

Sistema de reconocimiento automático de lenguaje señas colombiano mediante Kinect y Leap Motion

Manuel Alejandro López Barrera

Diego Andrés López Albán

Estudiantes del Programa de Ingeniería Mecatrónica

Universidad Mariana

Dagoberto Mayorca Torres

Asesor y Docente del Programa de Ingeniería Mecatrónica

Universidad Mariana

Resumen

Como parte del desarrollo del proyecto de grado, en este documento se presentan los avances de la implementación de una herramienta digital, como apoyo a la Inclusión social de personas con limitaciones fonético–auditivas y que emplean el lenguaje de señas colombiano (LSC) como su medio de comunicación; dicho aplicativo permite la traducción de este lenguaje a texto escrito mediante dispositivos ópticos de captura, como lo es Leap Motion, además de emplear la Plataforma 3D de Unity y estar codificado en C# en Visual Studio, cuyos avances son realizados de acuerdo con la metodología establecida.

Introducción

En la actualidad, los aspectos de inclusión social han recibido una mayor atención por parte de la sociedad, para esto la tecnología puede aportar, por ejemplo, en desarrollos como un traductor de señas.

El proyecto plantea la realización de un traductor de señas con un glosario limitado de palabras, tanto estáticas como dinámicas, donde se vean involucradas manos y brazos; sin embargo, como resultado preliminar se plantea un modelo de reconocimiento del alfabeto estático, el cual emplea únicamente manos, referente al lenguaje de señas colombiano a partir de la adquisición de datos con Leap Motion; el trabajo con dispositivo Kinect es dejado para

desarrollos posteriores que determinen las trayectorias de brazos. Por lo tanto, se realizará un desarrollo de software implementado con este dispositivo, donde se adquieren los datos capturados y se los clasifica empleando el algoritmo de Machine Learning, ya que es el que más se ajusta a nuestro propósito.

Fundamentos para el desarrollo del proyecto

La comunicación no verbal la emplean personas con limitaciones auditiva para poder comunicarse con su entorno, hecho por el cual existe el lenguaje de señas como forma de comunicación usado por población sordomuda; sin embargo, es difícil la comunicación con las personas hablantes, ya que la gran mayoría de la población no conoce y no logra entender este lenguaje, por lo tanto, es necesario contar con la ayuda de un intérprete, y en consecuencia se tienen muchos obstáculos para llevar una vida en plena conformidad, por la falta inclusión de esta comunidad, tanto social, educativa, deportiva y laboral.

En el Censo 2005 se reportaron 455.718 personas con una Limitación para oír sobre un total de 41'468.384 lo cual representa una tasa del 1,1% a nivel nacional. Bajo el supuesto que esa tasa se mantuviera en el tiempo, para el presente año habría alrededor de 543.614 personas con esa condición. (Tróchez, 2018, p. 2)

Como se nota, las personas con limitaciones auditivas van a seguir incrementando en Colombia, si bien, el gobierno nacional está haciendo los esfuerzos para la inclusión de personas sordomudas a la sociedad, todavía siguen existiendo problemas para la incorporación de esta población, y también existe el riesgo de que se los discrimine por no entender su forma de comunicación. Por ello, se busca desarrollar un sistema de reconocimiento de gestos para la comunicación entre la población hablante y sordomuda que emplee el LSC.

Avances en el desarrollo del proyecto

En cuanto al Leap Motion, se eligió porque presenta un nuevo sistema de seguimiento de gestos y posiciones con precisión submilimétrica (Weichert, Bachmann, Rudak & Fisseler, 2013). A diferencia de las soluciones multitáctiles estándar, este sensor de superficie se analiza para su uso en sistemas de interacción 3D estéreo realistas (Ching, E., 2012), especialmente en lo que respecta a la selección directa de objetos mostrados estereoscópicamente (Bruder, Steinicke & Stürzlinger, 2013).



Figura 1. Dispositivos de Captura.

Fuente: Propia.

Como se puede observar en la Figura 1, se ha acoplado el dispositivo Leap Motion a una banda para la cabeza, lo que permite realizar la captura de datos desde este punto de vista y así tener dentro del rango del dispositivo la zona donde se realizan los gestos referentes a dicho lenguaje.

Se plantea una interfaz demostrativa en la que se pueda visualizar la cámara proveniente del dispositivo Leap Motion, el renderizado de la mano del usuario y mensajes alusivos a la seña correspondiente a cada letra del Alfabeto Colombiano (Estático), la cual se encuentra realizada en la plataforma de desarrollo de Unity (ver Figura 2).



Figura 2. Interfaz de Usuario.

Fuente: Propia.

Para la toma de datos de entrenamiento se ha elegido las distancias que componen a la mano desde la palma hasta los diferentes puntos de los dedos, como se aprecia en la Figura 3, mediante el cálculo realizado con la Ecuación 1.

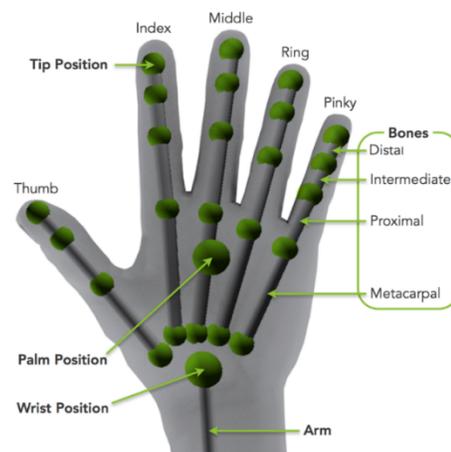


Figura 3. Positional Tracking.

Fuente: <https://blog.leapmotion.com/getting-started-leap-motion-sdk/>

Ecuación 1. Distancia entre palma y punto.

$$\sqrt{Dist_i = (P_{ix} - Q_x)^2 + (P_{iy} - Q_y)^2 + (P_{iz} - Q_z)^2}$$

En la Ecuación 1 se indica el cálculo para obtener los valores necesarios para ser incorporados en el modelo a realizar, donde “Pi” es la coordenada correspondiente a uno de los puntos entregados por Leap Motion y “Q” es la coordenada de la palma de la mano, en los ejes de referencia “X” “Y” “Z”, se realiza en los diferentes puntos de la mano.

Una vez adquiridos los datos necesarios para realizar la caracterización de la mano, se realiza una base de datos compuesta de 100 capturas por letra, con 22 letras, donde se incorporan las características explicadas anteriormente y se pone una etiqueta al final con la letra correspondiente, se obtiene una base de datos de 2.200 datos.

Tras tener dicha base de datos se procede a implementar un modelo de Machine Learning que permita realizar la identificación y clasificación de estos datos. Para la selección del algoritmo que mejor se ajuste al propósito buscado, se emplea el complemento de Visual Studio “Machine Learning Model Builder”, con el cual se encuentra un modelo óptimo basado en el algoritmo Stochastic Dual Coordinate Ascendent (SDCA).

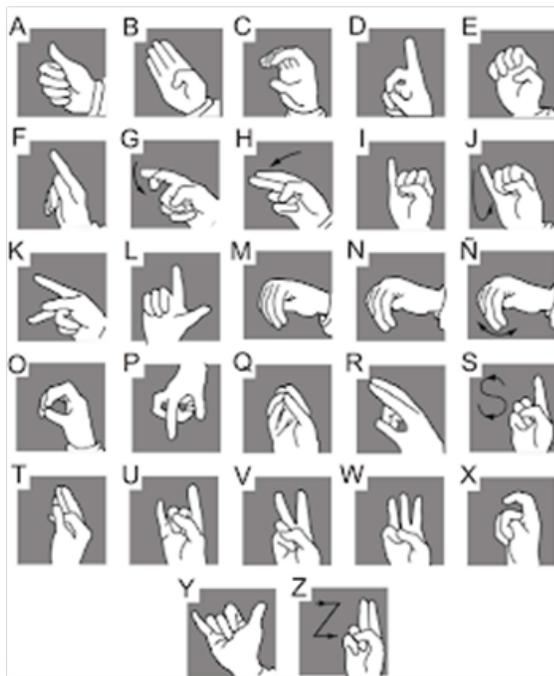


Figura 4. Dactilología LSC.

Fuente: <http://fonoaldia.unisucre.edu.co/2015/11/1er-seminario-taller-en-lengua-de-senas.html>

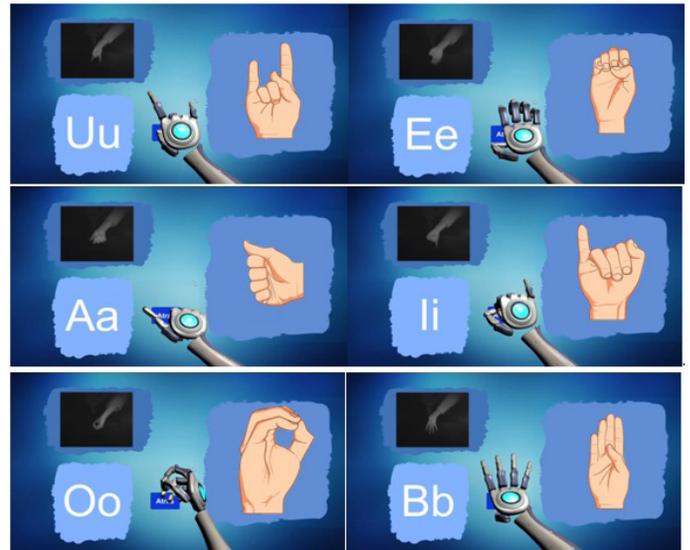


Figura 5. Mosaico 3x3 - Reconocimiento de algunas letras.

Fuente: Propia.

Se implementa el modelo obtenido en conjunto con la interfaz mostrada en la Figura 2, con lo cual se obtiene un software que permite reconocer las letras del abecedario, como lo mostrado en la Figura 5, la cual representa la interfaz en funcionamiento, donde se reconoce algunas de las letras del abecedario incorporado en la base de datos; además, la Figura 4 representa la Dactilología en LSC basado dicho software, de la cual cabe destacar que se implementan únicamente los de carácter estático, los cuales son 22 de las 27 letras.

Se puede visualizar la fiabilidad del modelo basado en la precisión obtenida de los diferentes métodos de validación cruzada, a partir de conjuntos de datos de entrenamiento y prueba de la base de datos adquirida (ver Figura 6).

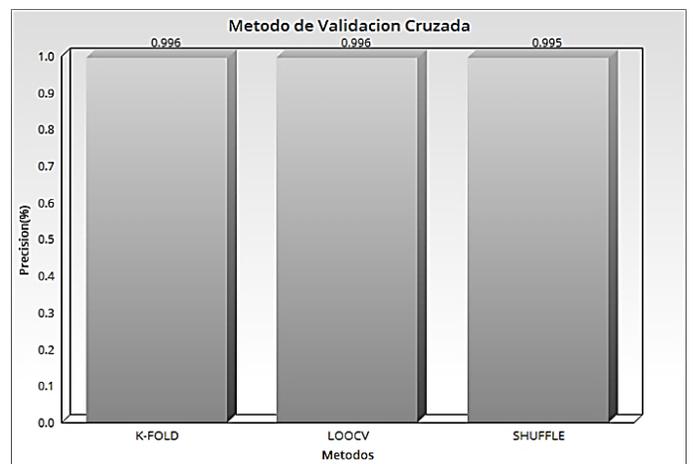


Figura 6. Porcentaje de aciertos del modelo con Validación Cruzada.

Fuente: Propia.

En la Figura 7 se muestran las fases de desarrollo del proyecto propuesto y las actividades realizadas (Azul) y también las que faltan por cumplir (Verde).

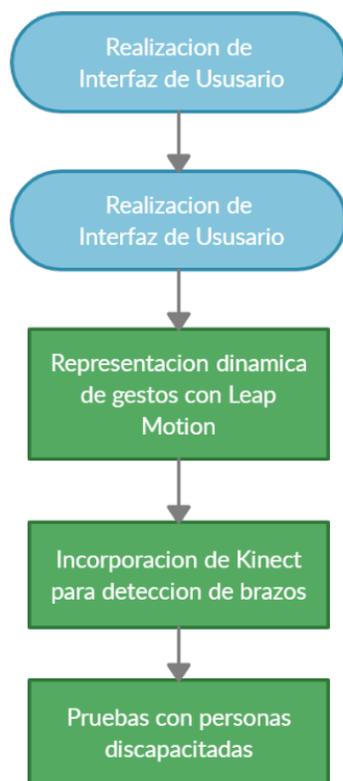


Figura 7. Metodología Propuesta.

Fuente: Propia.

Con base en el estado actual del proyecto, se puede concluir que la efectividad del modelo se debe en gran medida a tener dispositivos de captura precisos, por lo cual, se demuestra que el dispositivo de captura Leap Motion es óptimo para el reconocimiento automático de señas.

Un traductor de señas bien implementado podría ser de bastante utilidad en la sociedad, por ello, en futuros trabajos se puede implementar un vocabulario más amplio, el cual puede tener un gran impacto en la comunidad que utiliza lenguaje de señas.

Referencias

- Bruder, G., Steinicke, F., & Stürzlinger, W. (2013). Effects of visual conflicts on 3D selection task performance in stereoscopic display environments. *IEEE Symposium on 3D User Interfaces (3DUI)*. doi: 10.1109 / 3DUI.2013.6550207
- Ching, E. (2012). New Ways of Accessing Information Spaces Using 3D Multitouch Tables. *International Conference on Cyberworlds*, 144-150. doi: 10.1109 / CW.2012.27
- Tróchez, C. P. (2018). Servicios de diagnóstico e intervención de las pérdidas auditivas en Colombia: Un factor para el bienestar y calidad de vida.
- Weichert, F., Bachmann, D., Rudak, B., & Fisseler, D. (2013). Analysis of the Accuracy and Robustness of the Leap Motion Controller. *Revista Sensors*, 13(5), 6.380-6.393. doi: 10.3390/s130506380