

Voice-Controlled Artificial Handspeak System

Jonathan Gatti, Carlo Fonda, Livio Tenze and Enrique Canessa

Science Dissemination Unit (SDU),
The Abdus Salam International Centre for Theoretical Physics (ICTP), Trieste, ITALY

Abstract

A man-machine interaction project is described which aims to establish an automated voice to sign language translator for communication with the deaf using integrated open technologies. The first prototype consists of a robotic hand designed with OpenSCAD and manufactured with a low-cost 3D printer —which smoothly reproduces the alphabet of the sign language controlled by voice only. The core automation comprises an Arduino UNO controller used to activate a set of servo motors that follow instructions from a Raspberry Pi mini-computer having installed the open source speech recognition engine Julius. We discuss its features, limitations and possible future developments.

Keywords

Artificial Intelligence, Man-Machine Interaction, Open Technologies, 3D Printing

1. Motivation

Science Dissemination for the disabled is a new research topic with the noble ideal of facilitating the access to science and education to a minority of scholars by trying to offer them equal opportunities. There are a few interesting recent initiatives worth to mention including the project "A Touch of the Universe" of the Astronomical Observatory of Valencia for blind students [1], and the large action named "Astro vers Tous", which promotes astronomy in different directions for deaf people in hospitals and jails [2]. According to the World Health Organization [3], a 5% of the world population has disabling hearing loss (with more than 40 million people in Europe), often preventing their proper psychological and social development. To be affected by deafness generally implies the risk of exclusion and illiteracy, thus compromising communication skills and integration.

There are some rehabilitation techniques aimed to helping individuals affected by deafness, especially young people and children. These techniques are often expensive and not accessible to all. This situation becomes more critical in the developing countries [3]. In less severe cases, such methods involve the use of expensive hearing aids, or in the most extreme cases, the sign language (SL) is taught. However, this last solution may become inefficient when a person suffering from this handicap needs, for example, to communicate to a person without the knowledge of the SL, or need to travel abroad. There is not a universal SL yet [4]. There are variants such as the popular American (A)SL that depends on its society, culture, idiom, *etc.* Therefore, providing the opportunity to have an alternative, automated and configurable system capable of receiving information from anyone and anywhere would allow deaf people to explore new spaces of inclusion and participation.

Motivated by these needs and facts, we investigated the possibility for having an automated translator for communication with the deaf to help reduce language barriers. We have developed a

man-machine interaction project which aims to establish an automated voice to sign language translator for communication with the deaf using integrated open technologies. Our prototype consists of a robotic hand —designed with OpenSCAD [5] and manufactured with a low-cost 3D printer [6], which smoothly reproduces the alphabet of the sign language controlled by voice only. The core automation comprises an Arduino UNO micro-controller [7] used to activate a set of servo motors that follow instructions from a Raspberry Pi mini-computer [8] having installed the open source speech recognition engine Julius [9]. In this paper, we discuss its features, limitations and development as a mean to support deaf individuals. The hope is to create an innovative channel of communication based on the integration of different open and affordable technologies.

2. Automated Prototype

Our *Voice to SL Translator* consists of a robotic hand aiming to satisfy an important fundamental human need such as face to face communication. As shown in Fig. 1, the robotic hand has been designed using the OpenSCAD software [5] and is made out of PLA bioplastic by means of a low-cost 3D printer [6]. This 3D hand is printed with the FDM technology and the fingers are printed in one shot as single objects, without the need of any screws to ensemble its parts.



Figure 1: Modelling process: OpenSCAD design and 3D printout. Example of modular finger is on the right.

In this initial study the rotational movements needed to mimic the SL have been obtained using six servo motors (torque 9 kg*cm) for a precise control of angular motion and velocity of the fingers, as well as of the whole hand. The moving parts were linked to the servo motors by tiny nylon fish-lines, hidden along the arm. Instructions were sent to the servo motors using an internal Arduino UNO controller board and drivers [7]. Thanks to the open source Julius voice recognition framework [9], installed on a Raspberry Pi lightweight desktop computer [8], we were able to convert the voice input (received via a USB microphone) into text. And precise instructions were given (via the second USB port of Raspberry Pi) to the Arduino controlling the six servo motors, to position them at specific angles so that the fingers could move in correspondence to the single number or letter being pronounced (see our YouTube video example in [10]). The integrated prototype we envisioned is schematically shown in Fig. 2.

In this way we translated the human voice into actual signs using a 3D printed hand integrated with an electronic board able to command its fingers by voice inputs. In fact, thanks to the open

source software Julius (easily implemented into a Raspberry Pi [11]), we are able to translate reasonably well the human voice into text to a high precision for isolated letters and numbers (more than 90% of success). Due to its small size and low power requirements, the Raspberry Pi platform is adequate for the Julius open source software to execute. However, its small processor in some cases can slow down the speech recognition. Julius is a Large-Vocabulary Speech Recognition engine (LVCSR) which needs a language (vocabulary) and an acoustic model (grammars), which have to be loaded manually into the system.

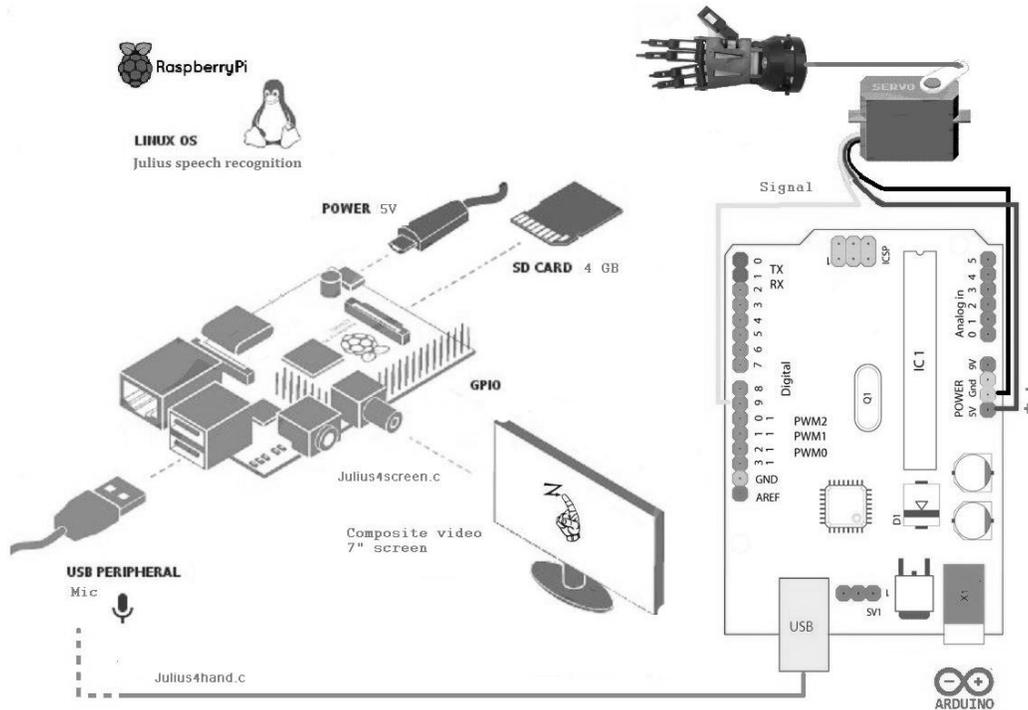


Figure 2: Diagram of the handspeak system based on open technologies (see video example in [10]).

An Arduino UNO controller is set to command six servo motors following instructions received by an USB connected Raspberry Pi mini-computer with the speech recognition decoder Julius. Peripherals also include a USB microphone and a RCA composite video.

Since Julius is not able to translate voice directly into a visual language, we wrote two C-language programs: *julius4hand.c* to drive the servo motors attached to the Arduino UNO controller and *julius4screen.c* to display a still image on a small video screen.

The Julius application provides speech recognition results via the standard output (STDO). The *julius4hand.c* code reads from the standard input (STDIN) and reacts to the recognized letters and numbers by sending the proper command to the Arduino ONE board. To achieve this, we created a generic transmission protocol to interact with Arduino. The connection between the Raspberry PI and the Arduino board consisted in a serial channel which can be easily driven by using the standard Linux system call. *Julius4hand.c* reads an external configuration file where the mapping between the recognized letters and numbers and their associated positions of servo motors are stored.

On the other hand, our *julius4screen.c* code waits on the STDIN for a recognized character, sent via the STDO by Julius, and calls the external tool *fbi* (which allows to display images directly on

the Linux frame buffer) to show up the correspondent image in SL. In this case the mapping configuration file contains the recognized characters and their associated images (.jpg) to be shown.

Finally, we developed the C code to manage the six attached servo motors, connected to the Arduino UNO board by exploiting the Servo.h library [7], provided by the standard Arduino framework. The code reads inputs from the USB serial connection and sets the required servos angles to move the hand to the required position. Servo motors are driven by using a PWM (Pulse Width Modulation) protocol.

3. Prototype Features and Development

To some extent, our simple, low-cost system is capable of reproducing most of the SL alphabet following the voice of an interlocutor as shown in Fig. 3 or see our video in [10]. It does not need an internet connection, neither the use of smartphones or fast computers, which makes it completely independent and easy to use especially in rural areas where connectivity is poor or absent. The plastic robotic hand in Figs. 1 and 3 is easily replicated with low-cost 3D printers. The whole integrated system consumes very little energy.

Our open device is different from alternative approaches designed for those people with hearing problems. In fact, among the available applications, there are those which can translate voice into text like Telecommunications Device for the Deaf (TDD). Alternative smartphone Apps include, *e.g.*, the *finger* App, which can play SL and can recreate gestures of one hand by typed text. Important to notice is that smartphones Apps can make massive use of the battery for long talks. A non-hearing person may also communicate by cell phones via SMS texting. There are also special systems, called Auditory Science Interact (AS), that transmit written text in streaming to the interlocutor (for a recent review see, *e.g.*, [12]). However, so far as we know, there does not exist an application that can translate the voice directly into SL as done with our first prototype. Vocally controlled robot arms, *e.g.*, the earlier VoiceBot, have been proposed to follow voice commands limited in their movements and actions, such as opening and closing fingers, holding objects or moving them from an initial position to a target [11,13-17]. Such robots do not have any specific purpose for the deaf people.

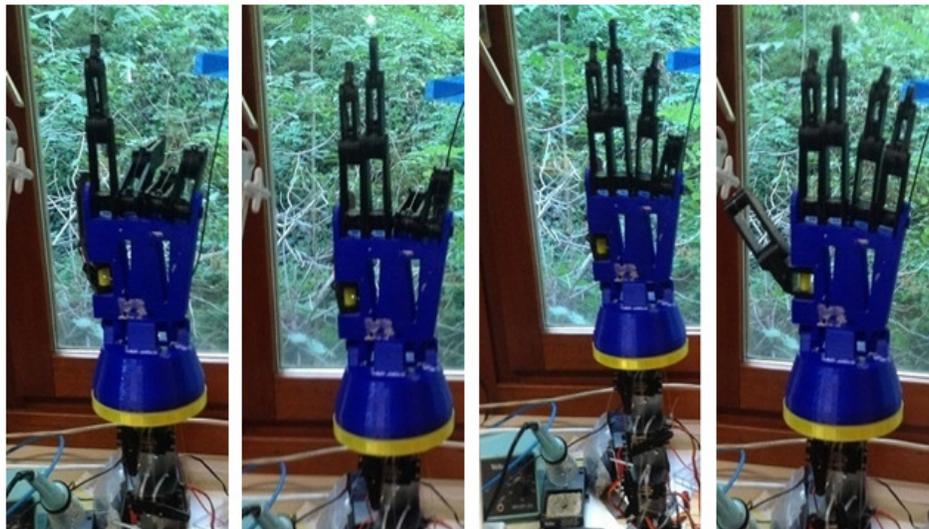


Figure 3: SL sequence for numbers 1,2,3 and 5 following the voice of an interlocutor (see video in [10]).

Robotic devices in general can also be useful for teaching purposes. By constructing a prototype,

many non-experts could use it to communicate with deaf people. Our device, in principle, could be able to reproduce any spoken language by properly configuring Julius. It could help to illustrate the alphabet and all the movements which are necessary to learn the SL, especially to children in schools, hospitals and within specialized associations. Our prototype may offer an interactive and fun way to learn SL by cognitive means, especially in the developing world where there are not enough facilities dedicated to the disabled.

The rapid expansion and the exponential improvement of the quality of manufacturing guarantee the possibility to offer customized systems on-demand in line with the modern "*Long Tail*" distribution model of, *e.g.*, Amazon.com. The plastic provides long durability and, if a piece of the printed hand eventually breaks, it can be easily reproduced by a low-cost 3D printer [6]. The authors intend the designs and codes of this research to be the first in a series of automated voice-controlled handspeak prototypes, released on the public domain [10].

4. Limitations

For single letters and numbers, the speed reaction of the fingers to the input voice can be increased with future improvements in hardware and software, and the servo motor activation sequences can be better implemented by a fine tuning configuration of the rotation angles (from 0 to 180 degree). The reproduction of more complex signs, such as the letters **Z**, **J**, requires larger and complex rotational and translational movements. The use of a USB directional microphone with an on/off switch can help to reduce any external noise. One major limitation of the current prototype is its inability to reproduce a speech (to express an ongoing action or describe a specific meaning). Still, a set of hand signs can spell out a word. The problem of the voice-controlled representation of concepts by a 3D robotic hand can, in principle, be solved by using an alternative PC/TV screen as also represented in Fig. 2 to display still images or even short videos associated to the speaker input.

We believe that with further research into voice-controlled robotic hands that mimics and reproduce the sign language (let it be 3D printed or, even displayed by sequences of images or videos on a small screen), a more advance communication tool could be created to allow deaf individuals to listen to anyone, anywhere, and thus facilitate their integration into society. The community working on open technologies is making fast progress to improve the performance, efficiency and durability of the compact Arduino UNO [7] and Raspberry PI [8] hardware.

Our first prototype device, even if incomplete and still in its infancy, sets for itself the future goal of following the voice of any speaker/lecturer and translating its speech into a particular SL alphabet configured for a specific country. Looking forward, we hope that one day an automated voice-controlled handspeak system will help to improve the quality of life and demolish the walls which separate the main population to those who have hearing problems and are isolated from access to, *e.g.*, scientific knowledge.

Conclusions

In this work we have introduced, and evaluated the feasibility of a voice to SL translator for letters and numbers with affordable open technologies (*i.e.*, 3D printing with OpenSCAD modelling, Arduino ONE controller and Raspberry Pi mini-computer). As far as we know, our results are the first instance of a voice-controlled robotic hand specially built with the aim to help deaf individuals. Our initial results demonstrate that such a possibility seems to be reasonable and feasible. To translate automatically a spoken language into a specific SL is certainly the most challenging target since SL is mainly a visual language also incorporating complex gestures, facial expressions, head movements and body language -work along these lines is in progress to

convey a lot of information, *e.g.*, within a few video frames.

Acknowledgements

Sincere thanks go to B. Nava and M. Zennaro (ICTP-ICT/4D) for their important contributions during the development of this project. Prof. F. Quevedo (ICTP Director) is also acknowledged for continuous support. One of us, Jonathan Gatti, participated in this work during his Internship at ICTP-SDU, in the Summer of 2013, assisted by Gaia Fior, collaborator of ICTP-SDU.

References

- [1] UNAWA non-profit project "A Touch of the Universe" at <http://astrokit.uv.es>
- [2] D. Proust in "Astronomy in Sign Language" at <http://www.youtube.com/watch?v=ksFhan9IKKU>
- [3] World Health Organization (WHO), 2013 Deafness and Hearing Loss Statistics, at <http://www.who.int/mediacentre/factsheets/fs300/en/>
- [4] Wikipedia on "Sign Languages relation with Spoken Languages" at http://en.wikipedia.org/wiki/Sign_language
- [5] OpenSCAD – The Programmers Solid 3D CAD Modeller at <http://www.openscad.org>
- [6] E. Canessa, C. Fonda and M. Zennaro Eds., open book in "Low-cost 3D Printing for Science, Education and Sustainable Development", Trieste, ICTP 2013, ISBN 9295003489
- [7] The Arduino Uno microcontroller board a <http://www.arduino.cc> Servo.h library available at <http://arduino.cc/en/reference/servo>
- [8] Raspberry Pi credit-card sized computer at <http://www.raspberrypi.org>
- [9] Julius speech recognition decoder developed by CSRC, Japan at <http://julius.sourceforge.jp>
- [10] ICTP-SDU voice2hand project at <http://sdu.ictp.it/3D/voice2hand.html> @sciencedissemination
YouTube video "Voice to Sign Language System" at <http://www.youtube.com/watch?v=J5whsEsGr4s>
- [11] YouTube video: "Speech Recognition using the Raspberry Pi" at http://www.youtube.com/watch?v=_U_v9tRD68k
- [12] Top 10 iPhone Apps for People Who Are Deaf or Hard of Hearing at <http://atcoalition.org/news/top-10-iphone-apps-people-who-are-deaf-or-hard-hearing>
- [13] Free, open-source 3D printing robotic hands are available at <http://www.thingiverse.com>
- [14] H. Dang and P. Allen, "Stable Grasping under Pose Uncertainty using Tactile Feedback", Autonomous Robots, August 2013.
- [15] R.R. Ma, L.U. Odhner and A.M. Dollar, "A Modular, Open-Source 3D Printed Underactuated Hand", Proceedings IEEE International Conference on Robotics and Automation (ICRA) Karlsruhe, Germany, May 6-10, 2013, pp. 2722-2728.
- [16] B. Leon, A. Morales and J. Sancho-Bru, "The Model of the Human Hand", Cognitive System Monographs 19 (2014) 123-173.
- [17] P.S. Ramaiah , M. Venkateswara Rao , G.V. Satyanarayana, "A Microcontroller Based Four Fingered Robotic Hand", Int. J. Art. Int. & App. (IJAIA) 2 (2011) 90-102.