

Prototipo embebido de la interpretación de patrones de movimiento en una mano para la reproducción de fonemas y palabras completas en el idioma Español de México

Trabajo Terminal No. ____ - ____

Alumnos: *Valle Martínez Luis Eduardo

Directores: Rodolfo Romero Herrera, Dr. Jesús Yaljá Montiel Pérez

*e-mail: lvallem1400@alumno.ipn.mx

Resumen - Creación de un sistema embebido basado en RaspberryPi que utiliza el SoC Micro Bit de la BBC como dispositivo sensor, y la tarjeta de sonido adaptada para sistemas RaspberryPi, WM8960 de WaveShare, para la reproducción de los fonemas. La tarjeta Micro Bit se coloca sobre el dorso de la mano permitiendo identificar los movimientos realizados con esta al aprovecharse el sensor de acelerómetro integrado, mientras la transmisión de su recopilación se logra mediante conexión Bluetooth. Los datos adquiridos son procesados con un algoritmo de clasificación basado en *Deep Learning* y DTW(*Dynamic Time Warping*), que permitiera calcular la coincidencia óptima entre la secuencia recopilada y las secuencias modelos correspondientes a cada fonema para su correcta identificación y finalmente reproducción sonora.

Palabras clave - Algoritmo DTW(*Dynamic Time Warping*), *Deep Learning*, Reproducción de voz, Sistemas Embebidos

1. Introducción

El alfabeto fonético internacional es un sistema de notación fonética creado por lingüistas con el propósito de establecer de forma regularizada, precisa y única, la representación de los sonidos del habla de cualquier lengua. Bajo este esquema alfabético, el español se caracteriza por contar con 22 fonemas descritos en este, y que se dividen en 19 consonánticos y 5 vocálicos.

Los 5 vocálicos corresponden a los fonemas /a/, /e/, /i/, /o/, /u/.

Los fonemas consonánticos a pesar de contabilizarse 19, en realidad son utilizados en el español de México únicamente 17, siendo estos 2 fonemas de uso común únicamente en regiones específicas: /θ/ (también comunemente utilizado en el idioma inglés) y /ʎ/ (también comunemente utilizado en el idioma italiano).

Fonema	Orden	Ortográfico
/m/	Nasal Bilabial Sonora	m
/n/	Nasal Alveolar Sonora	n
/ɲ/	Nasal Palatar Sonora	ñ
/p/	Oclusiva Bilabial Sorda	p
/b/	Oclusiva Bilabial Sonora	b
/t/	Oclusiva Dental Sorda	t
/d/	Oclusiva Dental Sonora	d
/k/	Oclusiva Velar Sorda	c+a,o,u ; qu+e,i ; k
/g/	Oclusiva Velar Sonora	g+a,o,u ; gu+e,i
/f/	Fricativa Labiodental Sorda	f
/s/	Fricativa Alveolar Sorda	s
/ʃ/	Fricativa Palatal Sonora	y ; hi+vocal ; ll
/x/	Fricativa Velar Sorda	j+a,e,i,o,u ; g+e,i
/tʃ/	Africada Postalveolar Sorda	ch
/r/	Vibrante Múltiple Alveolar Sonora	rr ; r-(posición inicial o precedida de l, n, s)
/r/	Vibrante Simple Alveolar Sonora	r(no inicial ni precedida de n, l, s)
/l/	Aproximante Lateral Alveolar Sonora	l

Figura 1: Los 17 fonemas básicos considerados en la implementación del sistema para su reproducción sonora

El idioma español a su vez como lenguaje escrito, cuenta con un total de 27 símbolos alfabéticos utilizados para la escritura de las palabras, además de 3 compuestos por 2 letras combinadas (rr, ch, ll) que también cuentan con un fonema específico asociado. Aún cuando no es homogéneamente aplicable la asociación de un fonema a cada letra, al menos 11 de los fonemas totales usados en español, tienen una correspondencia directa con una letra y el ch, cuando estas son pronunciadas oralmente.

Aunado a estas correspondencias, se hace notar que el símbolo h, letra muda en español, no requiere representación como patrón dentro del sencillo lenguaje motriz que se desarrollará, así como también se omitirá la asociación directa de algunas letras por múltiple fonética contextual a la posición y combinación de otras letras en las palabras.

De forma intuitiva para la mayor parte de la población (cerca del 95 % de la población que es alfabeta en México) la asociación de la pronunciación de las letras en las palabras, es natural mediante la relación en su representación escrita, siendo mucho más complejo un desglose directo en los fonemas que la compone, inclusive podría describirse como imposible por el desconocimiento del colectivo general de una representación estandarizada de los sonidos, marco que es propuesto por el Alfabeto Fonético Internacional. Con este hecho en mente, la ortografía de las palabras será utilizada como interfaz para describir la palabra, o directamente los sonidos, que el usuario desea expresar.

La creación de un sencillo lenguaje codificado en un contexto de patrones de movimiento, le servirá al usuario para indicar el sonido que desea sea reproducido.

Probablemente el referente más cercano e inmediato a este sencillo lenguaje con el que podría ser comparado, es el LSM (Lenguaje de Señas Mexicano), sin embargo, y aún cuando este trabajo tiene el objetivo común de permitir la comunicación principalmente de una comunidad con impedimento del habla, difiere primordialmente en algunos aspectos:

- El LSM es un lenguaje que requiere para su expresión el involucramiento de gestos con movimiento desde encima de la cabeza y hasta debajo de la cadera, incluyendo movimiento de manos, expresiones faciales y mirada intencional. Por su parte el lenguaje motriz que se propondrá se limita únicamente a 1 mano, donde se colocará el SoC para el sensado.
- La interpretación de ambos lenguajes se realiza en dominios sensoriales humanos distintos. Completamente visual para el LSM, y parcialmente visual para el lenguaje propuesto, pero principalmente auditivo.
- Otra diferencia importante se encuentra en la complejidad y relación con el idioma español. Mientras el LSM cuenta con una rica y compleja gramática, y vocabulario abundante, también existe desacoplado de la estructura, gramática y normas propias del español. El enfoque principal del lenguaje que se propondrá, es la expresión de palabras y letras del español, sobrepasando sus capacidades la formulación de normas gramaticales o vocabulario propio, fungiendo únicamente como un simplificado codificado motriz para la expresión sonora del español.
- Finalmente su implementación, aún cuando es común para las personas con impedimento del habla u afecciones relacionadas a la expresión oral, el LSM brinda mayores posibilidades para la población sordo-muda. Mientras tanto el prototipo de este trabajo podrá ser poco beneficioso para la población sordo-muda frente al LSM, pero propone una alternativa para la comunicación (dentro de los límites de este trabajo, aún unilateral) entre personas con impedimento del habla hacia la población con afecciones relacionadas a la percepción visual.

2. Objetivo

Objetivo general:

Crear un sistema embebido que opere en conjunto a un prototipo de guante sensor, utilizado para el muestreo de los patrones de movimiento en una mano, que realice el procesamiento y clasificación de los patrones, para la reproducción sonora del fonema o palabra completa en español asociados al movimiento ejecutado.

Objetivos específicos:

1. Ensamblar el prototipo de guante sensor implementando el SoC micro:bit como dispositivo microcontrolador y sensor del cambio de aceleraciones con el acelerómetro integrado. Añadiendo 1 sensor táctil usado para indicar el inicio y fin del muestreo de un movimiento.
2. Muestrear y preprocesar los patrones tridimensionales utilizando el SoC micro:bit.
3. Comunicar y transmitir los patrones muestreados a través de Bluetooth entre el SoC micro:bit y el mini PC RaspberryPi 4.
4. Clasificar los patrones de movimiento utilizando un modelo de *Deep Learning* y el algoritmo DTW (*Dynamic Time Warping*).
5. Concatenar las asociaciones de los movimientos para convertirlos en una cadena textual que pueda ser interpretada por el servicio en la nube de *Text-to-speech*
6. Reproducir el audio resultante del servicio en la nube mediante la tarjeta de sonido WM8960 de Waveshare.

3. Justificación

4. Productos o Resultados esperados

5. Metodología

6. Cronograma

7. Referencias

8. Alumnos y Directores

Luis Eduardo Valle Martínez.- Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta: 2015090780, Tel. 5566143276, email: lvallem1400@alumno.ipn.mx

Firma: _____

Rodolfo Romero.- Profesor de tiempo completo Laboratorio de posgrado Sistemas computacionales móviles. Candidato a Doctor en ciencias en Comunicaciones y Electrónica. Maestría en ciencias en Ingeniería electrónica. Ingeniería en comunicaciones y Electrónica. Área de trabajo Inteligencia Artificial y Procesamiento digital de señales. Tel. 5535216128, email: rromeroh@ipn.mx.

Firma: _____

Dr. Jesús Yaljá Montiel Pérez.- Profesor de tiempo completo adscrito al Laboratorio de Robótica y Mecatrónica del Centro de Investigación en Computación del Instituto Politécnico nacional. Doctor en comunicaciones y electrónica. Maestro en Ciencias en Ingeniería electrónica e Ingeniero Físico. Sus intereses son: la Inteligencia Artificial, sensores y robótica. Tel: 5524940919, Ext. IPN: 56665, email: yalja@ipn.mx

Firma: _____

CARÁCTER: Confidencial
FUNDAMENTO LEGAL: Artículo 11 Fracc. V y Artículos 108, 113 y 117 de la Ley Federal de Transparencia y Acceso a la Información Pública.
PARTES CONFIDENCIALES: Número de boleta y teléfono