

Sistema de interpretación de imágenes para el reconocimiento y traducción del lenguaje de señas

Dora María Calderón Nepamuceno, Gabriela Kramer Bustos,
Efrén González Gómez

Universidad Autónoma del Estado de México,
Centro Universitario Nezahualcóyotl,
México

{dmcalderonn, gkramerb, egonzalezg}@uaemex.mx

Resumen. El lenguaje de señas, es utilizado por las personas sordas, con la finalidad de comunicarse. Se compone de movimientos y expresiones realizadas especialmente con las manos, existe una gran ausencia de tecnologías al interpretar este lenguaje. Como sociedad es importante llevar a cabo iniciativas que mejoren la calidad de vida de la comunidad sordomuda del país. En el presente trabajo se muestra el proceso y la creación del diseño e implementación del sistema de reconocimiento de gestos, a través del lenguaje de programación en Python, junto con librerías como OpenCV (librería que permite detección de movimientos, reconocimiento de objetos a partir de imágenes), Numpy (Especializada en el cálculo numérico y el análisis de datos), Imutils (Funciones para realizar tareas básicas de procesamiento de imágenes de una manera más fácil). El sistema ha sido creado para realizar tareas básicas de procesamiento de imágenes de una manera más fácil, permite la visualización de la imagen adquirida y su traducción al lenguaje de señas.

Palabras clave: Python, procesamiento de imágenes, lenguaje de señas.

Image Interpretation System for Sign Language Recognition and Translation

Abstract. Sign language is used by deaf people to communicate. It is made up of movements and expressions made especially with the hands, there is a great absence of technologies when interpreting this language. As a society it is important to carry out initiatives that improve the quality of life of the country's deaf community. This paper shows the process and the creation of the design and implementation of the gesture recognition system, through the Python programming language, together with libraries such as OpenCV (library that allows movement detection, object recognition from images), Numpy (Specialized in numerical computation and data analysis), Imutils (Functions to perform basic image processing tasks in an easier way). The system has been created to perform basic image processing tasks in an easier way, it allows the visualization of the acquired image and its translation into sign language.

Keywords: Python, image processing, sign language.

1. Introducción

En la actualidad existen en el mundo millones de personas que sufren algún tipo de discapacidad y de las cuales sufren discriminación, por ejemplo, en algunos casos les es imposible acudir a la escuela y realizar ciertas actividades, o incluso para ellos es difícil conseguir empleo. Actualmente el lenguaje de señas, es la manera más efectiva y básica para la comunicación verbal de las personas sordas, con dificultad auditiva y para hablar.

Cuando hay personas que hablan diferentes idiomas y se quieren comunicar, utilizan gestos o movimientos, como, por ejemplo, gestos faciales y corporales, al realizarlos podemos comprender lo que quiere expresar la otra persona sin pronunciar una palabra; un simple movimiento de la boca, ojos, cejas, manos o cualquier extremidad, puede significar algo [1].

Las lenguas de los signos son habladas (en silencio) por cien millones de personas sordas en todo el mundo. En total hay al menos 138 idiomas de señas vigentes según el catálogo de Etnología, y muchos de ellos son lenguas oficiales (nacionales) u oficiales de la comunicación.

Existen personas que nacieron sordas y que incluso, no son capaces de leer. Además de los lenguajes que ya existen también hay alfabetos que se utilizan para deletrear letras (nombres, palabras raras, signos desconocidos, etc.).

Las personas que sufren este tipo de discapacidad tienen la necesidad de utilizar este sistema de señas, buscan la alternativa de aprender este tipo de lenguaje, por lo que se convierte en un reto memorizar cada una de las señas que conforman este sistema de comunicación, y sobre todo para las personas que no requieran descifrar algún mensaje.

Es por esto, que se requiere desarrollar una herramienta, que permita reconocer los gestos que son realizados por personas con discapacidad auditiva, con el fin de brindar instrumentos didácticos, que les facilite el aprendizaje para este tipo de lenguaje y lo puedan hacer de una manera más interactiva.

Por desgracia, aunque existan estas herramientas no todas las personas con esta discapacidad podrían obtenerlas, y esto hace que sigan dependiendo de la interpretación del lenguaje de señas mediante el apoyo de personas, para que tengan la facilidad de comunicarse con otras personas.

El objetivo de realizar este proyecto, surge por los problemas de comunicación que existe entre las personas sordas y el resto del mundo, y también por la discriminación que éstas reciben. Por ejemplo, hoy en día, es posible que la población sorda pueda acceder a instituciones educativas universitarias de forma virtual [3].

El uso de la tecnología ha permitido que la población discapacitada acceda a programas de la facultad de la Universidad Pedagógica y Tecnológica, de manera remota y con apoyo de traductores digitales.

Según estadísticas arrojadas por el Departamento Administrativo Nacional de Estadística, el 16.4% de los habitantes de la ciudad, tiene limitaciones permanentes para oír, es decir, que, de cada 100 personas con esta condición, 17 presentan algún tipo de discapacidad permanente auditiva [4]. “El lenguaje de señas se caracteriza por ser visual y corporal, es decir la comunicación se establece con el cuerpo, en un espacio determinado” [5, 6].



Fig. 1 Caracteres del lenguaje de señas.

Tabla 1. Materiales empleados para el desarrollo del sistema.

Librerías:	Descripción
Python	Es un lenguaje de alto nivel de programación, se utiliza para desarrollar aplicaciones de todo tipo, Además, se trata de un lenguaje multiplataforma de código abierto y, por lo tanto, gratuito, lo que permite desarrollar software sin límites.
OpenCV:	Es una librería para la detección de movimiento, reconocimiento de objetos, reconstrucción 3D a partir de imágenes, son sólo algunos ejemplos de aplicaciones
Numpy:	Biblioteca Especializada en el cálculo numérico y el análisis de datos, especialmente para un gran volumen de datos
Imutils:	Es creado para realizar tareas básicas de procesamiento de imágenes de una manera más fácil.
Cámara	Es el dispositivo empleado para la captura de imágenes.

2. Materiales y métodos

Para poder desarrollar el proyecto se necesitaron materiales y métodos para la implementación del desarrollo del sistema, a continuación, se realiza la descripción de los materiales y métodos empleados.

Entre estos trabajos, se encuentra la propuesta de García Incertis [1] que reconoce algunas letras del alfabeto de la Lengua de Signos; por su parte, Razo Gil [2] reconoce imágenes de la mayoría de las letras del alfabeto del Lenguaje de Señas. Los materiales que se utilizaron para la creación del sistema, son indicados en la Tabla 1, con el que cada uno tiene su propia descripción, por lo cual, es útil para el desarrollo de dicho sistema, y para que sea más sencilla la comprensión de usuarios que son ajenos a este desarrollo.

3. Desarrollo

Para el desarrollo de este sistema, se realizó el reconocimiento de las imágenes del alfabeto del lenguaje de señas, el cual, se enfocará en el cálculo de las variantes para poder identificar las imágenes capturadas e identificar las señales obtenidas y hacer su

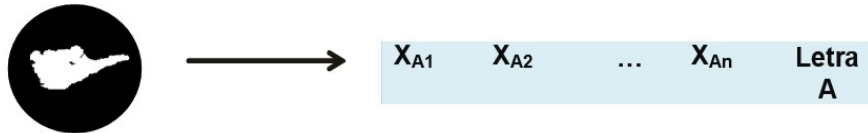


Fig. 3. Señal de la letra A.

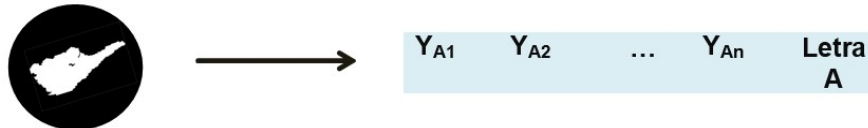


Fig. 4. Señal de la letra A con una variación.

correspondiente traducción. Se mostrarán las etapas, para llevar a cabo el sistema de reconocimiento de las imágenes:

1. Uso de una cámara digital.
2. Iluminaciones controladas.
3. Uso de imágenes estáticas.
4. Enfocado en transformaciones.
5. Las letras del alfabeto son realizadas con las manos.

3.1. Realización de la captura de la imagen

Se captura la imagen con una cámara digital, el formato de la imagen será jpg a color, la iluminación se dará por la lámpara del dispositivo, la imagen será convertida a una matriz.

3.2. Captura de la imagen

La corrección de la iluminación se realiza mediante un proceso morfológico, este proceso se usa para disminuir objetos claros sobre fondos oscuros. La imagen original es mejorada mediante la resta de la imagen de apertura.

3.3. Segmentación

De la imagen obtenida con la cámara se realiza un segmentado, en el proceso se realiza un umbral de reconocimiento.

3.4. Filtro morfológico

El filtro morfológico, facilita la interpretación de la imagen removiendo estructuras brillantes de la imagen.

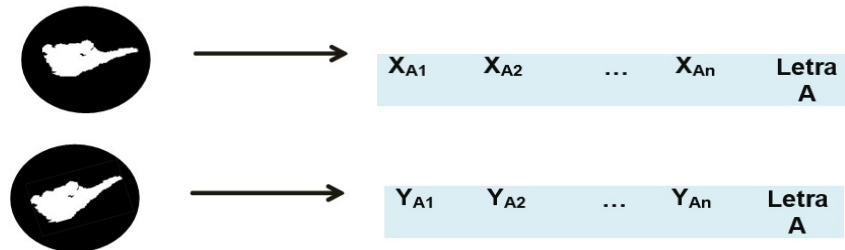


Fig. 5. Resultado del proceso en una matriz evolutiva.



Fig. 6. Imagen del resultado del proceso.

3.5. Arquitectura de las etapas del reconocimiento de imágenes

En la figura 2 se muestra la arquitectura del proceso del reconocimiento de una imagen.

3.5.1. Segmentación de la mano

Se realizan cálculos hasta encontrar uno o más objetos. La imagen procesada es una imagen binaria y los píxeles que contienen información tienen valor de 1.

3.5.2. Escalamiento de la mano

Transforma la imagen a una forma cuadrada, este proceso se realiza dando un ajuste a un tamaño definido de 200 x 200 píxeles. El paralelogramo obtenido se escala al tamaño con el fin de minimizar la variación de tamaño de las manos.

3.5.3. Matrices evolutivas

Se usa una matriz evolutiva para almacenar los patrones del conjunto de letras del alfabeto y hacer su identificación mediante un programa de computadora.

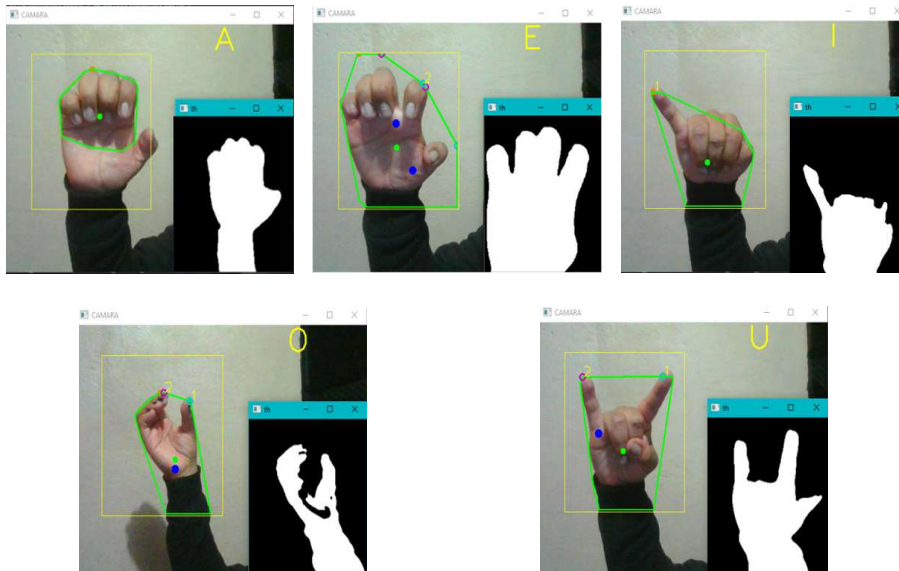


Fig. 7. Resultado del reconocimiento de las vocales en forma de texto.

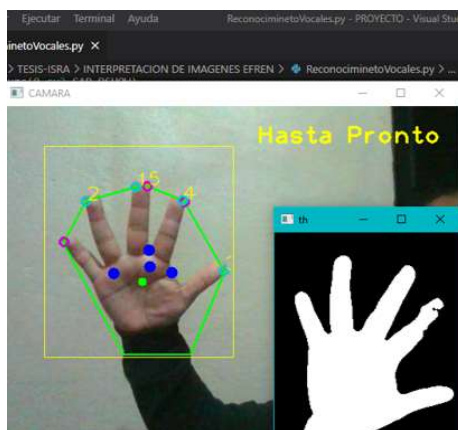


Fig. 8. Resultado del reconocimiento de Hasta Pronto (despedida) en forma de texto.

3.5.4. Ejemplo de matriz evolutiva

- Se procesa señal de la letra A como se muestra en la figura 3.
- Se procesa otra señal de la letra A con una variación cómo se muestra en la figura 4.
- El resultado se almacena con el significado de la letra A en la matriz evolutiva figura 5.

4. Resultados

Los resultados obtenidos con el sistema, reconocen las señas del alfabeto, de lenguaje de señas identificadas en imágenes capturadas mediante la cámara. Para realizar Las operaciones se requiere capturar la imagen realizando la señal del alfabeto que se desea. A la imagen se le aplica el proceso de umbralado dando como resultado la Figura 6.

El resultado del proceso final del reconocimiento de la imagen se muestra en forma de texto tal como se muestra en las figuras 7, donde se ejemplifica las 5 vocales y en la figura 8 se muestra la ejemplificación de una palabra.

5. Conclusiones y trabajo a futuro

El sistema de reconocimiento, alcanzó a distinguir las 5 vocales del alfabeto del lenguaje de señas que son: A, E, I, O, U. Para ello las principales características son extraídas de la imagen y estas deben ser detectables con el ruido y la iluminación que contenga. El uso de la matriz evolutiva y los procesos de normalización, permiten mayor similitud en la detección de las imágenes.

El reforzamiento disminuye patrones con los que se hace la comparación para el reconocimiento de las señas hechas con la mano. Con base en los experimentos y dado el parecido de algunas señas se detectó que el umbral debe de ser menor o igual al 10% para garantizar una adecuada relación entre los patrones.

Los cambios de iluminación se deben controlar ya que de ello depende la ejecución correcta de los resultados que obtengamos del sistema. OpenCV cuenta con una extensa documentación, así como la integración de varios lenguajes de programación, es una herramienta de software libre, con lo cual se permitió realizar dicho sistema.

Referencias

1. González-Riveros, C. G., Yimes-Inostroza, F. J.: Sistema de reconocimiento gestual de lengua de señas Chilena mediante cámara digital. Informe final del proyecto para optar al título profesional de ingeniero de ejecución en informática, Pontificia Universidad Católica de Valparaíso Facultad de Ingeniería Escuela de Ingeniería Informática (2016) opac.pucv.cl/pucv_txt/txt-0500/UCC0990_01.pdf
2. Rojas-Rojas, L. M., Arboleda-Toro, N., Pinzón-Jaime, L. J.: Caracterización de población con discapacidad visual, auditiva, de habla y motora para su vinculación a programas de pregrado a distancia de una universidad de Colombia. *Revista Electrónica Educare*, vol. 22, no. 1, pp. 1 (2017) doi: 10.15359/ree.22-1.6
3. Ballesta-Perez, J. L.: Diseño e implementación de sistema de interpretación. Obtenido de traducción de gestos asociados a preguntas, necesidades y saludos básicos del lenguaje de señas colombiano. Universidad de Córdoba Facultad de Ingenierías Programa de Ingeniería de Sistemas y Telecomunicaciones, pp. 1–109 (2015) <https://repositorio.unicordoba.edu.co/handle/ucordoba/280>
4. Visual studio code: Code editing, redefined. Edición de Código (2021) code.visualstudio.com/

Dora María Calderón Nepamuceno, Gabriela Kramer Bustos, Efrén González Gómez

5. Mejía, H.: Lengua de Señas Colombiana: Segundo Tomo, Santa Fe de Bogotá, Federación Nacional de Sordos de Colombia (1996) <https://fenascol.org.co/wp-content/uploads/2022/01/Tomo%20%20LSC%20Fenascol.pdf>
6. Seijas, L. M.: Reconocimiento de patrones utilizando técnicas estadísticas y conexionistas aplicadas a la clasificación de dígitos manuscritos. Tesis Doctoral, Universidad Nacional de Buenos Aires, Facultad de Ciencias Exactas y Naturales (2011) https://bibliotecadigital.exactas.uba.ar/collection/tesis/document/tesis_n4997_Seijas