June 11, 2019

Voice Signal Translation on Wireless-Communication Devices for the Speech and/or Hearing Impaired

Jian Yao

Voice Signal Translation on Wireless-Communication Devices for the Speech and/or Hearing Impaired

Abstract:

Hearing impaired and/or speech impaired individuals can experience difficulties in communicating with other people when utilizing voice transmission of wireless-communication devices, such as smartphones.

This publication describes techniques for aiding hearing impaired and/or speech impaired individuals in communicating during phone calls. The techniques include a translator application on a wireless-communication device that receives incoming voice signals, analyzes the voice signals, and generates a display (*e.g.*, digital text, sign language) on the user interface (UI) of the wireless-communication device for the user's visualization. Moreover, this publication describes techniques for aiding hearing impaired and/or speech impaired individuals, through use of a translator application on a wireless-communication device, to communicate with another individual utilizing voice transmissions. For instance, by providing an input (*e.g.*, typing a message utilizing an input device connected to the wireless-communication device or expressing language via gestures (*e.g.*, sign language) utilizing a video camera of the wireless-communication device) that is received by the translator application and converted to a voice signal that is then transmitted to the other individual. Additionally, the translator application can translate between different languages.

Keywords:

Translation, accessibility, assistive technology, communication, disability, hearing impaired, speech impaired, voice signal, voice transmission, sign language, user equipment, mobile device, application

Background:

The well-established means by which most humans can comprehend information is by use of their five senses (*e.g.*, sight, smell, taste, touch, hearing). Take for instance an ambulance speeding along down a congested road; flashing lights and blaring sirens collectively communicate one thing—an emergency. These lights and noises are perceived by bystanders' sight and hearing receptors, respectively, and communicate the information that there is an urgent situation. Perception of only one of these features, lights, or sirens, is still sufficient to convey the meaning. In other words, if the ability to see is absent, simply hearing the sirens can still warn that there is an accident.

There are many instances, however, where only one mode of communication is provided, such as stoplights and voice calls. Individuals who lack the ability to perceive or comprehend unilateral modes of communication are incapable of receiving the information contained in that communication form. The inability to hear or speak frequently makes otherwise conventional customs challenging. Take for example cellphone usage, the Pew Research Center (PRC) has found that 82% of Americans own cellphones. If a user is hearing impaired and/or speech impaired, a major utility of a cellphone, specifically the ability to make a call, is significantly more difficult. This is quite the inconvenience, for cellphones are an increasingly essential commodity for societal involvement. In fact, the PRC also recorded that the average adult cellphone owner makes and receives up to five voice calls a day.

Efforts to assist those with hearing and speech impairments have manifested in cellphone use by the conversion of voice signals to digital words and vice versa. This assistance may alleviate some of the difficulties associated with voice calls for the hearing impaired and/or speech impaired, but it is limited. It heavily relies on the hearing impaired and/or speech impaired person's ability to type grammatically well and fluently, such that a coherent audio message can be generated. Moreover, the typed message must be generated quickly, otherwise the other party to the call may think the lack of communication is due to poor connection. If the hearing impaired and/or speech impaired person had the additional option to respond by a display of gestures (*e.g.*, American Sign Language (ASL)), then this may provide not only a more convenient response method but also a hastier reply.

Therefore, it is desirable of an application to afford hearing impaired and/or speech impaired persons the option of receiving a voice signal in the form of a display of gestures or digital words, along with the choice of responding by means of text or a display of gestures and expecting these inaudible messages to be translated to a voice signal.

Description:

This publication describes techniques directed to aid hearing impaired and/or speech impaired individuals in communicating during a voice call. A wireless-communication device, such as a smartphone, includes a translator application (TA). The wireless-communication device performs operations under the direction of the TA to translate an incoming voice signal of a caller into either digital words or in the form of a display of gestures (*e.g.*, American Sign Language

(ASL)). This converted message will then be displayed on the user interface (UI). In addition, the TA enables the receiver to respond either by inputting the response message or in the form of a display of gestures. The selected message type can then be translated into a voice signal for the caller to receive. Finally, the TA can include a language translator that enables word translation between different languages.

Figure 1, below, illustrates an example wireless-communication device and the elements of the wireless-communication device that support the voice signal translation technique.

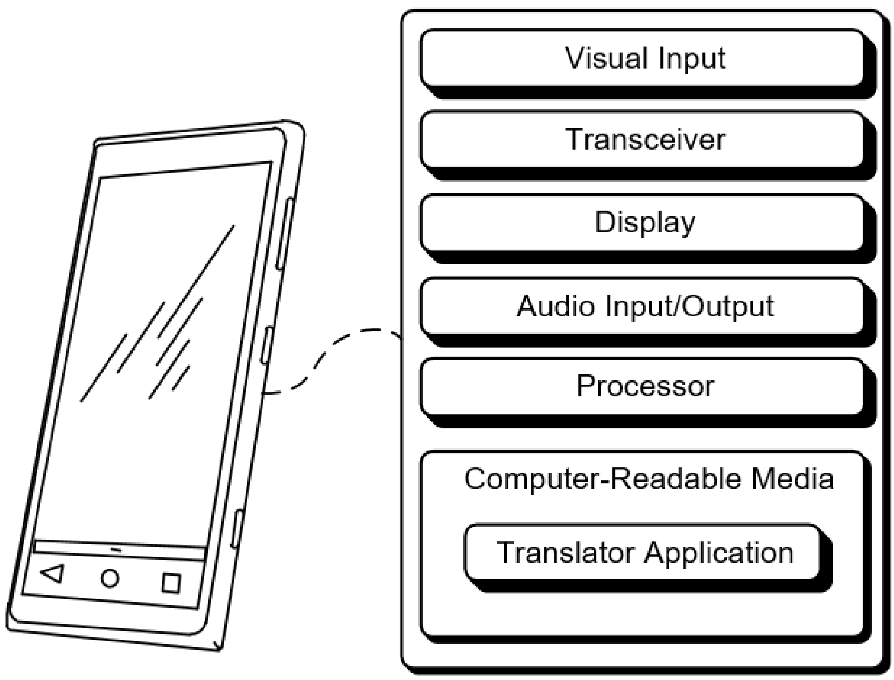

Figure 1

As illustrated, the wireless-communication device is a smartphone. The wireless-communication device includes a visual input (*e.g.*, a camera), transceiver (*e.g.*, a 4G LTE transceiver, a 5G NR transceiver) for transmitting data to, and receiving data from, the access point of the wireless network), a display (*e.g.*, a light emitting diode (LED) display or liquid crystal display (LCD)), and an audio input/output mechanism (*e.g.*, a microphone). The wireless-communication device may also contain means by which it connects to an external audio/video input/output (*e.g.*, wireless headphones, a Bluetooth device, an attached camera).

The wireless-communication device also includes a processor. Finally, the wireless-communication device includes a computer-readable medium (CRM) storing executable instructions of the TA. The CRM may include any suitable memory or storage device such as random-access memory (RAM). The TA, when executed by the processor of the wireless-communication device, causes the wireless-communication device to perform operations described within this document.

Human speech is composed of mechanical waves, often referred to as sound waves. These sound waves are an analog signal in the form of a sinusoidal wave. A wireless-communication device receives the analog signals by means of the audio input mechanism. The analog voice signal is translated into a digital signal by the wireless-communication device.

The transceiver in the wireless-communication device then transmits the digital signal via radio waves to an access point of a wireless network (*e.g.*, a cell tower, a base station). This access point transmits the information to a core network connected to an access point connected to the receiver's wireless-communication device. The transceiver of the recipient's wireless-communication device receives the message in the form of radio waves and translates the message to a digital signal. The digital voice signal is converted back to an analog signal by the audio output mechanism, such as a speaker, and is perceived as an audio voice message by the receiver.

The process described above is the conventional technique by which voice signals are generated and received by wireless-communication devices. During this process, the TA on a wireless-communication device can be executed by the processor upon receiving a digital voice signal. Depending on user selections, the digital voice signal, instead of or in addition to being translated to an analog signal, can be converted to digital text or in the form of a display of gestures by the TA.

Figure 3A and 3B, below, demonstrate what the TA would display on the UI of the recipient's wireless-communication device.

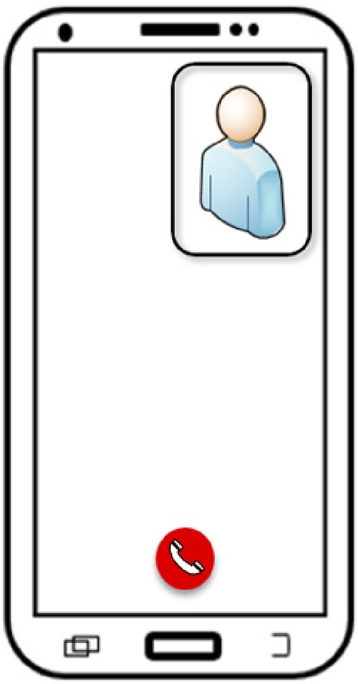

Figure 3A

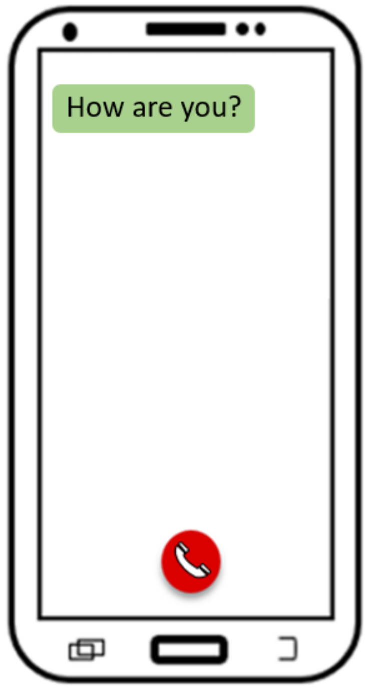

Figure 3B

If the user selects the option to visualize the voice call in gestures, then the TA can convert the digital voice signal received by the wireless-communication device's transceiver into gestures (*e.g.*, sign language). On the wireless-communication device's UI, an automated character could relay the message in the form of a display of gestures as illustrated in Figure 3A. If the user selects the option to visualize the voice call in digital text, then the TA can convert the digital voice signal received by the wireless-communication device's transceiver into text. Displayed on the wireless-communication device's UI would be digital text of the voice signal.

In response, the hearing impaired and/or speech impaired person can select to reply either by inputting a response message or displaying gestures. If the user opts to input a response message, then a keyboard will display on the wireless-communication device's UI for which the user can type the message or other input means will be provided to the user. If the user determines to respond via gesture input (*e.g.*, perform sign language), then the visual input mechanism, such as a camera, can analyze the performed gestures. Either method of communication can then be converted to an automated digital voice signal. Figure 4, below, demonstrates the two options the user can select to receive and reply by during a voice call.

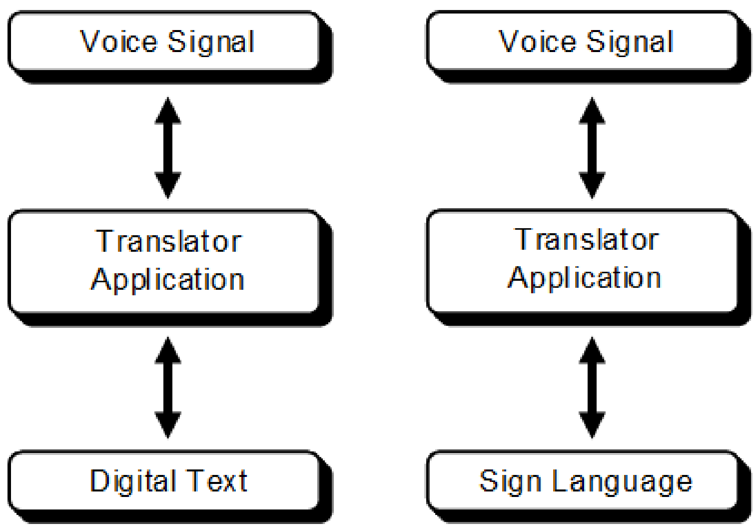

Figure 4

The wireless-communication device's transceiver can then transmit the automated digital voice signal to the connected access point, and, ultimately, the original caller will receive a response voice message.

In addition, the TA can translate between different languages. As soon as the digital voice signal is received by the wireless-communication device's transceiver, the TA executed by the processor can translate the digital voice signal to the user's preselected language. The user can respond either by speaking, typing, or displaying gestures and expect the message to be converted by the TA back to the original received language or another user-selected language.

Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs or features described herein may enable collection of user information (*e.g.*, information about a user's social network, social actions or activities, profession, a user's preferences, or a user's current location), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user, or a user's geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

In conclusion, the TA can convert incoming voice signals to either digital text or in the form of a display of gestures and generate the imagery on the UI for the recipient's observation. Moreover, a hearing impaired and/or speech impaired person can, by means of the TA, reply either by inputting a message or displaying gestures, and have the message converted to an automated digital voice signal. Additionally, the TA can translate between different languages. Such an assistant translator application is not available today.

References:

[1] Kaspar, Jakub. "Hear-o." Hear-o. Accessed May 28, 2019. https://www.iterait.com/hear-o.

[2] Batra, Kunal. "Deaftel Phone Service - Updated March 2012." Mobile Social Work. March 21, 2012. Accessed May 28, 2019. https://www.youtube.com/watch?v=S3D7drfjRtQ.