



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

SISTEMA DE POSGRADO
MAESTRÍA EN TELECOMUNICACIONES

TEMA:

MODELAJE DE UNA INTERFAZ COMPUTARIZADA
CONTROLADA POR VOZ PARA CUADRIPLÉJICOS QUE CONTRIBUYA
AL DESARROLLO DE APLICACIONES QUE PERMITAN COMANDAR
ELEMENTOS DE DOMÓTICA EN SU DOMICILIO.

AUTOR:

Ing. Wilmer Enrique Jiménez Vivanco

Trabajo de titulación previo a la obtención del grado de
Magister en Telecomunicaciones

TUTORA:

MSc. María Luzmila Ruilova Aguirre

Guayaquil, a los 3 días del mes Julio del año 2017



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

SISTEMA DE POSGRADO

CERTIFICACIÓN

Certificamos que el presente trabajo fue realizado en su totalidad por el Magíster Wilmer Enrique Jiménez Vivanco como requerimiento parcial para la obtención del Grado Académico de Magíster en Telecomunicaciones.

Guayaquil, a los 3 días del mes Julio año 2017

TUTORA

MsC. Luzmila Ruilova María

DIRECTOR DEL PROGRAMA

MsC. Romero Paz Manuel



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

SISTEMA DE POSGRADO

DECLARACIÓN DE RESPONSABILIDAD

YO, **Wilmer Enrique Jiménez Vivanco**

DECLARÓ QUE:

El trabajo de Titulación “**Modelaje de una interfaz computarizada controlada por voz para cuadripléjicos que contribuya al desarrollo de aplicaciones que permitan comandar elementos de domótica en su domicilio**”, previo a la obtención del grado Académico de Magíster, ha sido desarrollado en base a una investigación exhaustiva, respetando derechos intelectuales de terceros conforme las citas que constan en las páginas correspondientes. Consecuentemente este trabajo es de mi total autoría.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance científico del Trabajo de Titulación del Grado Académico en mención.

Guayaquil, a los 3 días del mes Julio año 2017

EL AUTOR

Ing. Wilmer Enrique Jiménez Vivanco



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

SISTEMA DE POSGRADO

AUTORIZACIÓN

Yo, **Wilmer Enrique Jiménez Vivanco**


Autorizo a la Universidad Católica de Santiago de Guayaquil, la publicación, en la biblioteca de la institución del Trabajo de Titulación de Maestría titulado: **“Modelaje de una interfaz computarizada controlada por voz para cuadripléjicos que contribuya al desarrollo de aplicaciones que permitan comandar elementos de domótica en su domicilio”**, cuyo contenido, ideas y criterios son de mi exclusiva responsabilidad y total autoría.

Guayaquil, a los 3 días del mes Julio año 2017

EL AUTOR

Ing. Wilmer Enrique Jiménez Vivanco

REPORTE DE URKUND



Lista de fuentes | Bloques

Documento [TesisMaestriaWilmerJimenez.docx](#) (D28160517)

Presentado 2017-05-12 19:11 (-05:00)

Presentado por wilmer_jv@hotmail.com

Recibido edwin.palacios.ucsg@analysis.orkund.com

Mensaje TesisMaestriaWilmerJimenez [Mostrar el mensaje completo](#)

1% de estas 26 páginas, se componen de texto presente en 2 fuentes.

+	Categoría	Enlace/nombre de archivo	
+		http://docplayer.es/7793265-T-e-s-i-s-...	-
+		http://hdl.handle.net/10915/21513	✓
+	Fuentes alternativas		
+	La fuente no se usa		

📊 🔍 🗨️ 👁️ ⬆️ ⬅️ ➡️ 🔄 Reiniciar 📄 Exportar 🔗 Compartir ⚙️

⚠️ 1 Advertencias

SISTEMA DE POSGRADO MAESTRÍA EN TELECOMUNICACIONES

TEMA DE TITULACIÓN:

MODELAJE DE UNA INTERFAZ COMPUTARIZADA CONTROLADA POR VOZ PARA CUADRIPLÉJICOS QUE CONTRIBUYA AL DESARROLLO DE APLICACIONES QUE PERMITAN COMANDAR ELEMENTOS DE

64% #1 Activo

Previa la obtención del Grado Académico de Magister en Telecomunicaciones

ELABORADO POR: Ing. Wilmer Enrique Jiménez Vivanco

Guayaquil, a los 10 días del mes Mayo del año 2017

SISTEMA DE POSGRADO

CERTIFICACIÓN Certificamos que el presente

Archivo de registro Urkund: Universidad C... 64%

Previa la obtención del Grado Académico de Magister en Telecomunicaciones

ELABORADO POR: Ing. Natalia Vanessa Romero Garcia

Guayaquil, a los 10 días del mes Noviembre año 2015

SISTEMA DE POSGRADO

Dedicatoria

Dedico el presente trabajo de titulación, a mi esposa e hijas quienes han estado a mi lado en cada uno de mis retos profesionales y a mis padres quienes sembraron en mí la semilla de la superación, la constancia y la responsabilidad.

Agradecimientos

Agradezco a la Universidad Católica Santiago de Guayaquil, por la oportunidad de especializarme con los excelentes docentes que en ella laboran, agradezco especialmente a mi tutora la Msc. Luzmila Ruiloba, quien me acompañó incondicionalmente en el desarrollo de mi trabajo de investigación.



UNIVERSIDAD CATÓLICA
DE SANTIAGO DE GUAYAQUIL

SISTEMA DE POSGRADO

TRIBUNAL DE SUSTENTACIÓN

f. _____

M.Sc Luzmila Ruiloba María

TUTORA

REVISORES:

MsC. Orlando Philco Asqui

MsC. Luis Córdova Rivadeneira

f. _____

M.Sc ROMERO PAZ MANUEL DE JESÚS

DIRECTOR DEL PROGRAMA

ÍNDICE GENERAL

ÍNDICE DE FIGURAS.....	XII
ÍNDICE DE TABLAS	XIV
Resumen.....	XV
Abstract	XVI

CAPÍTULO 1: DESCRIPCIÓN DEL PROYECTO DE INTERVENCIÓN. **17**

1.1. <i>Introducción.....</i>	17
1.2. <i>Antecedentes.....</i>	17
1.3. <i>Justificación del problema a investigar.....</i>	18
1.4. <i>Definición del problema</i>	18
1.5. <i>Objetivos.....</i>	18
1.5.1. <i>Objetivo General:.....</i>	18
1.5.2. <i>Objetivos específicos:.....</i>	19
1.6. <i>Hipótesis.....</i>	19
1.7. <i>Metodología de investigación.....</i>	19

CAPÍTULO 2: FUNDAMENTACIÓN TEÓRICA. **20**

2.1. <i>Cuadriplejia.....</i>	20
2.2. <i>Sistema vocal del ser humano.....</i>	21
2.4. <i>Modelos de Voz.....</i>	24
2.5. <i>Técnicas de procesamiento digital de señales.....</i>	25
2.5.1. <i>Transformada rápida de Fourier.....</i>	25
2.5.2. <i>Ventanas.....</i>	26
2.5.3. <i>Filtrado de una señal digital.....</i>	27
2.6. <i>Extracción de las características esenciales de la voz.....</i>	30

2.6.1. Detectores basados en predicción lineal: LPC (Linear Predictive coding).....	30
2.6.2. Coeficientes Cepstrales de MEL	30
2.7. Métodos de decisión basados en cuantificación vectorial	33
2.8. Software de procesamiento de señales-Matlab.....	36

CAPÍTULO 3: DISEÑO DE UN SISTEMA DE RECONOCIMIENTO DE VOZ. 37

3.1. Adquisición de la señal de voz	37
3.2. División de la señal de voz.....	38
3.3. Ventaneo de la señal.....	39
3.4. Transformada rápida de Fourier.....	40
3.5. Coeficientes Cepstrales.....	41
3.6. Método abreviado de MatLab.....	43
3.7. Etapa de decisión	44
3.8. Interfaz de usuario.....	45
3.8.1. Etapa de alarmas y mensajes	45
3.8.2. Etapa de configuración.....	45
3.8.3. Etapa de reconocimiento.....	47
3.8.4. Etapa de Control.....	49
3.9. Interfaz con una tarjeta serial.	51
3.10. Etapas del programa.	53
3.10.1. Etapa de inicialización.....	53
3.10.2. Etapa de grabado de patrones.....	54
3.10.3. Etapa de borrado.....	55
3.10.4. Etapa de reconocimiento.....	56

CAPÍTULO 4: PRUEBAS DE ROBUSTEZ DEL SISTEMA DE RECONOCIMIENTO DE VOZ. 58

4.1. Consideraciones iniciales.....	58
4.2. Resultados de la pruebas funcionales.....	59

4.3. <i>Análisis de resultados</i>	62
CAPÍTULO 5: CONCLUSIONES Y RECOMENDACIONES	64
5.1. <i>Conclusiones</i>	64
5.2. <i>Recomendaciones</i>	65
REFERENCIAS BIBLIOGRÁFICAS	66
GLOSARIO DE TERMINOS	69
ANEXOS	70

ÍNDICE DE FIGURAS

CAPÍTULO 2: FUNDAMENTACIÓN TEÓRICA

Figura 2.1: Cuadriplejia-paraplejia	21
Figura 2. 2: Modelo de generación de voz.....	24
Figura 2.3: Señal sinusoidal de 1KHz, (a) dominio del tiempo.....	26
Figura 2.4: Ventana de Hamming	27
Figura 2.5: Respuesta en frecuencia - Filtro pasa alto.....	28
Figura 2.6: Respuesta en frecuencia - Filtro pasa bajo.....	28
Figura 2.7: Respuesta en frecuencia - Filtro pasa banda.	29
Figura 2.8: Respuesta en frecuencia - Filtro elimina banda.....	29
Figura 2.9: Ejemplo de los Filtros de Mel.....	32
Figura 2.10: Software MatLab.....	36

CAPÍTULO 3: DISEÑO DE UN SISTEMA DE RECONOCIMIENTO DE VOZ

Figura 3.1: Señal de voz digitalizada.	38
Figura 3.2: Trama de 18mseg.....	39
Figura 3.3: Trama de 18mseg ventaneada.	40
Figura 3.4: FFT de la señal ventaneada.	41
Figura 3.5: Banco de filtros de Mel para $p=20$, $n=200$ y $F_s=11025$	42
Figura 3.6: Coeficientes cepstrales de una de las tramas.	43
Figura 3.7: Etapa de mensajes y alarmas-interfaz de reconocimiento de voz.	45
Figura 3.8: Etapa de configuración de la interfaz de reconocimiento de voz.	46
Figura 3.9: Etapa de configuración de la interfaz de reconocimiento de voz-parámetros.	46
Figura 3.10: Mensaje "Grabando" de la interfaz de reconocimiento de voz	47
Figura 3.11: Etapa de Reconocimiento de voz	48
Figura 3.12: Tarjeta Arduino utilizada	49
Figura 3.13: Etapa de control.....	50

Figura 3.14: Interfaz de reconocimiento de voz	51
Figura 3.15: Comienzo de la etapa de inicialización.....	53
Figura 3.16: Comienzo de la etapa de grabado de patrones.....	54
Figura 3.17: Estructura de DB.DAT	55
Figura 3.18: Inicio de la etapa de borrado	55
Figura 3.19: Inicio de la etapa de reconocimiento.....	56

CAPÍTULO 4: PRUEBAS DE ROBUSTEZ DEL SISTEMA DE RECONOCIMIENTO DE VOZ

Figura 4.1: Configuración para pruebas de robustez.....	59
---	----

ÍNDICE DE TABLAS

CAPÍTULO 2: FUNDAMENTACIÓN TEÓRICA

Tabla 2.1: Algoritmo k-means	35
------------------------------------	----

CAPÍTULO 3: DISEÑO DE UN SISTEMA DE RECONOCIMIENTO DE VOZ

Tabla 3.1: Muestreo de tiempos para pronunciar un comando de voz	37
--	----

Tabla 3.2: Comandos seriales para el control del sistema.....	52
---	----

CAPÍTULO 4: PRUEBAS DE ROBUSTEZ DEL SISTEMA DE RECONOCIMIENTO DE VOZ

Tabla 4.1: Pruebas de robustez con el usuario 1.....	60
--	----

Tabla 4.2: Pruebas de robustez con el usuario 2.....	60
--	----

Tabla 4.3: Pruebas de robustez con el usuario 3.....	61
--	----

Tabla 4.4: Pruebas de robustez con el usuario 4.....	61
--	----

Tabla 4.5: Resumen de las pruebas de robustez.....	62
--	----

Resumen

En el presente trabajo de investigación se inicia con una investigación acerca de un tema social bastante importante como es la discapacidad por paraplejia y cuadriplejia, y la dependencia que esta última genera hacia otras personas, identificando la necesidad de un sistema que disminuya esta dependencia mediante el control de elementos de domótica en el domicilio mediante comandos de voz.

Se continúa con la investigación de los métodos de reconocimiento de voz que existen en la actualidad, como es el caso de la extracción de los coeficientes cepstrales de la escala de Mel, para luego hacer una propuesta de una interfaz computarizada controlada por voz para cuadripléjicos que contribuya al desarrollo de aplicaciones, que permitan comandar elementos de domótica en su domicilio; en donde además explicamos los bloques de programa con el fin de que este proyecto sea fácilmente modificable y aplicable a cada caso en particular, más adelante pasamos a realizar las pruebas de robustez de la interfaz desarrollada con el fin de medir la efectividad que ofrece la misma.

Palabras clave: Reconocimiento, voz, MatLab, cuadriplejia, domótica.

Abstract

In this paper we start with an investigation about a very important social issue such as disability by paraplegia and quadriplegia, and the dependence that this generates on other people, identifying the need for a system that decreases this dependence through control of domotic elements in the home by means of voice commands.

We continue with the investigation of the speech recognition methods that exist currently, such as the extraction of cepstral coefficients through the Mel scale, and then we make a proposal for a voice-controlled computer interface for quadriplegic that contributes to the development of applications, that allow to command domotic elements in the domicile; Where we also explain the program blocks so that this project is easily modifiable and applicable to each particular case, later we do the robustness tests of the interface developed to measure the effectiveness it offers.

CAPÍTULO 1: DESCRIPCIÓN DEL PROYECTO DE INTERVENCIÓN.

1.1. Introducción.

En el presente documento se describe el modelaje de una interfaz computarizada controlada por voz para cuadripléjicos, la cual podría ser utilizada para desarrollar distintas aplicaciones de domótica para estas personas con capacidades especiales.

Se inicia identificando la problemática que existe, lo que permite hacer el planteamiento del proyecto de titulación, luego se continúa con el análisis del proceso de generación y reconocimiento de voz, identificando sus componentes y métricas. El presente proyecto comprende además el desarrollo de un algoritmo de reconocimiento de voz que permita hacer una simulación de la detección de los comandos de audio emitidos por las personas con cuádrupleja.

1.2. Antecedentes.

La cuádrupleja es una enfermedad en la cual la personas no pueden mover sus extremidades, en la mayoría de los casos se da debido a accidentes automovilísticos, caídas o por heridas con armas; se debe diferenciar la paraplejia de la cuádrupleja, ya que la primera es la inmovilidad de las extremidades inferiores y la segunda corresponde la inmovilidad total de las extremidades.

Matlab es un software con motores de procesamiento de señales digitales más importantes que hay en la actualidad, que entre otras cosas permite hacer estudios de métodos de procesamiento y reconocimiento de señales de voz.

1.3. Justificación del problema a investigar.

Los cuadripléjicos son personas que no pueden mover sus extremidades, lo que los dificulta poder presionar botones para comandar electrodomésticos, y dispositivos en sus domicilios. Esto provoca que exista dependencia para realizar estas tareas, y es común que los cuadripléjicos sientan depresión al tener esta dependencia hacia otras personas.

En el presente tema de titulación se propone investigar acerca de los sistemas de reconocimiento de voz basados en técnicas de extracción de coeficientes cepstrales, con el fin de identificar las características especiales de un mensaje de voz y con la investigación planteada se modelará una interfaz computarizada controlada por voz para cuadripléjicos, la misma que podría ser utilizada para el desarrollo de aplicaciones de domótica.

1.4. Definición del problema

La necesidad de modelar una interfaz computarizada controlada por voz para las personas cuadripléjicas que contribuya al desarrollo de aplicaciones que permita comandar elementos de domótica en su domicilio.

1.5. Objetivos

1.5.1. Objetivo General:

Modelar una interfaz computarizada controlada por voz para cuadripléjicos que contribuya al desarrollo de aplicaciones que permitan comandar elementos de domótica en su domicilio.

1.5.2. Objetivos específicos:

- ✓ Analizar las técnicas de reconocimiento de voz basadas en la extracción de coeficientes cepstrales.
- ✓ Desarrollar y simular un algoritmo de procesamiento y reconocimiento de comandos voz en MatLab.
- ✓ Evaluar la robustez de la aplicación mediante pruebas funcionales.

1.6. Hipótesis

El modelaje de una interfaz computarizada para cuadripléjicos controlada por voz permitiría disponer de una aplicación robusta de reconocimiento de comandos de audio, con la cual se podría desarrollar sistemas de domótica para personas que tienen estas capacidades especiales.

1.7. Metodología de investigación.

La presente trabajo de titulación utiliza una metodología de investigación empírica-analítica, en la cual se va modelar un algoritmo de procesamiento digital de señales, que permita reconocer comandos de voz, además de realizar las pruebas correspondientes para identificar la robustez del mismo.

CAPÍTULO 2: FUNDAMENTACIÓN TEÓRICA.

2.1. Cuadriplejia

Para definir la Cuadriplejia primero se va a definir la paraplejia; (Moreno-Fergusson & Rey, 2012, p. 83) manifiestan que la paraplejia consiste en lesiones de la médula Espinal que constituyen un evento devastador para las personas, ya que las secuelas repentinas y permanentes se ocasionan al nivel motriz autónomo y sensitivo del ser humano, además expresa que la incidencia de este tipos de lesiones se encuentran entre 10,4 y 83 por millón de habitantes por año, y es presentado en mayor número en hombres que en mujeres en edades entre 15 y 35 años, estos pueden ser ocasionados por accidentes automovilísticos, deportivos, heridas por arma de fuego, caídas entre otros; con mayor frecuencia esta lesión se presenta a nivel cervical que, toraco-lumbar y sacro, siendo más comunes, las lesiones torácicas y lumbares completas.

Estos cambios en el cuerpo son difíciles de afrontar ya que de un momento a otro pierden la sensibilidad del cuerpo imposibilitándose de caminar, esto incrementa la dependencia de las personas y el riesgo de caer en depresión.

(Moreno-Fergusson & Rey, 2012, p. 84) Manifiestan además que por tal razón las personas tienen que re-aprender a realizar sus actividades de una manera diferente y realizar cambios ambientales para mejorar los accesos a los lugares de su domicilio. Esto provoca un cambio completamente significativo entre la vida que llevaban y la vida que llevarán.

Como se ha descrito anteriormente la Paraplejia es la imposibilidad que tienen las personas para caminar, es decir incapacidad de las extremidades inferiores, mientras que la

cuadruplejia es más complicado, como lo establece (Patiño & Gualotuña, 2013, p. xvii) las personas tienen la imposibilidad de mover las cuatro extremidades, es decir piernas y brazos, esto sucede cuando la lesión en la médula espinal es más severa, en la mayoría de los casos la Cuadruplejia se presenta con la pérdida de la sensibilidad del cuello para abajo.

En la figura 2.1 (“NewYork-Presbyterian/Queens - Lesiones Agudas de la Médula Espinal,” n.d., fig. 2) se muestra las características de la Cuadruplejia y de la Paraplejia en una imagen que define estos dos términos de manera muy clara.

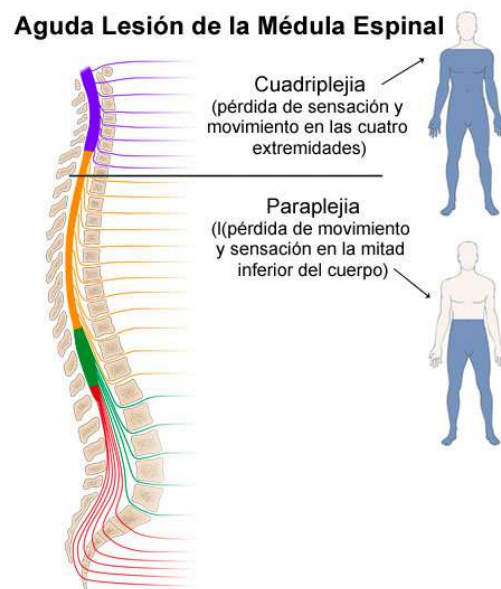


Figura 2.1: Cuadruplejia-paraplejia.

Fuente: (“NewYork-Presbyterian/Queens - Lesiones Agudas de la Médula Espinal,” n.d., fig. 2).

2.2. Sistema vocal del ser humano

La voz es una secuencia de sonidos, generada por el aparato fonador humano (o por una imitación de este mediante una máquina),

que codifica cierta información lingüística o un mensaje. (Cordovilla & Andrés, 2012, p. 7)

Conforme lo establece (Cordovilla & Andrés, 2012, pp. 7–9) la voz está compuesta por tres tipos de elementos:

- Los silencios.- son etapas que no se transmite voz, sin embargo conlleva información del mensaje
- Los sonidos sonoros.- es el espacio de tiempo que lleva el mensaje o las palabras
- Los sonidos sordos.- en esta etapa están todos ruidos asociados a la voz.

Los sonidos sonoros poseen una estructura periódica, por lo tanto contienen un frecuencia fundamental (PITCH), por lo tanto el espectro de los sonidos sonoros poseen una serie de armónicos separados un intervalo de frecuencia igual al pitch aproximadamente. Con lo anteriormente mencionado, se puede considerar que el resultado es un espectro formado por un tren de pulsos con una envolvente suave conforme lo menciona (Cordovilla & Andrés, 2012, p. 7)

(Cordovilla & Andrés, 2012, p. 7,8) establece que para hacer un análisis de las señales de voz se debe tomar en cuenta la magnitud y la posición de los picos de la mencionada envolvente. Los valles de la onda que envuelve la señal poseen menos importancia, debido a que su profundidad varía en función de los locutores

En cambio los sonidos Sordos tienen todo tipo de ruido producido por la misma voz humana, y en general varían con gran velocidad, concentrándose alrededor de los 2KHz. Esta información es más difícil de clasificar ya que se pueden mezclar con las consonantes s, f, p o k, conforme menciona (Cordovilla & Andrés, 2012, p. 8)

(Cordovilla & Andrés, 2012, p. 8) Manifiesta además que existe una relación entre los sonidos Sordos y Sonoros, esto es debido a varias razones; los sonidos sordos están al menos a 0.2 seg de un sonido sonoro, esto es antes y después del mismo; los sonidos sordos poseen una baja energía espectral en comparación de los sonidos sonoros.

2.3. Pitch.

(Wainschenker, Doorn, Castro, & Legrottaglie, 2003, p. 1), expresa que el pitch involucra una gran cantidad de tópicos de gran complejidad. Sin embargo sus funcionalidades hacen posible implementarlo en una gran variedad de algoritmos de adquisición. A lo largo del tiempo el Pitch se ha definido como la frecuencia fundamental de espectro de frecuencias del habla y ha sido asociado al movimiento que realiza la glotis en la generación del sonido. Sin embargo de cualquier forma que se lo defina es difícil su aplicación práctica debido a que la generación de sonidos en el glotis es cuasi-periódica

La frecuencia fundamental en algunos casos desaparece de la frecuencia sonora, lo que hace muy difícil de identificar. Esto ocurre cuando la energía del tracto vocal se concentra en ciertos armónicos. Sin embargo, no se pierde completamente, y se puede utilizar algunos armónicos para su rastreo. Cuando una persona esta hablando por un largo tiempo, provoca que esta variación no sea constante. Las variaciones de la glotis dependen en gran medida del gesto emocional de la persona, para el análisis de los sistemas de reconocimiento de voz, se vuelve de gran importancia el estudio de la frecuencia fundamental conforme lo analiza (Wainschenker et al., 2003, p. 1)

2.4. Modelos de Voz.

(Cordovilla & Andrés, 2012, p. 11) Muestran que la voz es una señal que proviene de la excitación de las cuerdas vocales mientras una persona habla, y en los momentos de silencio, esta señal puede ser reemplazada por ruido. Esta señal se puede considerar como una señal periódica, ya que posee un Pitch (Producto de la vibración de las cuerdas vocales)

Un modelo simplificado de voz se puede considerar con una división en dos partes, un aparato fonador encargado de generar cualquier tipo de sonidos y un bloque de generación de ruido como se puede observar en la figura 2.2 que lo muestran (Cordovilla & Andrés, 2012, p. 12).

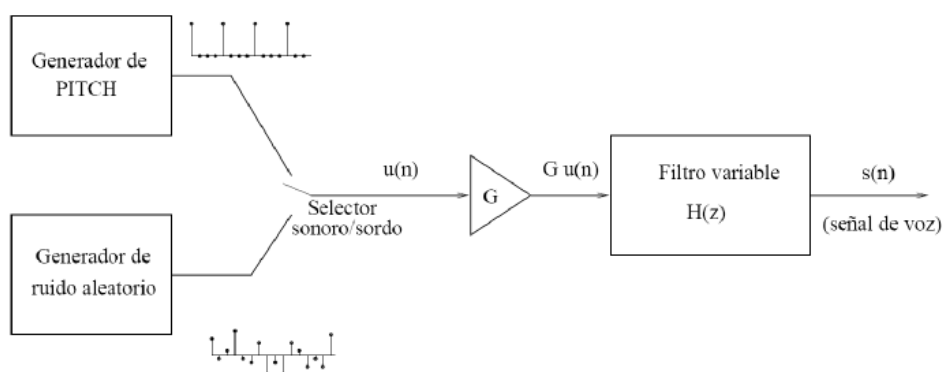


Figura 2. 2: Modelo de generación de voz.

Fuente:(Cordovilla & Andrés, 2012, p. 12).

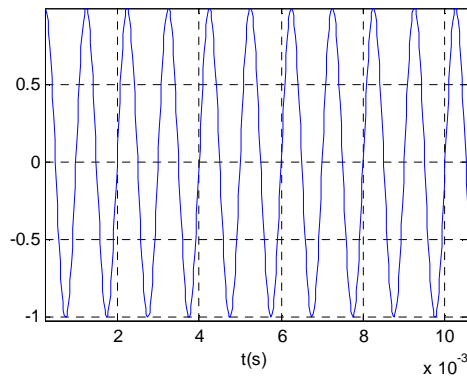
Los sonidos sonoros son representados con un generador de pitch, mientras que los sonidos sordos son representados mediante un generador de ruido, la ciencia de reconocimiento de voz consiste en detectar los sonidos sonoros, aislando la información de los sonidos sordos.

2.5. Técnicas de procesamiento digital de señales.

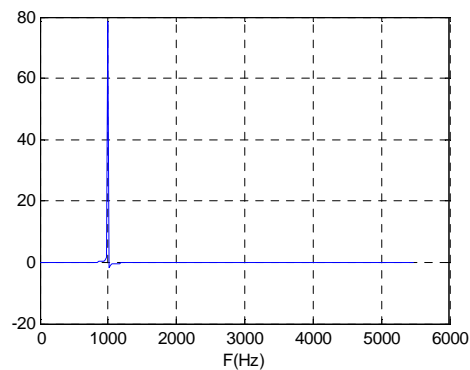
2.5.1. Transformada rápida de Fourier.

Es una herramienta matemática que permite pasar una señal del dominio del tiempo al dominio de la frecuencia, proviene de una variación de la transformada discreta de Fourier, y permite hacer un cálculo más sencillo teniendo como entrada una señal seno y coseno con distintas frecuencias. La transformada rápida de Fourier (FFT), utilizan muestras discretas y da como resultado un vector de números complejos, que representan la amplitud y fase de la señal resultante, para análisis de voz se toma en cuenta la mitad del vector de la amplitud.

Con el fin de ejemplificar en la figura 2.3 (a) se muestra una señal en el dominio del tiempo y en la figura 2.3 (b) se puede observar la misma señal en el dominio de la frecuencia obtenida mediante transformada rápida de Fourier.



(a)



(b)

Figura 2.3: Señal sinusoidal de 1KHz, (a) dominio del tiempo.

(b) dominio de la frecuencia.

Elaborado por: El Autor.

2.5.2. Ventanas.

Son funciones matemáticas que ayudan a segmentar una señal, esta tiene que ser escogida de tal manera que optimice nuestro procesamiento. Una de las familias de ventanas más utilizadas son las que se modelan mediante una señal sinusoidal, como el caso de la ventana de Hamming y la ventana de Hann, las cuales se representan en la siguiente fórmula conforme lo expresa (Segura & Elena, 2007, p. 6) :

$$W(n) = a - (1 - a) \cdot \cos\left(\frac{2 \cdot \pi \cdot n}{N - 1}\right)$$

Donde a toma el valor de 0.54 para hamming y 0.5 para Hann, además la función está definida para $0 \leq n \leq N-1$, siendo 0 para los otros casos conforme lo establece (Segura & Elena, 2007, p. 7); con el fin de ejemplificar en la figura 2.4 se muestra un ejemplo de la ventana de Hamming.

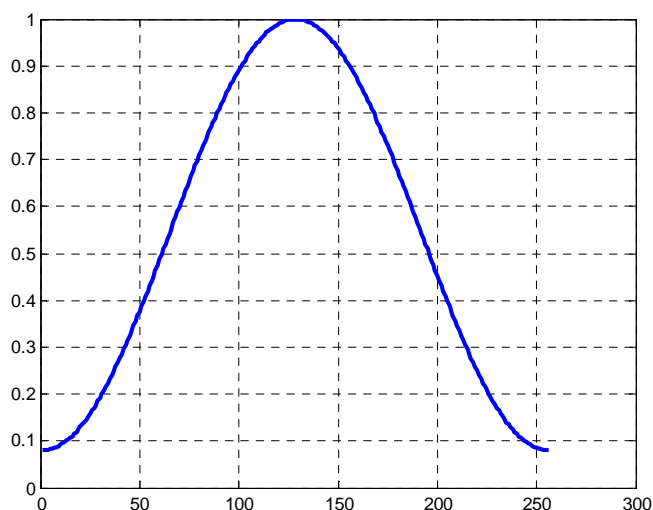


Figura 2.4: Ventana de Hamming

Elaborado por: El Autor.

2.5.3. Filtrado de una señal digital.

Esta técnica es muy útil cuando deseamos eliminar información o ruido que se encuentra en un determinado espacio espectral, la mayor ventaja de los filtros digitales frente a los filtros analógicos, es que los filtros digitales presentan un mayor acercamiento a un filtro ideal. Los filtros más conocidos, son los siguientes:

- **Filtro pasa Alto:** permite el paso de frecuencias altas, en la figura 2.5 se puede observar la respuesta en frecuencia de un filtro pasa Alto.

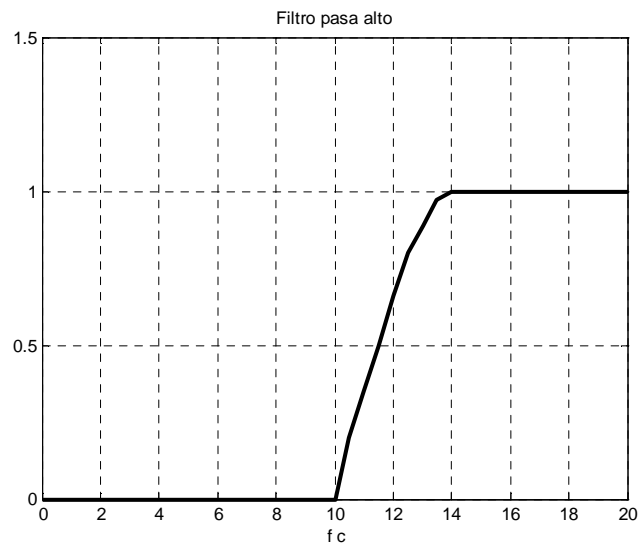


Figura 2.5: Resposta en frecuencia - Filtro pasa alto.

Fuente:(Jiménez Vivanco & Ochoa Coronel, 2010, p. 34).

- **Filtro pasa Bajo:** permite el paso de frecuencias bajas, en la figura 2.6 se puede observar la respuesta en frecuencia de un filtro pasa Alto.

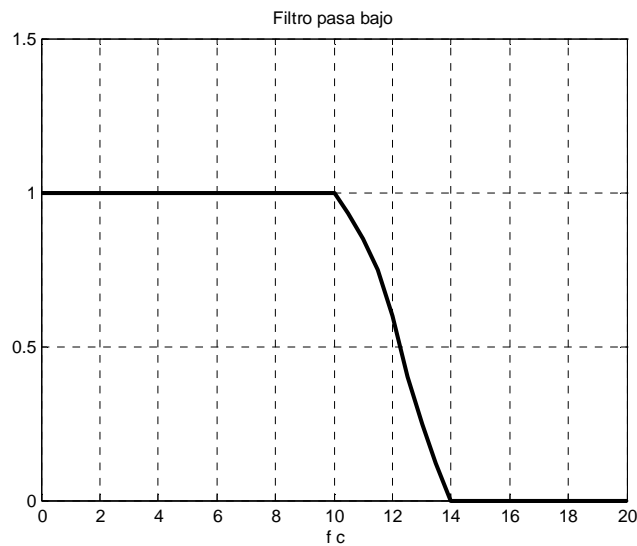


Figura 2.6: Resposta en frecuencia - Filtro pasa bajo.

Fuente:(Jiménez Vivanco & Ochoa Coronel, 2010, p. 35).

- **Filtro pasa banda:** permite el paso de un rango de frecuencias, en la figura 2.7 se puede observar la respuesta en frecuencia de un filtro pasa Alto.

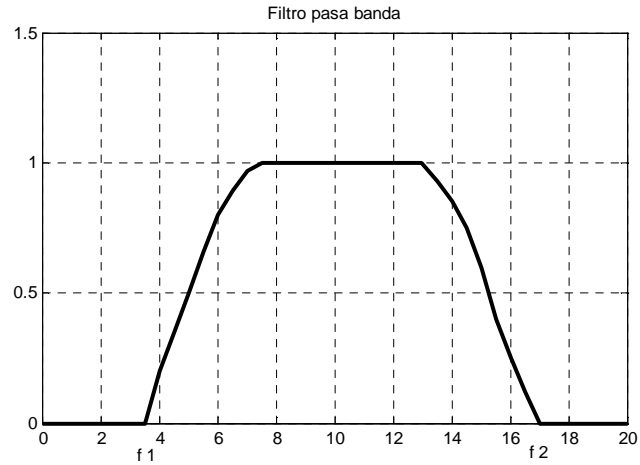


Figura 2.7: Respuesta en frecuencia - Filtro pasa banda.

Fuente:(Jiménez Vivanco & Ochoa Coronel, 2010, p. 35).

- **Filtro elimina banda:** elimina o atenúa un rango de frecuencias, en la figura 2.8 se puede observar la respuesta en frecuencia de un filtro pasa Alto.

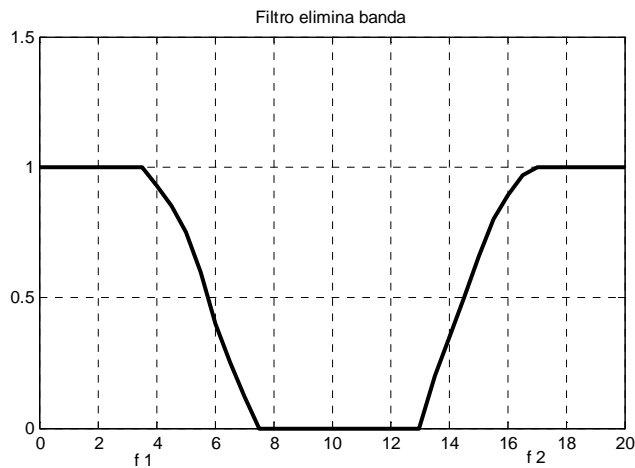


Figura 2.8: Respuesta en frecuencia - Filtro elimina banda.

Fuente:(Jiménez Vivanco & Ochoa Coronel, 2010, p. 36).

2.6. Extracción de las características esenciales de la voz

La extracción de las características esenciales de la voz consiste en transformar una señal en otra más pequeña que contenga información importante de la misma, de tal forma que pueda ser comparada e identificada, la señal resultante puede ser procesada con menor cantidad de recursos informáticos, existen diversos métodos de extracción de las características esenciales como los basados en predicción lineal y los basados en la extracción de coeficientes cepstrales mediante los filtros de Mel.

2.6.1. Detectores basados en predicción lineal: LPC (Linear Predictive coding).

Conforme menciona (Varela Serrano, 2011, p. 34) este método es utilizado cuando se necesita parametrizar la envolvente espectral de una señal con un pequeño número de coeficientes, de tal manera que la señal pueda ser reconstruida adecuadamente. Los modelos LPC son sencillos de aplicar y no tienen mayor complejidad para su implementación tanto en Hardware como en software, se ha tenido buenos resultados en sistemas de reconocimiento de Voz,

Antes de realizar un análisis LPC, la señal necesita pasar un pre-procesamiento de la señal, mediante un filtrado pre-énfasis, división en tramas y un ventaneo de la señal entramada, las técnicas utilizadas para el análisis LPC son las de autocorrelación; para la división en tramas se acostumbra a utilizar intervalos de tiempo de 20ms.

2.6.2. Coeficientes Cepstrales de MEL

Conforme menciona (Guajardo & Guillermo, 2014, p. 52) para extraer los coeficientes cepstrales de Mel, es necesario trabajar con una escala logarítmica basada en la percepción del pitch, además

establece que las fórmulas que relacionan la escala de Mel con la escala en Hz son las siguientes:

$$f_{in} = 700 \cdot \left[\exp\left(\frac{f_{mel}}{1127}\right) - 1 \right]$$

$$f_{mel}^{-1} = 1127 \cdot \ln \left[1127 \left(1 + \frac{f_{in}}{700} \right) \right]$$

La primera fórmula corresponde al paso de la escala de Mel a la escala en Hz, y la segunda fórmula lo contrario conforme lo establece (Guajardo & Guillermo, 2014, p. 53).

(Guajardo & Guillermo, 2014, p. 53) también menciona que las fórmulas para obtener el banco de filtros de Mel son las siguientes:

$$f_{bi} = \left(\frac{N}{FS} \right) \cdot f_{mel}^{-1} \cdot \left(f_{mel}(f_{low}) + i \cdot \frac{f_{mel}(f_{high}) - f_{mel}(f_{low})}{M + 1} \right)$$

$$H_i(k) = \begin{cases} 0 & \text{para } k < f_{bi-1} \\ \frac{(k-f_{bi-1})}{(f_{bi}-f_{bi-1})} & \text{para } f_{bi-1} < k < f_{bi} \\ \frac{(f_{bi+1}-k)}{(f_{bi+1}-f_{bi})} & \text{para } f_{bi} < k < f_{bi+1} \\ 0 & \text{para } k > f_{bi+1} \end{cases}$$

$i = 1, 2, 3, \dots, M$

Donde:

- N= Número de muestras
- f_{low} = Frecuencia inferior del banco de filtros
- f_{High} = Frecuencia superior del banco de filtros
- M= Número de filtros
- H= Matriz de filtros
- i= Índice del filtro

K= Número de muestra

En la figura 2.9 se muestra un ejemplo de los bancos de filtros de Mel, los cuales tienen una escala logarítmica, siendo mas angosto a bajas frecuencias y mas ancho en altas frecuencias.

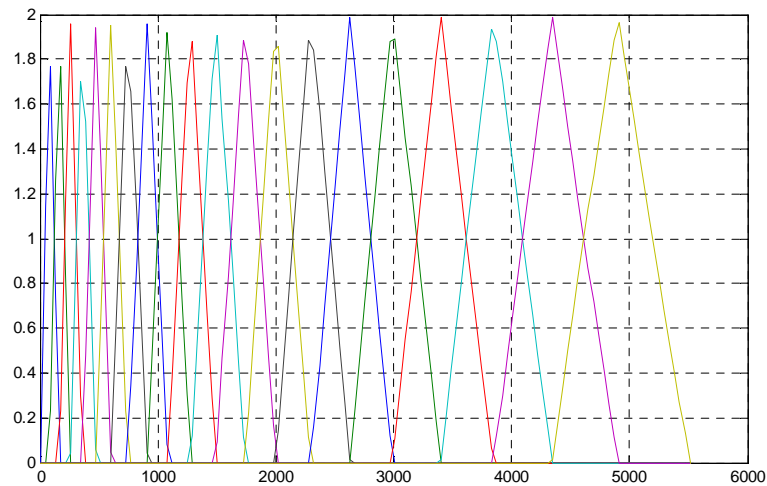


Figura 2.9: Ejemplo de los Filtros de Mel.

Elaborado por: El Autor.

Mediante los bancos de filtros de la escala de Mel se puede obtener el cepstrum de la señal, con la siguiente fórmula, conforme lo menciona (Guajardo & Guillermo, 2014, p. 53).

$$X_i = \log_{10} \left(\sum_{K=0}^{N-1} |X(k)| \cdot H_i(k) \right)$$

Donde:

$|X(k)|$ = Magnitud de la FFT.

X_i = Vector de energía

Con las ecuaciones mencionadas anteriormente (Guajardo & Guillermo, 2014, p. 54) establecen que la fórmula para obtener los coeficientes cepstrales de la escala de Mel es la siguiente:

$$C_j = X_i \cdot \cos\left(j \cdot \left(\frac{i-1}{2}\right) \cdot \frac{\pi}{M}\right)$$

2.7. Métodos de decisión basados en cuantificación vectorial

Uno de los principales inconvenientes en los sistemas de reconocimiento de voz, es escoger un método apropiado de cuantificación y decisión, existen métodos como, la medición de distancias punto a punto así como métodos más complejos basados en la cuantificación vectorial.

Para iniciar el proceso de cuantificación vectorial primero las señales tienen que estar separadas en tramas y haber sido ventaneadas, (Rodríguez, 2008, p. 13) indica que la cuantificación vectorial se encuentra definida por un libro de códigos llamados "CodeBook", siendo este un conjunto de vectores prototipo, y a cada vector se lo conoce como código de palabras CodeWord perteneciente a los CodeBook usando una medida de distorsión.

Medidas de distorsión.- en la generación de los CodeWord se requiere la minimización de la distorsión, entonces (Rodríguez, 2008, p. 13) indica que si asumimos que:

$$x = (x_1, x_2, \dots, x_d)^t \in R^d$$

es un vector d-dimensional cuyos componentes $(x_k, 1 \leq k \leq d)$ son números reales, para obtener el vector cuantificado, el vector x es mapeado con otro vector Z de amplitud discreta, por lo tanto:

$$z = q(x)$$

donde $q(x)$ en el operador de cuantización, z típicamente se encuentra definido por:

$$Z = \{z_j, 1 \leq j \leq M\}$$

Z y z_j son vectores d -dimensionales, entonces el conjunto de Z es el vector de CodeBook, M es el tamaño del CodeBook y z_j es el j -ésimo CodeWord, (Rodríguez, 2008, p. 14) además establece que para diseñar el CodeWord el espacio d -dimensional del vector aleatorio X tiene que estar dividido en M regiones o celdas $\{C_i, 1 \leq i \leq M\}$, donde cada C_i está vinculado con el CodeWord z_i conforme:

$$q(x) = z_i, \text{ si } x \in C_i$$

(Rodríguez, 2008, p. 14) indica además que cuando se comparan los vectores x con z_i , existirá un error de cuantización, esto se le denomina distorsión y se lo definirá como $d(x,y)$, para medir la calidad de la cuantización; entonces para prevenir esta calidad de la cuantización se puede redefinir la ecuación de la siguiente manera:

$$q(x) = z_i \text{ si } i = \operatorname{argmin} d(x, z_k)$$

El método más común para medir la distorsión es mediante a la distancia Euclidiana, entonces (Rodríguez, 2008, p. 14) además establece que la fórmula para medir la distorsión es:

$$d(x, z) = (x - z)^t \cdot (x - z) = \sum_{i=1}^d (x - z_i)^2$$

Para observar de una mejor manera la distorsión se puede introducir un peso diferente, esta ecuación se conoce como distancia de Mahalanobis y conforme lo establece (Rodríguez, 2008, p. 14) la ecuación está definida por:

$$d(x, z) = (x - z)^t \cdot \sum^{-1} (x - z)$$

Generación de los CodeBook.- Una de las partes fundamentales en los métodos de decisión es el diseño de los codebook, en donde se debe enfocar a minimizar la distorsión global todos los niveles de los vectores de cuantización, conforme lo menciona (Rodríguez, 2008, p. 14) existe un algoritmo conocido k-means que cumple estas condiciones, en el cual la idea fundamental es dividir el vector de entrenamiento en M agrupaciones, tomando en cuenta condiciones de optimización, el algoritmo es descrito en la tabla 2.1:

Tabla 2.1: Algoritmo k-means.

<p>Paso_1: <i>Inicialización.</i> Escoger un método adecuado para obtener el vector de cuantización inicial el cual requiere ser optimizado.</p>
<p>Paso_2: <i>Clasificación por el método del vecino más cercano.</i> Clasificar cada vector de entrenamiento $\{x_k\}$ dentro de cada una de las C_i celdas para escoger el más cercano codeword $z_i(x \in C_i, si d(x, z_i) \leq d(x, z_j) para todo j \neq i)$. Esta clasificación es llamada clasificación de mínima distancia.</p>
<p>Paso_3: Actualización de codebook. Actualizar el codeword de cada célula para calcular el centroide de los vectores de entrenamiento de cada célula de acuerdo a,</p> $\hat{z}_i = \arg \min_{z_j} \frac{1}{T} \sum_{x \in C_i} d(x, z_j), \hat{z}_i = cent(C_i, 1 \leq i \leq M)$
<p>Paso_4: Reiteración. Repetir los pasos 2 y 3 hasta que la nueva distorsión global d en la actual iteración sea menor a un cierto umbral fijado</p>

Fuente: (Rodríguez, 2008, p. 15).

2.8. Software de procesamiento de señales-Matlab.

Matlab es considerado como uno de los software más poderoso para cálculos matemáticos, en donde el procesamiento digital de señales puede ser realizado; entre las aplicaciones que se pueden realizar con Matlab, están las siguientes:

- Análisis Matemático.
- Creación de algoritmos.
- Adquisición de Datos.
- Simulación de prototipos.
- Visualización y análisis de datos.
- Gráficos científicos y de ingeniería
- Interfaz gráfica del usuario y desarrollo de aplicaciones.
- Sistema de simulación de sistemas con las respectivas funciones de transferencia.
- Procesamiento digital de señales.

El laboratorio Matricial (MatLab), en donde el fundamento principal es una matriz sin dimensiones, que permite resolver cálculos matriciales, los cuales son muy complejos para realizarlos manualmente, Matlab además contiene una serie de ToolBoxes que sirven para aplicaciones específicas, en la figura 2.10 se puede observar la imagen del software MatLab.

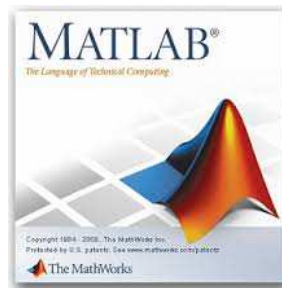


Figura 2.10: Software MatLab.

Fuente: (www.matlab.com).

CAPÍTULO 3: DISEÑO DE UN SISTEMA DE RECONOCIMIENTO DE VOZ.

3.1. Adquisición de la señal de voz.

El diseño de este proyecto considera la adquisición de la señal de voz por medio de un micrófono capacitivo conectado a una computadora, se utilizará la tarjeta de audio incorporada en la computadora para realizar la digitalización de la señal de voz, considerando que esta señal fluctúa entre los 100 Hz a los 3.5 KHz, se utilizará una frecuencia de muestreo de 11025 Hz lo cual cumple con el teorema de Nyquist visto en el capítulo anterior, donde:

$$F_s \geq 2 \cdot f_m$$

Para definir el tiempo de adquisición para cada comando de voz se han realizado varias pruebas y cuyos resultados se muestran en la tabla 3.1.

Tabla 3.1: Muestreo de tiempos para pronunciar un comando de voz.

	Persona 1 (seg)	Persona 2 (seg)	Persona 3 (seg)	Máximo (seg)
Luces	0.82	0.99	1.09	1.09
Puerta uno	1.17	1.34	1.57	1.57
Puerta dos	1.23	1.5	1.63	1.63
Válvula	1.2	1.04	1.56	1.56
Climatización	1.62	1.48	1.6	1.62
TV	0.7	0.8	0.85	0.85
Canal mas	1.31	1.25	1.01	1.31
Canal menos	1.21	1.4	1.51	1.51
Volumen mas	1.1	1.48	1.54	1.54
volumen menos	1.47	1.61	1.72	1.72
Tiempo máximo necesario				1.72

Elaborado por: El Autor.

En función de las pruebas realizadas se ha tomado la decisión de utilizar un tiempo de adquisición de 2 segundos; en la figura 3.1 se muestra la señal resultante digitalizada con una frecuencia de muestreo de 11025 Hz, a 8 Bits con una duración de 2 segundos.

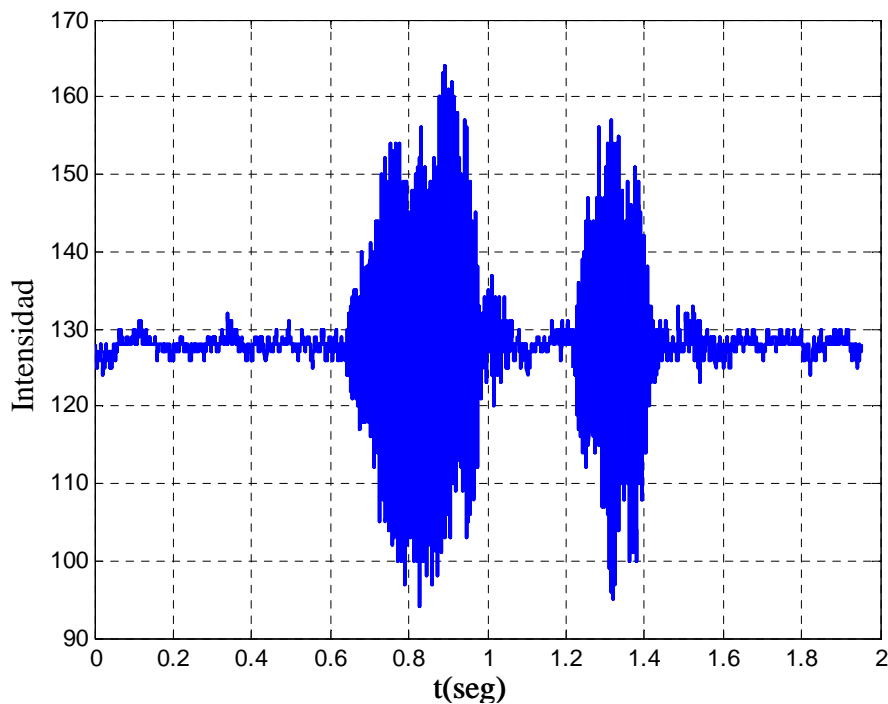


Figura 3.1: Señal de voz digitalizada.

Elaborado por: El Autor.

3.2. División de la señal de voz

Considerando que la frecuencia de muestreo es de 11025 Hz, se ha dividido en tramas de 200 muestras cada una, lo que da un tiempo total por trama de 18 mseg, además con el fin de extraer la energía de la señal en la siguiente etapa de ventaneo, se ha solapado las muestras en una cantidad de 100 para cada trama, lo que da un tiempo de solapamiento de 9 mseg; en la figura 3.2 se muestra una trama de la señal de voz.

Para calcular el total de tramas por señal de voz grabada se utiliza la siguiente fórmula:

$$N^{\circ} \text{ tramas} = \text{EnteroInf} \left[\frac{\text{Longitud del vector} - \text{muestras por trama}}{\text{muestras de solapamiento}} + 1 \right]$$

Donde:

Longitud del vector= cantidad de muestras de la señal de audio.

Muestras por trama= para el presente caso se ha tomado 200 muestras equivalente a 18 mseg

Muestras de solapamiento= Se ha tomado 100 muestras equivalentes a 9 mseg.

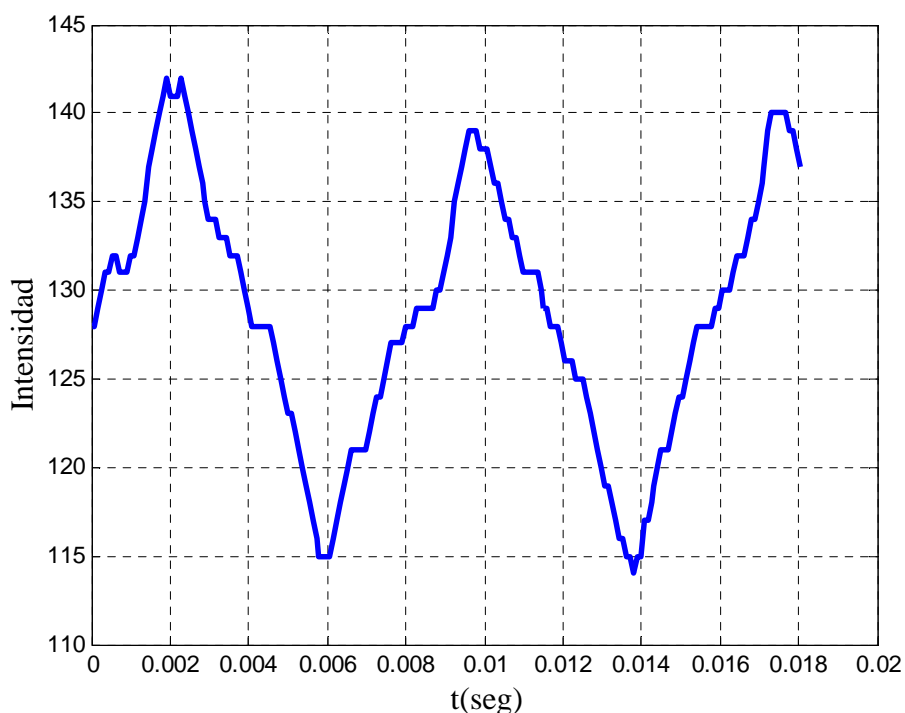


Figura 3.2: Trama de 18mseg.

Elaborado por: El Autor.

3.3. Ventaneo de la señal

Con el fin de extraer la energía central de cada una de las tramas obtenidas en el punto anterior, se ha aplicado una ventana de hamming con la cual se consigue un mejor resultado para el

procesamiento de señales de audio, la señal ventaneada es mostrada en la figura 3,3.

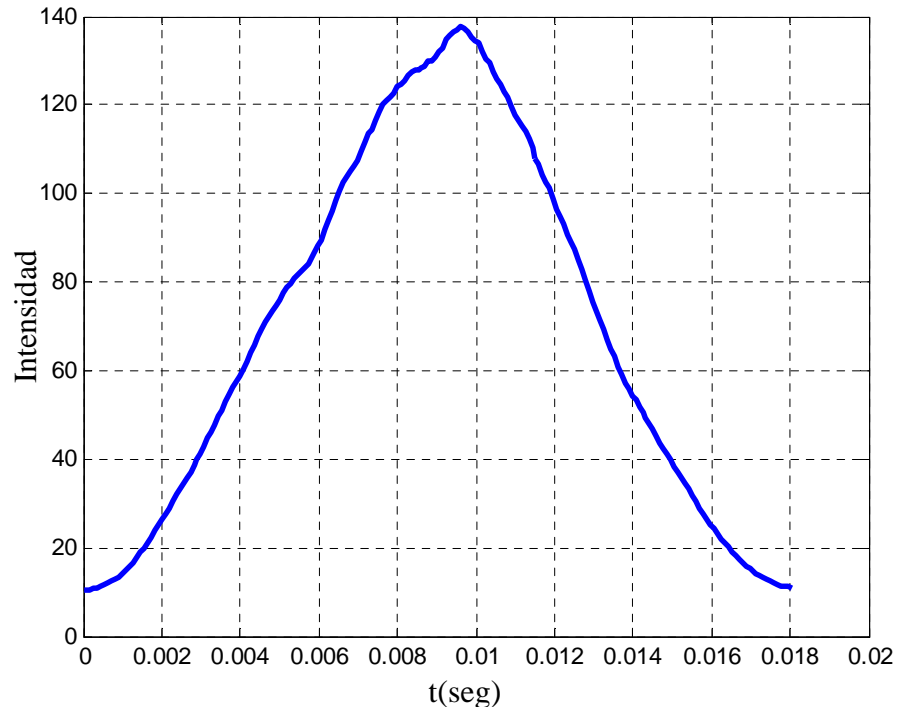


Figura 3.3: Trama de 18mseg ventaneada.

Elaborado por: El Autor.

3.4. Transformada rápida de Fourier.

Uno de los pasos previos para la extracción los coeficientes cepstrales de la escala de Mel, es pasar las señales de cada una de las tramas al dominio de la frecuencia mediante la transformada rápida de Fourier, para lo cual se aplica en MatLab la siguiente función:

`fft(X);`

En la figura 3.4 se muestra la transformada de Fourier correspondiente a la señal ventaneada de la trama mostrada en la figura 3.2

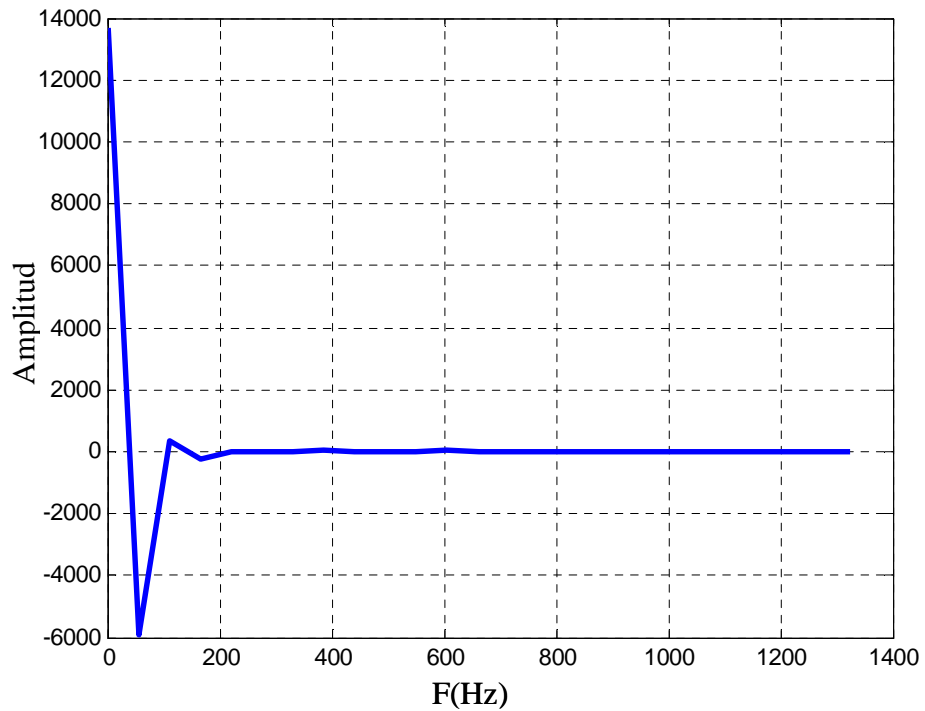


Figura 3.4: FFT de la señal ventaneada.

Elaborado por: El Autor.

3.5. Coeficientes Cepstrales

Antes de extraer los coeficientes cepstrales primero se necesita fabricar un banco de filtros de Mel, para lo cual existen varias técnicas, sin embargo para este proyecto se ha utilizado una función propia de MatLab:

$m = \text{melfb}(p, n, fs);$

donde:

p = cantidad de filtros a extraer.

n = números de muestras por trama

fs = frecuencia de muestreo.

Obteniendo el banco de filtros que se muestra en la figura 3.5.

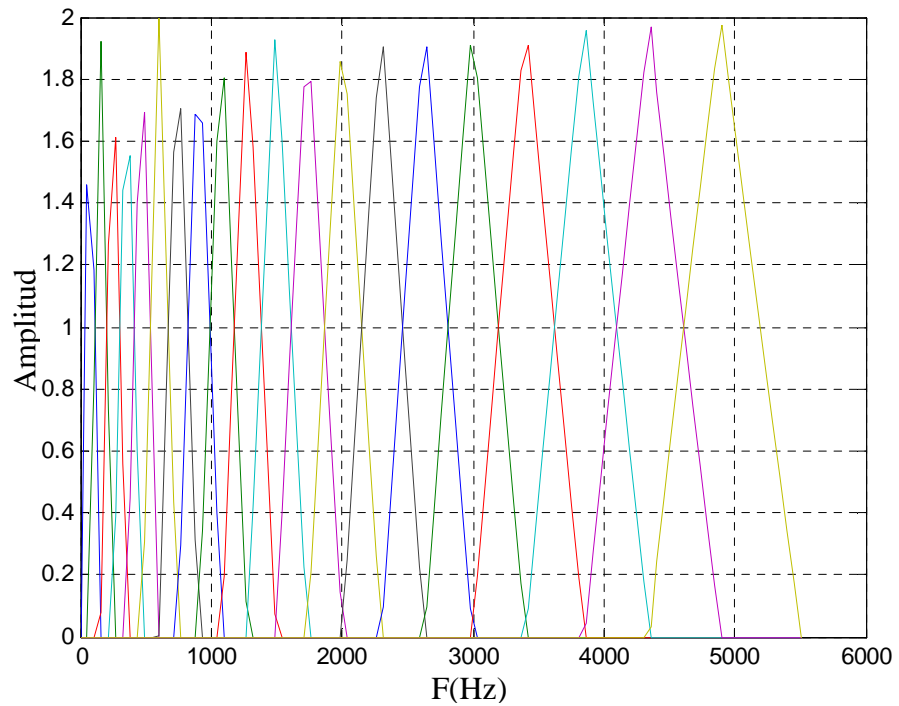


Figura 3.5: Banco de filtros de Mel para $p=20$, $n=200$ y $F_s=11025$.

Elaborado por: El Autor.

Los coeficientes cepstrales de la escala de Mel resultan al pasar cada una de las tramas en el dominio de la frecuencia por los filtros obtenidos anteriormente, en la figura 3.6 se muestra los coeficientes cepstrales de una de las tramas.

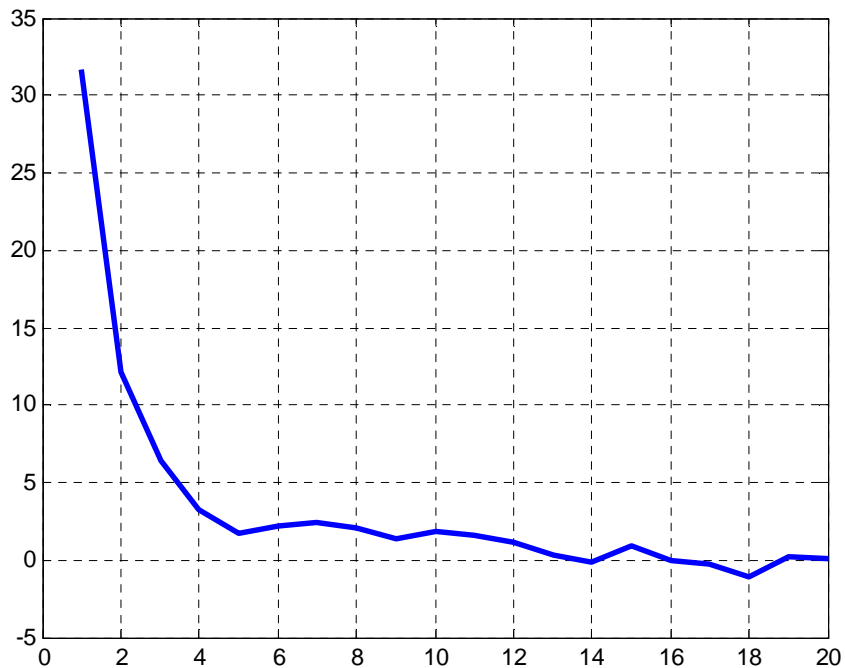


Figura 3.6: Coeficientes Cepstrales de una de las tramas.

Elaborado por: El Autor.

Con este procesamiento se obtiene una matriz por cada señal de voz, para nuestro proyecto el tamaño de la matriz es de 214 tramas de 20 valores cada una.

3.6. Método abreviado de MatLab.

Si bien es cierto los puntos 3.2 al 3.5 detallan los pasos a seguir para la extracción de los coeficientes cepstrales en cada una de sus etapas, en el presente trabajo de investigación también se realizó pruebas mediante un método abreviado que ofrece MatLab, con la función que se detalla a continuación obteniendo resultados similares:

Coeficientes=`melcepst(S,FS,W,NC,P,N,INC,FL,FH)`

Donde;

S Señal.

FS Frecuencia de Muestreo

W	Modo de las tramas (Para nuestro caso ventaneo de hamming)
NC	Número de coeficientes cepstrales
P	Número de filtros de la escala de MEL
N	Muestras de la trama
INC	Incremento de la trama
FL	Extremo inferior del filtro más bajo [default = 0]
FH	Extremo superior del filtro más alto [default = 0.5]

Para el presente trabajo de investigación la expresión en MatLab quedó de la siguiente manera.

```
Coeficientes= melcepst(s,Fs,'M',20,20, 200, 100);
```

3.7. Etapa de decisión

Una vez obtenidas cada una de las matrices de los coeficientes cepstrales de las señales de audio de referencia, se continúa con la fase de decisión para cual se calcula los CodeBook mediante el algoritmo de K-Means detallado en la tabla 2.1

Para el presente trabajo de investigación se utilizará 16 centroides de decisión conforme análisis detallado en el Anexo 1, valor con el cual se obtuvieron mejores resultados, además se ha establecido un umbral de error de 0.01

Con los CodeBook calculados, se realiza la medición de las distancias euclidianas entre la señal de voz entrante y cada uno de los CodeWord correspondientes a los patrones de señal previamente guardados, de la siguiente manera.

$$d = \sqrt{\sum (x_i - y_i)^2}$$

Con los resultados de las distancias entre la señal de voz y los CodeWord de cada uno de los patrones se procede a calcular la distancia mínima, con la cual se obtiene la señal reconocida.

3.8. Interfaz de usuario

La interfaz de la aplicación de reconocimiento de voz consta de cuatro etapas: etapa de alarmas y mensajes, etapa de configuración, etapa de reconocimiento de voz y etapa de control.

3.8.1. Etapa de alarmas y mensajes

La etapa de alarmas y mensajes mostrará al usuario las advertencias que se pudiesen ocasionar mientras se está usando la aplicación de reconocimiento de voz conforme se muestra en la figura 3.7

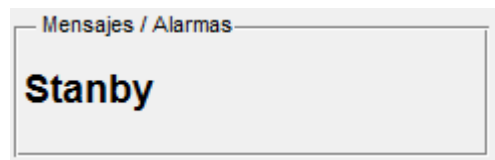


Figura 3.7: Etapa de mensajes y alarmas-interfaz de Reconocimiento de voz.

Fuente: Ing. Wilmer Jiménez

3.8.2. Etapa de configuración

Esta etapa deberá ser configurada con la asistencia de una persona no cuádrupléjica, en donde se podrán grabar los patrones de voz del usuario para cada uno de los comandos que se desea ejecutar como se puede ver en la figura 3.8.

Configuración

Identificador

Frase

Grabar Borrar Datos

Figura 3.8: Etapa de configuración de la interfaz de Reconocimiento de voz.

Elaborado por: El Autor.

En el cuadro “Identificador” se irá colocando de forma ascendente los números correspondientes a la secuencia de comandos de voz grabados; En la opción “Frase” se deberá colocar en letras el comando que se desea grabar, como se detalla en la figura 3.9.

Configuración

Identificador

Frase

Grabar Borrar Datos

Figura 3.9: Etapa de configuración de la interfaz de reconocimiento de voz-parámetros.

Elaborado por: El Autor.

Luego se deberá presionar el botón “Grabar” y en el cuadro de mensajes saldrá la palabra “Grabando”, momento en el cual el usuario deberá pronunciar el comando de voz en un tiempo de 2 segundos, el mensaje se muestra en la figura 3.10

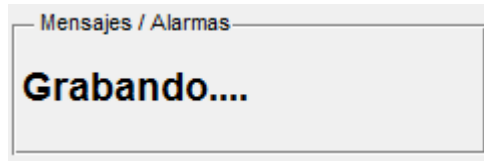


Figura 3.10: Mensaje “Grabando” de la interfaz de reconocimiento de voz.

Elaborado por: El Autor.

Este proceso se lo debe realizar con cada uno de los comandos de voz que se desea controlar, tomando en cuenta las siguientes consideraciones:

- No debemos duplicar los comandos de voz.
- No se requiere ingresar dos comandos para controlar un mismo elemento, es decir si la intención es abrir y cerrar una puerta, basta con decir “Puerta”, ya que el programa identificará la frase, y al pronunciarla por primera vez reconocerá como comando para abrir y al pronunciar por segunda vez reconocerá el comando para cerrar, con esto se logra optimizar el sistema disminuyendo los patrones de voz ingresados.
- Grabar los comandos con el mismo acento de voz con el cual se van a ejecutar la ordenes posteriormente.
- Los comandos de voz deben ser grabados por el usuario que va a utilizar el sistema.
- Utilizar comandos que se diferencien en su pronunciación.

3.8.3. Etapa de reconocimiento.

En esta etapa se puede iniciar el proceso de reconocimiento con el botón “Iniciar”, botón que tendrá que ser presionado por alguien que asista a la persona con cuádruplejia, una vez inicializado el sistema la persona con cuádruplejia podrá emitir los comandos de voz para

controlar los elementos de domótica que pudiesen ser conectados al sistema.

Con el botón reconocimiento apagado, se puede detener el sistema por motivos de mantenerlo en Stanby (Espera), mantenimiento o configuración; en el área gráfica se podrá visualizar la señal de voz y en la parte inferior se podrá observar la frase e identificador reconocido como se muestra en la figura 3.11.



Figura 3.11: Etapa de Reconocimiento de voz.

Elaborado por: El Autor.

Para que el usuario pueda iniciar un comando de voz es necesario una señal digital, para esto se ha conectado la interfaz de reconocimiento de voz a una tarjeta Arduino nano, con el fin de simular un sensor puesto cerca de la cabeza del usuario, con esto la persona con cuadriplejía podría iniciar los comandos de voz al hacer

un ligero movimiento de su cabeza, la tarjeta Arduino utilizada se muestra en la figura 3.12

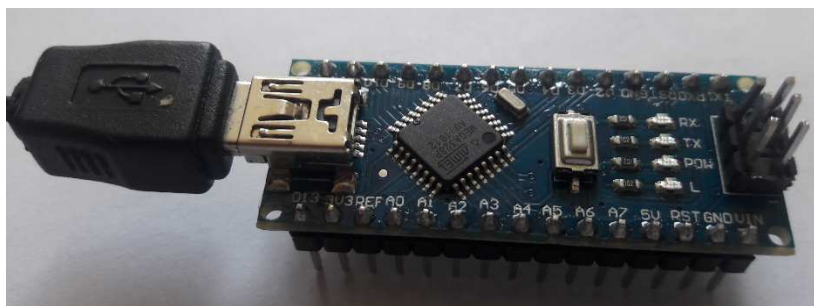


Figura 3.12: Tarjeta Arduino utilizada.

Fotografía tomada por: El Autor.

Se ha establecido una comunicación serial entre la tarjeta Arduino y MatLab, en donde la tarjeta Arduino envía un carácter “A” a MatLab, cada vez que se presiona el botón incorporado en la tarjeta, el alcance de este trabajo de investigación es el diseño de la interfaz de usuario para el reconocimiento de comandos de voz, sin embargo se deja establecida la comunicación con una tarjeta externa en donde se podrán colocar las salidas a los elementos de domótica según sea la aplicación específica que se desee desarrollar.

3.8.4. Etapa de Control.

Esta etapa le permite al usuario observar el estado de cada uno de los elementos programados conforme se muestra en la figura 3.13



Figura 3.13: Etapa de control.

Elaborado por: El Autor.

Los textos de los elementos se cargan automáticamente en función de las frases que fueron previamente grabadas en la etapa de configuración, los cuadros de colores representan el estado de los elementos de domótica, verde significa activado, rojo significa desactivado y negro significa que no se ha utilizado, el sistema se ha configurado para un máximo de 12 elementos con el fin de lograr efectividad en el mismo.

La interfaz de usuario completa se muestra en la figura 3.14 en donde se puede visualizar cada una de las etapas.

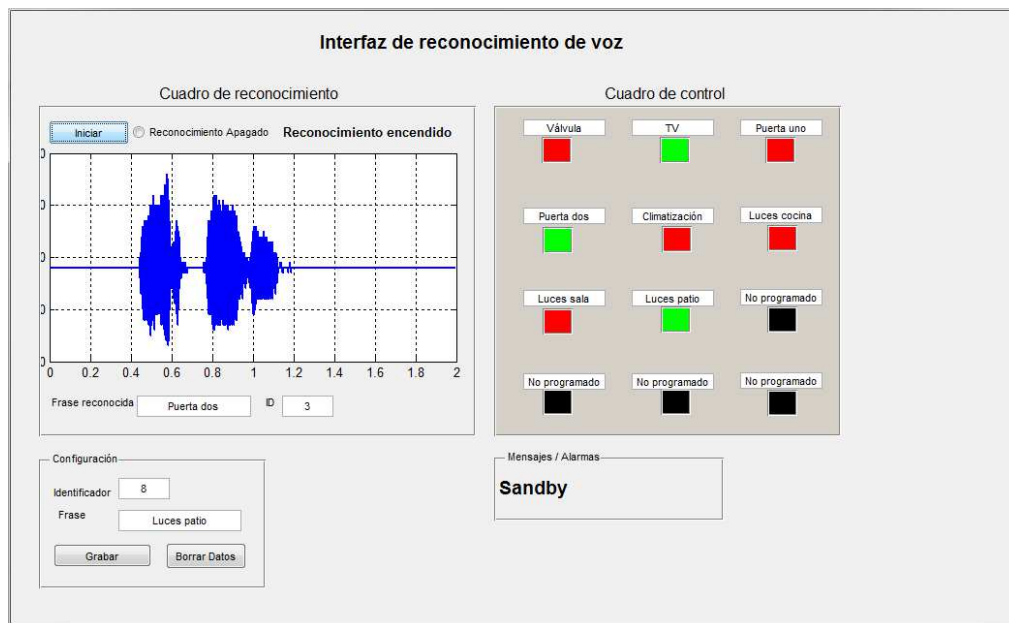


Figura 3.14: Interfaz de reconocimiento de voz.

Elaborado por: El Autor.

3.9. Interfaz con una tarjeta serial.

Como se mencionó anteriormente el proyecto comprende el desarrollo de una interfaz computarizada controlada por voz para cuadripléjicos, sin embargo se deja establecida la comunicación serial con una tarjeta externa para el control del sistema.

La configuración de la comunicación serial se la realiza a 9600 baudios, con un tiempo máximo de espera de 2 segundos y utilizando el puerto serial COM4 de la computadora, si el sistema utiliza otro puerto serial, este debe ser modificado en el programa.

En la tabla 3.2 se muestra los comandos seriales utilizados para el control del sistema tanto de transmisión como de recepción.

Tabla 3.2: Comandos seriales para el control del sistema.

Comando serial	Tipo	Función
A	Recepción	Iniciar el reconocimiento de un comando de voz
101	Transmisión	Activar la salida 1 (correspondiente al patrón de voz 1)
201	Transmisión	Desactivar la salida 1 (correspondiente al patrón de voz 1)
102	Transmisión	Activar la salida 2 (correspondiente al patrón de voz 2)
202	Transmisión	Desactivar la salida 2 (correspondiente al patrón de voz 2)
103	Transmisión	Activar la salida 3 (correspondiente al patrón de voz 3)
203	Transmisión	Desactivar la salida 3 (correspondiente al patrón de voz 3)
104	Transmisión	Activar la salida 4 (correspondiente al patrón de voz 4)
204	Transmisión	Desactivar la salida 4 (correspondiente al patrón de voz 4)
105	Transmisión	Activar la salida 5 (correspondiente al patrón de voz 5)
205	Transmisión	Desactivar la salida 5 (correspondiente al patrón de voz 5)
106	Transmisión	Activar la salida 6 (correspondiente al patrón de voz 6)
206	Transmisión	Desactivar la salida 6 (correspondiente al patrón de voz 6)
107	Transmisión	Activar la salida 7 (correspondiente al patrón de voz 7)
207	Transmisión	Desactivar la salida 7 (correspondiente al patrón de voz 7)
108	Transmisión	Activar la salida 8 (correspondiente al patrón de voz 8)
208	Transmisión	Desactivar la salida 8 (correspondiente al patrón de voz 8)
109	Transmisión	Activar la salida 9 (correspondiente al patrón de voz 9)
209	Transmisión	Desactivar la salida 9 (correspondiente al patrón de voz 9)
110	Transmisión	Activar la salida 10 (correspondiente al patrón de voz 10)
210	Transmisión	Desactivar la salida 10 (correspondiente al patrón de voz 10)
111	Transmisión	Activar la salida 11 (correspondiente al patrón de voz 11)
211	Transmisión	Desactivar la salida 11 (correspondiente al patrón de voz 11)
112	Transmisión	Activar la salida 12 (correspondiente al patrón de voz 12)
212	Transmisión	Desactivar la salida 12 (correspondiente al patrón de voz 12)

Elaborado por: El Autor.

Al salida 1 corresponde al primer patrón de voz guardado en el sistema, además es el primer elemento que se muestra en el cuadro de control, estos comandos seriales pueden ser utilizados con una tarjeta Arduino o cualquier otra interfaz serial configurada con 9600 baudios, y mediante los mismos se pueden comandar elementos de domótica en el domicilio.

3.10. Etapas del programa.

Con el objetivo de que la interfaz de reconocimiento de voz pueda ser utilizada para desarrollar proyectos de domótica para ayudar a las personas con cuádruplejía, a continuación se detalla cada una de las etapas del programa desarrollado en MatLab, este programa puede ser utilizado sin ninguna modificación conectando una tarjeta serial al puerto COM4 de la computadora y adaptándola a la aplicación deseada, sin embargo de existir alguna aplicación en la que se requiera la modificación de la parte estructural del programa se detalla cada uno de los bloques del mismo.

3.10.1. Etapa de inicialización

En esta etapa se realiza una lectura de las bases de datos de los estados de cada actuador y se muestra los resultados en el cuadro de control, con esta etapa logramos que el usuario pueda observar los estados de los actuadores apenas inicie la aplicación, en la figura 3.15 se muestra la parte de programa donde comienza la etapa de inicialización.

```
    % INICIALIZACION%%.
function varargout = Tesis_OutputFcn(hObject, eventdata, handles)
% varargout cell array for returning output args (see VARARGOUT);
% hObject handle to figure
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)
% INICIALIZACION%%
varargout{1} = handles.output;
load('BD.dat', '-mat');
```

Figura 3.15: Comienzo de la etapa de inicialización.

Elaborado por: El Autor.

3.10.2. Etapa de grabado de patrones.

La grabación de cada uno de los patrones de voz del usuario se lo realiza en esta etapa, para cual se utiliza una base de datos llamada BD.DAT, además los comandos de voz son grabados de manera independiente en la carpeta del programa, en la figura 3.16 se muestra la parte de programa donde comienza la etapa de inicialización.

```
%%ETAPA GRABADO DE PATRONES%%.  
function pushbutton1_Callback(hObject, eventdata, handles)  
% hObject    handle to pushbutton1 (see GCBO)  
% eventdata  reserved - to be defined in a future version of MATLAB  
% handles    structure with handles and user data (see GUIDATA)  
% Función grabar.
```

Figura 3.16: Comienzo de la etapa de grabado de patrones.

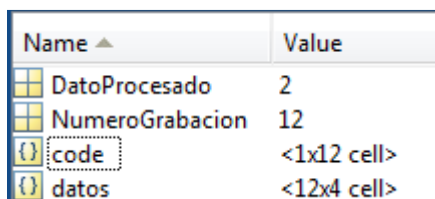
Elaborado por: El Autor.

La base de datos DB.DAT posee la siguiente información:

- **Datos.-** Contiene la siguiente información
 - Vector de la señal de audio
 - Identificador
 - Frase
 - Estado (Activado/Desactivado)
- **Code.-** contiene las matrices de los CodeBook de los patrones guardados.
- **NumeroGrabación.-** Variable auxiliar que indica la cantidad de patrones grabados.
- **DatoProcesado.-** Variable auxiliar que indica si se han extraído o no los CodeBook de los patrones guardados 1 significa que no han sido extraídos y 2 significa que ya han sido extraídos, con esto se optimiza el

funcionamiento del programa ahorrando tiempos de procesamiento.

En la figura 3.17 se muestra la estructura de la base de datos DB.DAT visualizada en MatLab.



Name	Value
DatoProcesado	2
NumeroGrabacion	12
code	<1x12 cell>
datos	<12x4 cell>

Figura 3.17: Estructura de DB.DAT

Elaborado por: El Autor.

3.10.3. Etapa de borrado.

Cuando se requiere borrar la base de datos de los patrones y volver a configurar el sistema se utiliza esta etapa, además cuando se tiene problemas con el reconocimiento de voz del usuario, se puede volver a grabar los patrones del mismo usuario para lo cual primero debemos borrar la base de datos, en la figura 3.18 se visualiza el inicio de la etapa de borrado.

```
%% ETAPA DE BORRADO %  
function pushbutton3_Callback(hObject, eventdata, handles)  
% hObject handle to pushbutton3 (see GCBO)  
% eventdata reserved - to be defined in a future version of MATLAB  
% handles structure with handles and user data (see GUIDATA)  
%funcion para borrar la base de datos
```

Figura 3.18: Inicio de la etapa de borrado

Elaborado por: El Autor.

3.10.4. Etapa de reconocimiento.

Esta es la etapa principal del programa, donde se realiza el procesamiento de las señales de voz y la respectiva comparación con los patrones previamente guardados, con el fin de obtener los resultados esperados, además en esta etapa se configura la comunicación serial con la tarjeta externa, en la figura 3.19 se muestra el inicio de la etapa de reconocimiento.

```
    % ETAPA DE RECONOCIMIENTO%
function pushbutton4_Callback(hObject, eventdata, handles)
% hObject      handle to pushbutton4 (see GCBO)
% eventdata    reserved - to be defined in a future version of MATLAB
% handles      structure with handles and user data (see GUIDATA)
```

Figura 3.19: Inicio de la etapa de reconocimiento.

Elaborado por: El Autor.

Los pasos que se siguen en esta etapa de programa son los siguientes:

- Apertura y configuración del puerto serial
- Espera de comando para iniciar el reconocimiento.
- Adquisición de la señal de voz
- Procesamiento de la señal
- Medición de las distancias euclidianas
- Obtención del resultado
- Envío de los comandos seriales hacia una tarjeta externa
- Gráfica de los resultados en la interfaz de usuario.

Para el correcto funcionamiento de esta etapa se utilizan las siguientes funciones:

- **melcepst:** Función propiedad de MatLab para la obtención de los coeficientes cepstrales.
- **CODEBOOK:** obtención de los CodeBook de cada una de las señales.
- **DistanciasEu:** medición de la distancia euclidiana entre dos vectores, esta es una función que se encuentra alojada en la carpeta de programa.
- **Gráficas:** Escritura en el panel de control del usuario.

CAPÍTULO 4: PRUEBAS DE ROBUSTEZ DEL SISTEMA DE RECONOCIMIENTO DE VOZ.

4.1. Consideraciones iniciales.

Para realizar las pruebas funcionales de la interfaz de reconocimiento de voz se ha escogido los siguientes comandos de voz, los cuales pueden cumplir las necesidades básicas de una persona con cuadriplejía:

1. Válvula (Abrir o cerrar una válvula de agua)
2. TV (Encender o apagar el TV)
3. Puerta uno (Abrir/cerrar una puerta)
4. Puerta dos (Abrir/cerrar una puerta)
5. Climatización (Encender/apagar la climatización)
6. Luces cocina (Encender/apagar luces)
7. Luces sala (Encender/apagar luces)
8. Luces Cuarto (Encender/apagar luces)
9. Ventana (Abrir/cerrar una ventana)
10. Silla (Activar/Desactivar el movimiento de la silla de ruedas)
11. Izquierda (Girar a la izquierda la silla de ruedas)
12. Derecha (Girar a la derecha la silla de ruedas)

Se ha configurado el sistema para cuatro usuarios diferentes, obteniendo la distribución que se observa en la figura 4.1, cabe mencionar que el sistema puede ser configurado con un usuario a la vez, luego de terminar las pruebas con usuario se realiza el borrado de la base de datos y se vuelve a configurar para el siguiente usuario.

Las pruebas fueron realizadas dentro de una casa con los niveles de ruido estándares que existen en los domicilios, simulando el escenario real que tendría la persona con cuadriplejía en su casa.

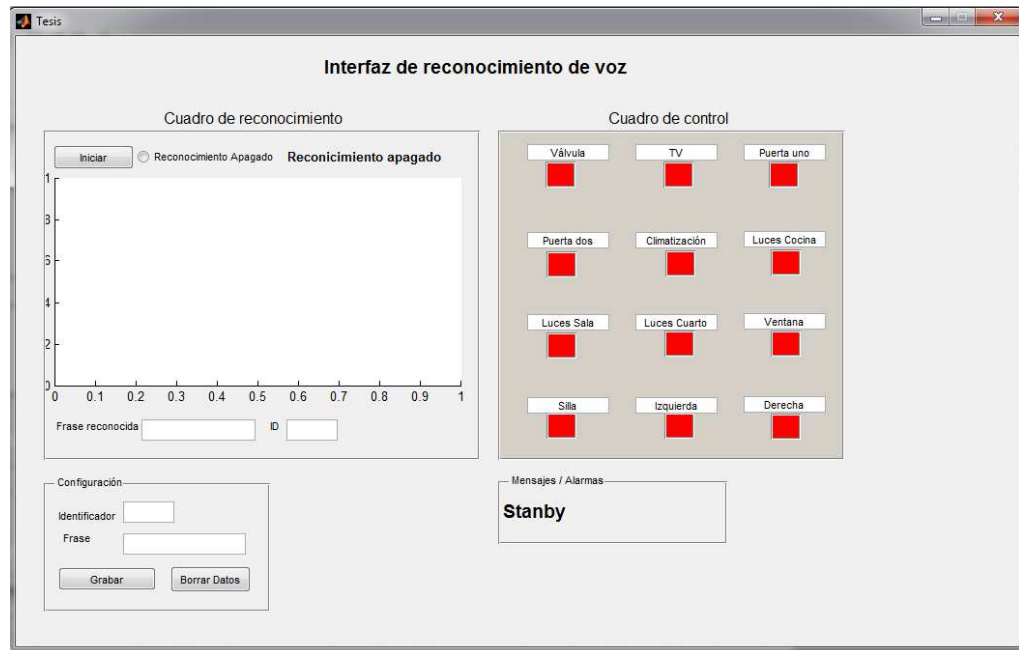


Figura 4.1: Configuración para pruebas de robustez.

Elaborado por: El Autor.

Para las pruebas funcionales se ha tomado a dos usuarios de sexo femenino y a dos usuarios de sexo masculino con el fin de obtener el porcentaje de efectividad más cercano a la realidad:

- Usuario 1: Hombre de 30 años de edad.
- Usuario 2: Hombre de 39 años de edad.
- Usuario 3: Mujer de 26 años de edad.
- Usuario 4: Mujer de 24 años de edad.

4.2. Resultados de la pruebas funcionales.

En las tablas 4.1, 4.2, 4.3 y 4.4 se detallan los resultados de las pruebas realizadas a los cuatro usuarios respectivamente, en donde se repite 3 veces el mismo comando, con el fin de verificar la robustez del mismo, los resultados pueden ser los siguientes:

Positivo: significa que la aplicación ha reconocido correctamente el comando de voz.

Negativo: significa que la aplicación no ha reconocido el comando de voz emitido por el usuario.

Tabla 4.1: Pruebas de robustez con el usuario 1.

Resultados de las pruebas del usuario 1				
Ítem	Comando	Intento 1	Intento 2	Intento 3
1	Válvula	Positivo	Positivo	Positivo
2	TV	Positivo	Positivo	Positivo
3	Puerta uno	Positivo	Positivo	Positivo
4	Puerta dos	Positivo	Positivo	Positivo
5	Climatización	Positivo	Positivo	Positivo
6	Luces cocina	Positivo	Positivo	Positivo
7	Luces sala	Positivo	Positivo	Positivo
8	Luces cuarto	Positivo	Positivo	Positivo
9	Ventana	Positivo	Positivo	Positivo
10	Silla	Positivo	Positivo	Positivo
11	Izquierda	Positivo	Positivo	Positivo
12	Derecha	Positivo	Positivo	Positivo

Elaborado por: El Autor.

Tabla 4.2: Pruebas de robustez con el usuario 2.

Resultados de las pruebas del usuario 2				
Ítem	Comando	Intento 1	Intento 2	Intento 3
1	Válvula	Positivo	Positivo	Positivo
2	TV	Positivo	Positivo	Positivo
3	Puerta uno	Positivo	Positivo	Positivo
4	Puerta dos	Positivo	Positivo	Positivo
5	Climatización	Positivo	Positivo	Positivo
6	Luces cocina	Positivo	Negativo	Positivo
7	Luces sala	Positivo	Positivo	Positivo
8	Luces cuarto	Positivo	Positivo	Positivo
9	Ventana	Positivo	Positivo	Positivo
10	Silla	Positivo	Positivo	Positivo
11	Izquierda	Positivo	Positivo	Positivo
12	Derecha	Positivo	Positivo	Positivo

Elaborado por: El Autor.

Tabla 4.3: Pruebas de robustez con el usuario 3.

Resultados de las pruebas del usuario 3				
Ítem	Comando	Intento 1	Intento 2	Intento 3
1	Válvula	Positivo	Positivo	Positivo
2	TV	Positivo	Positivo	Positivo
3	Puerta uno	Positivo	Positivo	Positivo
4	Puerta dos	Positivo	Positivo	Positivo
5	Climatización	Positivo	Positivo	Positivo
6	Luces cocina	Positivo	Positivo	Positivo
7	Luces sala	Positivo	Positivo	Positivo
8	Luces cuarto	Positivo	Positivo	Negativo
9	Ventana	Positivo	Positivo	Positivo
10	Silla	Positivo	Positivo	Positivo
11	Izquierda	Positivo	Positivo	Positivo
12	Derecha	Positivo	Positivo	Positivo

Elaborado por: El Autor.

Tabla 4.4: Pruebas de robustez con el usuario 4.

Resultados de las pruebas del usuario 4				
Ítem	Comando	Intento 1	Intento 2	Intento 3
1	Válvula	Positivo	Positivo	Positivo
2	TV	Positivo	Positivo	Positivo
3	Puerta uno	Positivo	Positivo	Positivo
4	Puerta dos	Positivo	Positivo	Positivo
5	Climatización	Positivo	Positivo	Positivo
6	Luces cocina	Positivo	Positivo	Positivo
7	Luces sala	Positivo	Positivo	Positivo
8	Luces cuarto	Positivo	Positivo	Positivo
9	Ventana	Positivo	Positivo	Positivo
10	Silla	Positivo	Positivo	Positivo
11	Izquierda	Positivo	Positivo	Positivo
12	Derecha	Positivo	Positivo	Positivo

Elaborado por: El Autor.

En la tabla 4.5 se detalla el resumen de las pruebas realizadas con los 4 usuarios diferentes, en donde se realizó un total de 144 pruebas, 36 con cada usuario, con lo cual se obtuvo 142 resultados

positivos y 2 resultados negativos de tal manera que se puede realizar el siguiente cálculo:

$$\%efectividad = \frac{142}{(142 + 2)} \cdot 100\%$$

$$\%efectividad = 98,6\%$$

Tabla 4.5: Resumen de las pruebas de robustez.

	Usuario 1	Usuario 2	Usuario 3	Usuario 4	Total
Resultados Positivos	36	35	35	36	142
Resultados negativos	0	1	1	0	2

Elaborado por: El Autor.

En función a los resultados anteriores se puede mencionar que el sistema de reconocimiento de voz desarrollado en el presente trabajo de investigación ofrece un 98,6% de efectividad.

4.3. Análisis de resultados.

La interfaz desarrollada en el presente trabajo de investigación presenta un 98,6% de efectividad, lo que significa que de 1000 comandos de voz 14 fallarían, sin embargo en las pruebas anteriores se pudo observar que los comandos que dieron resultados negativos están en el grupo de comandos con similitud entre sí, los cuales son los siguientes:

- Luces cocina
- Luces sala
- Luces cuarto

Estos tres comandos tienen una palabra en Común “Luces”, lo cual hace que los sistemas de reconocimiento de voz tengan dificultades para dar un resultado positivo, esto se puede mejorar al

modificar los comandos de voz, de tal manera que no contengan palabras en común; un ejemplo puede ser el siguiente:

- Cocina
- Sala
- Cuarto

Eliminando la palabra “Luces” conseguiremos que el sistema de reconocimiento de voz responda de una mejor manera conforme lo esperado.

CAPÍTULO 5: CONCLUSIONES Y RECOMENDACIONES.

5.1. Conclusiones

- Las personas con cuadriplejia se encuentran limitadas a interactuar con electrodomésticos, luces, puertas, ventanas y demás elementos en su domicilio, debido a la incapacidad de mover sus extremidades, por lo tanto se pudo identificar la necesidad de interactuar con dichos elementos mediante comandos de voz.
- El método utilizado para el reconocimiento de voz “Extracción de los coeficientes cepstrales de la escala de Mel”, en la presente interfaz presenta un 98,6% de efectividad, conforme las pruebas realizadas, dicho porcentaje de efectividad puede ser mejorado aún más al elegir comandos que difieran en su pronunciación.
- La interfaz desarrollada en el presente trabajo de investigación puede ser utilizada directamente al conectarle una tarjeta serial de salidas digitales al puerto COM4 con 9600 baudios de velocidad de comunicación y programando a la tarjeta para la aplicación específica, sin embargo el programa es fácilmente modificable si se requiere alguna aplicación en particular.
- La interfaz desarrollada en el presente trabajo de investigación ayudará a que las personas con cuadriplejia, disminuyan la dependencia hacia otras personas y puedan interactuar con sus electrodomésticos, luces, puertas, ventanas y demás elementos del domicilio.

5.2. Recomendaciones

- Aplicar la interfaz desarrollada para proyectos de domótica enfocado a personas con capacidades especiales, quienes necesiten controlar sus electrodomésticos, luces, puertas, ventanas o demás elementos en sus domicilios mediante comandos de voz.
- Fomentar la investigación de sistemas de reconocimiento de reconocimiento de voz, imágenes y ondas cerebrales, con el fin de desarrollar sistemas que ayuden a la sociedad.

REFERENCIAS BIBLIOGRÁFICAS

- Cordovilla, M., & Andrés, J. (2012). *Técnicas de reconocimiento robusto de la voz basadas en el pitch*. Granada: Universidad de Granada. Retrieved from <http://digibug.ugr.es/handle/10481/19962>
- Guajardo, G., & Guillermo, J. (2014, July 11). *Sistema de reconocimiento de voz usando perceptrón multicapa y Coeficientes Cepstrales de Mel*. (Thesis). Retrieved from <http://ri.uaq.mx/handle/123456789/1277>
- Jiménez Vivanco, W. E., & Ochoa Coronel, D. F. (2010). Estudio y diseño de factibilidad para la implementación de un laboratorio de procesamiento de señales y simulación para la Universidad Politécnica Salesiana Sede Cuenca. Retrieved from <http://localhost:8080/xmlui/handle/123456789/2206>
- Maya, J., Carrera, E., & Tobar, J. (n.d.). Detección del Pitch en MATLAB y mediante una FPGA. Retrieved August 4, 2016, from https://www.academia.edu/9818010/Detecci%C3%B3n_del_Pitch_en_MATLAB_y_mediante_una_FPGA
- Moreno-Fergusson, M. E., & Rey, M. C. del P. A. (2012). Cuerpo y corporalidad en la paraplejia: significado de los cambios. *Avances en Enfermería*, 30(1), 82–94.
- Moros, M. (2010). Caracterización de señales de precipitación mediante la transformada de Fourier y transformada Wavelet. Retrieved from <http://repository.javeriana.edu.co/handle/10554/6289>
- NewYork-Presbyterian/Queens - Lesiones Agudas de la Médula Espinal. (n.d.). Retrieved August 4, 2016, from

<http://www.nyhq.org/diw/Content.asp?PageID=DIW000770&language=Spanish>

Patiño, D., & Gualotuña, M. (2013). Participación de la familia en la atención domiciliaria al paciente del área de traumatología del Hospital Carlos Andrade Marín con paraplejia o cuadriplejia y su relación con la incidencia de complicaciones. Enero 2009 a Marzo 2011. Retrieved from <http://www.dspace.uce.edu.ec/handle/25000/1081>

Ramírez, M. (2013). T E S I S. Cuantificación Vectorial de Imágenes con base en Memorias Asociativas Extendidas. Retrieved August 27, 2016, from <http://docplayer.es/7793265-T-e-s-i-s-cuantificacion-vectorial-de-imagenes-con-base-en-memorias-asociativas-extendidas-universidad-tecnologica-de-la-mixteca.html>

Rodríguez, P. (2008). Sistema de reconocimiento del locutor basado en modelado no paramétrico. Retrieved from <http://tesis.ipn.mx:8080/xmlui/handle/123456789/15542>

Sánchez, M., & Manuel, C. (2010). Modelado de la cualidad de la voz para la síntesis del habla expresiva. *TDX (Tesis Doctorals en Xarxa)*. Retrieved from <http://www.tdx.cat/handle/10803/9145>

Segura, D. la V., & Elena, L. (2007). Diseño de un sintetizador de voz del idioma español hablado en México. Retrieved from 132.248.52.100:8080/xmlui/handle/132.248.52.100/1915

T E S I S. Cuantificación Vectorial de Imágenes con base en Memorias Asociativas Extendidas UNIVERSIDAD TECNOLÓGICA DE LA MIXTECA. (n.d.). Retrieved August 27, 2016, from <http://docplayer.es/7793265-T-e-s-i-s->

cuantificacion-vectorial-de-imagenes-con-base-en-memorias-asociativas-
extendidas-universidad-tecnologica-de-la-mixteca.html

Varela Serrano, O. (2011). *Técnicas de análisis, caracterización y detección de señales de voz en entornos acústicos adversos* (phd). E.T.S.I.

Telecomunicación (UPM). Retrieved from <http://oa.upm.es/11101/>

Wainschenker, R., Doorn, J. H., Castro, M., & Legrottaglie, C. F. (2003). Cálculo y análisis del pitch en señales sonoras de voz humana. Presented at the V

Workshop de Investigadores en Ciencias de la Computación. Retrieved from

<http://hdl.handle.net/10915/21513>

GLOSARIO DE TERMINOS

FFT. - Fast Fourier transform

LPC. - Linear Predictive coding

Fm. -Frecuencia de muestreo

Pitch.- Se ha definido como frecuencia fundamental de espectro de frecuencias del habla.

ANEXOS

Pruebas para definir el número de centroides a utilizar.

Con el fin de definir la cantidad adecuada de centroides a utilizar en la etapa de decisión se ha realizado pruebas con distintas frases y con distintos números de centroides, donde se ha tomado las siguientes consideraciones:

1. Se ha configurado la interfaz tal como se muestra en la Figura A1.1:



Figura A1.2: Configuración para pruebas de números de centroides.

Elaborado por: El Autor.

2. Se ha grabado un audio de voz con una palabra determinada (por ejemplo "Izquierda" para la prueba 1), y este mismo comando es evaluado con la configuración de 16, 32, 64 y 128 centroides.
3. Se ha evaluado las distancias euclidianas en cada uno de los casos, en donde la distancia mínima corresponde al patrón de la frase reconocida.

4. Se calculó el % de diferencia entre el ID reconocido y la distancia más cercana al ID reconocido, con la siguiente fórmula:

$$\%Diferencia = \frac{Dist.ID reconocido - Dist.ID cercano}{Dist.ID reconocido} \cdot 100\%$$

5. Se realizó las pruebas desde 16 hasta 128 centroides debido a que con 4 centroides se dieron errores de reconocimiento y con más de 128 centroides se obtienen tiempos de procesamiento muy altos.

Tabla A1.1: Prueba 1-Frase "Izquierda" /ID: 11.

Palabra:	Izquierda	ID:	11		
ID frase	Distancias desde la frase a los patrones				
	16 centroides	32 centroides	64 centroides	128 centroides	
1	0,73143	0,70925	0,65477	0,63797	
2	0,70747	0,66875	0,65376	0,64427	
3	0,75582	0,70381	0,65332	0,61042	
4	0,6474	0,60241	0,57886	0,55781	
5	0,60133	0,53567	0,5053	0,49356	
6	0,58996	0,52968	0,52039	0,50865	
7	0,59358	0,56426	0,53279	0,51962	
8	0,61483	0,58881	0,55273	0,5388	
9	0,65039	0,60898	0,57534	0,55917	
10	0,5272	0,48989	0,47992	0,47544	
11	0,50143	0,48354	0,46216	0,45433	
12	0,61259	0,59318	0,57131	0,5594	
Distancia a ID:11	0,50143	0,48354	0,46216	0,45433	
Distancia ID cercano	0,5272	0,48989	0,47992	0,47544	
%Diferencia	5,1%	1,3%	3,8%	4,6%	

Elaborado por: El Autor.

Tabla A1.2: Prueba 2-Frase “Climatización” /ID: 5.

Palabra:	Climatización	ID:	5		
ID frase	Distancias desde la frase a los patrones				
	16 centroides	32 centroides	64 centroides	128 centroides	
1	0,88924	0,87269	0,82464	0,79272	
2	0,98377	0,89925	0,88026	0,86709	
3	0,93203	0,85831	0,80656	0,77045	
4	0,81249	0,77162	0,72612	0,71514	
5	0,73512	0,68779	0,65829	0,63813	
6	0,76353	0,70263	0,67775	0,65671	
7	0,79219	0,79672	0,75508	0,73612	
8	0,80444	0,75975	0,73472	0,71641	
9	0,87562	0,8358	0,78079	0,77248	
10	0,82977	0,78279	0,74253	0,72595	
11	0,82725	0,78009	0,74701	0,72419	
12	0,79977	0,76535	0,75236	0,74351	
Distancia a ID:5	0,73512	0,68779	0,65829	0,63813	
Distancia ID cercano	0,76353	0,70263	0,67775	0,65671	
%Diferencia	3,9%	2,2%	3,0%	2,9%	

Elaborado por: El Autor.

Tabla A1.3: Prueba 3-Frase “Puerta Uno” /ID: 3.

Palabra:	Puerta Uno	ID:	3		
ID frase	Distancias desde la frase a los patrones				
	16 centroides	32 centroides	64 centroides	128 centroides	
1	0,67202	0,64713	0,61497	0,60429	
2	0,91397	0,84836	0,83176	0,80632	
3	0,58855	0,55252	0,53706	0,51956	
4	0,6143	0,57125	0,53741	0,52189	
5	0,71314	0,69235	0,6602	0,62577	
6	0,7329	0,67623	0,66476	0,64432	
7	0,6479	0,61528	0,59594	0,57962	
8	0,71413	0,6784	0,64413	0,62982	
9	0,66501	0,63225	0,60845	0,58998	
10	0,81897	0,75517	0,69507	0,66759	
11	0,78359	0,7175	0,68089	0,65446	
12	0,74418	0,7163	0,67854	0,65954	
Distancia a ID:3	0,58855	0,55252	0,53706	0,51956	
Distancia ID cercano	0,6143	0,57125	0,53741	0,52189	
%Diferencia	4,4%	3,4%	0,1%	0,4%	

Elaborado por: El Autor.

Tabla A1.4: Prueba 4-Frase “Luces Cuarto” /ID: 8.

Palabra:	Luces Cuarto	ID:	8		
ID frase	Distancias desde la frase a los patrones				
	16 centroides	32 centroides	64 centroides	128 centroides	
1	1,0335	1,0137	0,94185	0,92419	
2	1,2942	1,2095	1,1896	1,1655	
3	1,0374	0,95563	0,8813	0,82327	
4	0,9411	0,87426	0,82858	0,80582	
5	0,96821	0,93135	0,90131	0,87418	
6	0,95594	0,87351	0,85167	0,82203	
7	0,9189	0,82756	0,78781	0,76657	
8	0,87119	0,81657	0,7696	0,73511	
9	1,0958	1,0531	0,99387	0,95452	
10	1,1698	1,093	1,0176	1,0023	
11	1,1278	1,041	1,0017	0,96752	
12	1,0736	1,0242	0,98066	0,95323	
Distancia a ID:8	0,87119	0,81657	0,7696	0,73511	
Distancia ID cercano	0,9189	0,82756	0,78781	0,76657	
%Diferencia	5,5%	1,3%	2,4%	4,3%	

Elaborado por: El Autor.

Se ha podido identificar que la configuración con 16 centroides presenta una mayor separación porcentual entre la distancia euclidiana del ID reconocido y la del ID del patrón más cercano, con lo que presenta mayores ventajas en el momento del reconocimiento, evitando posibles errores en la etapa de decisión, por tal razón se utiliza la configuración de 16 centroides en el diseño de la interfaz.

DECLARACIÓN Y AUTORIZACIÓN

Yo, **Wilmer Enrique Jiménez Vivanco**, con C.C: # 070487990-7 autor del trabajo de titulación: Modelaje de una interfaz computarizada controlada por voz para cuadripléjicos que contribuya al desarrollo de aplicaciones que permitan comandar elementos de domótica en su domicilio, previo a la obtención del título de **Magister en Telecomunicaciones** en la Universidad Católica de Santiago de Guayaquil.

1.- Declaro tener pleno conocimiento de la obligación que tienen las instituciones de educación superior, de conformidad con el Artículo 144 de la Ley Orgánica de Educación Superior, de entregar a la SENESCYT en formato digital una copia del referido trabajo de titulación para que sea integrado al Sistema Nacional de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor.

2.- Autorizo a la SENESCYT a tener una copia del referido trabajo de titulación, con el propósito de generar un repositorio que democratice la información, respetando las políticas de propiedad intelectual vigentes.

Guayaquil, **3 de Julio de 2017**

f. _____

Nombre: **Wilmer Enrique Jiménez Vivanco**

C.C: **070487990-7**

REPOSITORIO NACIONAL EN CIENCIA Y TECNOLOGÍA

FICHA DE REGISTRO DE TESIS/TRABAJO DE TITULACIÓN

TÍTULO Y SUBTÍTULO:	Modelaje de una interfaz computarizada controlada por voz para cuádrupléticos que contribuya al desarrollo de aplicaciones que permitan comandar elementos de domótica en su domicilio		
AUTOR(ES)	Wilmer Enrique Jiménez Vivanco		
REVISOR(ES)/TUTOR(ES)	MSc. María Luzmila Ruilova Aguirre		
INSTITUCIÓN:	Universidad Católica de Santiago de Guayaquil		
FACULTAD:	Facultad de Educación Técnica para el Desarrollo		
CARRERA:	Telecomunicaciones		
TÍTULO OBTENIDO:	Magister en Telecomunicaciones		
FECHA DE PUBLICACIÓN:	3 de Julio de 2017	No. PÁGINAS:	73
ÁREAS TEMÁTICAS:	Procesamiento Digital de Señales, Programación, Telecomunicaciones, Electrónica Digital.		
PALABRAS CLAVES/ KEYWORDS:	Reconocimiento, voz, MatLab, cuádrupleja, domótica.		
RESUMEN/ABSTRACT (150-250 palabras):	<p>En el presente trabajo de investigación se inicia con una investigación acerca de un tema social bastante importante como es la discapacidad por paraplejía y cuádrupleja, y la dependencia que esta última genera hacia otras personas, identificando la necesidad de un sistema que disminuya esta dependencia mediante el control de elementos de domótica en el domicilio mediante comandos de voz.</p> <p>Se continúa con la investigación de los métodos de reconocimiento de voz que existen en la actualidad, como es el caso de la extracción de los coeficientes cepstrales de la escala de Mel, para luego hacer una propuesta de una interfaz computarizada controlada por voz para cuádrupléticos que contribuya al desarrollo de aplicaciones, que permitan comandar elementos de domótica en su domicilio; en donde además explicamos los bloques de programa con el fin de que este proyecto sea fácilmente modificable y aplicable a cada caso en particular, más adelante pasamos a realizar las pruebas de robustez de la interfaz desarrollada con el fin de medir la efectividad que ofrece la misma.</p>		
ADJUNTO PDF:	<input checked="" type="checkbox"/> SI	<input type="checkbox"/> NO	
CONTACTO CON AUTOR/ES:	Teléfono: 0995781804	E-mail: wilmer_jv@hotmail.com	
CONTACTO CON LA INSTITUCIÓN (COORDINADOR DEL PROCESO UTE)::	Nombre: Manuel Romero Paz		
	Teléfono: 0994606932		
	E-mail: mromeropaz@yahoo.com		
SECCIÓN PARA USO DE BIBLIOTECA			
Nº. DE REGISTRO (en base a datos):			
Nº. DE CLASIFICACIÓN:			
DIRECCIÓN URL (tesis en la web):			