

UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE TECNOLOGIA
CARRERA: ELECTRÓNICA Y TELECOMUNICACIONES



NIVEL LICENCIATURA
EXAMEN DE GRADO
TRABAJO DE APLICACIÓN

**“IMPLEMENTACIÓN DE UN CONTROL DE ACCESO EN BASE A UN
SISTEMA DE RECONOCIMIENTO DE VOZ EN MATLAB PARA UNA SALA DE
SISTEMAS EN BANCO UNIÓN”**

POSTULANTE: ALVARO HUAYCHO MAMANI

LA PAZ- BOLIVIA

2013

INDICE

1. INTRODUCCION	2
2. PLANTEAMIENTO DEL PROBLEMA	2
2.1. IDENTIFICACION DEL PROBLEMA	2
2.2. FORMULACION DEL PROBLEMA	2
3. OBJETIVOS	2
3.1. OBJETIVO GENERAL	2
3.2. OBJETIVOS ESPECIFICOS	3
4. JUSTIFICACION	3
4.1. ALCANCES Y LIMITES	3
5. FUNDAMENTACION TEORICA	3
5.1. MICROFONO	3
5.1.1. CARACTERISTICAS DE UN MICROFONO	3
5.1.1.1. DIRECTIVIDAD Y DIAGRAMA POLAR	3
5.1.1.2. SENSIBILIDAD	4
5.1.1.3. NIVEL DE RUIDO	5
5.1.1.4. RELACION SEÑAL RUIDO	5
5.1.1.5. RESPUESTA EN FRECUENCIAS	6
5.1.1.6. IMPEDANCIA	6
5.1.2. CLASIFICACION DE LOS MICROFONOS SEGÚN EL TRANSDUCTOR	6
5.1.2.1. RESISTENCIA VARIABLE	6
5.1.2.2. PIEZOELECTRICOS	7
5.1.2.3. ELECTROMAGNETICOS	7
5.1.2.4. ELECTROESTATICOS	8
5.1.3. CLASIFICACION DE LOS MICROFONOS SEGÚN SU DIRECTIVIDAD	9
5.2. CERRADURA ELECTROMAGNETICA	10
5.3. MATLAB	11
6. MARCO PRACTICO	12
6.1. OBTENCION DE LA INFORMACION MEDIANTE MICROFONO	12
6.1.1. MUESTREO Y CUANTIFICACION	13
6.1.2. FILTRO DE PRE ENFASIS	13
6.1.3. SEGMENTACION	13
6.1.4. EXTRACCION DE CARACTRISTICAS	14
6.1.5. PROGRAMA	14
6.2. CHAPA MAGNETICA	16
6.2.1. MONTAJE DEL ELECTROIMAN	16
6.3. PROGRAMA EN MATLAB	16
7. CONCLUSIONES	21
8. BIBLIOGRAFIA	22

IMPLEMENTACION DE UN CONTROL DE ACCESO EN BASE A UN SISTEMA DE RECONOCIMIENTO DE VOZ EN MATLAB PARA UNA SALA DE SISTEMAS EN BANCO UNION

1. INTRODUCCIÓN

Tomando en cuenta la importancia del habla, el presente trabajo de aplicación pretende crear una interacción entre una de las expresiones más importantes del hombre con la computadora, creando así un control de acceso a través de un sistema de reconocimiento de voz.

El sistema de reconocimiento de voz es una de las aplicaciones del procesamiento digital de señales de voz. El sistema consiste en obtener una señal de voz que permita reconocer de qué palabra se está hablando, consta de una interfaz gráfica que permite la interacción del usuario por medio un micrófono con la computadora, basados en los resultados de este sistema se puede observar cómo se plantea la base para el control de acceso a través del reconocimiento de voz.

2. PLANTEAMIENTO DEL PROBLEMA

2.1. IDENTIFICACIÓN DEL PROBLEMA

Debido a la gran importancia de los equipos que se instalan en una sala de sistemas de una agencia del banco Unión, el acceso a ellas debe ser restringido a un grupo determinado de personas, Jefe Operativo, Encargado de Seguridad y personal de Sistemas.

El ingreso a la sala de sistemas se debe realizar a través de un control de acceso, que sea capaz de reconocer el sonido de la voz de un grupo selecto de personas para permitir la apertura de la puerta.

2.2. FORMULACIÓN DEL PROBLEMA

¿Cómo establecer un óptimo sistema de reconocimiento de voz que brinde fiabilidad, calidad y seguridad en un control de acceso?

3. OBJETIVOS

3.1. OBJETIVO GENERAL

- Implementar un control de acceso en base a un sistema que realice el reconocimiento de señales de voz por medio de la interacción entre el usuario y la computadora.

3.2. OBJETIVOS ESPECÍFICOS

- Examinar algoritmos de procesamiento digital de voz que permitan un tratamiento sencillo de información de señales de voz.
- Obtener la interacción automática entre humano y computadora por medio de un sistema digital.

4. JUSTIFICACIÓN

4.1. ALCANCES Y LÍMITES

- Este trabajo de aplicación se puede implementar como control de acceso a distintos ambientes en los cuales el acceso a ellas sea limitado.
- Para una mejor implementación el programa efectuado en Matlab debe ser riguroso al momento de reconocer la voz.

5. FUNDAMENTACIÓN TEÓRICA

5.1. MICRÓFONO

El micrófono es un transductor que nos permite realizar la conversión entre las variaciones de presión y variaciones de nivel en una corriente eléctrica. A la hora de estudiar los diferentes tipos de micrófonos, podemos hacerlo, bien sea por su tipo de funcionamiento, o bien por la forma en que recoge el sonido, dado que no presentan la misma sensibilidad en todos los ángulos con respecto a la fuente sonora, forma que se representa por medio de un diagrama polar.

5.1.1. CARÁCTERÍSTICAS DE UN MICRÓFONO

5.1.1.1. DIRECTIVIDAD Y DIAGRAMA POLAR

La directividad de un micro indica la capacidad de captar el sonido procedente de algunas direcciones y de rechazar otras. La directividad máxima aparece en el eje y se representa por medio de un diagrama polar.

El diagrama polar de un micrófono refleja la sensibilidad con que es capaz de captar un sonido según el ángulo con que le incida este. Para determinar el diagrama polar de un micrófono, se utiliza una cámara anecoica (cámara aislada y que no tiene reverberación) en la que se coloca el micrófono y frente a él una fuente sonora que genera un tono a una frecuencia determinada. Teniendo el micrófono en el eje de 0° sobre la fuente sonora, se mide la tensión de salida del mismo. A esta tensión se le llama "tensión de referencia a 0 dBs" y se toma como tensión de

referencia. A continuación se va rotando el micrófono sobre su eje variando el ángulo de incidencia con respecto a la fuente sonora, y se van anotando los valores de tensión que obtenemos en su salida. En el Gráfico 1 podemos ver una muestra más clara de la forma en se realiza un diagrama polar de un micrófono.

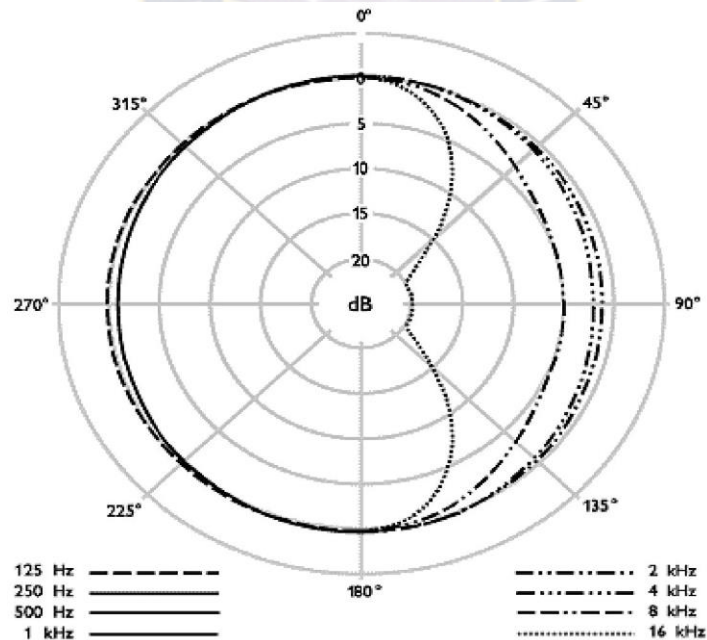


Figura 1. Diagrama polar de un micrófono

Utilizando este sistema hay que repetir la misma operación para diferentes frecuencias y así poder saber el comportamiento que tiene en varias bandas de frecuencias. También se puede realizar el diagrama polar mediante el sistema de espectrometría de retardo de tiempos, donde se realiza una medida de la respuesta en frecuencia del micrófono cada 10° y después se procesa obteniéndose los diagramas a las frecuencias deseadas.

Como hemos podido ver el diagrama polar de un micrófono nos da la información necesaria para saber de qué forma se va a comportar el micrófono con los sonidos dependiendo de dónde le vengan estos. Los diagramas polares se pueden dividir básicamente en tres, el omnidireccional, el bidireccional y el unidireccional (estos a su vez se dividen en cardioides, supercardioides e hipercardioides).

5.1.1.2. SENSIBILIDAD

La sensibilidad de un micrófono es la relación entre la tensión de salida del micrófono y el nivel de intensidad acústica que recibe. Normalmente se mide en decibelios referenciados a 1 voltio con

una presión de 1 dina/cm² y la señal de referencia usada es un tono de 1000 Hz a 74 dB SPL (Sonorous pressure level, o también nivel de presión sonora)

Como es lógico cuanto mayor sea la sensibilidad de un micrófono, mejor.

La sensibilidad del micrófono no influye en su calidad sonora, ni en su respuesta en frecuencia, únicamente es importante a la hora de su uso ya que un micrófono de baja sensibilidad nos fuerza, al utilizar un preamplificador para el micrófono, a utilizar un nivel mayor de ganancia de entrada para dicho micrófono, aumentando de esta manera el ruido de fondo que produce la electrónica de los preamplificadores.

Para las mismas condiciones si tenemos un micrófono con una sensibilidad mayor, necesitaremos menos ganancia en la entrada del preamplificador con lo que reduciremos el nivel de ruido de fondo. Puede parecer que esto no tiene excesiva importancia, y no la tiene cuando únicamente se utiliza un micrófono y lo que se trata de grabar o amplificar no es muy importante. Sin embargo cuando se utilizan muchos micrófonos, caso muy típico en grabaciones y actuaciones en directo, el nivel de ruido de fondo producido en cada canal se va sumando y el resultado puede ser realmente problemático, sobre todo cuando grabamos en soporte digital.

5.1.1.3. NIVEL DE RUIDO

El ruido propio de un micrófono es el que produce cuando no hay ninguna señal externa que excite el micrófono. Esta medida se realiza normalmente en una cámara anecoica y se especifica como una medida de presión sonora y por tanto en dB SPL, equivalente a una fuente sonora que hubiese generado la misma tensión de salida que el ruido producido por el micrófono.

Se puede considerar como excelente un nivel de ruido de 20 dBA SPL, como valor bueno sobre unos 30 dBA SPL, y como malo 40 dBA SPL.

5.1.1.4. RELACIÓN SEÑAL RUIDO

La relación señal ruido (S/R) representa realmente la diferencia entre el nivel SPL y el ruido propio del micrófono. Cuanto mayor sea la SPL y menor el ruido mejor será la relación señal ruido, y por contra si el nivel de SPL es menor y el ruido propio aumenta, la relación será menor y por tanto peor.

Cuanto mayor sea la relación señal ruido mejor.

Nos indica que porcentaje de la señal SPL está por encima del ruido de fondo. Si tenemos una SPL de 100 dB y un ruido propio en el micrófono de 30 dB, la relación señal/ruido será de 70 .dB

Para una señal de 100 dB una relación señal/ruido de 80 dB es muy buena y 70 dB es buena.

5.1.1.5. RESPUESTA DE FRECUENCIAS

La respuesta en frecuencia de un micrófono indica la sensibilidad del mismo a cada frecuencia. Como hemos visto al principio al hablar de los diagramas polares, los micrófonos no tienen la misma sensibilidad para cada ángulo de incidencia ni para cada frecuencia, por tanto es difícil conseguir una respuesta uniforme en todo el espectro.

Como es lógico hay que observar que la longitud de un sonido influye o tiene una relación en el comportamiento del diafragma según la relación de tamaño que haya entre ambos. Con todos los micrófonos se entrega una hoja con la curva de respuesta en frecuencia del micrófono, teniendo en un eje (x) la frecuencia de 20 Hz a 20 KHz y en el otro eje (y) los decibelios.

5.1.1.6. IMPEDANCIA

La impedancia en un micrófono es la propiedad de limitar el paso de la corriente, como ya sabemos se mide en Ohmios. Normalmente en los micrófonos se mide sobre una frecuencia de 1KHz y en micrófonos de baja impedancia, esta, suele valer 200 Ohmios.

Los micrófonos más habituales son los de baja impedancia, considerados hasta unos 600 Ohmios. También existen los de alta impedancia que suelen tener un valor tipo de 3000 Ohmios y más.

La diferencia entre uno y otro radica en que a la hora de conectar un cable para unirlo a la mesa de mezclas o al amplificador, los de baja impedancia al oponer poca resistencia a la corriente que circula, permiten utilizar cables de longitud muy grande mientras que los de alta impedancia al restringir de forma mayor el paso de la corriente, solo se pueden usar con cables de corta distancia.

5.1.2. CLASIFICACIÓN DE LOS MICRÓFONOS SEGÚN SU TRANSDUCTOR

5.1.2.1. RESISTENCIA VARIABLE

Basa su funcionamiento en las diferencias de presión existentes entre el exterior y el interior. Al producirse una variación de presión modifican su resistencia provocando una pequeña carga eléctrica. Fueron inventados por Ludge y Huges en 1878 y utilizados desde la invención del teléfono.

- **Micrófono de carbón:** una membrana metálica unida a un electrodo de carbono que está en contacto con un compartimento cerrado de varios centenares de gránulos de carbón, generalmente grafito o antracita. Al incidir la onda sonora en su superficie, produce un desorden en los gránulos de carbón y hace que varíe la resistencia de contacto entre ellos originando una tensión que es proporcional al desplazamiento de la membrana. Tienen una impedancia de unos 30-40 Ohmios y un importante nivel de ruido con respuesta bastante limitada. Estando entre los 250 Hz y los 300 kHz, producen una notable distorsión por lo que su utilización queda reducida al campo de la telefonía o trabajos de baja calidad debido a sus costes muy bajos.

5.1.2.2. PIEZO ELÉCTRICOS

Un elemento piezoeléctrico genera una carga eléctrica cuando se le somete a una deformación provocada por una presión sonora. Son sistemas controlados por elasticidad que funcionan a presión.

- **Micrófono de cristal:** utilizan cristales situados de tal forma que al ser golpeada su superficie por una onda sonora se doblan. Al someter al cristal a un esfuerzo mecánico genera un tipo de fuerza electromotriz proporcional al desplazamiento causado. El material utilizado comúnmente es la sal de Rochélie con problemas en los cambios de temperatura. También se utilizan los cristales de fosfato ácido de amonio que son más estables pero de menor sensibilidad. Son micrófonos incapaces de reproducir frecuencias muy altas debido a que los propios cristales no pueden vibrar de forma muy rápida. Sus posibilidades de respuesta oscilan entre los 80 y los 10.000 Hz, lo que limita mucho su utilización.
- **Micrófono de cerámica:** de funcionamiento similar a los de cristal, construidos con piezas de cerámica como el bario y el titanio. Son insensibles a la temperatura y a la humedad pero su uso se limita a la megafonía. Su sensibilidad está por debajo de los de cristal, con lo cual solo tienen aplicaciones en el campo de la megafonía pese a su bajo coste.

5.1.2.3. ELECTROMAGNÉTICOS

Basado en unos transductores en los que un conductor eléctrico, movido por la fuerza generada por la onda sonora, se desplaza dentro del campo creado por un imán.

- Micrófono dinámico o de bobina móvil: compuesto de una membrana muy ligera y flexible unida a una bobina que se desplaza en el campo magnético de un imán. Las vibraciones de la membrana se acompañan de desplazamientos en forma de vaivén de la bobina de forma que se induce en ella una corriente eléctrica. Son micrófonos de gran robustez, con una excelente respuesta que puede oscilar entre los 20 y los 20.000 Hz, muy utilizados en los estudios y de precios asequibles.
- Micrófono de cinta: funciona suspendiendo una cinta extremadamente delgada, de unas 20 micras que expuesta a las ondas sonoras por las dos caras, externa e interna de un campo magnético generado por un imán. Las vibraciones de la onda sonora producen en ella un movimiento de las líneas del campo magnético obteniendo en sus extremos una fuerza electromotriz de frecuencia proporcional a la onda incidente. Son micrófonos con una respuesta de frecuencias muy uniforme, de buena sensibilidad y normalmente bidireccionales. No obstante su coste de producción es caro y son muy sensibles a golpes de voz o viento que pueden quebrar su membrana.

5.1.2.4. ELECTROESTÁTICOS

Su funcionamiento se basa en la disposición de dos placas, una fija llamada terminal y la otra móvil que actúa de diafragma, y separada una de la otra 0.001 pulgada. Cuando la onda incide sobre la placa diafragma esta vibra y altera la separación entre ambas. La placa terminal está aislada del resto del micrófono y se le aplica una tensión de polarización. Cuando la membrana se desplaza, varía la capacidad apareciendo una tensión variable.

- Micrófono de condensador: en este tipo de micros constan de una cámara de aire a modo de condensador, donde la placa del diafragma suele ser de metal muy fino, quizás en forma de membrana de plástico cubierta de oro. La parte trasera tiene pequeñas cavidades, que sin ellas, la capa de aire entre las dos membranas sería demasiado rígida e impediría que el diafragma se moviera libremente. Para su funcionamiento hay que aplicar un voltaje de polarización entre el diafragma y la placa trasera (normalmente alimentación phantom). Son micrófonos con una excelente calidad con un proceso de fabricación preciso y complejo que los hace caros. Muy sensibles a la humedad y poco robustos pero sus características son estables en el tiempo.
- Micrófono electret: se basa en el mismo principio que los de condensador, diferenciándose de ellos en que la cápsula no necesita polarización. La cámara de aire del

condensador ha sido sustituida por un polímetro llamado electret. Dicho polímetro ha sido previamente polarizado durante su fabricación, es decir, se ha introducido en él una carga eléctrica permanente, que será la que suministre la energía electrostática para mantener el condensador cargado. El polímetro ahorra la corriente polarizadora, pero sigue siendo necesario el preamplificador-adaptador de impedancias, que requiere una alimentación de corriente continua, (1.5 a 12 voltios) que puede ser proporcionada por una pila. Poseen una buena respuesta de frecuencias pero son menos sensibles a frecuencias agudas que los de condensador, una alta impedancia y un bajo coste, lo que los hace muy útiles en entrevistas y locuciones.

5.1.3. CLASIFICACIÓN DE LOS MICROFONOS SEGÚN SU DIRECTIVIDAD

Como se mencionó con anterioridad, los micrófonos se pueden diferenciar según su directividad en tres categorías principalmente, omnidireccional, bidireccional y unidireccional.

En el caso del diagrama polar omnidireccional, tal y como su nombre lo indica, este recibe prácticamente con la misma sensibilidad cualquier sonido independientemente del punto donde proceda el mismo, su diagrama es por tanto circular.

El bidireccional presenta una gran sensibilidad en el frente, con un ángulo amplio, y una imagen simétrica en la parte posterior, o sea que es menos sensible a los sonidos que le llegan desde los laterales y más sensible a los que le llegan desde el frente y la parte posterior.

El micrófono unidireccional se puede clasificar como aquel que tiene una mayor sensibilidad a los sonidos que vienen de frente a la cápsula con un ángulo relativamente amplio. Este tipo de diagrama polar, se puede subdividir en tres que son, el cardiode, el supercardioide y el hipercardiode. Cada uno de ellos va presentando un diagrama polar cada vez más estrecho y por tanto se van haciendo más insensibles a los sonidos que les llegan desde la parte posterior así como del lateral.

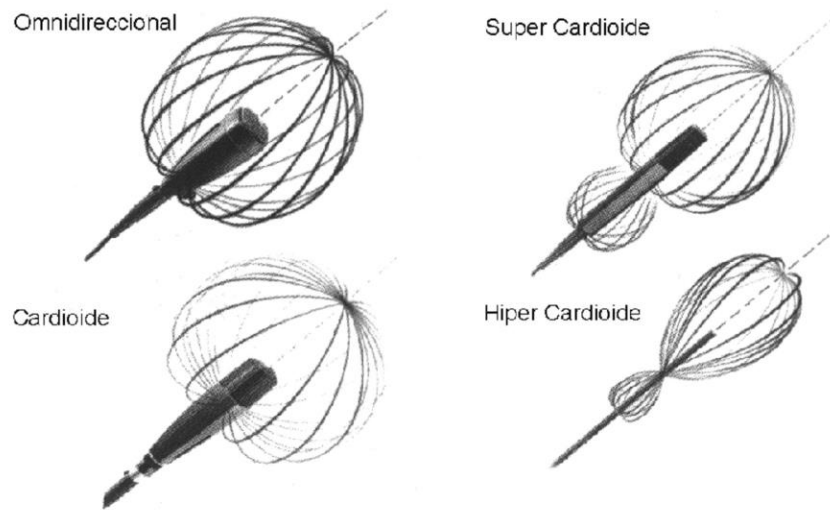


Figura 2. Directividad de los micrófonos

Un factor importante es que el micrófono, con un diagrama polar determinado, lo mantengan los más igualado posible en todas las frecuencias, dado que si no, se presentan coloraciones en el sonido debido al acercamiento o separación desde o hacia la fuente sonora. Si tenemos unos diagramas polares uniformes para diferentes frecuencias, sabremos que la respuesta en frecuencia del micrófono no variara en exceso según los ángulos de incidencia del sonido.

5.2. CERRADURA ELECTROMAGNÉTICA

Las cerraduras electromagnéticas son cerraduras eléctricas conformadas por un electroimán que se instala en el marco de la puerta, y una placa montada en la hoja de la misma. Cuando se encuentra energizada, la fuerza magnética del electroimán asegura la puerta. Estos elementos están diseñados para ser utilizados en diferentes tipos de puertas, según el volumen de tráfico peatonal y el nivel de seguridad que se requiera.

El electroimán es el elemento que crea un campo magnético al proporcionarle corriente eléctrica y consta de un núcleo o barra de hierro al que se enrolla un cable barnizado de cobre, creando una bobina. Si a ésta se le suministra corriente eléctrica el núcleo se convierte en un imán capaz de atraer objetos metálicos (hierro), perdiendo sus propiedades magnéticas al cortar la corriente.

Este efecto se ha aplicado a la seguridad para crear dispositivos electromagnéticos idóneos para controlar el estado y funcionamiento de puertas, manteniéndolas abiertas hasta que se activan los automatismos de alarma o manualmente.



Figura 3. Chapa Magnética

5.3. MATLAB

MATLAB es un lenguaje de alto desempeño diseñado para realizar cálculos técnicos. MATLAB integra el cálculo, la visualización y la programación en un ambiente fácil de utilizar donde los problemas y las soluciones se expresan en una notación matemática. MATLAB es un sistema interactivo cuyo elemento básico de datos es el arreglo que no requiere de dimensionamiento previo. Esto permite resolver muchos problemas computacionales, específicamente aquellos que involucren vectores y matrices, en un tiempo mucho menor al requerido para escribir un programa en un lenguaje escalar no interactivo tal como C o Fortran.

MATLAB se utiliza ampliamente en:

- Cálculos numéricos
- Desarrollo de algoritmos
- Modelado, simulación y prueba de prototipos
- Análisis de datos, exploración y visualización
- Graficación de datos con fines científicos o de ingeniería
- Desarrollo de aplicaciones que requieran de una interfaz gráfica de usuario (GUI, GraphicalUser Interface).

En el ámbito académico y de investigación, es la herramienta estándar para los cursos introductorios y avanzados de matemáticas, ingeniería e investigación. En la industria MATLAB es la herramienta usada para el análisis, investigación y desarrollo de nuevos productos tecnológicos.

La ventaja principal de MATLAB es el uso de familias de comandos de áreas específicas llamadas toolboxes. Lo más importante para los usuarios de MATLAB es que los toolboxes le permiten aprender y aplicar la teoría. Los toolboxes son grupos de comandos de MATLAB (archivos M) que extienden el ambiente de MATLAB para resolver problemas de áreas específicas de la ciencia e ingeniería. Por ejemplo, existen toolboxes para las áreas de Procesamiento Digital de Señales, Sistemas de Control, Redes Neuronales, Lógica Difusa, Wavelets, etc.

6. MARCO PRÁCTICO

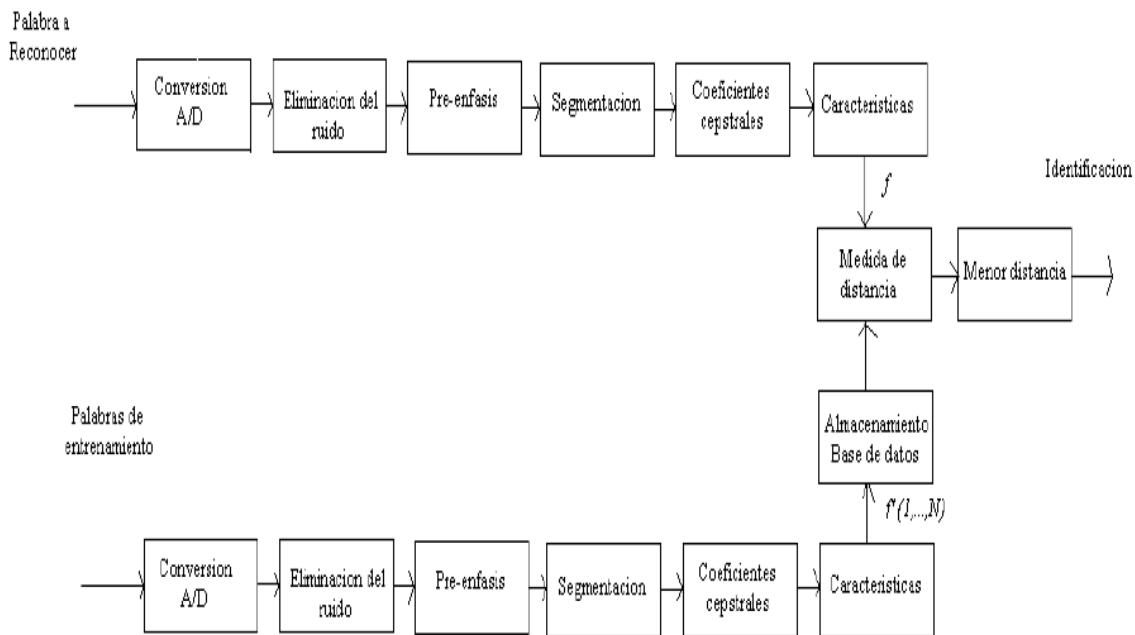


Grafico 4. Diagrama del sistema de reconocimiento de voz

6.1. OBTENCIÓN DE LA INFORMACIÓN MEDIANTE MICRÓFONO

Para grabar una señal de voz se hace la conversión de la señal analógica del micrófono en una señal digital por medio del conversor A/D en la tarjeta de sonido. Cuando un micrófono está operando las ondas de sonido hacen que vibre el elemento magnético del micrófono causando una corriente eléctrica hacia la tarjeta de sonido, donde el conversor A/D básicamente graba los voltajes eléctricos en intervalos específicos.

Hay dos factores importantes durante este proceso. Primero está la tasa de muestreo o que tan seguido los valores de voltaje son grabados. Segundo, son los bits por segundo, o que tan exactamente los valores son grabados.

Dado a que el habla es relativamente de bajas frecuencias (entre 100Hz- 8kHz), una frecuencia de muestreo de 16000 muestras/seg provee una mayor exactitud en la adquisición de la información, la frecuencia de nyquist.

La obtención de la información mediante micrófono en MATLAB se realiza con la función wavrecord(n,Fs) graba n muestras de la señal de audio, muestreadas a una frecuencia de Fs, utilizamos la frecuencia de 11025 Hz ya que se adapta bien a nuestro algoritmo y no se pierde información.

6.1.1. MUESTREO Y CUANTIFICACIÓN

Muestreo consiste en el proceso de conversión de señales continuas a señales discretas en el tiempo, es un paso para digitalizar una señal analógica. Este proceso se realiza midiendo la señal en momentos periódicos del tiempo.

En la cuantificación el valor de cada muestra de la señal se representa como un valor elegido de entre un conjunto finito de posibles valores.

6.1.2. FILTRO DE PRE - ENFASIS

Se aplica un filtro digital pasa altas de primer orden a la señal, para enfatizar las frecuencias altas de los formantes por dos razones, primero para que no se pierda información durante la segmentación, ya que la mayoría de la información está contenida en las frecuencias bajas, en segundo remueve la componente DC de la señal, aplanando espectralmente la señal. Uno de los filtros de pre-énfasis más utilizados tiene la ecuación:

$$H(z) = (1 - az^{-1})$$

6.1.3. SEGMENTACIÓN

La segmentación consiste en cortar la señal en segmentos de análisis. La señal de voz es asumida como estacionaria en estos segmentos.

Durante la segmentación los segmentos son guardados cada uno como la columna de una matriz, para el posterior procesamiento de la señal de voz.

Para el proceso una ventana de Hamming de 30ms es aplicada a la señal de voz, enfatizada previamente con el filtro de pre-énfasis. Con un desplazamiento típico 10ms entre cada ventaneo

6.1.4. EXTRACCIÓN DE CARACTERÍSTICAS

En el reconocimiento del habla, la señal de voz pre-procesada se ingresa a un nuevo procesamiento para producir una representación de la voz en forma de secuencia de vectores o agrupaciones de valores que se denominan parámetros, que deben representar la información contenida en la envolvente del espectro.

6.1.5. PROGRAMA

El programa tiene una interfaz gráfica, muy fácil de utilizar. Con una base de datos de prueba de 2 nombres, esta base de datos puede ser aumentada en cualquier cantidad. Los nombres en la base de datos son:

- Álvaro
- Cindy

El entorno grafico es como se muestra en la figura 5.

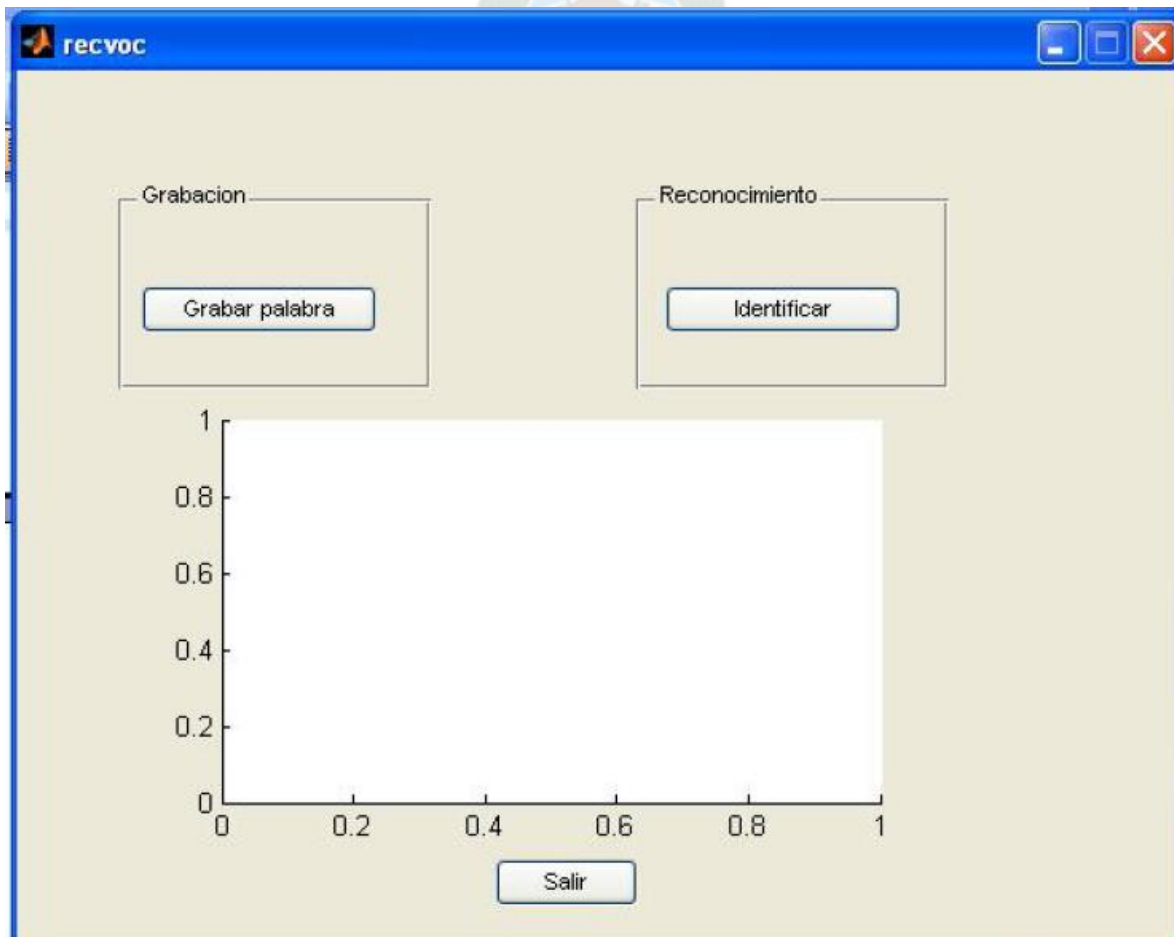


Figura 5. Interfaz grafica

Consta de 3 cuadros de selección, dos cuadros de mensaje y un plano donde serán graficadas las señales de voz en tiempo continuo y el espectro de frecuencias de la señal discreta ya procesada.

- Grabación

En este cuadro nos aparece la opción de grabar la palabra para ingresarla al reconocedor de voz, cuando la palabra es grabada aparece un cuadro de mensaje indicando que la grabación ha sido terminada. Por otra parte en el plano aparece una gráfica en el tiempo (eje X) y en amplitud (eje Y), de la señal de voz grabada por el usuario (véase figura 6).

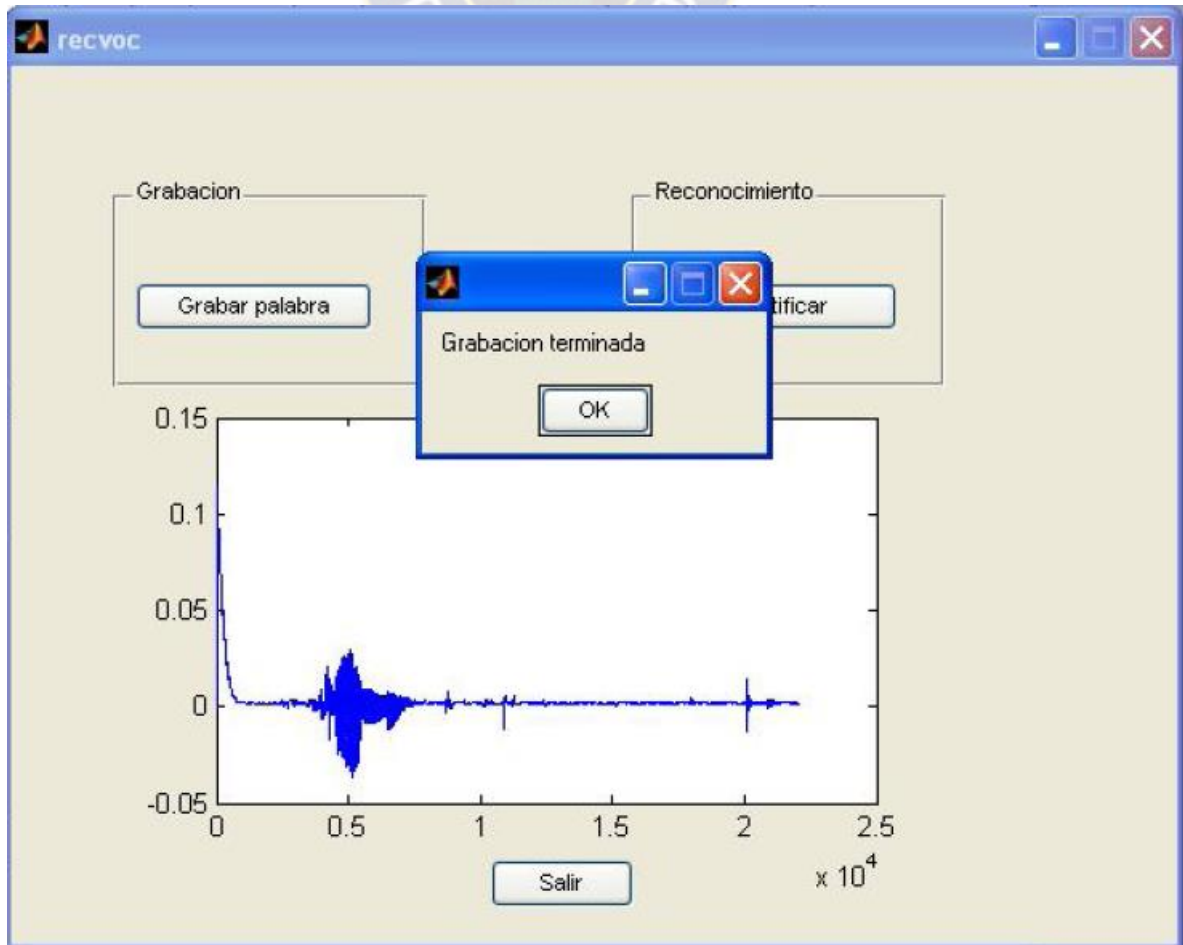


Figura 6 Modo de grabación

- Salir

Este botón nos permite salir completamente del Sistema de Reconocimiento de Voz.

- Reconocimiento

El cuadro de reconocimiento brinda los resultados del sistema, con un cuadro de mensaje indicando cual palabra ha sido reconocida y la palabra grabada es reproducida. Además, nos

da una representación gráfica en el plano con la información del espectro de frecuencias de la señal del usuario reconocido.

6.1.6. CHAPA MAGNETICA

6.1.6.1. MONTAJE DEL ELECTROIMAN

Para el montaje del electroimán debe tenerse en cuenta las siguientes recomendaciones:

- El contacto entre el electroimán y la placa debe permitir un ajuste perfecto al cierre de la misma.
- El electroimán debe quedar paralelo a la placa.
- El área seleccionada para sujetar al electroimán y la placa debe ser suficientemente fuerte para permitir una instalación segura.
- La superficie del marco a utilizarse debe estar completamente plana.

Para escoger entre un montaje horizontal o vertical, debe considerarse la fortaleza del marco.

Frecuentemente en marcos de aluminio se encuentra más fortaleza en el tendido vertical que en el horizontal y en esos casos para prevenir posibles roturas, obviamente es mejor la instalación vertical. Se puede también reforzar el marco superior con una platina. El instalador debe evitar colocar las cerraduras en soportes débiles que podrían disminuir la seguridad de la cerradura.

6.1.7. PROGRAMA EN MATLAB

```
% BASE DE DATOS
clc
Fs=11025;
y0=wavrecord(1*Fs,Fs,1);
soundsc(y0);
word0=['Alvaro'];
y0=chop_silencio(y0);
z0=enfasis(y0);
seg0=segmentos(z0);
c0=cepstrum(seg0);
f0=features(c0);
savealvaro.mat word0 f0 y0;
y1=wavrecord(1*Fs,Fs,1);
soundsc(y1);
word1=['cindy'];
y1=chop_silencio(y1);
z1=enfasis(y1);
seg1=segmentos(z1);
```


```

c1=cepstrum(seg1);
f1=features(c1);
savecindy.mat word1 f1 y1;
% FUNCIÓN ELIMINACIÓN DEL RUIDO
%chop_silencio
%Corta el silencio en la señal completa
%y=chop_silencio(s)
function y = chop_silencio(s)
len = length(s); % length del vector
avg_e = sum(s.*s)/len; %promedioseñalentera
THRES = 0.2;
y = [0];
for i = 1:80:len-80 % cada 10ms
seg = s(i:i+79); % segmentos
e = sum(seg.*seg)/80; % promedio de cada segmento
if( e> THRES*avg_e) % si el promedio energético es mayor que la señal completa por el valor
umbral
y=[y,seg(1:end)]; % almacena en y si no es eliminado como espacio en blanco
end;
end;
% FUNCIÓN FILTRO DE PRE-ÉNFASIS
%énfasis
%filtro de énfasis para la señal
%énfasis(s)
function [y]=énfasis(x)
b=[1 -0.95];
y=filter(b,1,x);
% FUNCIÓN SEGMENTACIÓN
%segmentos
%segmenta la señal en tramas de 30ms
%[segs]=segmentos(y)
function [segs]=segmentos(y)
len = length(y); %longitud del vector
num_segments = floor(len/80) -2; % redondeamos el número de segmentos
segs = zeros(num_segments,240); % matriz de segmentos
num_segments x 240
win = hamming(240)'; % ventana hamming de 240 puntos(30ms)
win=coefficients (numero de muestras en el analisis )
for i = 0:num_segments-1
inicio = i*80+1;
segs(i+1,1:240) = (y(inicio:inicio+239).*win); %multiplicamos por la
ventana %hamming
end;
% FUNCIÓN CEPSTRUM
% (EXTRACCIÓN DE LOS COEFICIENTES CEPSTRALES)
%cepstrum
%Encuentra los coeficientes cepstrales de cada segmento
%[cc]=cepstrum(seg)

```

```

function [c] = cepstrum(segs)
%programa para encontrar los coeficientes cepstrales de cada segmento
[M,N] = size(segs); % M filas, N columnas, para definir Mo=output frame
size(# de %frecuencias a las que se aplica FFT)
[c] = zeros(M,10); % iniciando el vector para guardar los coeficientes
cepstrales. P=10
%calculando los coeficientes cepstrales reales
for i=1:M
% f = fft(segs(i,:));
% m = abs(f);
% l = log(m+1e-5);
% in = real(iff(l));
% c(i,:)= in(1:10); o resumidamente
r = rceps(segs(i,:));
c(i,:)= r(1:10);
end
%
% normalizacion de los coeficientes cepstrales
sum_c = sum(c);
avg_c = sum_c/M;
%sustrayendole el minimo, solo las que estan arriba del promedio los fonemas
for i=1:M
c(i,:)= c(i,:)-avg_c; % smoothed spectrum
end
% FUNCIÓN FEATURES (DE CARACTERISTICAS)
%features
%Encuentra los coeficientes polinomiales de la expansión correspondientes al los coeficientes
cepstrum
%[ftr]=features(c)
function ftr=features(c)
[M,N] = size(c);
ftr = zeros(M-8,20);
j = (1:9); %col vector
p1 = repmat((j-5)',1,10); %crea una matriz llena de P1j=j-5
sum_p1 = sum((j-5).*(j-5)); % suma total
for i = 1:M-8
%caracteristicas de orden cero (coeficientes cepstrales)
ftr(i,1:10)=c(i,1:10);
%caracteristicas de primer orden
b = zeros(1,10);
for j = 0:8 % aqui se sacan los coeficientes polinomiales para cada segmento
b = b+ p1(j+1,:).*c(i+j,:); % numerador de la ecuacion
end;
ftr(i,11:20)=b/sum_p1; % dividiendolo en el denominador
end; %se obtienen los valores de b
% FUNCIÓN DE DISTANCIA
%distancia
%
```



```

%[dist]=distancia(ftr_a,ftr_b)
functiontot_dist = spv_dis(ftr_a,ftr_b)
[m_a,n_a]=size(ftr_a);
[m_b,n_b]=size(ftr_b);
%establece the guide and slave
if(m_a<m_b )
guide = ftr_a;
slave = ftr_b;
m = m_b;
n = m_a;
else
guide = ftr_b;
slave = ftr_a;
m = m_a;
n = m_b;
end;
%computa la matriz de distancia
dist = zeros(m,n);
for i = 1:m
for j = 1:n
dist(i,j) = sqrt(sum(((guide(j,:)-slave(i,:)).*(guide(j,:)-slave(i,:)).^2));
end;
end;
end;
% INICIALIZACION
functionvarargout = recvoc(varargin)
gui_Singleton = 1;
gui_State = struct('gui_Name', mfilename, ...
'gui_Singleton', gui_Singleton, ...
'gui_OpeningFcn', @recvoc_OpeningFcn, ...
'gui_OutputFcn', @recvoc_OutputFcn, ...
'gui_LayoutFcn', [] , ...
'gui_Callback', []);
if nargin&&ischar(varargin{1})
gui_State.gui_Callback = str2func(varargin{1});
end

ifnargout
[varargout{1:nargout}] = gui_mainfcn(gui_State, varargin{:});
else
gui_mainfcn(gui_State, varargin{:});
end
% PROGRAMA PRINCIPAL
functionrecvoc_OpeningFcn(hObject, eventdata, handles, varargin)
loadgeno.mat;
loadrosa.mat;
loadflavio.mat;
loadandres.mat;
load 'handel';

```



```

wavwrite(y,'handel');
rec=reconociendovoz(y0);
handles.voz_grabada=y;
handles.feature_datos=rec
handles.feature_f0=f0;
handles.feature_f1=f1;
handles.feature_f2=f2;
handles.feature_f3=f3;
handles.feature_y0=y0;
handles.feature_y1=y1;
handles.feature_y2=y2;
handles.feature_y3=y3;
handles.output = hObject;
% Update handles structure
guidata(hObject, handles);
% --- Outputs from this function are returned to the command line.
functionvarargout = recvoc_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)
% Get default command line output from handles structure
varargout{1} = handles.output;
% --- Executes on button press in Grabar.
functionGrabar_Callback(hObject, eventdata, handles)
% hObject handle to Grabar (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)
% obtener la palabra del microfono
Fs = 11025; % SamplingFrequency (Hz)
n = 20;
Nseconds = 2; % largo de la señal de voz
y = wavrecord(Nseconds*Fs, Fs, 1)';
rec1=reconociendovoz(y);
handles.voz_grabada = y;
handles.feature_datos=rec1
plot(y);
msgbox('Grabacion terminada');
guidata(hObject, handles);
%load rec.mat;
%handles.feature_voz=f;
% --- Executes on button press in identificar.
functionidentificar_Callback(hObject, eventdata, handles)
% hObject handle to identificar (see GCBO)
% eventdata reserved - to be defined in a future version of MATLAB
% handles structure with handles and user data (see GUIDATA)
rec1=handles.feature_datos;
y0=handles.feature_y0;

```



```

y1=handles.feature_y1;
y2=handles.feature_y2;
y3=handles.feature_y3;
f0=handles.feature_f0;
f1=handles.feature_f1;
f2=handles.feature_f2;
f3=handles.feature_f3;
d1=distancia(f0,rec1)
d2=distancia(f1,rec1)
d3=distancia(f2,rec1)
d4=distancia(f3,rec1)
if (d1 < d2) & (d1 < d3) & (d1 < d4)
plot(abs(fft(y0)));
soundsc(y0);
msgbox('Usted dijo: alvaro');
else if (d2 < d1) & (d2 < d3)& (d2 < d4)
plot(abs(fft(y1)));
soundsc(y1);
msgbox('Identificado: cindy')
end
end
% --- Executes on button press in pushbutton3.
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)
exit;
% FUNCIÓN RECONOCIENDOVOZ
functionrec=reconociendovoz(y)
% load data
y=chop_silencio(y);
z=enfasis(y);
seg=segmentos(z);
c=cepstrum(seg);
f=features(c);
rec=f;
save('rec.mat','f','y');

```



7. CONCLUSIONES

- El reconocimiento de voz es una de las aplicaciones del procesamiento digital de señales que permite interacción entre seres humanos y computadoras.
- Mediante Matlab se logró implementar un control de acceso para una sala de sistemas en base al reconocimiento de voz.

8. BIBLIOGRAFÍA

- Micrófono

<http://es.wikipedia.org/wiki/Mic%C3%B3fono>

- Gomez, J.C. Procesamiento digital de Señales de Voz. Apuntes: Modelos de producción de Voz.

http://www.eie.fceia.unr.edu.ar/%7Eprodivoz/apuntes_index.html

- MATLAB RB2010Help

