

Sistema de reconocimiento y clasificación de señas para el lenguaje español

Jorge Cerezo¹, Griselda Saldaña¹, Mario M. Bustillo², Apolonio Ata²,
J. Andrés Vázquez², Beatriz Bernabé², Gerardo Martínez²

¹ Universidad Tecnológica de Puebla,
Ingeniería en Tecnologías para la Automatización, México

² Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación, México

{jorge.cerezo, griselda.saldana}@utpuebla.edu.mx,
bustillo1956@hotmail.com,
{apolonio.ata, andrexsol, beatriz.bernabe}@gmail.com,
gguzman@cs.buap.mx

Resumen. En este trabajo se presenta un sistema para reconocimiento y clasificación de las letras que componen el lenguaje de señas en español, diseñado para apoyar en el entrenamiento de personas sordomudas y facilitar su comunicación con el resto de las personas. Se basa en un guante de bajo costo que captura el movimiento de la mano, el cual contiene un acelerómetro en cada dedo para detectar posición, y está conectado a una tarjeta de adquisición de datos. La información de los sensores se envía de forma inalámbrica a una computadora que contiene una interface desarrollada en LabVIEW, donde se genera una base de datos de símbolos. Para el reconocimiento de las letras, se aplicó un tratamiento estadístico a los datos en dicha base obteniendo una precisión superior al 96% independientemente del usuario.

Palabras clave: Lenguaje de señas, clasificación automática, guante.

Recognition and Classification of Sign Language for Spanish

Abstract. In this paper it is presented a computational system for recognition and classification of letters for sign language in Spanish, designed for helping deaf-mute people to communicate with other people. A low-cost glove that capture the hand movements has been constructed. This one contains an accelerometer for each finger which allows to detect its position by using an acquisition data card. Sensor information is sent wirelessly to a computer having a software interface, developed in LabVIEW, in which the symbols dataset is generated. For the automatic recognition of letters, we have applied a statistical treatment to the dataset obtaining an accuracy greater than 96%.

Keywords. Signs language, machine learning, glove.

1. Introducción

El reconocimiento de gestos ha recibido atención de muchas áreas de investigación tales como interacción humano-computadora, la realidad virtual, la tele-manipulación y el procesamiento de imágenes. Otra área de aplicación es la interpretación del lenguaje de señas [1]. Entre los tipos gestos, el lenguaje de señas es uno de los más estructurados, usualmente cada gesto está asociado a un significado predefinido. Por otra parte, la aplicación de fuertes reglas de contexto y gramática hace que el lenguaje de señas sea más fácil de reconocer [2].

De acuerdo a la tecnología de sensado empleada para capturar los gestos, existen dos aproximaciones principales para el reconocimiento de señas, una basada en técnicas de visión [3], donde se sigue el movimiento de la mano y se interpreta la seña correspondiente [4, 5] y otra basada en guantes [6] que cuentan con sensores que capturan la rotación y movimiento de la mano y los dedos [7]. El reconocimiento en base a sensores tales como acelerómetros y giroscopios ofrecen las siguientes ventajas: a) dado que los sensores de movimiento no se ven afectados por el entorno, el reconocimiento es más adecuado que el reconocimiento basado en visión en entornos complejos b) están unidos a un usuario, proporcionando así una mayor cobertura, y c) las señales se pueden adquirir de forma inalámbrica [8].

En trabajos previos se han utilizado de manera exitosa guantes como elementos para reconocimiento de señas [10], en [1] se presenta un sistema para reconocimiento de las 23 letras del lenguaje vietnamita, empleando un guante con acelerómetros MEMS, cuyos datos se transforman a ángulos relativos entre los dedos y la palma. Para el reconocimiento de las letras utiliza un sistema de clasificación basado en lógica difusa. En [11] se reporta un guante basado en acelerómetros y sensores mioeléctricos, los cuales permiten detectar de forma automática el punto inicial y final de segmentos significativos de los símbolos mediante la intensidad de los sensores mioeléctricos. Para obtener el resultado final, utiliza árboles de decisión y modelos ocultos de Markov. La funcionalidad del sistema se demuestra al clasificar los 72 símbolos del lenguaje de señas chino.

En este trabajo se presenta la implementación de un sistema entrenador del lenguaje de señas del alfabeto en español para sordomudos. Consta de un dispositivo tipo guante con un acelerómetro conectado a cada dedo. Las salidas de los sensores pasan a una tarjeta de adquisición que envía los datos de forma inalámbrica a una computadora donde reside una interface en LabVIEW.

Los datos recolectados se mantienen en una base de datos de símbolos, donde a diferencia de [7] esta información se clasifica empleando un método estadístico. Una vez que se discriminan los símbolos sin ambigüedad, el sistema puede utilizarse para entrenamiento de personas sordomudas, quienes desde otra interface en LabVIEW, pueden realizar cada una de las letras del alfabeto en español y comprobar si lo hacen de forma correcta.

El resto del documento está organizado de la siguiente manera. En la sección 2 se presenta una descripción del sistema, haciendo énfasis en la implementación del guante y el funcionamiento de los sensores. En la sección 3 se presentan el mecanismo de clasificación de los datos. En la sección 4 se presentan las pruebas realizadas al sistema así como algunos resultados obtenidos para finalmente en la sección 5 presentar las conclusiones y trabajo futuro.

2. Descripción del sistema

El sistema consta de tres elementos, un guante instrumentado con acelerómetros analógicos que puede enviar información de forma inalámbrica, y dos programas en labVIEW, el primero para la captura de muestras y el segundo para el entrenamiento de personas en la realización de símbolos. Los programas cuentan con una interface gráfica que es muy intuitiva y permiten a cualquier usuario interactuar con el sistema.

2.1. Construcción del guante

El diseño del guante se basa en la utilización de acelerómetros, en este caso los ADXL335 ya que son de bajo costo y consumen poca potencia. Dichos acelerómetros proporcionan una medida de la posición de los dedos en tres ejes con un formato serial (x, y, z), los acelerómetros del guante proporcionan datos crudos que se envían a la tarjeta de adquisición en un formato de vector y se envían a la computadora central a través de un dispositivo Xbee.



Fig. 1. Estructura del guante.

2.2. Captura de las muestras

La computadora tiene desarrollado un programa en LabVIEW que se utiliza para capturar los datos correspondientes a cada una de las letras del alfabeto en español y almacenarlo en una base de datos. Para ello se recurrió a un grupo de 25 personas sordomudas quienes realizaron un total de 50 veces cada letra. La interface de usuario realizada para la captura de las muestras se observa en la figura 2.

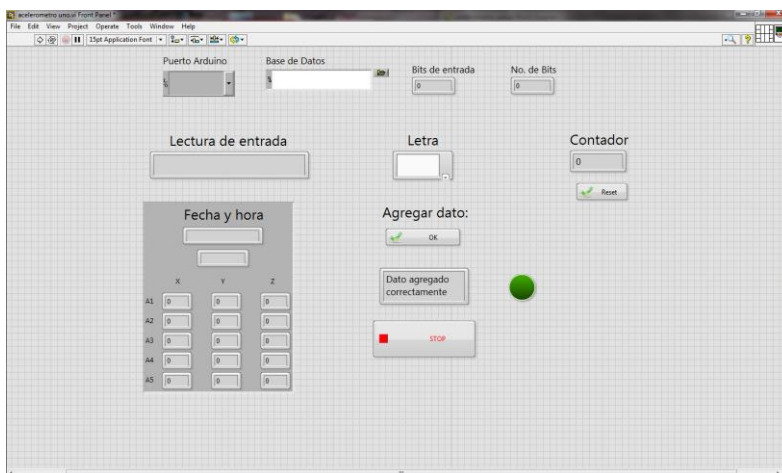


Fig. 2. Interface de usuario para la captura de datos.

Fuera de línea, se utilizan estos datos para realizar un proceso de clasificación donde cada subclase corresponde a una letra en particular. Para la operación en línea el usuario a entrenar accede a otra interface de usuario donde se le indica si es que está realizando de forma adecuada cada letra.

El usuario ejecuta una letra y a continuación se hace una lectura de los datos del guante, la información se compara con la información obtenida del sistema entrenador. Una vez que se reconocen cada una de las lecturas X, Y, Z de cada acelerómetro, se muestra la letra correspondiente en la pantalla, y así el usuario puede corroborar que la realiza de manera correcta y repite el proceso con un nuevo símbolo. Si lo desea, puede proceder a formar una palabra.

3. Clasificación

Una vez capturado los datos, podemos utilizarlos para construir un modelo de clasificación que pueda ser utilizado posteriormente para identificar una señal y asociarla automáticamente a una determinada letra. Las lecturas X, Y y Z que se obtuvieron de cada uno de los cinco acelerómetros se utilizan como características para la construcción del modelo de clasificación. En particular, hemos hecho experimentos con los siguientes tres clasificadores:

- a. J48: Es un clasificador del tipo de árbol de decisión: El algoritmo J48 es una implementación del algoritmo C4.5, uno de los algoritmos de minería de datos que más se ha utilizado en multitud de aplicaciones.
- b. SMO: El nombre proviene del inglés “Sequential minimal optimization”, y es un algoritmo para solucionar el problema de programación cuadrática que surge durante el entrenamiento de las máquinas de soporte vectorial. Fue inventado en 1998 por John Platt [12] y es ampliamente usado en la actualidad.
- c. El perceptrón multicapa es una red neuronal artificial (RNA) formada por múltiples capas, esto le permite resolver problemas que no son linealmente separables, lo cual es la principal limitación del perceptrón (también llamado perceptrón simple). Los resultados obtenidos en los experimentos se muestran en la siguiente sección.

4. Pruebas y resultados

En esta sección se describen características asociadas al corpus de entrenamiento, así como la metodología de evaluación y los resultados obtenidos.

4.1. Conjunto de datos

En la Tabla 1, se puede observar el número de muestras que han sido capturadas para cada uno de los símbolos considerados en el corpus de entrenamiento. El número mínimo de muestras es de 47 para la letra ‘m’, mientras que el máximo número de muestras fue de 96 para la letra ‘f’. La media de muestras es de 55.12. En total se capturaron 1,378 muestras.

Tabla 1. Cantidad de muestras por cada símbolo del alfabeto.

Letra	Muestras	Letra	Muestras	Letra	Muestras
a	58	i	51	q	50
b	51	j	48	r	65
c	57	k	48	s	66
d	51	l	49	t	56
e	50	m	47	u	53
f	96	n	59	v	51
g	51	o	56	w	51
h	48	p	66	x	50
				y	50

4.2. Metodología de evaluación

El proceso de evaluación considera el uso del corpus de entrenamiento para validar la exactitud en la identificación de las letras, usando los tres modelos de clasificación automática planteados con anterioridad.

Se divide cada conjunto de muestras de cada letra en 10 particiones y se ejecutan diez iteraciones usando un 90% de los datos para entrenamiento y el 10% restante para pruebas en un proceso denominado 10-fold cross-validation y leave-one out.

Los resultados obtenidos en los tres clasificadores, así como la discusión de dichos resultados se presentan en la siguiente sección.

4.3. Resultados obtenidos

En la Tabla 2 se muestran los resultados obtenidos para cada uno de los clasificadores. Como puede observarse, es el clasificador basado en perceptrón multicapa el que obtiene los mejores resultados, con una exactitud superior al 97%. De las 1,378 muestras clasificadas, solamente clasificó incorrectamente un total de 36 instancias, dando como resultado un error del 2.61%. De hecho, supera en 5 puntos porcentuales al clasificador SMO y en 8 puntos porcentuales al clasificador J48.

Estos resultados muestran que el grado de exactitud es elevado y suficiente para el proceso de clasificación de letras basado en lenguaje de señas.

Tabla 2. Comparativa de resultados obtenidos entre los tres clasificadores.

	J48		SMO		Perceptrón multicapa	
	Cantidad	Porcentaje	Cantidad	Porcentaje	Cantidad	Porcentaje
Instancias clasificadas						
Correctamente	1,227	89.04%	1,276	92.60%	1,342	97.39%
Incorrectamente	151	10.96%	102	7.40%	36	2.61%

Es necesario, sin embargo, realizar un análisis de los tiempos de ejecución necesarios por cada algoritmo para construir el modelo de clasificación, a fin de verificar su pertinencia de uso en sistemas de tiempo real. En la Tabla 3 se muestran dichos resultados.

Tabla 3. Comparativa de tiempos de construcción del modelo de clasificación.

	J48	SMO	Perceptrón multicapa
Tiempo (segundos)	0.19	2.45	17.86

Como puede observarse, el nivel de exactitud es inversamente proporcional a los tiempos de ejecución necesarios para construir el modelo. En realidad, los casi 18 segundos necesarios por el clasificador basado en perceptrón multicapa no resulta ser prohibitivo para construir un modelo de clasificación. De hecho, los tiempos de

evaluación de las instancias de prueba son en milésimas de segundos para cualquiera de los tres clasificadores probados.

5. Conclusiones

En este trabajo se presentó un guante basado en acelerómetros que permite el entrenamiento de personas sordomudas para escribir las letras del alfabeto en español. El tratamiento de los datos se realizó de manera estadística, permitiendo precisión en el proceso de clasificación y hace al sistema ser independiente del usuario y permite la detección de símbolos aun si éstos no tienen una forma perfecta.

Los experimentos realizados con tres métodos de clasificación automática muestran que la precisión obtenida en la identificación de los símbolos es mayor al 89%. En particular, el algoritmo basado en redes neuronales utilizando un perceptrón multicapa obtuvo el mejor resultado, con una exactitud del 97%. Como trabajo futuro se pretende incluir un sintetizador de voz para producir palabras tras detectar un conjunto válido de símbolos del alfabeto.

Agradecimientos. Los autores agradecen al PRODEP por el apoyo para la realización de este trabajo.

Referencias

1. Bui, T. D., Nguyen, L. T.: Recognizing Postures in Vietnamese Sign Language with MEMS Accelerometers. *IEEE Sensors Journal*, 7 (5), 707–712 (2007)
2. Starner, T., Pentland, A.: Real-time american sign language recognition from video using hidden Markov models. MIT Media Lab, Perceptual Computing Group, Cambridge, MA, Tech. Rep, 375, 265–270 (1995)
3. Mitra S., Acharya T.: Gesture recognition: A survey. *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.* 37 (3), 311–324 (2007)
4. Tan, U. X., et al.: Estimating Displacement of Periodic Motion with Inertial Sensors. *IEEE Sensors Journal*, 8 (8), 1385–1388 (2008)
5. Park, J.W., Hyun, S. D., Lee, C. W.: Real-time Finger Gesture Recognition. *HCI 2008, Korea*, 847–850 (2008)
6. Mäntyjärvi, J., Kela, J., Korpipää, P., Kallio, S.: Enabling fast and effortless customisation in accelerometer based gesture interaction. In: *Proc. 3rd Int. Conf. Mobile Ubiquitous Multimedia*, 25–31 (2004)
7. Hernandez-Rebollar, J. L., Lindeman, R. W., Kyriakopoulos N.: A Multi-Class Pattern Recognition System for Practical Finger Spelling Translation. In: *Proceedings of the Fourth IEEE International Conference on Multimodal Interfaces*, 185–190 (2002)
8. Ji-Hwan K., Tae-Seong K., Kyung H.: 3-D Hand Motion Tracking and Gesture Recognition Using a Data Glove. *IEEE International Symposium on Industrial Electronics (ISIE 2009)*, Seoul Olympic Parktel (2009)
9. Dipietro, L., Sabatini, A. M., Dario, P.: A Survey of Glove-Based Systems and Their Applications. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 38 (4), 461–482 (2008)
10. Swee T. T., et. Al.: Malay Sign Language Gesture Recognition system. *International Conference on Intelligent and Advanced Systems* (2007)

Jorge Cerezo Sánchez, Griselda Saldaña González, Mario Mauricio Bustillo Díaz, et al.

11. Zhang, X., et al.: A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors. *IEEE Transactions on Systems, Man, And Cybernetics-Part A: Systems and Humans*, 41(6), 1064–1076 (2011)
12. Platt, J.: Sequential Minimal Optimization: A Fast Algorithm for Training Support Vector Machines. *CiteSeerX* 10.1.1.43.4376 (1998)