

Computing for deaf people

The race to teach sign language to computers

Paying attention to what deaf users actually want is vital



Mar 4th 2021

Using a computer used to mean bashing away at a keyboard. Then it meant tapping on a touchscreen. Increasingly, it means simply speaking. Over 100m devices powered by Alexa, Amazon's voice assistant, rest on the world's shelves. Apple's offering, Siri, processes 25bn requests a month. By 2025 the market for such technology could be worth more than \$27bn.

One group, though, has been left behind. The World Health Organisation counts 430m people as deaf or hard of hearing. Many use sign languages to communicate. If they cannot also use those languages to talk to computers, they risk being excluded from the digitisation that is taking over everyday life.

Many have tried to teach computers to understand sign language. There have been plenty of claims of breakthroughs in recent years, accompanied by so-called solutions ranging from haptic gloves that capture the wearer's finger movements to software that detects distinct hand shapes. Many of these have won acclaim while alienating the very people for whom they are ostensibly designed. "The value for us basically is zero," says Mark Wheatley, the executive director of the European Union of the Deaf (eud).

It is easy to see why. Gloves are intrusive, as are similar technological solutions such as body-worn cameras. Both require users to adapt to the needs of hearing people. Hand-shape recognition, while useful, cannot by itself handle the full complexity of sign languages, which also rely on facial expressions and body movements. Some projects have been touted as offering cheap alternatives to human interpreters in places like hospitals, police stations or classrooms, where the cost of even small errors can be very high.

But things are improving. Research groups, which increasingly include deaf scientists, are asking how technology can best serve deaf people's interests. Students of sign languages are compiling databases, known as corpora, full of examples of how the languages are used. Programmers are trying to turn them into useful products.

As with spoken languages, sign languages—of which the world has several hundred—possess their own grammars, idioms and dialects. Again like spoken languages, the hard-and-fast rules of grammar books do not really capture the subtleties of everyday usage. Single signs can be shorthand for complex ideas. Like speakers, signers often take shortcuts, such as representing two-handed signs with a single hand. They set up reference points within their signing space which can be vital for meaning. Correctly interpreting all this is much harder than recognising spoken syllables or written letters.

Generating data is tricky, too. Research led by a team at Microsoft, a big computing firm, and published in 2019, estimated that a typical publicly available corpus of a spoken language consists of around a billion words from as many as 1,000 different speakers. An equivalent data-set in a sign language might have fewer than 100,000 signs from just ten people. Besides large numbers, a good corpus also needs variety. This means conversations between native signers of diverse backgrounds, dialects and levels of fluency. Because deaf people more often have physical disabilities than do those with unaffected hearing, representing those with restricted fluency of movement is important.

Thomas Hanke, a researcher at the University of Hamburg, has, along with his colleagues, assembled a sign-language library containing about 560 hours of conversations, and which includes many dialects found in Germany. Originally, Dr Hanke asked participants in the project to travel to Hamburg. But while in the city, many volunteers started incorporating local signs into their communications. That skewed the data. Now, he says, he travels to his participants instead, and has been criss-crossing the country in a mobile studio for the best part of two years.

Collecting data, though, is the easy bit. Computers are slow learners, and must be told explicitly what each example means. That requires annotating everything—every movement, facial expression and subtlety of emphasis. This takes time, and lots of it. After eight years, Dr Hanke has only 50 hours of video which he is confident are annotated correctly.

Microsoft's researchers are using crowdsourcing to improve the amount and quality of data available. Danielle Bragg and her colleagues at the firm's campus in Massachusetts are developing a smartphone version of "Battleship", a game in which each player tries to sink their opponent's ships by indicating locations on a grid. In Dr Bragg's version, each grid square is associated with specific signs. Players not only generate signing data of their own, but also confirm the meaning of signs made by their opponents.

Privacy is a particular concern, since collecting sign-language data requires recording participants' faces rather than just their voices. When Dr Hanke tried to record people's gestures anonymously, their idiosyncratic signing techniques were so distinctive they could still be identified. Dr Bragg plans to use facial filters, or to replace faces with artificially generated alternatives. That will interfere with the quality of the data, but she hopes that lower quality will be made up for by greater quantity.

If enough data can be gathered, researchers with a good understanding of deaf culture and machine learning can achieve impressive results. The 25-person team at SignAll, a Hungarian firm, includes three deaf people and claims to be one of the biggest in the field. The firm's proprietary database contains 300,000 annotated videos of 100 users using over 3,000 signs from American Sign Language (asl), one of the most widespread. It was collected with help from Gallaudet University in Washington, dc, the only university which caters specifically for deaf students.

SignAll's software can recognise asl, though not yet at the speeds at which native signers communicate. Its current product, SignAll 1.0, can translate signs into written English, allowing a hearing interlocutor to respond with the help of speech-to-text software. But it relies on pointing three cameras at a signer wearing special motion-tracking gloves—a significant burden.

That may soon change. Zsolt Robotka, SignAll's boss, says the firm hopes to offer a glove-free option. It is also putting the finishing touches to a product that works with a single camera on a smartphone. If that technology can be integrated into other apps, it could allow deaf people to use their phones to do things like searching for directions, or looking up the meanings of unknown signs, without needing to resort to the written form of a spoken language.

Crossing the valley

For the moment, Dr Robotka's emphasis is on translating sign language into text or speech. Translating in the other direction poses greater difficulties, one being how to generate visual representations of sign language. The standard approach has been to use computer-generated avatars. But many fall into the "uncanny valley", a concept from computer graphics in which artificial humans fall just short enough of verisimilitude that they instead look eerie and disturbing.

Bridging the valley would permit widespread two-way communication. Creating smartphone apps that can recognise a range of European sign languages, and translate back and forth between these and oral speech, is one aim of two new multinational academic consortia: the Signon project, and the Intelligent Automatic Sign Language Translation project, also known as easier. Both are working with the eud, which represents 31 national associations across the continent.

Signon is targeting British, Dutch, Flemish, Irish and Spanish sign languages and, with the exception of Flemish, their hearing equivalents. Working with several European universities, it aims to solve three problems. One is to improve the machine-learning algorithms that recognise signs and their meaning. Another is to work out how best to interpret sign languages' distinctive grammars. Finally, it will try to create better avatars. easier, of which Dr Hanke's team at Hamburg is one of 14 partners, has similar goals: namely sign language recognition, robust two-way translation and avatar development.

Money and attention are always welcome. But previous attempts to automate the translation of sign language have too often been directed at making life convenient

for those with normal hearing rather than truly trying to help the deaf. This time, observers hope that a more sensitive approach will yield more useful products. “It’s a wonderful opportunity for us,” says Mr Wheatley of the eud. “We’ve got no time for cynicism.”

Source:

<https://www.economist.com/science-and-technology/2021/03/04/the-race-to-teach-sign-language-to-computers> (Accessed 13 September 2021)