is for power supply,Vee pin for contrast control, RS pin for register select In which the data mode= 0, command mode =1 , RW pin is used to read and write the data, E (Enable pin) must be 1 to do read and write function, Pin 15 & pin 16 is used for the backlight supply.

Figure 4 illustrates the Raspberry Pi camera, as an embodiment of the present subject matter. Using the CSI port to connect a camera board, the camera system is controlled by the libcamera software stack. The libcamera stack has emerged, allowing third-party image sensors and Image Signal Processors to be integrated (ISPs). The drivers are configured to make the components of the Raspberry Pi imaging system work with libcamera in this article, focusing in particular on the processes of calibrating and adjusting the ISP to operate well with different image sensors. Furthermore, the libcamera integration avoids using any of the chip vendor's proprietary control techniques (Broadcom). Rather, Raspberry Pi provides its own control algorithms as open source code that customized. These algorithms operate directly on the chip's ARM cores.

The Pi camera module, is with a ribbon wire that must be attached to the Pi's CSI (Camera Serial Interface) port. The Pi is setup to enable inputs from Camera after connecting the hardware. The interfacing options are used to enable camera. After that, reboot the Pi and your camera module should be ready to use

In an example implementation of the present subject matter, a customised sign language system was devised by linking a set of hand gestures to a set of phrases. The system comprises a Raspberry Pi 4 with the features and specification:

- The CPU of Quad core cortex (64 bit) @ 1.5 Ghz .

- A GPU H264 (1080p60 decode, 1080p30 encode) OpenGL ES 3.0 graphics, H.265 (4kp60 decode)
- RAM of 4 GB.
- Operation Voltage – 5V With 3A.
- 28 GPIO ports.
- Bluetooth and Wi-fi connectivity.
- A operating temperature of about 0 to 50 degrees.
- Ports:
    1) 2 Mini HDMI Ports (4k supported).
    2) 2 x 2.0 and 2 x 3.0 USB Ports.
    3) Power supply - USB - C port.
    4) SD Card slot.
    5) LAN Port.
    6) Audio output Jack.
    7) PoE enable.
    8) 2 lane MIPI DSI Display port.
    9) 2 lane MIPI CSI Camera port.

The system further comprises a LCD display. The 16x2 LCD can display 16 characters per line on each of its two lines. Each character is presented in a 5x7 pixel matrix on this LCD. The 224 distinct characters and symbols can be displayed on the 16 x 2 dot matrix display. It is based on the HD 44780 Micro controller (Hitachi). It displays all alphabets, numbers, symbols that users custom build.

Features & Specification of 16 x 2 LCD Display:

- It operates at the voltage between 4.7 V to 5.3 V.
- The backlight colour is of blue or green.
- It works in both 4 and 8 bit modes.
- Pixel box of each character is of 5 x 8 pixel.
- Font size is of 0.125 x 0.200.
- It displays characters in 16 rows.
- It shows characters in 2 columns.
- It has 16 LCD Pins

The Raspberry Pi Camera Algorithm and Tuning Guide is for Raspberry Pi users who have an image sensor (camera) connected to the Raspberry Pi's CSI (Camera Serial Interface) camera port, such as one of the official Raspberry Pi camera boards using version 1 (Omnivision OV5647) or version 2 (Sony imx219) sensors, or the High Quality Camera (aka. the HQ Cam, based on the Sony imx477). Using the CSI port to connect a camera board. The camera system will be controlled by the libcamera software stack. The libcamera stack has emerged, allowing third-party image sensors and Image Signal Processors to be integrated (ISPs). write drivers to make the components of the Raspberry Pi imaging system work with libcamera in this article, focusing in particular on the processes of calibrating and adjusting the ISP to operate well with different image sensors. Furthermore, the libcamera integration avoids using any of the chip vendor's proprietary control techniques (Broadcom). Rather, Raspberry Pi provides its own control algorithms as open source code that users can view and modify. These algorithms operate directly on the chip's ARM cores.

Integrating Pi camera with Raspberry Pi:

The Pi camera module comes with a ribbon wire that is attached to the Pi's CSI (Camera Serial Interface) port. The Pi is set up to enable camera after connecting the hardware. The configuration window is opened using sudo raspi-config. Then camera is enabled under interfacing options. After that, the Pi is reboot and camera module is usable.

Features of Raspberry Pi camera:

- It has 5MP colour camera without mic.
- It supports both Pi A and B models.
- It is Omnivision 5647 Camera Module
- It has the resolution of 2592 * 1944
- It Supports the resolution of 1080p, 720p and 480p
- Light weight and portable (3g only).

TRAINING DATA

It consists of 70% of data.

In our dataset consist of **27,455** Images**.**

TEST DATA

It consists of 30% of data.

Data: **7172** Images.

a) Customized Dataset:

The location identified is a site seeing place, the user history is indicating meal time of the user, the real time conversation could be “Do you need anything?”, external factors is hot weather, and the person addressed is an employee in the place.

Created personal datasets with corresponding hand gestures.

1 - "I need food"

2 - "I need water"

3 - "I need help"

4 - "How are you"

5 - "what is your name"

CNN algorithm are chosen the best model , implemented in the raspberry pi 4. Raspberry Pi camera is interfaced with raspberry pi 4.The hand gestures is captured through pi camera, processed with CNN model and corresponding voice command audible through speaker or earphone. In the figure 5 (a), hearing-impaired person shown the gesture (1) to pi-camera and the output will be heard to ordinary person in Earphone - **I need food**.

In figure above 5(b), hearing-impaired person shown the gesture (2) to pi-camera and the output will be heard to ordinary person in Earphone - **I need water.** In figure 5(c), hearing-impaired person shown the gesture (3) to pi-camera and the output will be heard to ordinary person in Earphone - **I need help.** In figure 5(d), hearing-impaired person shown the gesture (4) to pi-camera and the output will be heard to ordinary person in Earphone - **How are you.** In figure 5(e) hearing-impaired person shown the gesture (5) to pi-camera and the output will be heard to ordinary person in Earphone **- What is your name.**

Figure 6 is the Voice to Text Conversion Experimental Setup, as an embodiment of the present subject matter. In figure 6 (a), when an ordinary person speak the output will be displayed to hearing-impaired person. In figure 6(b), an ordinary person spoke the output was displayed in LCD Display - **Have a Nice Day.**

There is no need of any costly hardware device or computationally intensive processing of data to obtain the communication message. The system is cost effective as only simple components are used which are easily accessible to the users. The scope of the system is limited to the standard phrases which have been mapped to the gestures. However, the option of customizability of the system configuration makes it user friendly and adaptable to any unique situation which the user is facing. It will be extremely valuable to the dumb deaf people greater options for vocally & hearing impaired persons.

In a preferred embodiment of the present subject matter, system for communication between a vocal and auditory impaired person and a normal person comprises a Raspberry Pi 4 unit, LCD Display, Raspberry Pi camera, a speaker, a mic and a location sensor. The Raspberry Pi 4 unit is configured to generate a set of standard phrases based on the inputs from the location sensor, the user history, the real time conversation, external factors, and the person addressed; and to map the standard phrase to a set of hand gestures. The LCD display to display the standard phases and the transcribed speech of normal person. The Raspberry Pi camera to record a hand gesture. The location sensor to identify the locality of the user. The mic to record the speech of normal person. The speaker for audible output of deciphered hand gesture. The standard phrases are customisable by the user and override the automatically generated standard phrases. The Raspberry Pi camera and location sensor are interfaced with raspberry pi 4 which is the processor of the system. The Raspberry Pi comprises an input output header, USB ports, Camera port, SD card port, audio port, display port, wireless and Ethernet units. The camera port is used for interfacing with the Raspberry Pi camera and receiving the inputs of hand gesture images. The speaker and the LCD unit are connected to the respective audio and display ports. The user based inputs for customisation of the standard phrases is through any of the input ports. The location sensor identifies a place and generates an input to a data modelling module in the Raspberry PI processor. The LCD display unit displays all alphabets, numbers, and custom made symbols; and comprises small pins on a printed small PCB board used for connection to the micro controller in the Raspberry Pi.

The preferred method, for communication between a vocal and auditory impaired person and a normal person, comprises identifying the locality of the user using the location sensor. Generating a set of standard phrases based on the inputs from the location sensor, the user history, the real time conversation, external factors, and the person addressed. The mapping of each of the set of standard phrases to an unique hand gesture and recording a hand gesture using Raspberry Pi camera. Deciphering the hand gestures in the classifier module and the mapped standard phrases are identified. The text is sent to speaker for output of the communication message to the normal person. The speech of normal person captured by mic is converted to text; and sent to the display for viewing by the hearing impaired person.

A person skilled in the art will appreciate that one or more of the elements may well be combined into a single functional element. Alternatively, certain elements may be split into multiple functional elements. Elements from one embodiment may be added to another. The order of the processes described herein may be changed and not limited to the manner described herein. Moreover, the actions of any flow diagram need not be implemented in the order shown. The scope of the embodiments is by no means limited by those specific examples.

We Claim

1. A system for communication between a vocal and auditory impaired person and a normal person, wherein the system comprises:
   a Raspberry Pi 4 unit, LCD Display, Raspberry Pi camera, a speaker, a mic and a location sensor;

   wherein

   the Raspberry Pi 4 unit is configured

   to generate a set of standard phrases based on the inputs from the location sensor, the user history, the real time conversation, external factors, and the person addressed;

   to map the standard phrase to a set of hand gestures;

   the LCD display to display the standard phases and the transcribed speech of normal person;

   the Raspberry Pi camera to record a hand gesture;

   the location sensor to identify the locality of the user;

   the mic to record the speech of normal person; and

   the speaker for audible output of deciphered hand gesture.

2. The system as claimed in claim 1, wherein the standard phrases are customisable by the user and override the automatically generated standard phrases.

3. The system as claimed in claim 1, wherein the Raspberry Pi camera and location sensor are interfaced with raspberry pi 4 which is the processor of the system.

4. The system as claimed in claim 1, wherein the Raspberry Pi comprises an input output header, USB ports, Camera port, SD card port, audio port, display port, wireless and Ethernet units.

5. The system as claimed in claim 4, wherein

   the camera port is used for interfacing with the Raspberry Pi camera and receiving the inputs of hand gesture images;

   the speaker and the LCD unit are connected to the respective audio and display ports;

   user based inputs for customisation of the standard phrases is through any of the input ports.

6. The system as claimed in claim 1, wherein the location sensor identifies a place and generates an input to a data modelling module in the Raspberry PI processor.
7. The system as claimed in claim 1, wherein the LCD display unit displays all alphabets, numbers, and custom made symbols; and comprises small pins on a printed small PCB board used for connection to the micro controller in the Raspberry Pi.
8. A method for communication between a vocal and auditory impaired person and a normal person, wherein the method comprises
   a. identify the locality of the user using the location sensor;
   b. generate a set of standard phrases based on the inputs from the location sensor, the user history, the real time conversation, external factors, and the person addressed;
   c. mapping each of the set of standard phrases to an unique hand gesture;
   d. record a hand gesture using Raspberry Pi camera;
   e. decipher the hand gestures in the classifier module and identify the mapped standard phrases;

f. send the text to speaker for output of the communication message to the normal person;

g. converting speech of normal person to text; and sent to the display for viewing by the hearing impaired person.

Dated this 4th day of September 2023

V. Sudarkodi(IN/PA 3959)
Authorised Agent for Applicant