

Diseño de un sistema de reconocimiento de gestos no móviles mediante el procesamiento digital de imágenes

Digital image processing applied on static sign language recognition system

Betsy Villa^{1*}, Valeria Valencia², Julie Berrio³

^{1,2}Ingeniera Mecatrónica, Grupo de investigación de Ingeniería Mecatrónica, Universidad Autónoma del Caribe, Barranquilla, Colombia

³Master en Ingeniería Electrónica, Universidad del Valle, Cali, Colombia
Email: betsy.villa@uautonoma.edu.co

Recibido: 13/11/2017

Aceptado: 1/06/2018

Cite this article as: B. Villa, V. Valencia, J. Berrio, "Digital image processing applied on static sign language recognition system", *Prospectiva*, Vol 16, N° 2, 41-48, 2018.

RESUMEN

Lengua de Señas Colombiana o L.S.C. es el lenguaje de señas autóctono utilizado por las personas sordas para comunicarse en Colombia. Se compone de movimientos y expresiones realizadas con diferentes partes del cuerpo, especialmente las manos. En Colombia, hoy en día, hay una gran ausencia de tecnologías encaminadas al aprendizaje e interpretación de éste; por ende, es un compromiso social, llevar a cabo iniciativas que promuevan la mejora de la calidad de vida de la población sordomuda del país. En este artículo se muestra el proceso de diseño e implementación de un sistema de reconocimiento de gestos no móviles usando como principal herramienta de desarrollo el entorno de Matlab, donde se implementó el algoritmo SIFT ("Scale Invariant Feature Transform"). El sistema permite la visualización de la imagen adquirida y su traducción al lenguaje de señas colombiano respectivamente. A través de la identificación de puntos clave en las imágenes y su posterior comparación con patrones SIFT en la base de datos del sistema es posible obtener la traducción. La herramienta realiza el reconocimiento de las 20 letras no móviles de este conjunto, implementando una interfaz gráfica en Matlab para una mejor visualización, fácil acceso y uso por parte del usuario. Se comprueba una mejor respuesta del sistema mediante la utilización de un estandarizador de imagen, en este caso, un guante quirúrgico, generando un porcentaje de fiabilidad del 77%, y la necesidad del uso del mismo. Se plantea para continuidad de una segunda fase del proyecto la optimización del sistema aplicando métodos de redes neuronales en tiempo real, generando un mayor impacto para las necesidades actuales de la población colombiana.

Palabras clave: Lenguaje de Señas Colombiano; Matlab; Procesamiento de imágenes; Reconocimiento de gestos; Reconocimiento estadístico; Segmentación de imágenes; SIFT.

ABSTRACT

Sign language L.S.C. (Lengua de Señas Colombiana) is a native language which chiefly uses manual communication to convey meaning. This can involve simultaneously combining hand shapes, movement and orientation of the hands, arms or body, and even facial expressions to convey a speaker's ideas. Currently in Colombia, there is an absence of technology focus on teaching and interpreting this language; for this reason, it's interesting and a social commitment to carry out initiatives that promote life quality improvement for the country's deaf-mute population. In this article, the design and implementation process of a static hand gesture recognition system is shown, for this task we used Matlab as computing environment and the Scale Invariant Features Transform (SIFT) method to extract characteristics from the image. Our system allows the acquired image visualization and its corresponding translation to the colombian sign language. Through key points identification and their comparison with SIFT features in the system data base makes possible to retrieve the translation. The system can recognize 20 static letters from the Colombian Sign Languages, a graphical interface was implemented in Matlab that provides better visualization, simple access to the system and high usability. A better response of the system is noticed when a standardized element of the image is used, in our case, a surgical glove. As future work, we propose to apply neural networks to improvement of the tool, and a real time implementation, which can generate a greater impact for the current needs of the colombian population.

Key words: Colombian Sign Language; Matlab; Image processing; Gesture recognition; Static recognition; Image segmentation; SIFT.

1. INTRODUCCIÓN

En la actualidad los lenguajes de señas son el método efectivo para la comunicación verbal de las personas sordas, con dificultades auditivas y de las personas que tienen dificultades para hablar. No existe un lenguaje de señas universal, y casi cada país tiene su propia lengua de señas nacional. Todos los lenguajes de señas usan signos visual-cinéticos para la comunicación humano-humano combinando gestos manuales con articulación de labios y mímicas faciales. También poseen una gramática específica y simplificada que es muy diferente de la de las lenguas habladas acústicas.

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 humana en algunos países como Estados Unidos, Finlandia, la República Checa, Francia, la Federación de Rusia (desde 2013), etc. De acuerdo con las estadísticas de las organizaciones médicas, alrededor del 0,1% de la población de cualquier país es absolutamente sorda y la mayoría de esas personas se comunican únicamente por lenguaje de señas.

Muchas personas que nacieron sordas incluso no son capaces de leer. Además de los lenguajes de las lenguas de conversación, también hay alfabetos que se utilizan para deletrear letras (nombres, palabras raras, signos desconocidos, etc.), letra por letra [1].

En este artículo se presenta el desarrollo de un sistema de reconocimiento de gestos no móviles del lenguaje de señas colombiano, mediante el procesamiento digital de imágenes. El Sistema se divide principalmente en cuatro fases: adquisición de imagen, procesamiento de la imagen, reconocimiento de la imagen y muestra de resultado.

Según estadísticas arrojadas por el Departamento Administrativo Nacional de Estadística (DANE), en Colombia el 6,4% de los habitantes tiene una limitación permanente; de este grupo el 17,3% de las personas, tienen limitaciones permanentes para oír, es decir, que, de cada 100 colombianos con limitación, 17 presentan algún tipo de discapacidad permanente auditiva [2].

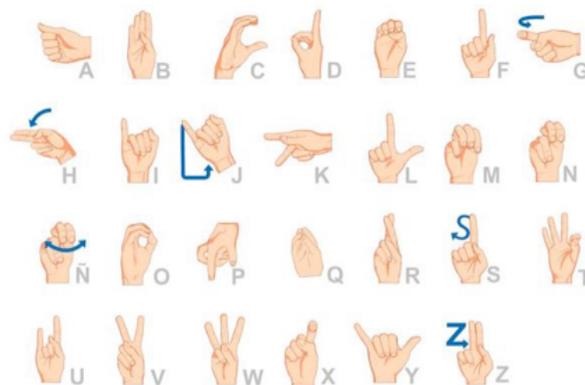
La comunicación mediante el lenguaje de señas es casi imprescindible en el día a día de un ser humano que sufre de limitación auditiva, pues a partir de ella puede entrar a relacionarse con las personas que le rodean y a la vez con su entorno, por lo que se hace un reto aprender cada una de las señas que conforman su sistema de comunicación, y más aún para las personas sin discapacidad al momento de lograr descifrar el mensaje proporcionado por esta población. Es por esto que surge la necesidad de desarrollar una herramienta que permita reconocer los gestos que son realizados por personas con discapacidad auditiva, con el fin de

brindar una herramienta didáctica que permita a las personas que están aprendiendo este tipo de lenguaje (discapacitados o no discapacitados); lo puedan hacer de una manera más interactiva y amigable con su entorno.

“El lenguaje de señas se caracteriza por ser visual y corporal, es decir la comunicación se establece con el cuerpo en un espacio determinado” [3].

El Lenguaje de Señas Colombiano (LSC) fue reconocido en el año 1996 y es el lenguaje oficial utilizado por la comunidad sorda del país. El LSC consta de 27 señas que representan a cada una de las letras del alfabeto, mediante la unión de varias señas se puede lograr expresar palabras, frases e ideas completas. En la implementación del sistema solamente se utilizan 20 señas no móviles del LSC debido a que esta investigación se basa principalmente en el análisis y tratamiento de imágenes. La figura 1 contiene todos los caracteres que representan el alfabeto en el LSC.

Figura 1. Caracteres del lenguaje de señas colombiano [3].
Figure 1. Colombian sign languages [3].



Por otro lado, el procesamiento de imágenes consiste en todos los métodos y técnicas implementados para extraer información de la imagen. La segmentación es una de esas técnicas la cual básicamente busca separar el fondo del objeto de interés. Las imágenes en su mayoría están compuestas por zonas que tienen características similares (nivel de gris, textura, momentos, etc.). Generalmente estas zonas corresponden a objetos de la imagen. La segmentación de una imagen consiste en la división o partición de la imagen en varias zonas o regiones homogéneas y disjuntas a partir de su contorno, su conectividad, o en términos de un conjunto de características de los píxeles de la imagen que permitan discriminar unas regiones de otras. Entre las características utilizadas en la segmentación encontramos: los tonos de gris, la textura, los momentos, la magnitud del gradiente, la dirección de los bordes, entre otras [4].

Actualmente existen varios enfoques de reconocimiento de objetos, de los cuales los más conocidos son:

- Reconocimiento estadístico: se basa en la determinación y uso de funciones de probabilidad [5].
- Reconocimiento sintáctico: analiza la estructura de los objetos, por ejemplo: el esqueleto [6].
- Reconocimiento de redes neuronales: constituye la técnica más reciente y trata de imitar el funcionamiento de los sistemas biológicos [7].

En el desarrollo del sistema de reconocimiento de gestos no móviles se utilizó el reconocimiento estadístico mediante el cual, para la detección de gesto de la mano, fue necesario hallar los puntos clave característicos de cada imagen con el fin de poder compararlos con la imagen de entrada y finalmente si existe una coincidencia mostrar la imagen de salida obtenida. Este método se conoce como SIFT (Scale Invariant Features Transform) introducido por David Lowe en 1999, es considerado una de las mejores técnicas de procesamiento de imágenes, ya que la información obtenida de la imagen se transforma en coordenadas invariantes de escala, rotación y luminosidad. Lo primero que se realiza es obtener un conjunto de puntos de la imagen, los cuales serán denominados “puntos clave”, de acuerdo con el paso por cada etapa, el número de puntos claves se irá reduciendo y quedarán los más importantes para ser usados en la comparación. Su idea principal es la transformación de la imagen a una representación compuesta de “puntos de interés”.

Esos puntos contienen la información característica de la imagen que luego son usados para la detección de muestras [8].

El algoritmo se realiza mediante 4 etapas:

1. Construcción de Pirámides de Scale-Space: en esta etapa se representa la imagen en diferentes escalas y tamaños. Mediante la función de diferencia gaussiana se identifican los posibles puntos de interés que son invariantes a escala y orientación [9].
2. Localización de puntos clave: en esta etapa se localizan los puntos que se mantienen invariantes en escala. Esto se realiza mediante la comparación de cada pixel con su pixel vecino. Los puntos clave son seleccionados en base a las medidas de su estabilidad [10].
3. Asignación de orientación: en esta en esta etapa se asigna a cada punto clave una orientación de acuerdo con las direcciones del gradiente y a la zona que rodea dicho punto [11].
4. Descriptor de puntos clave: esta etapa consiste en calcular un descriptor para la región de la imagen local que sea fácilmente identificable, sin embargo,

tan invariable como sea posible a las variaciones restantes, tales como el cambio en la iluminación o el punto de vista 3D [12].

Durante los últimos años se han realizado trabajos relacionados con sistemas de reconocimiento de señas, como el de Ansari (2017), el cual reconoce gestos con diversas técnicas de extracción de características y máquinas de vectores de soporte como un clasificador. Propone un enfoque híbrido entre SIFT y HoG (histograma orientado de gradientes) combinados como una característica entre esas encontramos SIFT [13].

El sistema de reconocimiento y traducción del lenguaje de señas desarrollado por Raziq y Latif (2016), implemento un dispositivo USB, que tiene la capacidad de obtener una descripción precisa de la mano y su movimiento mediante la técnica de leap motion. El sistema está compuesto por módulos de comunicación, preparación. El módulo de preparación permite utilizar los datos recibidos para entrenar el sistema y el módulo de comunicación captura el movimiento y envía la información al algoritmo que permite detectar y reconocer el signo para luego convertirlo a texto [14].

Aquilla y col. (2015), realizaron el proyecto de “Reconocimiento de caracteres del alfabeto dactilológico mediante redes neuronales artificiales”, por medio de las imágenes adquiridas por una cámara digital, procesan y segmentan esta imagen de entrada mediante técnicas de tratamiento digital para luego ser enviadas a un clasificador basado en Redes Neuronales Artificiales (RNA) que permite identificar el carácter mostrado [15].

En el trabajo de Rodríguez y col. (2014), se realiza un prototipo traductor de señales manuales a texto legible utilizando una cámara Kinect. El sistema es capaz de tomar los puntos de una escena captada por medio del sensor de profundidad de Kinect, aplicar un filtro para eliminar el ruido y posteriormente mostrar el Mesh o Maya que reconstruya la imagen de la escena en 2D, estableciendo la diferencia de distancia con el cambio de color. La aplicación desarrollada en Open CV proyecta la localización de la mano y luego toma una captura del gesto manual para finalmente esta pueda ser comparada con la base de datos de los gestos y cuente con su traducción respectiva [16].

Betancur y col. (2013), proponen un sistema integrado de hardware y software para el reconocimiento de lenguaje dactilológico. El hardware consiste en un sistema inalámbrico adherido a un guante, el cual posee un conjunto de sensores que permite capturar las señales generadas por los movimientos gestuales de la mano. El software trabaja con un modelo de adaptación de redes neuronales que permite la identificación de las vocales que conforman el lenguaje

de señas colombiano. Se obtuvo un resultado del 78% de reconocimiento de las vocales en el sistema integrado [17].

En el trabajo documentado en Priego-Pérez (2012), se presenta un sistema para reconocer la información contenida en imágenes del lenguaje de señas. Su autor establece dos etapas: en la etapa de reconocimiento utiliza la cámara de un dispositivo Kinect2, mediante la cual se adquiere nuevos patrones que posteriormente se comparan con los ya almacenados en la base de conocimiento del sistema durante la etapa de aprendizaje. En general, el enfoque se basa en una estimación de similitud entre la imagen adquirida y la del patrón almacenado, considerando únicamente valores de similitud que superen el 90% [18].

Chiguano y col. (2011), desarrollaron un sistema que traduce el lenguaje de señas e implementan un entrenador de este. El sistema fue desarrollado en LabView, concluyendo que el desempeño del sistema depende en gran medida de las condiciones del ambiente bajo el cual se realice la adquisición de la imagen, puesto que este último al ser controlado, arrojaba una imagen más limpia facilitando el tratamiento de la misma, pero reduciendo la flexibilidad de la aplicación para su uso en cualquier condición [19].

Kelly y col. (2010), realizaron un sistema de reconocimiento de señas en el lenguaje de señas americano. El sistema obtenía la imagen de una cámara digital para luego realizar la ubicación de la mano y segmentación de esta imagen mediante un algoritmo robusto que permitía obtener el contorno de la mano. Mediante una función de tamaño se convertía esta imagen digital en una binaria y esta información se almacenó en una matriz que contenía el contorno del gesto [20].

El trabajo de Razo-Gil y col. (2009), se considera únicamente aquellas letras que se representan en el alfabeto dactilológico sin la necesidad de ejercer movimiento, mediante un método de valor umbral para el procesamiento y segmentación de las imágenes que logra separar unos objetos de otros con el objetivo de identificar las regiones de interés de acuerdo con la postura en la que se encuentre la mano del sujeto. El objetivo de este enfoque es conseguir extraer características relevantes que se puedan medir [21].

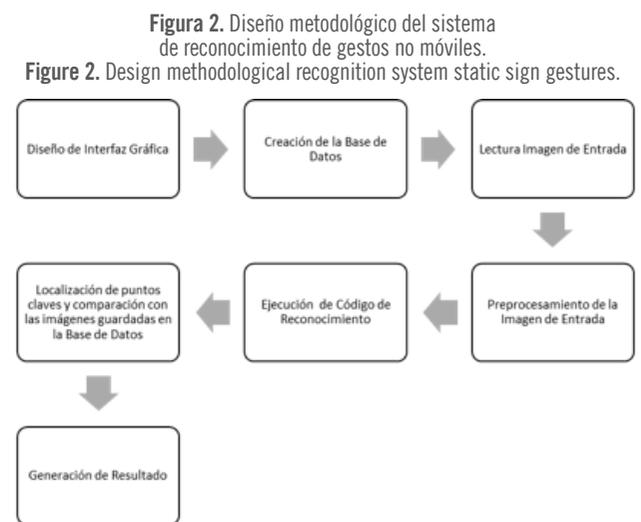
Como se puede observar en los trabajos analizados, es importante iniciar con un tratamiento de imágenes, para luego continuar con el proceso de reconocimiento y utilización del sistema, con el fin de garantizar resultados y minimizar probabilidades de error en la aplicación diseñada [21].

2. METODOLOGÍA

En la figura 2 se muestra todas las etapas

utilizadas para el diseño y la ejecución del sistema de reconocimiento de gestos no móviles. Como entorno de programación y diseño de la interfaz se utilizó Matlab, primero se realizó la captura de imágenes y la creación de la base de datos del sistema; la cual está conformada por 60 fotografías de tres patrones distintos. La captura de imágenes de prueba se da a partir de la cámara digital Canon PowerShot A4000 IS Azul, la cual nos brinda una alta calidad de las fotografías capturadas teniendo en cuenta sus 16 Megapíxeles y Zoom Óptico 8x.

La imagen de entrada del sistema es sometida a un preprocesamiento para disminuir el ruido y también se hace un cambio en las dimensiones para disminuir el costo computacional, quedando en un tamaño de 640x480 píxeles.



2.1 Interfaz gráfica

La interfaz se desarrolló por medio de GUIDE (Graphical User Interface Development Environment) del entorno de Matlab, el cual consta de una ventana, que se distribuye de la siguiente forma:

1. Encabezado: nombre del proyecto, estudiantes, y universidad.

2. Captura de Imagen: caja para la visualización de la imagen a tomar por parte de la cámara, y la cual posee los siguientes comandos:

2.1 Encender cámara: comando utilizado para que la cámara aparezca en el cuadro principal donde se encuentra el logo más grande la Universidad Autónoma del Caribe.

2.2 Apagar cámara: comando utilizado para dejar de utilizarla cámara.

2.3 Capturar: comando utilizado para realizar captura de imágenes.

2.4 Abrir imagen: comando utilizado para abrir las imágenes capturadas.

3. Visualización de caracteres del Lenguaje de Señas Colombiano e imagen capturada de entrada al algoritmo principal.

4. Visualización de imagen de salida.

5. Información del proyecto.

Figura 3. Distribución de la interfaz gráfica para el reconocimiento de gestos de la mano.
Figure 3. Distribution of the user graphic interface for the recognition of hand sign.



2.2 Función de reconocimiento

La función de reconocimiento recibe la imagen de entrada la convierte en escala de grises y mediante la técnica de SIFT los puntos de interés de la imagen de entrada para que estos puedan ser comparado con los puntos de interés de las imágenes que se encuentran en la base de datos del sistema, al final la función si encuentra dos imágenes con puntos de interés iguales o similares significa que ha encontrado la señal y letra correspondiente al lenguaje de señas colombiano. Si la función no encuentra ninguna relación entre los puntos de interés de la imagen de entrada y las que se encuentran en la base de datos simplemente el sistema arrojará que no ha sido posible identificar la seña y que lo intente nuevamente.

2.3 Scale Invariant Features Transform (SIFT)

SIFT es un método propuesto por David Lowe, en donde una imagen se transforma la información en coordenadas invariantes de escala y rotación. El algoritmo permite extraer puntos estables y de interés de una imagen. El conjunto de estos puntos de interés representa descripciones características que son utilizadas para identificar un objeto en otras imágenes [8].

Los puntos de interés SIFT contienen información en una región circulas de una imagen, y contienen cuatro parámetros importantes: el centro de coordenadas X e Y, la escala (radio de la región) y el ángulo de orientación expresado en radianes.

La secuencia del cálculo de los puntos de interés SIFT es dividida en 6 etapas:

1. Construcción de un espacio de escala: una imagen con varios detalles incluye mucha

información que debe reducirse; para ello, se utiliza un conjunto de operaciones como: suavizado Gaussiano, generación de octavas de la imagen (reducir el tamaño en un octavo) y aplicación del suavizado en las octavas. Los conjuntos de nuevas imágenes son utilizados en la siguiente etapa.

2. Aproximación del método LoG (Laplaciano Gaussiano): los puntos de interés SIFT se encuentran principalmente en los bordes y esquinas de una imagen, que a su vez pueden obtenerse tras aplicar el Laplaciano Gaussiano a la imagen suavizada; no obstante, este proceso es costoso computacionalmente. Lowe [8] demostró que el Laplaciano Gaussiano puede aproximarse a la Diferencia de Gaussianos de dos imágenes obtenidas a partir de la reducción de Octavas de una misma imagen. De este modo, se consigue un conjunto de puntos de interés iniciales invariantes a los cambios de escala.

3. Detección de los puntos de interés: se obtienen los puntos de interés a partir de los máximos y mínimos de la Diferencia de Gaussianos. Lowe [8] propone buscarlos alrededor de los 26 vecinos de un pixel a partir de la segunda imagen de Octava, ya que es la primera imagen en la que los 26 vecinos existen.

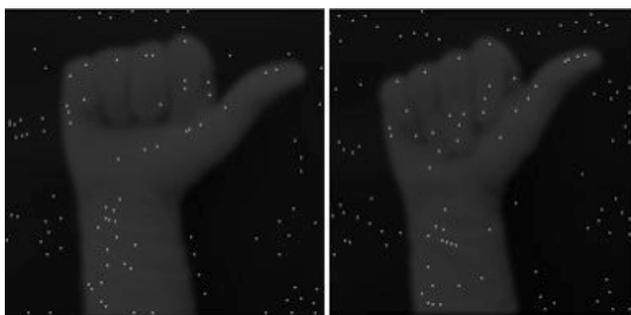
4. Eliminación de puntos de interés con características de bajo contraste y ubicados en esquinas: consta de dos posibles criterios de eliminación: El primero consiste en que, si la magnitud de la intensidad de un pixel en el DoG es menor a determinado valor, es rechazado. El segundo calcula los valores de dos gradientes perpendiculares en un punto de interés, si las gradientes son pequeñas (región plana), o si una es grande y la otra pequeña (borde) son eliminadas.

5. Asignación de la orientación para los puntos de interés: en esta etapa se asigna una orientación a los puntos de interés anteriores para otorgarles la propiedad "Invariancia a la Rotación". Se calcula la dirección y la magnitud de las gradientes de cada punto de interés, se forma un histograma de orientación, y la orientación más prominente de la región es asignada al punto de interés.

6. Generación de las características SIFT: finalmente, con la rotación y escala invariantes, se realiza una última representación que facilite la identificación de las características de los puntos de interés [22].

Figura 4. Imagen en base de datos vs imagen de entrada, puntos característicos obtenidos.

Figure 4. Image in database vs. input image, characteristic points obtained.



3. RESULTADOS Y DISCUSIÓN

En esta sección se comparten los resultados obtenidos durante el proceso de creación, organización, elaboración e implementación del proyecto.

3.1 Recolección de base de datos

Realizamos la base de datos teniendo en cuenta únicamente las 20 señas no móviles del lenguaje de señas colombiano. Todas las señas fueron realizadas con la mano derecha por 4 personas diferentes, cada seña fue fotografiada en fondo negro 3 veces. Por lo tanto, cada persona aportó 60 imágenes que formaron parte de nuestra base de datos.

En nuestra primera base de datos se realizó la captura de la mano derecha sin importar la tez del color de piel, sin embargo, se determinó que era necesario regular esta variable y por esto se realizó una segunda base de datos con la captura de las señas usando un guante quirúrgico de color blanco.

Las imágenes se guardaron en formato de archivo “.jpg”. Las imágenes contenidas en la base de datos son: A, B, C, D, E, F, I, K, L, M, N, O, P, Q, R, T, U, V, W, Y.

3.2 Sistema de reconocimiento

Durante la realización de pruebas al sistema de reconocimiento de gestos no móviles del lenguaje de señas colombiano, se ejecuta inicialmente el algoritmo principal, en el cual se encuentra el código desarrollado para la Interfaz de Usuario y en el que a su vez se lee la imagen de entrada que puede estar guardada en el ordenador, o capturada por medio de la cámara del mismo con dimensiones 640x480.

Una vez elegida la imagen, ésta se despliega en la parte superior de la interfaz, como se observa en la figura 5.

Posteriormente se ejecuta el proceso de reconocimiento, pulsando un botón llamado de la misma forma “Reconocimiento”, con el fin de encontrar los puntos claves o característicos de la imagen de entrada y las imágenes que se encuentran en la base de datos, y rea-

Figura 5. Visualización de imagen de entrada.

Figure 5. Input image display.



Figura 6. Resultado obtenido luego de realizar el proceso de reconocimiento.

Figure 6. Result obtained after performing the recognition process.



lizando una comparación donde se determina cuál es la imagen encontrada en la base de datos que se asocia a la imagen ingresada al sistema.

3.3 Resultado sistema reconocimiento: prueba 1

En la primera prueba se realizó la captura de las señas sin cubrir la mano con ningún elemento. Tres personas de diferente tez realizaron todas las señas no móviles del lenguaje de señas colombianos tres veces para el reconocimiento de la misma. La prueba fue realizada utilizando la cámara que se encontraba en el ordenador y el fondo de la imagen de entrada era negra iluminado con luz natural todas las capturas fueran realizadas mediante los mismos parámetros.

El porcentaje de error en esta primera prueba fue de un 54%, las señas no estaban siendo reconocidas satisfactoriamente, las condiciones de iluminación y el color de tez de la piel de los participantes afecto la fiabilidad del sistema. Solamente fueron reconocidas con un porcentaje del 70% o mayor las letras B, C, M, N, V, W, Y. En la figura 7, se encuentran documentados el porcentaje de reconocimiento y error de cada seña.

3.4 Resultados sistema de reconocimiento: prueba 2

Durante la segunda prueba se capturaron imágenes

Figura 7. Resultado prueba 1.
Figure 7. Test result 1.

N°	Imagen	Letra	Porcentaje de Reconocimiento	Porcentaje de Error
1		A	40%	60%
2		B	100%	0%
3		C	70%	30%
4		D	50%	50%
5		E	50%	50%
6		F	50%	50%
7		I	40%	60%
8		K	60%	40%
9		L	20%	80%
10		M	70%	30%
11		N	70%	30%
12		O	50%	50%
13		P	10%	90%
14		Q	20%	80%
15		R	50%	50%
16		T	60%	40%
17		U	40%	60%
18		V	70%	30%
19		W	80%	20%
20		Y	80%	20%

utilizando un guante quirúrgico, con el fin de estandarizar el sistema, para minimizar los errores arrojados en la primera prueba. Tres personas de diferente tez realizaron todas las señas no móviles del lenguaje de señas colombianos tres veces para el reconocimiento de la misma. La prueba fue realizada utilizando la cámara que se encontraba en el ordenador y el fondo de la imagen de entrada era negro, las condiciones de luz eran las óptimas y todas las pruebas fueran realizadas mediante estos mismos parámetros. Los resultados obtenidos arrojaron un porcentaje de fiabilidad del 77%, en esta prueba las señas de las letras A, C, D, K, M, P, R, T, U obtuvieron un porcentaje de reconocimiento mayor del 90% incluso algunas reconocidas en su totalidad. En la figura 8, se encuentran documentado el porcentaje de reconocimiento y error de cada seña para la segunda prueba.

Figura 8. Resultado prueba 2.
Figure 8. Test result 2.

N°	Imagen	Letra	Porcentaje de Reconocimiento	Porcentaje de Error
1		A	90%	10%
2		B	60%	40%
3		C	100%	0%
4		D	90%	10%
5		E	100%	0%
6		F	30%	70%
7		I	70%	30%
8		K	100%	0%
9		L	60%	40%
10		M	100%	0%
11		N	30%	70%
12		O	60%	40%
13		P	100%	0%
14		Q	70%	30%
15		R	90%	10%
16		T	100%	0%
17		U	100%	0%
18		V	70%	30%
19		W	40%	60%
20		Y	60%	20%

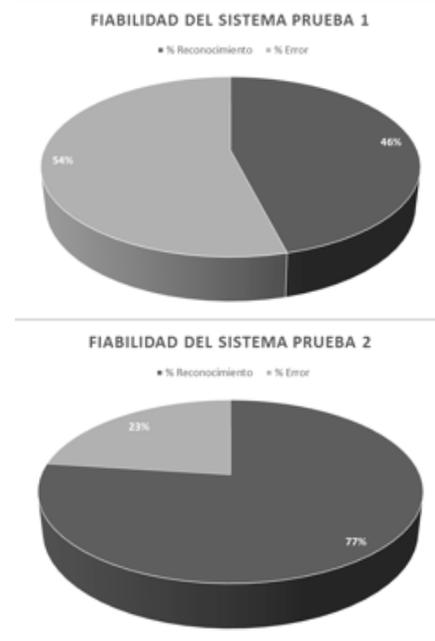
3.5 Análisis comparativo

Los sistemas que implementan procesamiento de imagen digital requieren que la captura de las imágenes tenga condiciones de ambiente adecuado, en la primera prueba se utilizó luz natural, este factor y además el realizar las señas sin tener en cuenta la tez de la piel disminuyó la eficiencia del sistema de

reconocimiento de señas no móviles. En la primera prueba se obtuvo un porcentaje de éxito del 54% el sistema solamente identificaba la mitad de las señas.

En la segunda prueba se utilizó una luz blanca frontal para iluminar correctamente las señas a realizar y además la mano se cubrió con un guante de color blanco. Al implementar estos dos cambios se obtuvo un porcentaje de éxito de 77%, las señas de las letras A, C, D, E, I, K, L, M, O, P, Q, R, T, U mejoraron su porcentaje de fiabilidad en comparación a la primera prueba; solo las señas de las letras B, F, N, W, Y mostraron una disminución en el proceso de reconocimiento.

Figura 9. Gráfica del porcentaje de fiabilidad del sistema.
Figure 9. Graph of the percentage of reliability of the system.



4. CONCLUSIONES

Actualmente, la implementación de aplicaciones utilizando el procesamiento digital de imágenes es de gran auge, debido al sin número de análisis que se pueden realizar por medio de la captura de imágenes o videos, para solventar distintas necesidades.

El algoritmo de reconocimiento se encarga de extraer características intrínsecas de las imágenes de entrada y base de datos, que resulten ser claves al momento de realizar la comparación que permita determinar la seña correspondiente al Lenguaje de Señas Colombiano. Además, se encarga de recibir una imagen de una seña del Lenguaje de Señas Colombiano dada por el usuario, y hallar con el mejor grado de certeza posible, en una base de datos.

Para el reconocimiento preciso y fiable, las características extraídas de la imagen de entrada deben ser detectables incluso en los cambios en la escala de la imagen, el ruido y la iluminación. Se encuentra una mejor respuesta del sistema, con la utilización de un

elemento que estandarice la mano con respecto a los factores de color, utilizando un guante, en este caso, tipo quirúrgico. Las condiciones de luz deben ser adecuadas para obtener un resultado preciso.

REFERENCIAS

- [1] A. Karvop, I. Kipyatkova, M. Zelezny, "Automatic Technologies for Processing Spoken Sign Languages", *Procedia Computer Science*, 81, 201-207, 2016. Departamento Administrativo Nacional De Estadística, (2006). Censo General 2005: Discapacidad personas con limitaciones permanentes [Internet], Santa Fe de Bogotá, DANE. Disponible desde: < <https://www.dane.gov.co/files/censo2005/discapacidad.pdf>> [Acceso 27 de noviembre 2014]
- [2] H. Mejía, *Lengua de Señas Colombiana: Segundo Tomo*, Santa Fe de Bogotá, Federación Nacional de Sordos de Colombia, 1996.
- [3] A. Naranjo, (2014). Aplicación de DSP's para la Transcripción de Lenguaje de Señas a Texto. Tesis de Pregrado, Universidad Técnica Ambato.
- [4] L. Seijas, (2011). Reconocimiento de patrones utilizando técnicas estadísticas y conexionistas aplicadas a la clasificación de dígitos manuscritos. Tesis de Doctorado, Universidad de Buenos Aires.
- [5] C. Platero, *Apuntes de visión artificial*, Departamento de Electrónica, Automática e Informática Industrial, Editor Universidad Politécnica de Madrid, 2009.
- [6] D. Paredes, (2009). Seguimiento y Caracterización de Componentes del Rostro para la Detección de Expresiones Faciales. Tesis de Maestría, Centro Nacional de Investigación y Desarrollo Tecnológico.
- [7] D. Lowe (1999, Sept.). Object Recognition from Local Scale-Invariant Features. Presentado en The Proceedings of the 7th IEEE International Conference on Computer Vision. Vol. 2. IEEE [En línea]
- [8] Disponible: <http://ieeexplore.ieee.org/document/790410?reload=true&tp=&arnumber=790410> M. Brown, D. Lowe. "Invariant Features from Interest Point Groups", *BMVC*, 4, University of British Columbia, 2002.
- [9] D. Lowe, "Distinctive image features from scale-invariant key points", *International journal of computer vision*, 60, 91-110, 2004.T. Lindeberg (2012). Citing reference: "Scale invariant feature transform." Scholarpedia. Disponible: http://www.scholarpedia.org/article/Scale_Invariant_Feature_Transform Y. Meng, B. Tiddeman (2006). Citing reference: "Implementing the Scale Invariant Feature Transform(SIFT) Method." Department of Computer Science University of St. Andrews. Disponible: <https://pdfs.semanticscholar.org/c373/01c2e0fe7ca8974c9e5e236bdb9a19c2cda.pdf>
- [10] Ansari, F. J. (2017). "Hand Gesture Recognition using fusion of SIFT and HoG with SVM as a Classifier", *International Journal of Engineering Technology Science and Research*, 4(9), 913-922, 2017.
- [11] N. Raziq, S. Latif, "Pakistan Sign Language Recognition and Translation System using Leap Motion Device", *Data Engineering and Communications Technologies*, 1, 895-902, 2017.
- [12] D. Auquilla, K. Palacio-Baus, V. Saquicela, (2015). Reconocimiento de caracteres del alfabeto dactilológico mediante redes neuronales artificiales: Un enfoque experimental. Tesis de Doctorado, Universidad Cuenca.
- [13] C. Rodríguez, J. Pineda, D. Sánchez, (2014). Prototipo traductor de señales manuales a texto legible utilizando Kinect. Tesis de Pregrado, Fundación Universitaria de San Gil.
- [14] D. Betancur, M. Vélez, A. Peña. "Traducción automática del lenguaje dactilológico de sordos y sordomudos mediante sistemas adaptativos", *Revista Ingeniería Biomédica*, 7 (13), 18-30, 2013.
- [15] F. Priego-Pérez, (2012). Reconocimiento de Imágenes del Lenguaje de Señas Mexicano. Tesis de Maestría en Ciencias de la Computación, Instituto Politécnico Nacional, Centro de Investigación en Computación.
- [16] C. Rodríguez, E. Fabián, N. Moreno, (2011). Diseño e implementación de un sistema traductor de lenguaje de señas de manos a un lenguaje de texto mediante visión artificial en un ambiente controlado. Tesis de Pregrado, Escuela Politécnica Nacional.
- [17] D. Kelly, J. McDonald, C. Markham, "A person independent system for recognition of hand postures used in sign language", *Pattern Recognition Letters*, 31(11), 1359-1368, 2010.
- [18] L. Razo-Gil, G. Salvador-Calderón, R. Barrón-Fernández, (2009). Sistema Traductor Para el Reconocimiento del Alfabeto Dactilológico. Tesis Maestría, CIC-IPN: Centro de Investigación en Computación, Laboratorio Inteligencia Artificial.
- [19] A. Andreopoulos, J. K. Tsotsos, "50 Years of object recognition: Directions forward", *Computer Vision and Image Understanding*, 117(8), 827-891, 2013.