

Desarrollo de un sistema de comunicación silenciosa dual basado en habla subvocal y Raspberry Pi

Dual silent communication system development based on subvocal speech and Raspberry Pi

Desenvolvimento de um sistema de comunicação silenciosa dual baseado em fala subvocal e Raspberry Pi

Fecha de recepción: 29 de febrero de 2016
Fecha de aprobación: 18 de abril de 2016

José Daniel Ramírez-Corzo*
Luis Enrique Mendoza**

Resumen

Presenta una metodología novedosa para establecer una comunicación silenciosa dual basada en habla subvocal; para ello se desarrollaron dos sistemas electrónicos que registran las señales bioeléctricas que llegan al aparato fonador, generadas al momento de realizar el proceso de lectura silenciosa por el individuo. Estos sistemas están basados en tres etapas fundamentales; la primera es la de *adquisición*, encargada de extraer, acondicionar, codificar y transmitir las señales electromiográficas del habla subvocal hacia la segunda etapa, denominada de *procesamiento*; en esta etapa, implementada en un sistema Raspberry Pi, se desarrollaron los procesos de almacenamiento, acondicionamiento, extracción de patrones y clasificación de palabras, utilizando técnicas matemáticas como: Entropía, análisis Wavelet y Máquinas de Soporte Vectorial de Mínimos Cuadrados, implementadas bajo el entorno libre de programación Python; finalmente, la última etapa del sistema se encargó de comunicar inalámbricamente los dos sistemas electrónicos, utilizando 4 clases de señales, para clasificar las palabras *hola*, *intruso*, *¿hola cómo estás?* y *tengo frío*.

Adicionalmente, en este artículo se muestra la implementación del sistema para el registro de señales de habla subvocal. El porcentaje de acierto promedio general es de 72.5 %. Se incluyen un total de 50 palabras por clase, es decir, 200 señales. Finalmente, se pudo demostrar que usando una Raspberry Pi es posible establecer un sistema de comunicación silenciosa a partir de las señales del habla subvocal.

Palabras clave: comunicación silenciosa; entropía; habla subvocal; MSV (Máquinas de Soporte Vectorial); Raspberry Pi; Wavelet.

* Universidad de Pamplona (Pamplona-Norte de Santander, Colombia). daniel.ramirez@unipamplona.edu.co.

** M. Sc. Universidad de Pamplona (Pamplona-Norte de Santander, Colombia). luis.mendoza@unipamplona.edu.co.

Abstract

This paper presents a novel methodology to develop a silent dual communication based on subvocal speech. Two electronic systems were developed for people's wireless communication. The system has 3 main stages. The first stage is the subvocal speech electromyographic signals *acquisition*, in charge to extract, condition, encode and transmit the system development. This signals were digitized and registered from the throat and sent to an embedded a raspberry pi.

In this device was implemented the processing, as it is called the second stage, which besides to store, assumes conditioning, extraction and pattern classification of subvocal speech signals. Mathematical techniques were used as Entropy, Wavelet analysis, Minimal Squares and Vector Support Machines, which were applied in Python free environment program. Finally, in the last stage in charge to communicate by wireless means, were developed the two electronic systems, by using 4 signal types, to classify the words: *Hello, intruder; hello how are you? and I am cold* to perform the silent communication.

Additionally, in this article we show the speech subvocal signals' recording system realization. The average accuracy percentage was 72.5 %, and includes a total of 50 words by class, this is 200 signals. Finally, it demonstrated that using the Raspberry Pi it is possible to set a silent communication system, using subvocal. speech signals.

Keywords: entropy; Raspberry Pi; silent communication; SVM (Support Vector Machines), subvocal speech, Wavelet.

Resumo

Apresenta-se uma metodologia nova para estabelecer uma comunicação silenciosa dual baseada em fala subvocal; para isso, foram desenvolvidos dois sistemas eletrônicos que registram os sinais bioelétricos que chegam ao aparelho fonador, gerados no momento de realizar o processo de leitura silenciosa pelo indivíduo. Estes sistemas estão baseados em três etapas fundamentais; a primeira é a de aquisição, encarregada de extrair, acondicionar, codificar e transmitir os sinais eletromiográficos da fala subvocal até a segunda etapa, denominada de processamento; nesta etapa, implementada em um sistema Raspberry Pi, desenvolveram-se os processos de armazenamento, acondicionamento, extração de padrões e classificação de palavras, utilizando técnicas matemáticas como: Entropia, análise Wavelet e Máquinas de Suporte Vetorial de Mínimos Quadrados, implementadas sob o ambiente livre de programação Python; finalmente, a última etapa do sistema se encarregou de comunicar com tecnologia wireless os dois sistemas eletrônicos, utilizando 4 classes de sinais, para classificar as palavras: *olá, intruso, olá como estás? e tenho frio.*

Adicionalmente, neste artigo mostra-se a implementação do sistema para o registro de sinais de fala subvocal. A porcentagem de acerto médio geral é de 72.5 %. Incluem-se um total de 50 palavras por classe, ou seja, 200 sinais. Finalmente, pôde-se demonstrar que usando uma Raspberry Pi é possível estabelecer um sistema de comunicação silenciosa a partir dos sinais de fala subvocal.

Palavras chave: comunicação silenciosa; entropia; fala subvocal; MSV (Máquinas de Suporte Vetorial); Raspberry Pi; Wavelet.

I. INTRODUCCIÓN

Actualmente, el estudio del Habla Subvocal (HSV) ha venido creciendo de manera abrumadora, debido a las grandes aplicaciones que se han logrado desarrollar en áreas como la ingeniería, las comunicaciones militares y el sector salud, entre otras [1-3]. El habla subvocal se define como el registro y estudio de las señales electromiográficas que se producen a partir de las vibraciones en las cuerdas vocales, movimientos de la garganta y la lengua, generadas durante el proceso de lectura silenciosa por el individuo [1]. Estas señales, en su representación física, son una mezcla de señales bioeléctricas de diferentes fuentes, en las cuales se encuentra inmersa información relevante y no relevante relacionada con palabras que se perciben auditivamente; es decir que, el habla subvocal se registra sin la producción física del sonido.

Una configuración diferencial de electrodos conectados en la superficie de la piel a la altura de la garganta, se encarga de captar las señales electromiográficas que se generan en la corteza cerebral, viajando por el sistema nervioso central hasta llegar al aparato fonador. Estas señales representan la intención del habla por parte del individuo antes de ser articuladas en voz alta [1-4]. La información contenida en estas señales es tan alta que con un adecuado tratamiento de estas se puede lograr que se reproduzca la voz sin que la persona hable, es decir, que el habla subvocal es la codificación de los pensamientos que llegan a las cuerdas vocales a través de las señales nerviosas y luego se convierten así en sonidos, como es mencionado en [5].

En la actualidad, el habla subvocal está siendo enfocada en aplicaciones en el sector salud, permitiendo la comunicación de personas con dificultades en la producción de señales audiblemente. Así mismo, el HSV puede servir en un futuro como guía para análisis en fonoaudiología, tales como tragar, deglutir y respirar. Esto evidencia el gran campo que puede abordar el habla subvocal. Son pocos los trabajos que se han desarrollado usando HSV, ya que los sistemas de registro de las señales deben ser de gran sensibilidad, debido a que las características de amplitud y frecuencias oscilan alrededor de $\pm 2 \mu\text{V}$ y 500 Hz de frecuencia máxima [6]. Otro punto importante es que las señales son contaminadas con facilidad por ruidos externos, ya que cualquier movimiento de las partes cercanas a los electrodos ubicados en la zona de la garganta hacen que se produzca una respuesta en el

sistema; por esto es importante eliminar información no relevante, aplicando filtrados analógicos de orden elevado.

En cuanto a los desarrollos que han presentado otros autores, se destacan [1, 7-11], que registran las señales del habla subvocal para controlar computadores, sillas de ruedas y escribir. En estos trabajos se han aplicado diferentes técnicas matemáticas, como el análisis de Fourier y la transformada discreta del coseno, en donde se buscan las componentes de frecuencia más representativas de cada grupo de datos. Sin embargo, los resultados conseguidos permiten concluir que es de gran importancia mejorar los resultados [12]. Por otro lado, se han utilizado técnicas de inteligencia artificial, como redes neuronales, para clasificar un conjunto de datos. Existen otros trabajos que presentan un estado del arte de habla subvocal muy interesantes [12], y claramente se evidencia que son pocos los aportes encontrados.

Finalmente, este artículo presenta un aporte significativo y novedoso en el cual se lograron implementar, en un sistema embebido Raspberry Pi (sistema embebido capaz de procesar información con la misma capacidad de una computadora) y bajo programación Python, técnicas de procesamiento basadas en el plano tiempo- frecuencias, extracción de patrones y de clasificación, como la transformada Wavelet discreta, que es una técnica de análisis de multiresolución que permite descomponer una señal en múltiples señales de diferentes frecuencias; Entropía de Shannon, que permite obtener un valor de la importancia que tienen ciertos conjuntos de datos con respecto a otros, y Máquinas de Vector Soporte de Mínimos Cuadrados, que es una técnica que permite clasificar un número de grupos a partir de un hiperplano y los vectores soporte conseguidos. Todo el desarrollo realizado permitió comunicar dos personas de forma inalámbricamente a partir de las señales del HSV.

II. METODOLOGÍA

El objetivo principal para establecer una comunicación silenciosa dual entre dos personas se basa en el desarrollo de dos sistemas electrónicos, encargados de adquirir, acondicionar, codificar, procesar, clasificar y transmitir de forma bidireccional los patrones extraídos a partir de las señales del HSV registradas. Este objetivo sugiere tratar de cambiar el concepto

de comunicación auditiva empleado hoy día, por un tipo de comunicación en la que no sea necesaria la producción física del sonido, lo cual sería un reto de gran envergadura. El propósito es realizar la extracción de los patrones más relevantes del habla a partir de las señales y lograr un sistema de comunicación “telepática” que pueda más adelante solventar algunos problemas de comunicación a personas en entornos ruidosos o en personas que tengan problemas en sus cuerdas vocales. De esta manera, la comunicación silenciosa basada en habla subvocal en este artículo se resume en el registro, procesamiento, extracción de patrones y la clasificación de los patrones extraídos, usando un sistema electrónico que es capaz de traducir las respuestas de clasificación en sonidos que fueron previamente registrados en el sistema.

A. Adquisición del habla subvocal

Para la adquisición de las señales de habla subvocal se usó electromiografía no invasiva, estableciendo la zona de la superficie del cuello para la ubicación de los electrodos; además, se usó un electrodo de referencia ubicado en el hueso mastoideo, detrás de la oreja, como se observa en la Figura 1. Dicha zona se acondicionó previamente, rasurando y limpiando con alcohol isopropílico al 70 %, con el fin de reducir su impedancia y así mejorar el proceso de adquisición. Se usó una configuración diferencial de electrodos de plata cloruro de plata $-Ag/AgCl-$, en la parte inferior de la garganta [1, 11, 13].



FIG. 1. Posicionamiento de los electrodos para la detección del habla subvocal.

Definidas las zonas de detección y ubicados los electrodos, se procedió al registro de las señales electromiográficas, usando el sistema de adquisición desarrollado, que contempla, básicamente, tres etapas fundamentales. La primera de ellas se encarga de amplificar las señales provenientes de los electrodos; se implementó el circuito de amplificación de señales electromiográficas EMG propuesto por el fabricante del amplificador de instrumentación INA 128 [14]; este circuito toma las dos señales proveniente de los electrodos y después amplifica su diferencia; es decir, cualquier señal que se origina a gran distancia del sitio de detección aparecerá como una señal común, mientras que las señales cercanas a la superficie de detección son diferenciadas y amplificadas a un factor de 1000. Las señales obtenidas de la etapa de amplificación fueron acondicionadas debido a que los electrodos captan una mezcla de señales biológicas no deseadas, como, por ejemplo, señales de electrocardiográficas ECG y respiración; entonces, dependiendo de la ubicación de los electrodos se podrían encontrar señales que hacen parte del HSV;

del mismo modo, se puede presentar ruido proveniente del movimiento de los cables, entre otros [6]; por estas razones, para lograr registros limpios de señales del habla subvocal es necesario realizar un proceso de filtrado. En la segunda etapa del sistema se realizó un filtro analógico activo Butterworth para banda de octavo orden que atenuó las frecuencias de la señal que estuvieran por debajo de 30 Hz y superiores a 500 Hz [6]. Finalmente, la tercera etapa del sistema se encarga de codificar la señal proveniente del filtro analógico, teniendo como referencia el teorema de muestreo de Nyquist. En el caso de los músculos que se relacionan con el HSV, la frecuencia con el armónico más grande para electromiografía de superficie está en el rango de 400-450 Hz [6]. Se utilizó el convertor análogo digital de referencia ADS7813 el cual codificó las señales a una frecuencia de muestreo de 2 kHz [15]. Este convertor cuenta en su estructura interna con el protocolo de comunicación Serial Peripheral Interface SPI, que permitió transferir muestras codificadas de 16 bits de resolución hacia el sistema embebido Raspberry pi.

B. Sistema Raspberry Pi

En la Raspberry PI se realizó la adquisición, almacenamiento, acondicionamiento y extracción de patrones, así como el proceso de clasificación y el enlace de comunicación dual. La información enviada por cada una de las tarjetas contiene el resultado del proceso de clasificación según la señal de entrada. Todo el cálculo y el tiempo que se requieren para llevar a cabo estos procedimientos se logran porque la Raspberry pi es un sistema embebido diseñado para realizar funciones que permiten aportar inteligencia empleando procesadores digitales (CPU) en formato microprocesador, microcontrolador o procesador digital de señales DSP [16]. La tarjeta Raspberry PI es un dispositivo que cuenta con todas las características de un miniordenador personal que proporciona un procesador ARM, memoria RAM DDR, una GPU, puertos USB/UART y puertos de entradas y salidas digitales GPIO, audio, salida de video HDMI y composite y slot para tarjeta SD, y tiene un entorno de propósito general de programación de alto nivel, denominado Python, el cual cuenta con varias características generales que lo hacen especial, por ser fácil de leer y simple de implementar, además de ser un código abierto (de libre uso) [16]. Los scripts de Python pueden ser procesados y ejecutados en tiempo real. Así, este sistema es idóneo para desarrollar todos los procesos necesarios para la extracción de patrones del habla subvocal.

C. Acondicionamiento y extracción de patrones

Las señales provenientes del sistema de adquisición se almacenaron en vectores de longitud de 20.000 puntos en la Raspberry pi; en la Figura 2a se observa la señal EMG registrada en el sistema. Posteriormente,

se estableció una metodología de programación en Python para el acondicionamiento y extracción de patrones característicos; en primera instancia, se identificó la zona en donde se encuentra la mayor cantidad de energía de la señal, es decir, la zona donde se encuentra la información más relevante; para ello se calculó el valor de entropía, ecuación (1), en pequeñas ventanas de longitud fija de 100 puntos S_i de la señal original, esto con el fin, de determinar la parte de la energía relevante que aporta mayor información, y así identificar la zona de interés. Los datos obtenidos de este análisis se almacenaron en un nuevo vector de longitud de 200 puntos. En la Figura 2b se observa el resultado de este análisis.

$$E = \sum_{i=1}^n S_i^2 \times \text{LOG}_2 S_i^2 \quad (1)$$

$$S_i = X_I \times \varphi(n) \quad (2)$$

Donde:

- $\varphi(n) = A(t - k_i)$ Función escalón unitario de amplitud uno.
- X_I = Señal electromiográfica
- S_i = Ventana

La duración de $\varphi(n)$ queda definida como:

$$k_1^i \leq t \leq k_2^i$$

- $k_1^i = \{0, 100, 200, 300, \dots, L-100\}$
- $k_2^i = \{100, 200, 300, 400, \dots, L\}$
- $i = 1, 2, 3, 4, \dots, L/n, L$ (número de puntos de la señal)

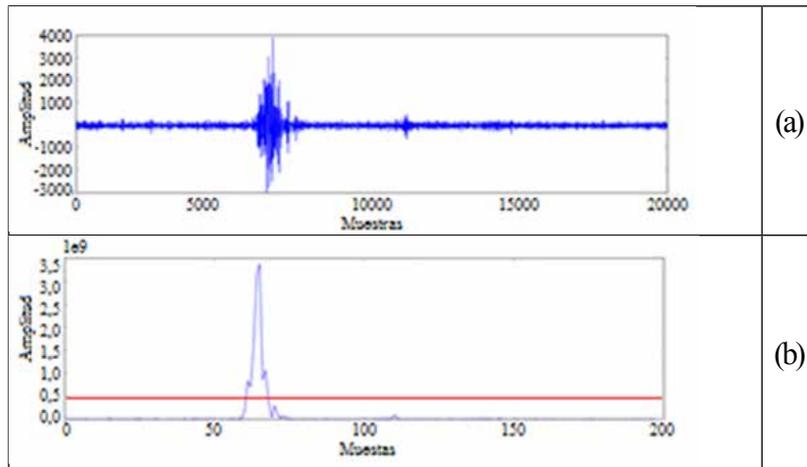


FIG. 2. (a) Señal electromiográfica; (b) Señal de entropía.

El lóbulo que se observa en la Figura 2b indica en dónde se encuentra la mayor concentración de energía de la señal. Para identificar la zona de interés en la señal de entropía, se trazó un umbral para hallar el inicio de la actividad subvocal. Nótese que la zona de actividad pasiva de la señal no supera la escala de los 0,5 de amplitud; por ende, este fue el umbral seleccionado y se define a partir de la ecuación (3). Identificado el punto de inicio de la actividad subvocal del vector de entropía, se escaló en un factor de 100, y a partir de este nuevo valor, se buscó identificar el punto de inicio de la zona de interés en la señal original, a partir del cual se recortó una ventana de longitud de 8000 puntos, donde se encuentra la mayor información relevante de la señal EMG.

$$U = (0.5) \times A_{max} \quad (3)$$

Donde:

- U = Umbral de detección
- A_{max} = Pico máximo del vector de entropía de la señal EMG

La zona de interés compuesta de 8000 muestras se filtró digitalmente, haciendo uso de la Transformada Wavelet Discreta (TWD), ecuación (4) [17]. Se empleó una Wavelet madre Daubechies ‘db3’, con un primer nivel de descomposición, en el cual se seleccionó el coeficiente de aproximación, con el propósito de eliminar componentes de frecuencias altas. La representación matemática de la TWD, se observa en la ecuación (4).

$$C_{a,b} = \sum_{n=a}^b X(n) \varphi \left[\frac{n-a}{b} \right] \quad (4)$$

Donde:

- $X(n)$ = Señal original
- $\varphi \left[\frac{n-a}{b} \right]$ = Wavelet madre
- a = Coeficiente de traslación
- b = Coeficiente de escalado

Al coeficiente de aproximación Wavelet se le determinó de nuevo el valor de la entropía, en pequeñas ventanas de 200 puntos S_k . La nueva señal de entropía se almacenó en vector de 18 muestras, como observa en la Figura 3.

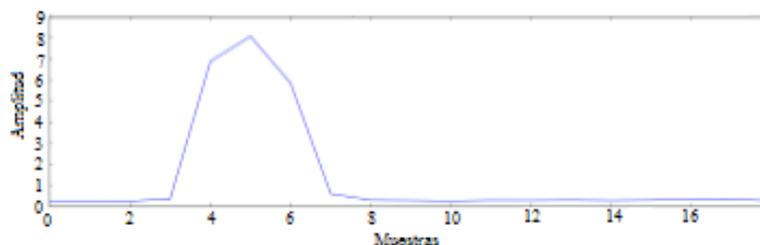


FIG. 3. Análisis entropía de la zona de interés.

Finalmente, se normalizó cada una de las señales, para así obtener una señal de valor máximo unitario de amplitud. El proceso de normalización se define como:

$$V_n = X_i + abs(\max(x)) \quad (5)$$

Donde:

- $\max(x)$ = Valor máximo de la señal de entropía de la señal
- V_n = Vector de entropía normalizado
- X_i = Señal de entropía sin normalizar

D. Clasificación de los patrones de las señales subvocales

Finalizados los procesos de acondicionamiento y extracción de las características de las señales del habla subvocal, se realizó el proceso de clasificación, que permite decidir a qué clase pertenece una nueva señal que ingrese al sistema; esto se logró entrenando una Máquina de Soporte Vectorial (MSV) en la Raspberry Pi. La MSV es una técnica de clasificación, en este trabajo se usó un entrenamiento supervisado [10]. La ecuación (7) modela el proceso de clasificación de la SVM.

$$f(x) = \text{sign} \left(\sum_{i=1}^m y_i \alpha_i (x \cdot x_i) + b \right) \quad (7)$$

Donde:

- $(x \cdot x_i)$ = Hyperplano óptimo Kernel
- y_i = Etiquetas de los grupos
- α_i = Coeficiente de Lagrange
- b = Coeficiente de normalización

Para la matriz de entrenamiento se registró una base de datos a partir de las señales normalizadas de entropía a las palabras ¿Hola cómo estás?, Hola, Intruso y Tengo frío; en la Figura 4 se observan los patrones para cada una de estas clases. La base de datos se conformó de 5 señales para cada patrón, y cada patrón se compuso de 18 muestras. La matriz de entrenamiento se compone de 20 filas (señales) y 18 columnas (muestras de cada señal). A cada clase de entrenamiento se le asignó su respectiva etiqueta, generando un vector de longitud 20. Almacenada la matriz de entrenamiento y el vector

de etiquetas en la Raspberry pi, se entrena la SVM con el 100 % de los datos generados, para finalmente ingresar nuevas señales al sistema en tiempo real y así determinar el porcentaje de acierto de clasificación.

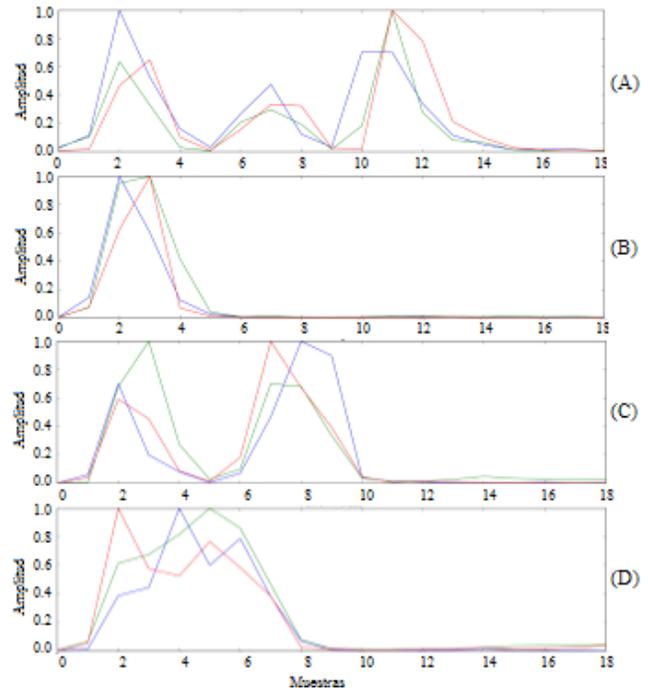


FIG. 4. Patrones de entrenamiento de las SVM para hola cómo está (A), hola (B), tengo frío (C) e intruso (D).

E. Comunicación inalámbrica

Una vez realizado el proceso de clasificación, se procedió a establecer la comunicación inalámbrica punto a punto entre los dos sistemas, con ayuda de tecnología Zigbee, gracias a sus dos módulos XBee de la serie 2; estos fueron configurados en modo de operación transparente, para establecer el enlace de comunicación entre los dos sistemas. La Raspberry Pi del primer sistema envía al puerto serial UART el patrón; usando el módulo de transmisión XBee, se almacena en el buffer de entrada; luego, se transmite a una banda libre de radio frecuencia RF de 2.4GHz hacia el otro módulo XBee del segundo sistema. El módulo receptor captura y almacena en el buffer de salida el paquete RF recibido y lo envía a la Raspberry Pi del segundo sistema, el cual se encargará de decodificar dicha información. El modo Transparente viene por defecto en los módulos XBee para la comunicación punto a punto. Tanto para la transmisión como para la recepción de datos con los módulos Xbee se

deben configurar parámetros como la velocidad de transmisión de 9600 Baud a 8 bits de datos sin bit de paridad y un bit de parada.

F. Decodificación

Enviada y entregada inalámbicamente al segundo sistema la clase a la que pertenece la señal registrada, se almacena, para ser comparada con tablas preestablecidas que muestran las señales de audio previamente grabadas. En este trabajo se utilizaron 4 clases: la primera es para Tengo frío; la segunda, para ¿Hola cómo estás?; la tercera, para Intruso, y, finalmente, para Hola; es decir, el sistema permite, según el patrón recibido, generar una respuesta auditiva reproducida por la salida de audio de la Raspberry Pi.

III. RESULTADOS Y DISCUSIÓN

En el sistema de adquisición de las señales del habla subvocal implementado se encuentran condensadas las etapas de adquisición y amplificación, filtrado analógico, conversión analógica a digital y comunicación SPI para el intercambio de datos entre la tarjeta y la Raspberry Pi.

En cuanto a la etapa de procesamiento, se entrenó la MSV utilizando la base de datos de 20 señales con 5 señales de cada clase. El test de efectividad se realizó ingresando en tiempo real 50 señales subvocales para cada uno de los patrones seleccionados, es decir, 200

señales para el proceso de validación. A continuación, se menciona cada una de las técnicas utilizadas para el proceso de validación.

- Señal cruda (sc): Se entrenó la MSV con una matriz de 5 señales de cada palabra con longitud de cada una de 20 000 puntos.
- Señal normalizada (n): Se entrenó la SVM con una matriz de 5 señales de cada palabra con normalizando la señal sc.
- Zona de interés (z1): Se entrenó la SVM con una matriz de 5 señales de cada palabra con la zona de interés normalizada de sc. Con una longitud de 8000 puntos por cada una.
- Wavelet primer coeficiente de aproximación (w1): Se entrenó la SVM con una matriz de 5 señales de cada palabra aplicando w1 a la señal z1 logrando una longitud de 4000 puntos por cada palabra.
- Entropía cada 400 puntos (e400): Se entrenó la SVM con una matriz de 5 señales de cada palabra aplicando entropía a cada 400 puntos de la señal w1.

En la Tabla 1 se observan los porcentajes de error obtenidos en el proceso de clasificación, haciendo uso de las técnicas mencionadas anteriormente.

TABLA 1
RESUMEN PORCENTAJE DE ACIERTO DE CADA TÉCNICA DE PROCESAMIENTO

	Técnica	Señal	% de acierto	Cantidad de datos
1	SC	Hola	0%	50
		Hola cómo estás	0%	
		Tengo frío	0%	
		Intruso	100%	
2	SC + N	Hola	100%	
		Hola cómo estás	0%	
		Tengo frío	0%	
		Intruso	0%	
3	SC + ZI + N	Hola	100%	
		Hola cómo estás	0%	
		Tengo frío	0%	
		Intruso	4%	
4	SC + ZI + W1 + N	Hola	100%	
		Hola cómo estás	0%	
		Tengo frío	6%	
		Intruso	5%	
5	SC + ZI + W1 + E400 + N	Hola	100%	
		Hola cómo estás	70%	
		Tengo frío	65%	
		Intruso	80%	

Como se observa en la Tabla 1, los mejores porcentajes de clasificación se lograron aplicando wavelet a la zona de interés de las señales, seleccionando el coeficiente de detalle, análisis de entropía y normalización. La efectividad del algoritmo de clasificación varía dependiendo de factores como la disposición de los electrodos, la humedad relativa, la calidad de los sensores, la preparación de la piel en el sitio de registro y de la persona testada [8].

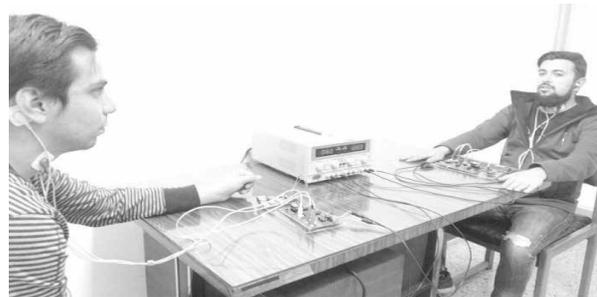


FIG. 5. Validación del sistema de comunicación silenciosa.

A. Validación sistema de comunicación

La prueba de comunicación silenciosa se realizó con dos personas de sexo masculino, de 25 y 24 años de edad, en sesiones de registro durante el transcurso de un día y con situaciones de contaminación auditiva normal. El montaje del sistema de comunicación se observa en la Figura 5.

Es importante destacar que el sistema debe calibrarse si se desea utilizar con otros pacientes, ya que las características electromiográficas de cada sujeto varían dependiendo de su fisiología. El porcentaje de clasificación para cada una de las personas que hacen parte del proceso de comunicación se observa en la Figura 6. La línea continua hace referencia a la persona 1, y la línea punteada a la persona 2.

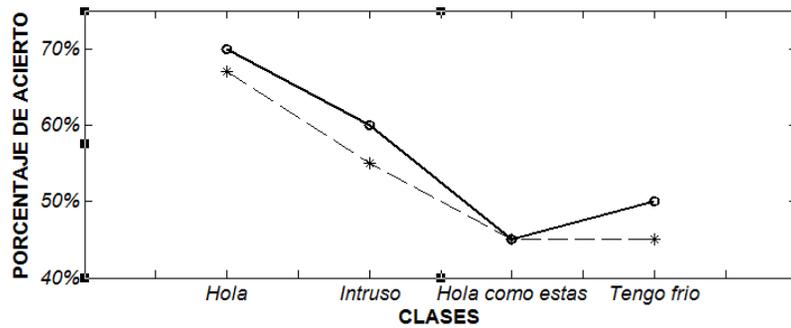


FIG. 6. Resultados de efectividad.

La Tabla 2 resume la efectividad de las pruebas, realizando la comunicación inalámbrica de los patrones característicos entre los dos sistemas desarrollados; el

porcentaje de efectividad es calculado a partir de la ecuación (8).

TABLA 2
EFECTIVIDAD PORCENTUAL DEL ACIERTO DE CADA UNA DE LAS CLASES

EFECTIVIDAD				
Palabra	Hola (1)	Intruso (2)	Hola cómo estás (3)	Tengo frío (4)
% efectividad Persona 1	100%	80%	50%	60%
% efectividad Persona 2	94%	70%	50%	50%

$$P(\%) = \left(\frac{N_a}{N_m} \right) \times 100 \quad (8)$$

Donde:

- $P(\%)$ = Porcentaje de clasificación
- N_a = Número de aciertos
- N_m = Número de muestras

El promedio de clasificación para validación del sistema de comunicación de la persona 1 obtuvo un acierto del 72.5 %; para la persona 2 se obtuvo un acierto del 66 %, lográndose demostrar así que se pueden procesar señales EMG usando sistemas embebidos, como Raspberry Pi.

IV. CONCLUSIONES

El habla subvocal demuestra ser una herramienta con mucho potencial, ya que las respectivas señales no solo pueden utilizarse para comunicaciones silenciosas, sino para otras aplicaciones, como control de sistemas mecatrónicos, comunicaciones espaciales

y aplicaciones en salud. Los resultados demostraron que es posible comunicar inalámbricamente dos personas usando señales de habla subvocal con un sistema embebido de bajo costo. Aplicando las diferentes etapas planteadas, se logró un porcentaje de clasificación de 72.5 %, que muestra que es importante mejorar el proceso de clasificación, además que el rendimiento de la tarjeta permitió evidenciar que es posible aumentar el número de palabras en el sistema de comunicación. Por otro lado, desarrollar la aplicación bajo el entorno de programación python permite tener un sistema multiplataforma que logra migrar a otros sistemas embebidos con las mismas características originales.

Finalmente, se presentan resultados novedosos, que serán usados para nuevas investigaciones, para de esta manera desarrollar aplicaciones enfocadas a problemas de afasia motrices. Así, se pudo demostrar que es posible realizar un sistema de comunicación dual (dos personas) y que en futuras investigaciones puede ser aplicado a personas con problemas del habla, de movilidad, y ofrecer una oportunidad de comunicación en personas con problemas de afasias motrices.

En cuanto a las herramientas matemáticas, se logró combinar máquinas de aprendizaje con wavelet y entropía, buscando el patrón más específico para la clasificación de las palabras; es decir, la extracción de patrones fue fundamental para el desarrollo de la aplicación.

AGRADECIMIENTOS

Al grupo de Ingeniería Biomédica, de la Universidad de Pamplona (GIBUP); a su Semillero de investigación de Ingeniería Biomédica y Telecomunicaciones (SIIBTEL), y al ing. German Arley Portilla y los docentes Hernando José Velandia y Luis Muñoz Bedoya, quienes contribuyeron en forma significativa a la elaboración de este artículo.

REFERENCIAS

- [1] C. Jorgensen and K. Binsted, "Web Browser Control Using EMG Based Sub Vocal Speech Recognition," in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, p. 294c, Jan. 2005. DOI: <http://dx.doi.org/10.1109/HICSS.2005.683>.
- [2] J. Mendes, R. Robson, S. Labidi, and A. Barros, "Subvocal Speech Recognition Based on EMG Signal Using Independent Component Analysis and Neural Network MLP," in *Congress on Image and Signal Processing*, vol. 1, pp. 221-224, May. 2008. DOI: <http://dx.doi.org/10.1109/cisp.2008.741>.
- [3] H. Curtis and L. F. Petrinovich, "Treatment of Subvo-cal Speech During Reading," *Journal of Reading*, vol. 12 (5), pp. 361-368, Feb. 1969.
- [4] E. N. Gamma, D. Amaya, and O. L. Ramos, "Revisión de las tecnologías y aplicaciones del habla subvocal," *Ingeniería*, vol. 20 (2), pp. 277-288, 2015. DOI: <http://dx.doi.org/10.14483/udistrital.jour.reveng.2015.2.a07>.
- [5] J. A. Gutiérrez, E. N. Gamma, D. Amaya, and O. F. Avilés, "Desarrollo de interfaces para la detección del habla sub-vocal," *Tecnura*, vol. 17(37), pp. 138 - 152, Jul. 2013. DOI: <http://dx.doi.org/10.14483/udistrital.jour.tecnura.2013.3.a12>.
- [6] R. Merletti and P. Philip A, *Electromyography Physiology, Engineering, and Noninvasive Applications*, 2004.
- [7] J. Peña Rodriguez and L. E. Mendoza, *Adquisición y procesamiento de señales electromiográficas basadas en habla subvocal*, Pamplona, Dec. 2010.
- [8] L. E. Mendoza, J. Peña, L. A. Muñoz-Bedoya, and H. J. Velandia-Villamizar, "Procesamiento de señales provenientes del habla subvocal usando Wavelet Packet y Redes Neuronales," *Tecno. Lógicas*, vol. Edición Especial 2013, pp. 655-667, Oct. 2013.
- [9] I. Ishii, S. Takemoto, T. Takaki, M. Takamoto, K. Imon, and K. Hirakawa, "Real-time laryngoscopic measurements of vocal-fold vibration," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 6623-6626, Aug. 2011. DOI: <http://dx.doi.org/10.1109/IEMBS.2011.6091633>.
- [10] G. Chau and G. Kemper, "One Channel Subvocal Speech Phrases Recognition Using Cumulative Residual Entropy and Support Vector Machines," *IEEE Latin American Transactions*, vol. 13 (7), pp. 2135-2143, Jul. 2015. DOI: <http://dx.doi.org/10.1109/TLA.2015.7273769>.
- [11] C. Jorgensen, D. D. Lee, and S. Agabont, "Sub Auditory Speech Recognition Based on EMG Signals," *Proceedings of the International Joint Conference on Neural Networks*, vol. 4, pp. 3128-3133, Jul. 2003. DOI: <http://dx.doi.org/10.1109/ijcnn.2003.1224072>.
- [12] L. E. Mendoza, J. Peña, and J. L. Ramón, "Electro-myographic patterns of sub-vocal Speech: Records and classification," *Revista de Tecnología*, vol. 12 (2), pp. 35-41, Jul. 2013.
- [13] SENIAM, [en línea]. Available: <http://www.seniam.org/>. [acceso: 09/02/2016].
- [14] Texas Instrument, [en línea]. Available: <http://www.ti.com/lit/ds/symlink/ina128.pdf>. [acceso: 23/09/2015].
- [15] TEXAS INSTRUMENT, [en línea]. Available: <http://www.ti.com/lit/ds/symlink/ads7813.pdf>. [acceso: 23/09/2015].
- [16] Raspberry Pi, [en línea]. Available: <https://www.raspberrypi.org/education/>. [acceso: 01/01/2014].
- [17] C. Sidney, R. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms: A Primer*, Pearson, 1997.