

Re: Sign Language Interpreter Project

Chris Haulmark <chris@sigd.net>

Wed 9/21/2022 5:34 PM

To: Adubi, Bolu <bolu.adubi@ku.edu>

Hello!

In order to find out what you can do for the Deaf community, I appreciate that you reached out to an authentic representative. As long as the contributions do not diminish our voices, anyone who can bridge the communication gap between the Deaf community and those outside the community will always be appreciated by the community.

This email is intended to open a door for you and anyone else on your team to peek into our community, so that we might better understand how we might work together on a project like yours.

Despite the fact that communication barriers have existed for a few millennia among the Deaf community, there have been many ideas like yours to overcome them over the last few decades with technological advancements. Among these are cameras that are attempting to understand sign language and gloves that are supposed to interpret sign language into written or spoken language.

The concept of interpreting sign language into speech is gimmicky at best, and is not capable of actually interpreting sign language into speech. I will explain why these ideas do not work and why they will probably continue to fail for some time to come.

If a camera were focused only on the motions and shapes of the hands of an individual using sign language, it would be mediocre at best at interpreting sign language into accurate speech.

Although the points are applicable to all signed and spoken languages, I will use American Sign Language for the sign language example and English for the spoken language example in this email. American Sign Language is represented in English using words in all capitals called Gloss, which you call when you write one language in another. Gloss is what you call one language written in another, and it is used to represent American Sign Language in English.

The camera is to capture and interpret the movement of the hand joints and optionally velocity changes to create vector-like patterns that are then matched with a database of hand shapes and movement patterns to find the corresponding English equivalent. As a result, a gesture corresponds directly to a spoken word/phrase. If one signs I WILL GO HOME then the computer interpret this signing to "I will go home" and if one signs using proper American Sign Language grammar, WILL GO HOME I then the computer will say "Will go home I". There will be a need for this later on.

It is possible, from my perspective as a computer engineer, that an AI system, such as an expert system or neural network, could create a one-to-many relationship using fuzzy logic or contextual information, but I have not seen a device such as this demonstrate this. For the purposes of this email, I will assume these do not exist to any significant degree.

What is a sign within the American Sign Language?

The hand shape, position, movement, [non-manual markers \(NMM\)](#), and context are the five properties which determine what a sign means in American Sign Language. To enhance or change the meaning of signs, non-manual markers are actions and movements made with something other than the hands. Consequently, a hand shape and movement made on one's forehead would mean a different thing than the same hand shape and movement made on one's chin or chest. In [FATHER](#), [MOTHER](#), and [FINE](#), the same hand shape is used with different positions on the signer's body. However, the same hand shape done in a certain movement with or without an NMM would mean something different which are the signs of [NOT-YET](#) and [LATE](#).

Next, let's consider how questions are asked in American Sign Language.

To distinguish between sentences and questions in spoken language, we use inflection. For example, if you say "You're hungry." and "You're hungry?" you might notice that the inflection at the end of "hungry" changes even though the words remain the same. NMMs, specifically eyebrow movement, are used to create these "inflections" in American Sign Language. The signs for "How old are you?" and "You're old." are identical because you would sign in American Sign Language "OLD YOU", but their meanings are obviously quite different because of the differing NMMs.

You can already tell the camera isn't going to be capturing true signs. Only three of the five properties are being captured by the camera, which means most of the information is being discarded. These cameras would not be able to distinguish between the examples given above, and so we see already a huge limitation to the devices. But let's move on.

In American Sign Language, what isn't a sign?

A classifier is a gesture that resembles a sign, but does not possess any of the properties of a true sign, and is used in a pantomime-like manner to convey its meaning via common understanding. As an example, if I were to extend my hand toward the table in the form of a C hand shape, and pantomime raising something to my lips and drinking, one might reasonably conclude that I was indicating drinking from a glass. As I would begin the same motion, but invert my hand and allow my gaze to fall to the floor as I did so, one might reasonably conclude that I was pouring something from a glass. A computer is unable to understand these signals because they are not true signs because with these examples, the classifier lacked a defined movement and position. Also it is because the signals are not strictly definable in a pattern matching algorithm. The only reason that these two examples would be meaningful to humans is because we have a general understanding of what a glass is and how it is used, as well as the ability to visualize a glass in my hand as I make the gestures.

Any signed conversation can be dominated by classifiers. A sign language description of a haircut, for example, would require several classifiers, non-manual markers, and pantomimes that these devices would miss, as well as contextual understanding that even a reasonably complex neural network could not provide.

YESTERDAY STORE I GO BUY-BUY MILK MANY COOKIE CANDY

A common misconception is that signed languages are manual versions of spoken languages. The American Sign Language is not English. It is not just the vocabularies that differ, but the grammar as well. In English, the title of the section would read: "I bought milk, cookies, and candies yesterday at the store." There are many similarities, but there are also many differences. In the sense that words are arranged in a specific order to form sentences for English but sign languages are not verbal

languages. They're visual languages, more like interpreting a painting hung on a wall for someone who doesn't see it than reading a paragraph out of a book. Any direct interpretation device would completely miss out on these significant differences, since the structure of the language itself allows meaning to be expressed in ways that are impossible in spoken languages.

Putting it simply, there is no technology that can interpret a sign language into speech. The idea that it will exist within our lifetimes seems almost inconceivable. A camera would never be able to capture enough information to make an accurate interpretation even if it did. In spite of the ability to capture finger, hand, arm, shoulder motion, body shifts, facial expressions, and all other NMMs, the device would still not be able to interpret sign language due to the fact that a device would need to be able to visualize, empathize, and extract information from common understanding the same way humans do. Unless the AI singularity is achieved, a computer will not be able to interpret signed and spoken languages fully and meaningfully. These current or similar devices are only capable of interpreting, at best, a very small portion of sign language and in very limited circumstances. Communication in any signed language is incomplete without expression and emotion. There would be a loss of body shifting. There would be a loss of indirect noun references in a most likely manner. An actual signed conversation would be impossible to understand due to the loss of too much information.

If such ideas are successful, there may be acceptable results for non-Deaf people, but a false sense of understanding could be conveyed. The non-Deaf parents of Deaf children should learn and use sign language in their homes. We don't want them to think that using those innovative devices to communicate with their children is a shortcut. People who are Deaf deserve to have their "voices" understood in full context with humanity through their emotions, body language, and intellect.

Even though many Deaf people have been deprived of their languages, they are mostly monolingual and speak only American Sign Language. I hope your project will be shifted towards being more supportive of these monolingual Deaf people by interpreting voice and text into sign language so that videos containing and not containing captioning can be understood by Deaf people. This is possible because the five properties of the American Sign Language can be displayed visually based on the context of these spoken or written languages.

I hope this email has been somewhat helpful for you and your team.

All my best,

Chris Haulmark

On Wed, Sep 21, 2022 at 1:55 PM Adubi, Bolu <bolu.adubi@ku.edu> wrote:

Apologies, I accidentally hit enter before completing the email.

My name is Bolu Adubi and I'm taking a group-focused computer science class where we work on a project for a year. My group members and I decided to work on software that could interpret sign language. We were wondering if you'd be available to meet sometime in the next two weeks (in person or over Zoom) to address some questions we have. None of us on the team know a lot about sign language but we're very interested in learning more about its structure and rules so we can design a model that can be used to interpret it.

Thank you,
Bolu Adubi and EECS 581 Team 1

From: Adubi, Bolu

Sent: Wednesday, September 21, 2022 1:53 PM

To: chris@sigd.net <chris@sigd.net>

Subject: Sign Language Interpreter Project

Good afternoon Mr. Haulmark,

I got your email from a previous professor of mine (Kimberly Anderson). She suggested I reach out to you concerning a project I'm working on.