Sign Out Loud

Isaac Reath · Gino Notto · Jason Gregory · Eric Ernst · Todd Lock

Table Of Contents

Problem Statement

What we need in the modern world is a way for us to facilitate communication between people who do not know sign language and people who can only use sign language as a way to communicate. We want to bring a voice to those people who are unable to have one enabling, for the first time ever, real-time communication between those who cannot speak and those who do not know sign language; making the world a better, more interconnected and open place for the deaf and mute. This could be easily solved using a real-time computer vision based sign language translation tool. The system should use some sort of video camera, be it a Microsoft Kinect or a simple phone or webcam, and a computer to do a translation. The system should watch for different sign language forms and build a real-time translation of the words that the user is signing. Our target audience will be the mute and deaf, as they are the people who would get the most individual use out of this product. However, as a result of the focus on the deaf and mute, hospitals and assisted living centers would also be a key audience for this product. This camera-based translation tool is a low-cost, easy to use, and a no-brainer for a system in the modern world. With

basic optimizations, we could make this accessible to nearly all deaf and mute for a small price.

Current solutions in Sign Language Recognition

There are several solutions out there currently that solve the problem of Sign Language Recognition, and looking at these is a great way to see what other people have done right, what other people have done wrong, and where there is a need for new features. We will look specifically at 3 different solutions to this problem to see what they have to offer.

1: Kinect Sign Language Translator

http://research.microsoft.com/en-us/collaboration/stories/kinect-sign-language-translator.aspx

This is a solution to the same problem as ours, Kinect based sign-language recognition. This was developed by researchers in China, developed by several schools including the Chinese Academy of Sciences, Beijing Union University, and Microsoft Research Asia.

Something most don't know about sign language is that there are different languages. One of the useful features of this translator is its ability to recognize Chinese and American Sign Language, and be able to convert back and forth between the two and others.

This system also has an extensive animation system to show virtual humans performing the signs to help people learn. In communication mode, the avatar signs the

words of the speaking person, and the software translates the signing persons word's into voice. This a great way to help deaf people communicate and gave us good ideas for how our system could work.

2: Dicta-Sign

http://www.dictasign.eu/

Dicta-sign is a very interesting European project developed to help deaf people use the web. Many websites are hard for deaf people to use because they rely very heavily on written language, something deaf people can struggle with. This seems counter-intuitive since they could still read everything, but deaf people often find written language hard to understand without being able to phonetically assign meaning to letters. Dicta-sign provides a number of functions that are related to our goal but significantly different.

One thing they created was a search tool that uses signs in order to search. They also created a tool which is able to display a corresponding sign in several sign languages, which is very useful for translation. They have also developed a sign wiki was created to help deaf people learn about sign language and differences between languages.

These goals aren't identical to ours, but have strong similarities, such as the need to convert to convert signs to words, and designing applications that are intuitive for both deaf people to use.

3: Automatic Sign Language Detection

https://sites.google.com/site/autosignlan/home

This was a solution created by people in a similar situation to ours, it was developed by students taking a class. This was a higher level class on pattern recognition, their website shows how they developed their system in detail, and gives us a good look at specific algorithms they have chosen, and how they implemented it. This look at the technical side of their implementation is what makes this a valuable resource.

Goals and Objectives

Technology has grown quite miraculously over the past few decades and at a ridiculously fast pace. Slowly and surely we are being able to have the capability of producing a technological product or application that can pretty much do anything. Smartphones, smart watches, tablets, the list goes on; these products are some of what the average person uses in their everyday lives.

However, technology doesn't just come out and be what people expect it to be. Every year, we go through the process of receiving a brand new application or product in our lives that tends to receive feedback at how poorly designed their product turned out to be. Feedback is helpful though because it gives the company another chance to improve their product and satisfy their customers. Customers do enjoy buying from a company who listens to the feedback they receive.

This brings me to a major point in the technology advancement that involves products being evolved overtime. This is for the satisfaction of their customers, whether it's

being improved by the same company who produced it, or another company that wishes to improve on an already created idea of a project. This has been the case throughout the whole technology advancement. We keep branching off of already thought out ideas and making them better than before.

One specific problem that has been currently worked on over the past 30 years and still needs improving is giving a voice to the deaf and mute. Currently, the system that is able to attempt giving a voice for them makes it difficult for anyone who does not know sign language. A company should adjust this system so it allows for signs to be translated directly to text and then from text to spoken communication.

The high-end goal/objective should be to use a camera-based system to allow fast detections of hand gestures made in sign language to generate spoken communication with the help of an application. Preferably, a kinect should be used because it is one of the most well-known devices to detect human movements. With the help of this application, we will be able to improve communication between those who cannot speak and those who cannot use sign language. This could also open up a new path if we can make the voice commands from the application interact with programs.

The quality of life of those who are deaf and mute will be improved because they will be able to have a real-time communication with those who do not know sign language. They will also have an improved voice than ever before. Communication for the majority of people is not difficult. It should be the same way for the deaf and mute.

Functional Requirements

Preface

The purpose of the system is to facilitate communication between those who are deaf or hearing impaired and those who aren't but cannot sign. That being said, there are several features in place to allow this. The system needs to allow a deaf person to start an internet video/audio chat with someone as well as be able to understand sign language. It must

also be able to convert the sign language into text to send to the other person. It will also have several other key features that will be outlined below.

The system will be broken down into several modules. Each module will then be broken down into sub-requirements. The modules are as followed:

1. User Interface
    a. Kinect / Webcam
2. Video Processing module
3. Text-to-Speech module
4. Speech-to-Text module
5. Server/Client interface

User Interface

1. Kinect Interface - the kinect interface will connect the user with the system. The user will be able to sign in front of the kinect. The kinect will then send the input to the video processing module
2. Webcam Interface - the webcam interface will allow the user to send both video and audio to the server interface to be sent to the client.
3. Text Interface - there will also be a text interface that will allow users to type to each other. This will be implemented at both ends of the chat.

Video Processing

1. The video processing module will take input from the kinect and output it to either the server module or the text-to-speech module.
2. It must take tracking information from the kinect and covert the hand gestures to text.

Text-to-speech

1. The text-to-speech module will allow vocalization of signs to speech.
2. It will have localized voices for regions.

Speech-to-text

1. The speech to text interface will read audio input from the user and convert it to text.

Server/Client

1. The server/client module will set up a connection between two users.
2. The module will handle any networking and facilitate video,audio, and text communication between two users.

User Profile/Deployment Enviornment

There is quite the demand for our product. With this machine it is beneficial to everyone that is deaf or even knows a deaf person from family member to therapist. There are so many different users that will benefit from this. It's even possible to respond to your online gaming friends and online gaming community.

Going to the doctor can be quite a task for a deaf person. Any misunderstanding or misinterpreted data can lead to serious frustration, anger, and even cause the wrong diagnosis. With our technology the one of our customers can talk to their physician with

ease. If your physician doesn't under stand sign language all they have to do is read the screen that is attached to the kinect and it will spell out for them clearly what you are saying. This also works if you are a doctor or therapist or whatever your occupation may be. It will make it easier for your hearing disabled customers to interact with you without having to have an interpreter there. This opens up your customer market to be able to help disabled customers who would normally look for a person of your profession that is equipped with their needs.

The kinect was originally made for gaming and physical interactive games. Well with our technology it doesn't just have to be for that. One big part in games is to be able to hear what is going on especially when you are playing online with your friends, or even with the online community. Our product can take the incoming messages and display them across the screen so you know what your online companions are saying. What was a difficult task before is now easy so you can tell your team when they are in trouble, or if you are in need of assistance.

Not everyone in our customer's family and friends know how to sign or understand sign language. Our product will open up this communication and bring them together. Every family member with a kinect can download our software so that when family comes to visit everyone can communicate. With our software no one will feel left out at family gatherings and even if our intended user is younger will be able to communicate with everyone, and vice versa.

Just about everyone can benefit from our software, even if they are not the ones that are hearing impaired. From opening your business up to people that would be weary of

going because of communication problems to getting closer with your family, everyone can use our software. No one with a hearing disability has to go on thinking that they can't have a full conversation with anyone, except for technically illiterate, but that's another story.