

RECONOCIMIENTO DE SIGNOS DEL LENGUAJE DE SEÑAS USANDO FIRMAS DE CORRELACIÓN

José Francisco Solís Villarreal

Profesor del Centro Universitario UAEM
Valle de Teotihuacán, Santo Domingo de
Aztacameca, Axapusco, Estado de México

Oscar Espinoza Ortega

Profesor del Centro Universitario UAEM
Valle de Teotihuacán, Santo Domingo de
Aztacameca, Axapusco, Estado de México

David Martínez Martínez

Profesor del Centro Universitario UAEM
Valle de Teotihuacán, Santo Domingo de
Aztacameca, Axapusco

All content in this magazine is licensed under a Creative Commons Attribution License. Attribution-Non-Commercial-Non-Derivatives 4.0 International (CC BY-NC-ND 4.0).



Resumen: Se presenta un esquema para el reconocimiento del lenguaje de señas usando técnicas de extracción de parámetros basados en firmas de correlación rotacional que toman como información de entrada imágenes en dos dimensiones y generan señales en una dimensión. Uno de los principales problemas del reconocimiento de lenguaje de señas es la extracción adecuada de la Región de Interés (ROI por sus siglas en inglés), este trabajo contempla el uso de un sistema de visión por computadora que evita la implementación de sensores electrónicos y marcadores visuales o físicos que limitan la practicidad de generar los gestos del lenguaje de señas, permitiendo así una expresión fluida y natural del individuo que se esté analizando. Se eligieron los signos de los dígitos del 1 al 9 del lenguaje de señas empleando un perceptrón multicapa con validación cruzada alcanzado un 100%.

Palabras clave: Sistema de visión artificial, reconocimiento de lenguaje de señas, firmas de correlación, redes neuronales artificiales.

INTRODUCCIÓN

El lenguaje de señas es el principal medio de comunicación de la gente sordomuda, cabe destacar que el lenguaje de señas no es universal, existen varios y son diversos, cambian en cada región debido a una serie de costumbres y tradiciones. El lenguaje de señas es un conjunto de movimientos de los dedos, manos, brazos, posturas corporales y gestos faciales (Aberni et al. 2020, García y Sánchez, 2020, Toygar et al. 2020, Wang y Qin, 2020 y Wu et al. 2020).

Desafortunadamente para la población sordo muda la gran mayoría de personas en el mundo desconocen los lenguajes de señas, esto es un problema mayor para los no hablantes vocalmente debido a que solo pueden comunicarse libremente entre ellos mismos y sus familiares más cercanos.

Frecuentemente sucede la necesidad de entablar comunicación entre un sordomudo y una persona que no habla el lenguaje de señas, por mencionar unos ejemplos se puede señalar el caso de un sordomudo tramitando la credencial de elector, abriendo una cuenta de banco, llevar a cabo un registro en una empresa privada o en un sector gubernamental, todas estas actividades y muchas otras se llevan a cabo generalmente en acompañamiento de familiar cercano de la persona sordomuda quien hace la traducción del lenguaje de señas al lenguaje verbal.

La dependencia que tienen los sordomudos de un traductor para poder hacer trámites de todo tipo impiden que las personas sordomudas tengan privacidad, es decir, el traductor (un familiar muy cercano regularmente) se va a enterar de todos los asuntos privados del sordo mudo como estados de cuenta de un banco y demás información sensible que no se puede en estos casos mantenerse como privada, además los sordomudos se limitan de hacer ciertos trámites donde prefieren sacrificar el beneficio de dicha transacción que hacer del conocimiento al traductor de algunas situaciones privadas.

Este trabajo se desarrolla con el fin de reducir la dependencia del traductor para los sordomudos, es decir, la idea en general es que este tipo de desarrollos puedan llegar algún día a hacer que los sordomudos no necesiten de un traductor para hacer ciertos tipos de trámites, por ejemplo que en el banco se pueda tener un sistema que realice la traducción del lenguaje de señas al lenguaje hablado de manera natural, esta es la meta de estas investigaciones.

Durante toda la elaboración de la investigación se procuró evitar el uso de marcadores especiales de color como guantes, puntos, pelotas, bolas u otros, así mismo se evitó usar dispositivos electrónicos pegados

o asegurados en los dedos, manos, brazos o cuerpo para que el individuo pueda expresar los signos del lenguaje de señas libremente y de manera natural.

BASE DE DATOS

El lenguaje de señas consiste en un conjunto de más de treinta signos estáticos, estos incluyen los signos del alfabeto, los números y algunas palabras especiales, para este trabajo se eligieron los nueve dígitos (del uno al nueve), como se puede observar en la figura 1.

SISTEMA DE VISIÓN PROPUESTO

El sistema propuesto en esta investigación es el reconocimiento de dígitos del lenguaje de señas, sin el uso de un fondo especial, guantes o dispositivos electrónicos, para lograr estos objetivos, fue necesario utilizar un cambio en el espacio de color y una umbralización

para lograr una segmentación adecuada de la información de la mano. Para poder reducir la cantidad de parámetros que contiene cada imagen se empleó una técnica de extracción de firmas de correlación y con esta información se alimentó a una red neuronal artificial para poder alcanzar un porcentaje de reconocimiento de patrones alto.

Para este trabajo se propuso realizar un cambio del espacio de color, pasando del RGB al HSI, mediante las expresiones de transformación

$$I = \frac{1}{3}(R + G + B)$$

$$S = 1 - \frac{3}{(R + G + B)} [\min(R, G, B)]$$

$$H = \cos^{-1} \left\{ \frac{\frac{1}{2} [(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\}$$

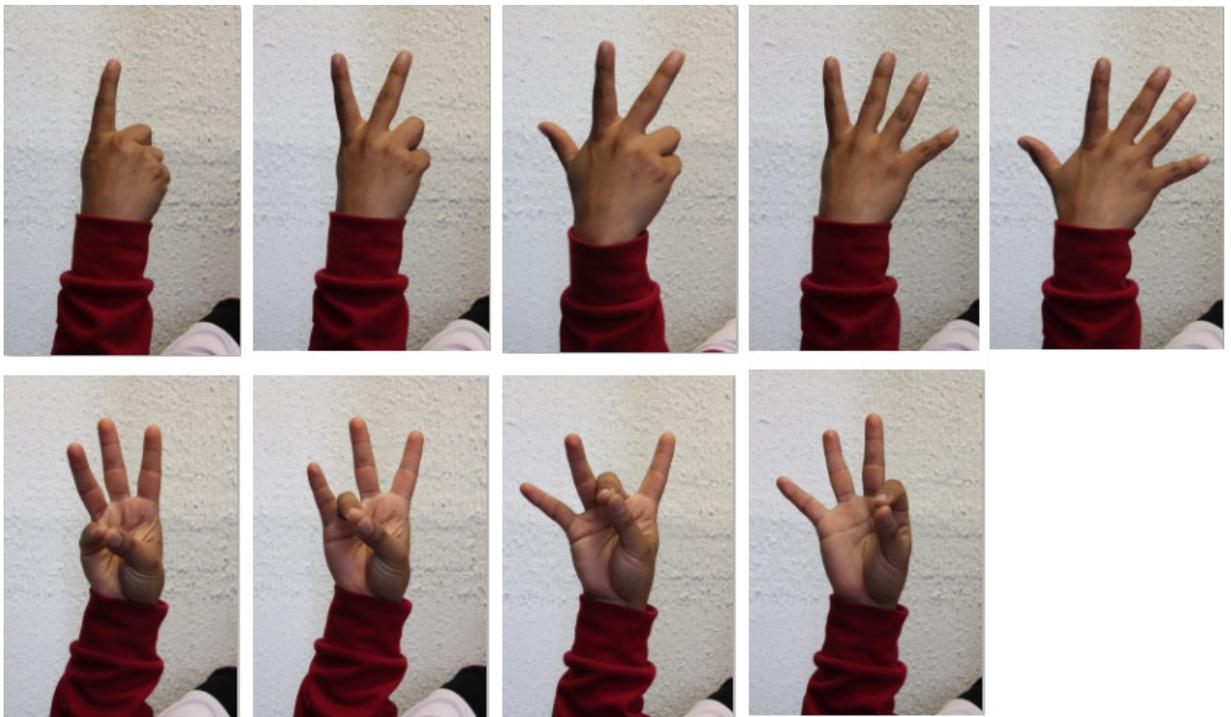


Figura 1. Nueve dígitos del lenguaje de señas del uno al nueve de izquierda a derecha y de arriba hacia abajo.

A partir de estos valores se aplicó un umbral para poder extraer la región de interés (ROI), que consiste en la zona de la mano (palma o dorso, y dedos). Los resultados se pueden observar en la figura 2.



Figura 2. Dígito uno en RGB (izquierda).
Dígito uno en HSI umbralizado (derecha).

En la figura 3 se esquematiza el sistema de visión propuesto para poder lograr el reconocimiento de los dígitos del lenguaje de señas, primeramente se observa la captura de imágenes mediante un sensor CCD que adquiere y representa las imágenes digitales

en los canales RGB, posteriormente se realiza una conversión de RGB a HSI, una vez terminada la conversión se realizan procesos de procesamiento digital de imágenes para poder segmentar la información contenida en el signo expresado, es decir la mano que incluye la piel del dorso o la palma y los dedos.

Una vez que se tiene la imagen procesada y segmentada, se prosigue a calcular la firma de correlación, para lograr esto se puede utilizar la expresión siguiente (Padilla et al. 2004)

$$g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x + i, y + j)h(i, j) didj$$

Este operador puede ser utilizado para calcular las firmas de correlación rotacional, el cual utiliza la función de la imagen $f(x, y)$ y una versión rotada de la misma $h(i, j)$, se puede calcular una serie de versiones rotadas con valores de $\theta = 0, 1, 2, 3, \dots, 360$, cuando se calculan los valores de correlación, se extraen los 360 valores máximos normalizados para obtener así una firma de correlación rotacional como se puede observar en la figura 4.

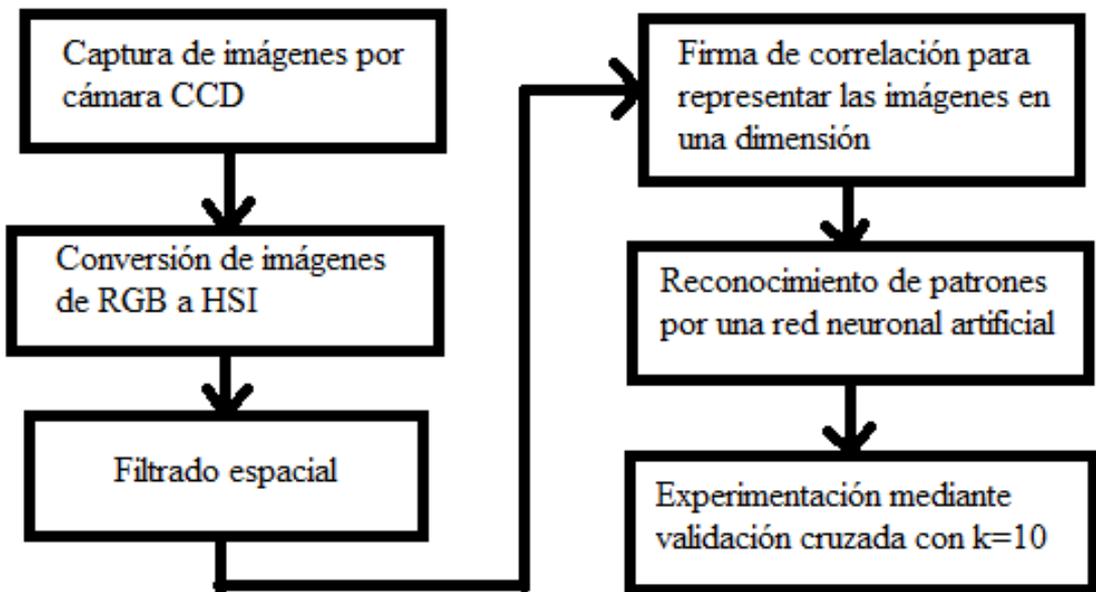


Figura 3. Modelo de sistema de visión por computadora para el reconocimiento de los dígitos del lenguaje de señas.

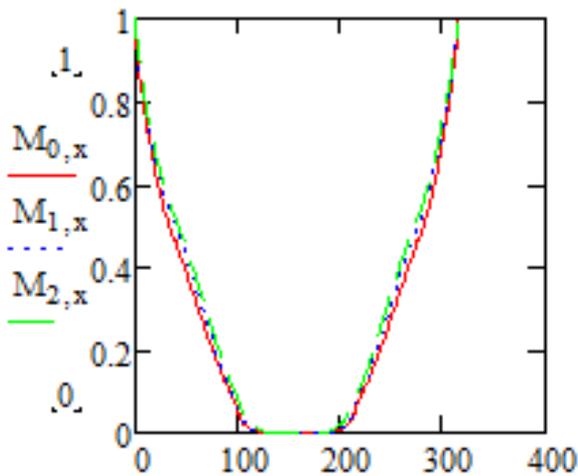


Figura 4. Firmas de correlación rotacional de 3 versiones del dígito uno.

Para el reconocimiento de los dígitos se utiliza el modelo del perceptrón multicapa, el cual se diseñó con trescientos sesenta neuronas en la capa de entrada para ser entrenada con los valores de las firmas de correlación, para la capa de salida se tuvieron nueve neuronas, las cuales se entrenaron para detectar cada uno de los dígitos de la base de datos de entrenamiento.

COMENTARIOS FINALES

Este trabajo aborda el problema del reconocimiento de signos del lenguaje de señas, particularmente los dígitos, se abordó primeramente cambiando el espacio de color de RGB a HSI y mediante técnicas de procesamiento digital de imágenes se logró la segmentación de la región de interés (mano), se aplicó una firma de correlación rotacional para reducir una dimensión del problema (de dos dimensiones a una), con esta última información se utilizó un perceptrón multicapa para lograr la clasificación correcta de los patrones.

Cabe destacar que se utilizó un esquema de validación cruzada para obtener resultados de clasificación confiables. Con los resultados alcanzados en este trabajo se puede afirmar

que con este conjunto de técnicas es suficiente para poder lograr el reconocimiento de dígitos bajo las condiciones encontradas en el presente experimento, como trabajo a futuro se plantea la experimentación añadiendo a más individuos para poder hacer pruebas con personas distintas. El porcentaje de reconocimiento alcanzado con el sistema de visión por computadora propuesto es del 100% con diez divisiones utilizando el esquema de validación cruzada.

REFERENCIAS BIBLIOGRÁFICAS

Las referencias bibliográficas se deben presentar por orden alfabético de primer autor: “El uso del método XZY ha resultado muy favorable en sistemas como el que propuesto por Wiley y Cabrera (2004). Otros autores (Puebla Romero et al. 2007 y Washington y Frank, 2000) prefieren el uso de las derivadas de Thomas. No fue sino hasta que Etxeberri y Blanco Gorrichoa (2007) propusieron sus radicales ideas que...” Nótese que el artículo donde aparece Puebla Romero tiene tres autores y por esa razón se usa la abreviación latina *et al* (del latín, “y los demás”. Al final de este manuscrito mostramos la forma de citar las referencias.

REFERENCIAS

Urcid, G., Padilla-Vivanco, A., Cornejo, A. and Baez, J. (2004) Correlation-Based Rotational Operator Signature of Planar Binary Objects. Proceedings of SPIE5558, Applications of Digital Image Processing XXVII, 2 November 2004, 87. <http://dx.doi.org/10.1117/12.558744>

Wu, W., S. Elliot., S. Lin, S. Sun y Y. Tang “Review of palm vein recognition” IET Biometrics, 9, 2020. <https://doi.org/10.1049/iet-bmt.2019.0034>

Wang, P. y H. Qin “Palm-vein verification based on U-Net” IOP Conference Series: Materials Science and Engineering, 806, 2020. <https://doi.org/10.1088/1757-899X/806/1/012043>

Toygar, O., F. Babalola y Y. Bitirim “FYO: A Novel Multimodal Vein Database With Palmar, Dorsal and Wrist Biometrics” IEEE Access, 8, 2020. <https://doi.org/10.1109/ACCESS.2020.2991475>

García R. y R Sánchez “Vein Biometric Recognition on a Smartphone”. IEEE Access, 8 2020. <https://doi.org/10.1109/ACCESS.2020.3000044>

Aberni Y., L. Boubchir y B. Daachi. “Palm vein recognition based on competitive coding scheme using multi-scale local binary pattern with ant colony optimization”. Pattern Recognition Letters, 136, 2020. <https://doi.org/10.1016/j.patrec.2020.05.030>