

	UNIVERSIDAD SURCOLOMBIANA GESTIÓN SERVICIOS BIBLIOTECARIOS						
	CARTA DE AUTORIZACIÓN						
CÓDIGO	AP-BIB-FO-06	VERSIÓN	1	VIGENCIA	2014	PÁGINA	1 de 2

Neiva, 14 de julio de 2017

Señores

CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN

UNIVERSIDAD SURCOLOMBIANA

Ciudad

El (Los) suscrito(s):

Nicolás Charry Moreno, con C.C. No. 1075273116,

David Eduardo Clavijo Osorio, con C.C. No. 1075268257,

autor(es) de la tesis y/o trabajo de grado o _____

titulado: Diseño e Implementación de un prototipo de brazalete de detección de sonidos de alarma para personas sordas e hipoacúsicas.

presentado y aprobado en el año 2017 como requisito para optar al título de

Ingeniero Electrónico;

Autorizo (amos) al CENTRO DE INFORMACIÓN Y DOCUMENTACIÓN de la Universidad Surcolombiana para que con fines académicos, muestre al país y el exterior la producción intelectual de la Universidad Surcolombiana, a través de la visibilidad de su contenido de la siguiente manera:

- Los usuarios puedan consultar el contenido de este trabajo de grado en los sitios web que administra la Universidad, en bases de datos, repositorio digital, catálogos y en otros sitios web, redes y sistemas de información nacionales e internacionales "open access" y en las redes de información con las cuales tenga convenio la Institución.
- Permita la consulta, la reproducción y préstamo a los usuarios interesados en el contenido de este trabajo, para todos los usos que tengan finalidad académica, ya sea en formato Cd-Rom o digital desde internet, intranet, etc., y en general para cualquier formato conocido o por conocer, dentro de los términos establecidos en la Ley 23 de 1982, Ley 44 de 1993, Decisión Andina 351 de 1993, Decreto 460 de 1995 y demás normas generales sobre la materia.
- Continúo conservando los correspondientes derechos sin modificación o restricción alguna; puesto que de acuerdo con la legislación colombiana aplicable, el presente es un acuerdo jurídico que en ningún caso conlleva la enajenación del derecho de autor y sus conexos.

De conformidad con lo establecido en el artículo 30 de la Ley 23 de 1982 y el artículo 11 de la Decisión Andina 351 de 1993, "Los derechos morales sobre el trabajo son propiedad de los autores", los cuales son irrenunciables, imprescriptibles, inembargables e inalienables.

Vigilada Mineducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.



UNIVERSIDAD SURCOLOMBIANA
GESTIÓN SERVICIOS BIBLIOTECARIOS



CARTA DE AUTORIZACIÓN

CÓDIGO

AP-BIB-FO-06

VERSIÓN

1

VIGENCIA

2014

PÁGINA

2 de 2

EL AUTOR/ESTUDIANTE:

Firma:

Nicolás Charry Mosero

EL AUTOR/ESTUDIANTE:

Firma:

Vigilada Mineducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.

	UNIVERSIDAD SURCOLOMBIANA GESTIÓN SERVICIOS BIBLIOTECARIOS				   		
DESCRIPCIÓN DE LA TESIS Y/O TRABAJOS DE GRADO							
CÓDIGO	AP-BIB-FO-07	VERSIÓN	1	VIGENCIA	2014	PÁGINA	1 de 4

TÍTULO COMPLETO DEL TRABAJO:

Diseño e implementación de un prototipo de brazalete de detección de sonidos de alarma para personas sordas e hipoacúsicas

AUTOR O AUTORES:

Primero y Segundo Apellido	Primero y Segundo Nombre
Charry Moreno	Nicolás
Clavijo Osorio	David Eduardo

DIRECTOR Y CODIRECTOR TESIS:

Primero y Segundo Apellido	Primero y Segundo Nombre
Soto Otálora	Agustín

ASESOR (ES):

Primero y Segundo Apellido	Primero y Segundo Nombre

PARA OPTAR AL TÍTULO DE: Ingeniero Electrónico

FACULTAD: Ingeniería

PROGRAMA O POSGRADO: Electrónica

CIUDAD: Neiva

AÑO DE PRESENTACIÓN: 2017

NÚMERO DE PÁGINAS: 83

Vigilada mieducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.

	UNIVERSIDAD SURCOLOMBIANA GESTIÓN SERVICIOS BIBLIOTECARIOS						
	DESCRIPCIÓN DE LA TESIS Y/O TRABAJOS DE GRADO						
CÓDIGO	AP-BIB-FO-07	VERSIÓN	1	VIGENCIA	2014	PÁGINA	2 de 4

TIPO DE ILUSTRACIONES (Marcar con una X):

Diagramas X Fotografías X Grabaciones en discos Ilustraciones en general Grabados
 Láminas Litografías Mapas Música impresa Planos Retratos Sin ilustraciones
 Tablas o Cuadros X

SOFTWARE requerido y/o especializado para la lectura del documento:

MATERIAL ANEXO:

PREMIO O DISTINCIÓN (En caso de ser LAUREADAS o Meritoria):

PALABRAS CLAVES EN ESPAÑOL E INGLÉS:

<u>Español</u>	<u>Inglés</u>	<u>Español</u>	<u>Inglés</u>
1. Redes Neuronales	Neural Networks	6. <u> </u>	<u> </u>
2. DSPIC	DSPIC	7. <u> </u>	<u> </u>
3. sonidos de emergencia	emergency sounds	8. <u> </u>	<u> </u>
4. Sordera	deafness	9. <u> </u>	<u> </u>
5. Hipoacusia	hearing loss	10. <u> </u>	<u> </u>

RESUMEN DEL CONTENIDO: (Máximo 250 palabras)

El respectivo proyecto de investigación consiste en el desarrollo e implementación de un dispositivo electrónico (brazalete) capaz de discriminar sonidos de alerta o emergencia, funcionando como un sistema de ayuda para las personas con sordera o hipoacusia, además de manejar una interfaz para el usuario bastante amigable con características adicionales como de Reloj, alarma, entre otras. Para el proceso de discriminación de sonidos y el sistema de detección de alertas se utilizó como herramienta el entrenamiento de Redes Neuronales (RNA), con ayuda del Toolbox Neural Networks de Matlab. En estas aplicaciones, el uso de redes neuronales para el reconocimiento automático de patrones sonoros ha tenido bastante éxito. Las características que presentan el reconocimiento de patrones acústicos hacen que este problema sea una aplicación factible para implementarse como una aplicación para el microcontrolador DSPIC33EP2656GP502 de Microchip. Ensayos previos demuestran una eficiencia mayor del 75 % en la identificación de tres tipos de sonidos de emergencia (ambulancia, policía, bomberos).

	UNIVERSIDAD SURCOLOMBIANA GESTIÓN SERVICIOS BIBLIOTECARIOS				   		
CÓDIGO	AP-BIB-FO-07	VERSIÓN	1	VIGENCIA	2014	PÁGINA	3 de 4

ABSTRACT: (Máximo 250 palabras)

Analysis process of discriminating sounds and alert detection system it was used as a tool, the training of Neural Networks (ANN), using the MATLAB Neural Networks Toolbox In these applications, the use of neural networks as a tool for the sound patterns automatic recognition has had quite success. The acoustic patterns recognition characteristics make of this analysis a posible application to be implemented for DSPIC33EP256GP502 Microchip microcontroller. Preliminary tests show a higher efficiency of 75% in identifying three types of sounds of emergency (ambulance, police, fireman).

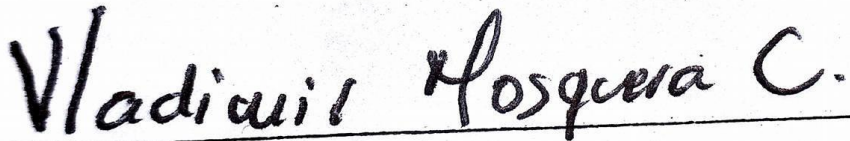
APROBACION DE LA TESIS

Nombre Presidente Jurado:

Firma:

Nombre Jurado: Vladimir Mosquera Cerquera

Firma:



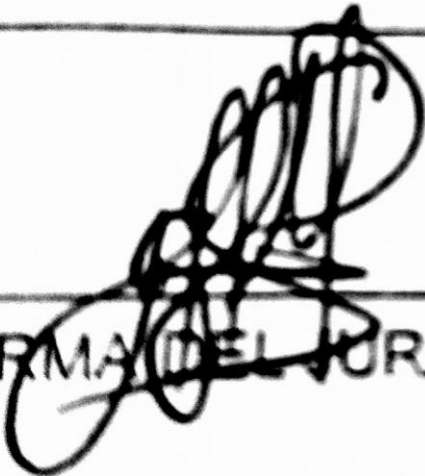
Vigilada mieducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.

	UNIVERSIDAD SURCOLOMBIANA GESTIÓN SERVICIOS BIBLIOTECARIOS					   	
	DESCRIPCIÓN DE LA TESIS Y/O TRABAJOS DE GRADO						
CÓDIGO	AP-BIB-FO-07	VERSIÓN	1	VIGENCIA	2014	PÁGINA	4 de 4

Nombre Jurado: José de Jesús Salgado Patrón

Firma:



FIRMA DEL JURADO

Vigilada mieducación

La versión vigente y controlada de este documento, solo podrá ser consultada a través del sitio web Institucional www.usco.edu.co, link Sistema Gestión de Calidad. La copia o impresión diferente a la publicada, será considerada como documento no controlado y su uso indebido no es de responsabilidad de la Universidad Surcolombiana.

**DISEÑO E IMPLEMENTACIÓN DE UN PROTOTIPO DE BRAZALETE DE DETECCIÓN
DE SONIDOS DE ALARMA PARA PERSONAS SORDAS E HIPOACÚSICAS**

**NICOLÁS CHARRY MORENO COD: 2010192028
DAVID EDUARDO CLAVIJO OSORIO COD: 2010192355**

**UNIVERSIDAD SURCOLOMBIANA
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA ELECTRÓNICA
NEIVA-HUILA**

**DISEÑO E IMPLEMENTACIÓN DE UN PROTOTIPO DE BRAZALETE DE DETECCIÓN
DE SONIDOS DE ALARMA PARA PERSONAS SORDAS E HIPOACÚSICAS**

**NICOLÁS CHARRY MORENO COD: 2010192028
DAVID EDUARDO CLAVIJO OSORIO COD: 2010192355**

**TRABAJO DE GRADO PARA OPTAR AL TITULO DE :
INGENIERO ELECTRONICO**

**DIRECTOR:
AGUSTIN SOTO OTALORA
INGENIERO ELECTRONICO**

**UNIVERSIDAD SURCOLOMBIANA
FACULTAD DE INGENIERÍA
PROGRAMA DE INGENIERÍA ELECTRÓNICA
NEIVA-HUILA
2017**

NOTA DE ACEPTACION

Firma del Director del Proyecto

Firma del Primer Jurado

Firma del Segundo Jurado

Neiva, Mayo 2017

DEDICATORIA Y AGRADECIMIENTOS

“Los tiempos difíciles pasan, pero los hombre fuertes permanecen...” -- Anónimo

*“Creo Firmemente que el mejor y mayor momento de cualquier hombre, su logro más grande y su mayor satisfacción, es aquel momento sublime en que después de haber trabajado arduamente con todo su empuje, esfuerzo, dedicación y corazón a favor de una causa noble, se encuentra exhausto en el campo de batalla, ¡ Victorioso !”
-- Vince Lombardi*

El esfuerzo, la dedicación y el compromiso son las bases del éxito, pero la motivación y la confianza en sí mismo deben ser el pan de cada día, deben ser inquebrantables. Como Trabajo de grado, comprendimos que el ingeniero llega a la sociedad para hacer grandes cambios, para dejar Huella. Nuestro trabajo no solo recorrió la enseñanza académica de un trabajo de investigación e implementación, sino las virtudes que como profesional debemos desarrollar, y hoy representan el primer paso para dejar una Huella que se haga imborrable.

Agradecemos en primera medida A Dios, A nuestros padres y Nuestras Familias, claramente ellos tienen el valor más importante para que hubiéramos alcanzado este logro. Al Ingeniero Agustín Soto Otálora Director del Proyecto, Al ingeniero Vladimir Mosquera y a las diferentes personas que con su conocimiento, cordialidad y asesoría nos guiaron para atar los cabos sueltos que en un momento parecían difíciles de hallar. Finalmente pero no menos importantes, a los diferentes Ingenieros Electrónicos de la Universidad Surcolombiana que como docentes dieron a través del tiempo el valor agregado a nuestra formación, así como nuestros compañeros de formación y amigos.

TABLA DE CONTENIDO

	PAG.
RESUMEN.....	11
ABSTRACT.....	11
1. INTRODUCCIÓN.....	14
2. OBJETIVOS.....	15
OBJETIVO GENERAL.....	15
OBJETIVOS ESPECIFICOS.....	15
3. MARCO TEORICO.....	16
3.1 REDES NEURONALES ARTIFICIALES.....	16
3.1.1 Elementos de una Red Neuronal.....	17
3.1.1.1 Unidad de Procesamiento (Neurona).....	18
3.1.1.2 Conexión Entre Neuronas	18
3.1.1.3 Función de Activación o Salida.....	19
3.1.2 Clasificación de las Redes Neuronales.....	20
3.1.2.1 Topología de las RNA.....	20
3.1.2.2 Mecanismos de Aprendizaje.....	21
3.2 HIPOACUSIA.....	25
3.3 CORRELACIÓN CRUZADA.....	26
3.3.1 Métodos para la correlación.....	27
3.3.2 Propiedades de la correlación.....	27
3.4 MFCC (COEFICIENTES CEPSTRALES DE MEL).....	29
3.4.1 Procedimiento de extracción de los MFCC.....	31
3.4.1.1 Banco de Filtros de Mel.....	31
3.4.1.2 Cepstrum.....	33
4. METODOLOGIA.....	35
4.1 ALCANCE.....	35

4.2. ANÁLISIS, CARACTERIZACIÓN Y ACONDICIONAMIENTO DE LAS SEÑALES DE EMERGENCIA.....	36
4.2.1. Etapa de Filtrado.....	37
4.2.1.1 Filtro Pasa-Bajo.....	38
4.2.1.2 Filtro Pasa-Alto.....	41
4.3 ANÁLISIS DE MÉTODOS DE EXTRACCIÓN DE CARÁCTERÍSTICAS DE LAS SEÑALES DE EMERGENCIA USANDO MATLAB.....	44
4.3.1 Transformada Discreta de Fourier.....	44
4.3.2. Cepstrum.....	46
4.3.3 Coeficientes Ceptrales de Mel (MFFC's).....	47
4.3.4 Espectrograma.....	48
4.4 ANÁLISIS DE MÉTODOS DE DISCRIMINACIÓN DE LAS SEÑALES DE EMERGENCIA USANDO MATLAB.....	49
4.4.1 Redes Neuronales para el reconocimiento de patrones	49
4.4.2 Redes Neuronales por mapas auto-organizados.....	51
4.5 SELECCIÓN DEL MEJOR MÉTODO PARA LA IDENTIFICACION Y DISCRIMINACION DE LAS SEÑALES DE EMERGENCIA.....	52
4.5.1 Transformada de Fourier y Red por reconocimiento de patrones.....	52
4.5.2 Cepstrum y Red por reconocimiento de Patrones.....	54
4.5.3 Coeficientes Ceptrales de Mel y Correlación Cruzada.....	54
4.5.4 Espectrograma y Red por mapas auto-organizados (SOM).....	56
4.5.5 Discusión y Selección del método más adecuado.....	58
4.6 IMPLEMENTACIÓN DEL MÉTODO DE IDENTIFICACIÓN Y DISCRIMINACIÓN EN UN CHIP PROCESADOR DIGITAL DE SEÑALES.....	59
4.6.1 Conversión Analógica a digital.....	60
4.6.2 Implementación Filtro Digital.....	60
4.6.3 Implementación de la Transformada de Fourier.....	62
4.6.4 Implementación de Ventana de Hamming.....	62
4.6.5 Implementación Red Neuronal.....	64
4.7 DISEÑO PROTOTIPO DEL BRAZALETE.....	64

4.7.1 Interfaz Visual.....	65
4.7.1.1 Visualización Hora-Fecha.....	67
4.7.1.2 Menú y configuración.....	68
4.7.2 Interfaz Mecánica.....	71
4.7.2.1 Batería.....	71
4.7.2.2 Motor Vibrador.....	72
4.7.2.3 Modulo de Carga de la Batería.....	72
4.7.2.4 Micrófono.....	73
4.7.3 Diseño del Prototipo del Brazalet.....	74
5. CONCLUSIONES.....	77
5.1 Discusión acerca de los resultados obtenidos en MATLAB	77
5.2 Discusión acerca de los métodos de identificación y discriminación propuestos.....	77
5.3 Discusión acerca de la implementación de los algoritmos de identificación y detección en el sistema embebido.....	78
6. RECOMENDACIONES.....	80
BIBLIOGRAFIA.....	81
ANEXOS.....	83

LISTA DE TABLAS

	PÁG.
Tabla 1. Funciones: escalón, semi-lineal y sigmoidal.....	19
Tabla 2. Vector Entrada 1x12288.....	52
Tabla 3. Matriz Clasificación 3x12288.....	52
Tabla 4 .Resultados de las Pruebas de los Métodos de extracción de características e identificación.....	57
Tabla 5. Características del Dispostivo PIC18F4685.....	65

LISTA DE ECUACIONES

	PÁG.
Ecuación 1 .salida de la neurona.....	18
Ecuación 2 .salida de la neurona al tiempo $t+1$	18
Ecuación 3. Función tangencial.....	19
Ecuación 4. Variación de pesos para aprendizaje supervisado.....	21
Ecuación 5 .La regla de Hebb en su forma más elemental.....	23
Ecuación 6. Correlación de dos funciones discretas.....	25
Ecuación 7. Correlación de dos funciones continuas.....	25
Ecuación 8. Propiedad de la Correlación (Asociatividad).....	26
Ecuación 9. Propiedad de la Correlación (Distributividad).....	26
Ecuación 10. Propiedad de la Correlación (Homogeneidad).....	27
Ecuación 11 Propiedad de la Correlación (Incomutatividad).....	27
Ecuación 12. Propiedad de la Correlación (Identidad).....	27
Ecuación 13. Propiedad de la Correlación (Reflexion).....	27
Ecuación 14. Propiedad de la Correlación (Invarianza Temporal).....	27
Ecuación 15. Propiedad de la Correlación (Relación con la Convolucion).....	28
Ecuación 16. Ecuación de Cepstrum.....	32
Ecuación 17. Modelo de Producción de voz.....	33
Ecuación 18. Función de Transferencia del Filtro pasabaja.....	37
Ecuación 19. Frecuencia de Corte analógica del filtro.....	38
Ecuación 20. Transformación Bilineal.....	38
Ecuación 21.Relacion de frecuencias analógicas y digitales.....	38
Ecuación 22.Frecuencia de muestreo del filtro digital.....	39
Ecuación 23. Función de transferencia del filtro pasa-alto.....	41
Ecuación 24. Transformada discreta de Fourier.....	43
Ecuación 25. Ecuación de Cepstrum.....	45
Ecuación 26. Ecuación Espectrograma.....	47
Ecuación 27. Ecuación Normalización.....	50
Ecuación 28. Ventana de Hamming.....	61

LISTA DE FIGURAS

	PÁG.
Figura 1. Esquema de trabajo de una red neuronal.....	18
Figura 2. Arquitectura de una red con conexiones hacia delante. 3 tipos de capas de entrada, ocultas y de salida.....	20
Figura 3. Diagrama esquemático de un sistema de entrenamiento supervisado.....	22
Figura 4. Grafica de Mel/Hz.....	28
Figura 5. Esquema para la Obtención de los MFCC.....	29
Figura 6. Diagrama del procedimiento de extracción de características mediante MFCC para cada bloques de muestras $v(n)$	30
Figura 7. Barras que indican las frecuencias de los centros de filtros triangulares.....	31
Figura 8. Banco de Filtros triangulares de MFCC.....	31
Figura 9. Sistema de Sirena de Perifoneo “ELECTRONIC SIREN SPRINT”.....	35
Figura 10. Espectrograma de la señal de ambulancia.....	35
Figura 11. Espectrograma de la señal de Policía.....	36
Figura 12. Espectrograma de la señal de Bomberos.....	36
Figura 13. Características del Filtro Pasabaja diseñado.....	36
Figura 14. Esquemático de un filtro pasabaja	37
Figura 15. Respuesta en frecuencia del filtro pasabaja análogo	38
Figura 16. Respuesta en frecuencia del filtro digital pasabaja	39
Figura 17. Esquemático del filtro pasa-alto análogo diseñado	40
Figura 18. Esquemático de un filtro pasa-alto	40
Figura 19. Respuesta en Frecuencia del filtro pasa-alto análogo.	41
Figura 20. Respuesta en frecuencia del filtro digital pasa-alto.	42
Figura 21. Extracción de Características mediante FFT.	44
Figura 22. Enventanado de un Frame correspondiente a la señal de Ambulancia.....	44
Figura 23. FFT de 512 pts de la señal de Ambulancia.....	45
Figura 24. Extracción de Características mediante Cepstrum	45
Figura 25. Extracción de características por medio de los MFCC's	46
Figura 26. Señal y MFCC's obtenidos de la sirena de Ambulancia	47
Figura 27. Extracción de Características por Espectrograma	47
Figura 28. Espectrograma Señal de Ambulancia	48
Figura 29. Función Tansig	49
Figura 30. Función Softmax	49
Figura 31. Distribución de neuronas en la capa de salida de una Red SOM	51
Figura 32. Método extracción de características e identificación DFT y Red Patternet.....	52
Figura 33. Topología red Patternet	53
Figura 34. Método extracción de Características e identificación Cepstrum y Red patternet.....	53
Figura 35. Método extracción de Características e identificación MFCC's y Correlacion Cruzada	53
Figura 36. Método extracción de Características e identificación Espectrograma y Redes SOM	55
Figura 37. Topología primeras 31 SOM.....	55
Figura 38. Topología SOM final.....	55
Figura 39. Sistema neuronal de clasificación.....	55
Figura 40. Tratamiento de un Buffer de 0,5 seg de Señal y Espectrograma Resultante.....	56
Figura 41. Proceso Previo de obtención y tratamiento de la información para su compresión...	59
Figura 42. Proceso de Obtención de la Señal	59
Figura 43. Espectro de la Señal de Prueba	59

Figura 44. Espectro de la Señal de Prueba Filtrada	60
Figura 45. Espectro Calculado de la Señal de Prueba Filtrada en el DSPIC.....	61
Figura 46. FFT de 512 Puntos para el Espectro Unilateral y Ventana de Hamming para la Señal de Prueba sin Filtrar realizada por el DSPIC	62
Figura 47. Espectrograma Señal de Emergencia Bomberos obtenida con el DSPIC.....	62
Figura 48. Diagrama de Bloques del Hardware del dispositivo	63
Figura 49. Nokia 6100LCD PCF8833.....	64
Figura 50. Pines DS1302	66
Figura 51. Visualización Hora-Fecha	66
Figura 52. Visualización Menú Principal	67
Figura 53. Visualización Configuración Hora	67
Figura 54. Visualización Notificación de la Hora Configurada	68
Figura 55. Visualización Configuración Alarma (izq), Visualización Notificación de la Hora Configurada (der).....	68
Figura 56. Visualización Configuración Fecha	68
Figura 57. Visualización Notificación de la Fecha Configurada	69
Figura 58. Visualización Cronómetro	69
Figura 59. Visualización Menú de Opciones	69
Figura 60. Visualización Configuración de Ajuste de Contraste y % de carga de la Batería.....	70
Figura 61. Visualización de las tres alertas de Emergencia	70
Figura 62. Batería de Lithium 3.7 V – 1000Ma	71
Figura 63. Motor Vibrador.....	71
Figura 64. Modulo Cargador de Batería TP4056.....	71
Figura 65. Micrófono SPQ1410.....	72
Figura 66. Pista del Circuito y Esquemático Final.....	73
Figura 67. Esquemático Final del diseño industrial del prototipo del brazalete	74
Figura 68. Fotos del dispositivo implementado	75

RESUMEN

El respectivo proyecto de investigación consiste en el desarrollo e implementación de un dispositivo electrónico (brazalete) capaz de discriminar sonidos de alerta o emergencia, funcionando como un sistema de ayuda para las personas con sordera o hipoacusia, además de manejar una interfaz para el usuario bastante amigable con características adicionales como de Reloj, alarma, entre otras.

Para el proceso de discriminación de sonidos y el sistema de detección de alertas se utilizó como herramienta el entrenamiento de Redes Neuronales (RNA), con ayuda del Toolbox Neural Networks de MATLAB. En estas aplicaciones, el uso de redes neuronales para el reconocimiento automático de patrones sonoros ha tenido bastante éxito. Las características que presentan el reconocimiento de patrones acústicos hacen que este problema sea una aplicación factible para implementarse como una aplicación para el microcontrolador DSPIC33EP256GP502 de Microchip. Ensayos previos demuestran una eficiencia mayor del 75 % en la identificación de tres tipos de sonidos de emergencia (ambulancia, policía, bomberos).

Palabras claves:Redes Neuronales,DSPIC ,sonidos de emergencia,Sordera,Hipoacusia.

ABSTRACT

This research analyzes the development and implementation of an electronic device (bracelet) able to discriminate sounds and analyze emergency, functioning as a support system for people with deafness or hearing loss, as well as managing a user Interface pretty friendly with additional features like Clock, alarm, among others.

Analysis process of discriminating sounds and alert detection system it was used as a tool, the training of Neural Networks (ANN), using the MATLAB Neural Networks Toolbox In these applications, the use of neural networks as a tool for the sound patterns automatic recognition has had quite success. The acoustic patterns recognition characteristics make of this analysis a possible application to be implemented for DSPIC33EP256GP502 Microchip microcontroller. Preliminary tests show a higher efficiency of 75% in identifying three types of sounds of emergency (ambulance, police, fireman).

Keywords: Neural Networks, DSPIC, emergency sounds, deafness, hearing loss.

1. INTRODUCCIÓN

¿Por qué las Poblaciones Especiales corren Riesgo? Las poblaciones especiales corren riesgo por una diversidad de razones: La disminución de la movilidad, salud, visión y capacidad auditiva son factores que pueden limitar la capacidad de una persona para tomar las medidas rápidas que se necesitan para escapar durante una emergencia, por ejemplo en caso de un incendio.

Dependiendo de las limitaciones físicas, muchas de las medidas que un individuo puede tomar para protegerse de cualquier tipo de peligro que ponga en riesgo su integridad física pueden necesitar la ayuda de un cuidador, vecino, o alguien de afuera. Ante estas situaciones planteadas la Universidad Surcolombiana en cumplimiento de uno de sus deberes primordiales el cual es prestar servicio a la comunidad a través de la investigación y el constante desarrollo de su comunidad a cargo, que es la región Huilense y en segunda instancia la comunidad Nacional, busca desarrollar un proyecto que de solución a estas problemáticas de una manera concreta y especial, más específicamente a los obstáculos de comunicación que viven a diarios las personas sordas o hipoacúsicas. A continuación, se verá el desarrollo de un proyecto que busca mejorar la calidad de vida de este determinado grupo de población, incluyendo en su diario vivir el uso de un dispositivo electrónico que permita alertar de situaciones de emergencia y al mismo tiempo se convierta en una herramienta cotidiana de perfil discreto y amable al usuario buscando mejorar y dar un plus de calidad de vida utilizando los conocimientos de la Ingeniería como herramienta principal.

2. OBJETIVOS

GENERAL

Diseñar e implementar un prototipo de brazalete de detección de sonidos de alarma (sirena de ambulancia, sirena de carro de policía, sirena de carro de bomberos, pitos de carros, ruidos fuertes) para personas sordas y establecer su utilidad en la reacción ante situaciones de emergencia

ESPECÍFICOS

- Analizar distintas técnicas y arquitecturas de sistemas de reconocimientos de sonidos ambientales, de esta forma vislumbrar el diseño de un sistema que permita obtener la versión digitalizada de las señales sonoras distinguidas y su posterior procesamiento para la extracción de sus características según los parámetros acústicos seleccionados.
- Explorar y registrar métodos que aborden lo correspondiente a la clasificación de sonidos, tales como modelos estadísticos, redes neuronales o modelos basados en patrones que conlleven al diseño de un sistema discriminador de sonidos que opere en tiempo real.
- Desarrollar el prototipo del brazalete teniendo en cuenta la portabilidad, tamaño, comodidad y resistencia e integrarlo con los circuitos ya diseñados
- Establecer el correcto funcionamiento del dispositivo diseñado, mediante pruebas en diferentes ambientes tanto de interiores como exteriores.
- Establecer la utilidad del dispositivo mediante pruebas realizadas junto a la población con discapacidad, a partir de estas verificar la reacción del usuario ante situaciones de emergencia y su opinión del dispositivo en cuanto a funcionamiento y portabilidad.

3. MARCO TEORICO

En este apartado se argumenta de manera breve y concisa las bases teóricas de los principales temas a en el desarrollo del proyecto de grado. Se comenzaran con la explicación de las redes neuronales y continuando en su orden respectivo: La Hipoacusia, el proceso de adquisición de datos, la correlación Cruzada y los MFCC's.

3.1 REDES NEURONALES ARTIFICIALES

Para el desarrollo de un método de discriminación de sonidos adecuado para el dispositivo, se consideraron diferentes opciones, una de ellas las Redes Neuronales Artificiales (RNA o ANN por sus siglas en inglés).

Un concepto bastante claro y conciso de que es una red neuronal es el propuesto por Hech-Nielsen que predica lo siguiente¹ : “ Una red Neuronal artificial es un sistema de computación, hecho por un gran número de elementos simples elementos de proceso muy interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entrad externas”.

Las Redes Neuronales permiten la búsqueda de una solución a problemas complejos, no como una secuencia de pasos, sino como la evolución de unos sistemas de computación inspirados en el cerebro humano, y dotados por tanto de cierta “inteligencia”, los cuales no son sino la combinación de elementos simples de proceso interconectados, que operando de forma paralela en varios estilos y consiguen resolver problemas relacionados con el reconocimiento de formas o patrones, predicción, codificación, control y optimización entre otras aplicaciones.

Las RNA al margen de “parecerse” al cerebro presentan una serie de características propias del cerebro. Por ejemplo las RNA aprenden de la experiencia, generalizan de ejemplos previos a ejemplos nuevos y abstraen las características principales de una serie de datos².

- **Aprender:** adquirir el conocimiento de una cosa por medio del estudio, ejercicio o experiencia. Las ANN pueden cambiar su comportamiento en función del entorno. Se les muestra un conjunto de entradas y ellas mismas se ajustan para producir unas salidas consistentes.

- **Generalizar:** extender o ampliar una cosa. Las RNA generalizan automáticamente debido a su propia estructura y naturaleza. Estas redes pueden ofrecer, dentro de un margen,

¹ Pino Raúl; Gómez Alberto; De Abajo Nicolás. Introducción a la Inteligencia Artificial: Sistemas Experto, Redes Neuronales Artificiales y Computación Evolutiva. Oviedo, España. 2001. Servicios de publicaciones Universidad de Oviedo(Ed).

² Basogain Xabier Olabe. Redes Neuronales y sus Aplicaciones. Bilbao, España. 2006. Escuela Superior de Ingeniería de Bilbao(Ed).

respuestas correctas a entradas que presentan pequeñas variaciones debido a los efectos de ruido o distorsión.

- **Abstraer:** aislar mentalmente o considerar por separado las cualidades de un objeto. Algunas ANN son capaces de abstraer la esencia de un conjunto de entradas que aparentemente no presentan aspectos comunes o relativos.

También se destacan las ventajas tan importantes que tiene trabajar con las redes neuronales, ventajas que permiten un campo de trabajo amplio y efectivo para los diferentes fines como identificación de procesos, detección de fallos en sistemas de control, modelación de dinámicas no lineales entre otras aplicaciones. Estas ventajas son³:

- Aprendizaje adaptativo. Capacidad de aprender a realizar tareas basadas en un entrenamiento o experiencia inicial.

- Auto-organización. Una red neuronal puede crear su propia organización o representación de la información que recibe durante la etapa de aprendizaje

- Tolerancia a fallos. Gracias a poseer la información distribuida o vía de información redundante la destrucción parcial de una red puede conducir a una degradación de su estructura; sin embargo algunas capacidades de la red se pueden retener, incluso sufriendo daños considerables.

- Capacidad de generalización. Ante la entrada de datos nuevos es capaz de producir resultados coherentes de acuerdo con la naturaleza del problema para el cual han sido entrenadas.

- Operación en tiempo real. El computo neuronal puede realizarse en paralelo, bien via software o mediante maquinas especiales para obtener esta ventaja(hardware conexionista o masivamente paralelo)

3.1.1 Elementos de una red neuronal

Las redes neuronales están compuestas de unos elementos principales para su organización y funcionamiento, todos son iguales de importantes e indispensables para un correcto proceso en la red que se vaya a implementar, un breve concepto de algunos de estos. Se extraen los siguientes elementos⁴:

- Conjunto de unidades de procesamiento (Neuronas)

- Un estado de activación de cada neurona (Salida de la unidad y_i)

- El patrón de conectividad de las neuronas. Cada conexión definidas por un peso w_{ij} , en donde determinan el efecto que tiene la neurona i sobre la neurona j .

³ Pose Gestal Marcos. Introducción a las redes de neuronas artificiales. España. 2008. Universidade da Coruña (Ed).

⁴ Martin del Brío Bonifacio; Serrano Cinca Carlos. Fundamentos de las redes neuronales artificiales: Hardware y software. Zaragoza, España. 1995.Universidad de Zaragoza (Ed)

- regla de propagación, la cual determina la entrada efectiva s_j a una neurona desde sus entradas externas.
- Una función de activación F_j , la cual determina el nuevo nivel de activación basado en la entrada efectiva $s_j(t)$ y el nivel actual $y_j(t)$
- Una entrada externa (bias, umbral) θ_j para cada neurona, que tiene el efecto de rebajar la entrada de la red al rango de la función de activación.
- Un método para inferir la información: Regla de aprendizaje.

3.1.1.1 Unidad de procesamiento (Neurona)

Cada unidad lleva a cabo un trabajo simple: recibe entradas de fuentes externas y usa estas para calcular la señal de salida que se propaga a otras unidades. Otra función es el ajuste de los pesos sinápticos (la fuerza del acoplamiento sináptico o conexiones entre neuronas) estos se modelan como parámetros numéricos, $w_j = (w_{1j} + w_{2j}, \dots, w_{nj})$ o vector de pesos.

Se distinguen 3 tipos de neuronas⁵:

- De entrada: recibe señales desde el entorno
- De salida: Envían señales fuera del sistema
- Ocultas: no tienen contacto exterior; generan la representación interna de la información.

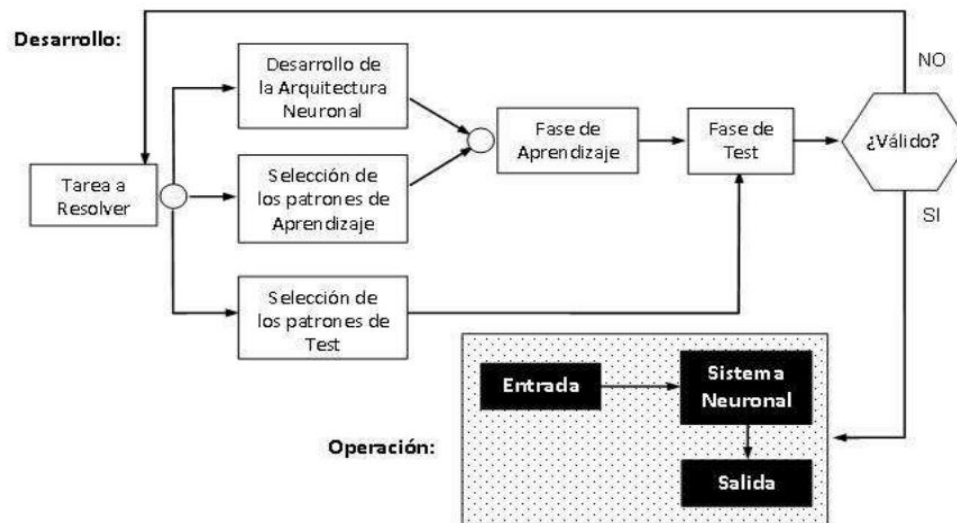
3.1.1.2 Conexión entre Neuronas

En la mayoría de los casos se asume que cada unidad da una contribución aditiva a la neurona a la cual se conecta. La entrada a la neurona j será S_j mas el bias o umbral θ_j . En un principio, se puede considerar este umbral, como el valor a superar para que la neurona se active.

Se puede tener sinapsis excitadoras o inhibitoras. Cuando el valor de activación de la neurona x_i , multiplicada por los pesos, w_{ij} , es positivo, entonces la unidad i sirve para activar la unidad j . el valor de activación de la neurona es dado por una función $f(S_j)$, llamada función de activación o función de transferencia⁶.

^{5 y 6} Pose Gestal Marcos. Introducción a las redes de neuronas artificiales. España. 2008. Universidade da Coruña (Ed).

Figura 1. Esquema de trabajo de una red neuronal.



Fuente. Gestal Pose Marcos. Introducción a las redes de neuronas artificiales .2008. Universidade da Coruña.

3.1.1.3 Función de activación o de salida

En una neurona biológica, la señal puede transmitirse o no a otras neuronas, en algunos casos, si se sobrepasa un cierto potencial, la señal se transmite, caso contrario no se transmite. La forma más simple de modelar este proceso es mediante una regla, de tal forma que si se sobrepasa el umbral, la neurona se active, caso contrario no, este sería el caso más simple, pero no el mejor⁷.

Se trabaja una función, la misma que da el efecto de la entrada total sobre la activación de la unidad. Es decir, una función f que tome la entrada total S_j y la activación actual y_i y produzca un nuevo valor de activación de la unidad j .

$$y_j = \sum_i w_{ij}(t) x_i(t) + \theta_j(t) \quad \text{Ecuación 1}$$

$$y_j(t+1) = \sum_i f_j(y_j(t), S_j(t)) \quad \text{Ecuación 2}$$

Una función de activación sencilla es tomar simplemente la suma de las entradas a la neurona S_j este será el valor de activación de la neurona. Otra función de activación común es probar si la suma S_j es mayor que un valor umbral (mínimo) entonces la señal se propaga (la neurona se activa, usualmente con un valor de activación igual a 1) de otra forma la neurona permanece en el mismo estado (usualmente con un valor de activación

^{7 y 8} Pose Gestal Marcos. Introducción a las redes de neuronas artificiales. España. 2008. Universidade da Coruña (Ed).

igual a 0, por ejemplo esto se simula con la función escalón). Otras funciones de transferencia típicas son: Función Lineal y función sigmoidal ⁸

Tabla 1. Funciones: escalón, semi-lineal y sigmoidal

$\varphi(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ -1 & \text{si } x < 0 \end{cases}$	$\varphi(x) = \begin{cases} -1 & \text{si } x < -c \\ 1 & \text{si } x > c \\ a. & x \text{ en otro caso} \end{cases}$	$\varphi(x) = \frac{1}{(1 + e^{-ax})}$
---	--	--

Otra función usada comúnmente es la función tanh, su salida está en el rango de -1 a 1 y puede ser escrita como

$$y = \frac{2}{(1 + e^{(-2x)}) - 1} \quad \text{Ecuación 3}$$

Aunque las funciones de activación tangente hiperbólica y sigmoidal son las más empleadas, las posibilidades son más amplias. Por ejemplo, la posibilidad de descomponer cualquier función en series de Fourier lleva a la conclusión que podría emplearse una función de activación sinusoidal.

3.1.2 Clasificación de las redes Neuronales

Las redes neuronales pueden ser clasificadas o caracterizadas desde cuatro aspectos diferentes, siguiendo a Hilera y Martínez⁹:

- Topología
- Mecanismo de aprendizaje.
- Tipo de Asociación realizada entre la información de entrada y salida.
- Forma de representación de la información de entrada y salida.

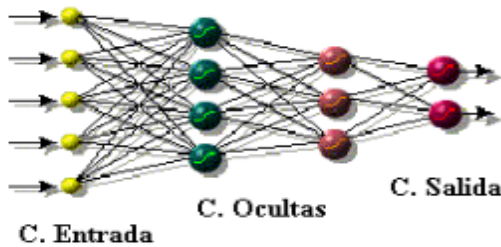
3.1.2.1 Topología de las RNA

La topología o arquitectura de una red consiste en la organización y disposición de las neuronas en la red formando capas de neuronas, en este sentido los parámetros fundamentales para esta distinción son:

⁹ HILERA, J.R. ; MARTÍNEZ, V.J.. "REDES NEURONALES ARTIFICIALES: Fundamentos, modelos y aplicaciones". .1995. RA-MA Editorial

- Numero de capas
- Numero de neuronas por capa
- Grado de conectividad “y”
- Tipo de conexiones entre neuronas.

Figura 2. Arquitectura de una red con conexiones hacia delante. 3 tipos de capas de entrada, ocultas y de salida.



En una red neuronal artificial, las neuronas están conectadas entre si, formando capas o niveles como se ilustra en la figura.

Fuente. Isasi, P. y Galván, I. Redes de neuronas artificiales. Un enfoque práctico. 2008. Pearson Education. Madrid.

Existen tres tipos de capas. Entrada, Ocultas y de Salida.

*De entrada: codifica las variables que definen el problema

*Ocultas: Internas a la red, no tienen contacto con el entorno exterior. Su importancia radica en que generan la representación de la información, en el sentido de representar todas las conexiones posibles entre todas las capas.

*De Salida: Transfieren información de la red hacia el exterior.

Por el número de capas se distinguen entre las redes monocapa y multicapa.

3.1.2.2 Mecanismos de aprendizaje

El aprendizaje, no es más que modificar los pesos en respuesta a la información de entrada. Esto implica:

Destrucción de las conexiones:

$$W_{ij} = 0$$

Modificación de las conexiones:

$$W_{ij}(t + 1) \neq W_{ij}(t)$$

Creación de conexiones:

$$W_{ij} \neq 0$$

En el contexto de redes neuronales se distingue entre la etapa de aprendizaje que se refiere al proceso por el cual la red modifica los pesos ante la información de entrada. Una vez que la red ha aprendido, esta está lista para realizar la tarea para la cual se la entrenó (a esta etapa se le denomina de funcionamiento). Normalmente en la mayoría de las redes se distingue entre

etapa de aprendizaje y etapa de funcionamiento. En función de la forma de aprendizaje, se puede clasificar a una red como¹⁰:

- Redes ON-LINE: la Red “aprende” durante su funcionamiento
- Redes OFF-LINE: se distingue entre una fase de entrenamiento o aprendizaje y una de funcionamiento

Reglas de aprendizaje

- **Aprendizaje supervisado:** Se realiza mediante un entrenamiento controlado por un agente externo (supervisor, maestro), que determina la respuesta que debería generar la red a partir de una entrada determinada. Implica que se debe diseñar un mecanismo que ajuste los pesos en función de cuan parecida es la salida de la red a la salida deseada. En una red de este tipo los pesos pueden evolucionar con una ecuación de la forma¹¹:

$$dw_j = ux_j(t) (y(t) - yd_j(t)) \quad \text{Ecuación 4}$$

$0 < u \leq 1$ Regula la velocidad de aprendizaje

Entre las formas de aprendizaje más conocidos dentro d este tipo están:

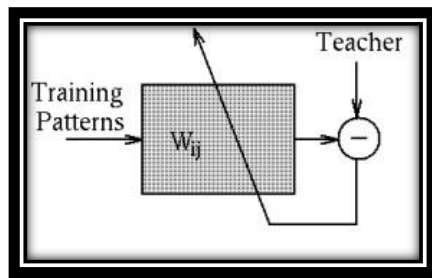
- Por corrección de error: Consiste en ajustar los pesos de las conexiones de la red en función de la diferencia entre los valores deseados y obtenidos en la salida de la red; es decir en función del error cometido en la salida. Se involucran como trabajo en este tipo la Regla de aprendizaje del perceptrón, Regla delta o del mínimo error cuadrado (LMS, Widrow- Hoff) y la regla delta generalizada o algoritmo de retropropagación de error. Algunos ejemplos de redes por corrección de Error son: El perceptrón, Adaline/Madaline y Backpropagation¹².

¹⁰ Matich Damián Jorge. Redes Neuronales: Conceptos Básicos y Aplicaciones. Rosario, Argentina. 2001. Universidad Tecnológica Nacional-Facultad Regional Rosario.

¹⁵ Aguilar Jorge. Introducción a las Redes Neuronales Artificiales. Material extraído de la web (http://www.flacsoandes.edu.ec/comunicacion/aaa/imagenes/publicaciones/pub_25.pdf)

- Aprendizaje Estocástico: Consiste en Realizar cambios aleatorios en los Pesos W_{ij} y evaluar su efecto a partir del objetivo deseado y de la distribución de probabilidades. Por lo general estos métodos parten de principios físicos que se dan en los sólidos (Termodinamica).El objetivo es llegar a un estado de mínima energía, evaluada con alguna función de energía. Redes que usan este tipo de algoritmo son Maquina de Boltzman y Maquina de Cauchy¹³.

Figura 3. Diagrama esquemático de un sistema de entrenamiento supervisado



Fuente. Material extraído de la web (http://html.rincondelvago.com/redes-neuronales_1.html)

- Aprendizaje por Refuerzo: Más lento que el de corrección de error, no se pretende exactitud en la salida del sistema, de modo que la función del supervisor se reduce a excitar o inhibir la actividad neuronal, mediante una señal de refuerzo que indica si la señal obtenida a la salida de la red se ajusta a la deseada (éxito =1, fracaso =-1), y en función de ello se ajustan los pesos en base a un mecanismo de probabilidades¹⁴.

- **Aprendizaje No Supervisado:** las redes con aprendizaje no supervisado (también conocido como autosupervisado) no requieren influencia externa para ajustar los pesos de las conexiones entre sus neuronas. La red no recibe ninguna información por parte del entorno que le indique si la salida genera en respuesta a una determinada entrada es o no correcta. Estas redes deben encontrar las características, regularidades, correlaciones o categorías que se puedan establecer entre los datos que se presenten en su entrada. Existen varias posibilidades en cuanto a la interpretación de la salida de estas redes, que dependen de su estructura y del algoritmo de aprendizaje empleado. En cuanto a los algoritmos de aprendizaje no supervisado, en general se suelen considerar dos tipos, que dan lugar a los siguientes aprendizajes¹⁵:

^{16 y 17} Aguilar Jorge. Introducción a las Redes Neuronales Artificiales. Material extraído de la web (http://www.flacsoandes.edu.ec/comunicacion/aaa/imagenes/publicaciones/pub_25.pdf)

¹⁴ Aguilar Jorge. Introducción a las Redes Neuronales Artificiales. Material extraído de la web (http://www.flacsoandes.edu.ec/comunicacion/aaa/imagenes/publicaciones/pub_25.pdf)

¹⁵ Matich Damián Jorge. Redes Neuronales: Conceptos Básicos y Aplicaciones. Rosario, Argentina. 2001. Universidad Tecnológica Nacional-Facultad Regional Rosario.

- Aprendizaje Hebbiano: Esta regla de aprendizaje es la base de muchas otras, la cual pretende medir la familiaridad o extraer características de los datos de entrada. El fundamento es una suposición bastante simple: si dos neuronas N_i y N_j toman el mismo estado simultáneamente (ambas activas o ambas inactivas), el peso de la conexión entre ellas se incrementa. Las entradas y salidas permitidas a la neurona son: $\{-1,1\}$ o $\{0,1\}$ (neuronas binarias). Esto puede explicarse porque la regla de aprendizaje de Hebb se originó a partir de la neurona biológica clásica, que solamente puede tener dos estados: activa o inactiva¹⁶

$$\Delta w_{ij} = y_i * y_j \quad \text{Ecuación 5}$$

- Aprendizaje competitivo y cooperativo: En estas redes las neuronas compiten y cooperan unas con otras con el fin de llevar a cabo una tarea dada. Cuando se presenta a la red cierta información de entrada, solo una de las neuronas de salida o un cierto grupo de las neuronas de salida o un cierto grupo de neuronas, se activan. Quedando anuladas las demás (valores de respuesta mínimos). Se orienta a la clusterización o clasificación de los datos de entrada. La característica principal es que si un patrón nuevo se determina que pertenece a una clase reconocida previamente, entonces la inclusión de este nuevo patrón a esta clase matizará la representación de la misma. Si el patrón de entrada se determinó que no pertenece a ninguna de las clases reconocidas anteriormente, entonces la estructura y los pesos de la red neuronal serán ajustados para reconocer la nueva clase.

¹⁶ Aguilar Jorge. Introducción a las Redes Neuronales Artificiales. Material extraído de la web (http://www.flacsoandes.edu.ec/comunicacion/aaa/imagenes/publicaciones/pub_25.pdf)

3.2 HIPOACUSIA

Para Definir la hipoacusia primero se debe empezar con el concepto sobre la pérdida de audición. La **pérdida de audición** es uno de los problemas de salud crónicos más comunes, que afecta a personas de todas las edades, en todos los segmentos de la población y de todos los niveles socioeconómicos. La pérdida de audición afecta aproximadamente a 17 de cada 1000 niños y jóvenes menores de 18 años. La incidencia aumenta con la edad: aproximadamente 314 de cada 1000 personas mayores de 65 años sufre pérdida de audición¹⁷. Ésta puede ser hereditaria o puede ser el resultado de una enfermedad, traumatismo, exposición a largo plazo al ruido, o medicamentos. La pérdida de audición puede variar desde una leve, pero importante disminución de la sensibilidad auditiva, a una pérdida total.

Sabiendo esto, se define como la **hipoacusia, sordera o deficiencia auditiva**, un trastorno sensorial que consiste en la incapacidad para escuchar sonidos, y que dificulta el desarrollo del habla, el lenguaje y la comunicación¹⁸. Uno de cada 300 niños nace con una deficiencia auditiva, y uno de cada 1.000 con una sordera profunda bilateral; el 74% de los niños que presentan retraso en el lenguaje padecen hipoacusia, y muchos niños con retraso psicomotor y alteraciones de la conducta tienen una hipoacusia leve que no ha sido diagnosticada.

¹⁷ Material extraído de la web (https://es.wikipedia.org/wiki/P%C3%A9rdida_de_audici%C3%B3n)

¹⁸ Material extraído de la web (<http://www.webconsultas.com/hipoacusia/hipoacusia-761>)

3.3 CORRELACIÓN CRUZADA

La correlación cruzada fue el método estadístico alternativo con el cual se trabajó para tratar el tema de la identificación de características, patrones y la discriminación de los sonidos. A parte de las redes neuronales también se utilizó este método, para realizar pruebas con ambos métodos y poder identificar cual sería el más efectivo y práctico a la hora de definir uno de los dos.

¹⁹ Frecuentemente, en el procesado digital de señales se necesita cuantificar el grado de interdependencia entre dos procesos o la similitud entre dos señales, tal cantidad es arrojada por la correlación cruzada. La correlación se suele utilizar en:

- La detección e identificación de señales.
- Detectores de desplazamiento.

En procesamiento de señales, la **correlación cruzada** (o a veces denominada “covarianza cruzada”) es una medida de la similitud entre dos señales, frecuentemente usada para encontrar características relevantes en una señal desconocida por medio de la comparación con otra que sí se conoce. Es función del tiempo relativo entre las señales, a veces también se la llama *producto escalar desplazado*, y tiene aplicaciones en el reconocimiento de patrones y en criptoanálisis. Dadas dos funciones discretas f_i y g_i la correlación cruzada se define como:

$$(f \star g)_i \stackrel{\text{def}}{=} \sum_j f_j^* g_{i+j}$$

Ecuación 6.

Donde la sumatoria se realiza sobre valores enteros de j apropiados; y el asterisco está indicando el conjugado. Para el caso de dos funciones continuas $f(x)$ y $g(x)$ la correlación cruzada se define como:

$$(f \star g)(x) \stackrel{\text{def}}{=} \int f^*(t)g(x+t) dt$$

Ecuación 7.

La correlación cruzada tiene una naturaleza similar a la convolución de dos funciones. Difiere en que la correlación no involucra una inversión de señal como ocurre en la convolución.

¹⁹ Ibarra Manzano Oscar .Notas Académicas Curso Procesamiento Digital de Señales.Mexico.2007
Universidad de Guanajuato(Ed).

3.3.1 Métodos para la correlación

²⁰Los métodos para trabajar la correlación cruzada más comunes y prácticos que existen son los siguientes:

- a) Por formula
- b) Cinta deslizante
- c) Matricial
- d) Malla
- e) Producto

El primer método, por fórmula, es el más complicado y es implementado solamente para su uso en computadoras. El método de la cinta deslizante facilita la comprensión visual de la mecánica que sigue la correlación. Los métodos de multiplicación y malla son algoritmos rápidos para el uso de humanos. El método matricial es usado para el planteamiento de ecuaciones que modelan algún sistema.

Sin embargo aunque existen estos métodos se trabajará la correlación cruzada por medio del software MATLAB.

3.3.2 Propiedades de la Correlación

Las siguientes son las propiedades más utilizadas a la hora de trabajar con la correlación cruzada. Al respecto, siempre debe recordar que aunque la correlación es una operación lineal no es conmutativa

- Asociatividad

$$\mathcal{F}(n) ** \mathcal{G}(n) ** \mathcal{H}(n) = \mathcal{F}(n) ** [\mathcal{G}(n) ** \mathcal{H}(n)] = [\mathcal{F}(n) ** \mathcal{G}(n)] ** \mathcal{H}(n) \quad \text{Ecuación 8.}$$

- Distributividad

$$\mathcal{F}(n) ** \mathcal{G}(n) + \mathcal{F}(n) ** \mathcal{H}(n) = \mathcal{F}(n) ** [\mathcal{G}(n) + \mathcal{H}(n)] \quad \text{Ecuación 9.}$$

En la aplicación de la propiedad de distributividad, se está dando prioridad a la correlación frente a la suma. Esta propiedad, declarada implícitamente, puede demostrarse fácilmente si se desarrolla la suma de correlación.

²⁰ Ibarra Manzano Oscar .Notas Académicas Curso Procesamiento Digital de Señales.Mexico.2007
Universidad de Guanajuato(Ed).

- Homogeneidad

$$A\mathcal{F}(n) ** \mathcal{G}(n) = A[\mathcal{F}(n) ** \mathcal{G}(n)] \quad \text{Ecuacion 10.}$$

$$\mathcal{F}(n) ** B\mathcal{G}(n) = B[\mathcal{F}(n) ** \mathcal{G}(n)]$$

$$A\mathcal{F}(n) ** B\mathcal{G}(n) = AB[\mathcal{F}(n) ** \mathcal{G}(n)]$$

En la aplicación de la propiedad de homogeneidad, se está dando prioridad a la multiplicación por una constante frente a la correlación.

- Inconmutatividad

La correlación no es conmutativa, esto es, dada:

$$\mathcal{F}(n) ** \mathcal{G}(n) = \mathcal{H}(n) \quad \text{Ecuacion 11.}$$

Se tiene a permutación de los operandos genera la secuencia reflejada

$$\mathcal{G}(n) ** \mathcal{F}(n) = \mathcal{H}(-n)$$

- Identidad

$$\mathcal{F}(n) ** \delta(n) = \mathcal{F}(n) \quad \text{Ecuacion 12.}$$

- Reflexión

$$\delta(n) ** \mathcal{F}(n) = \mathcal{F}(-n) \quad \text{Ecuación 13.}$$

- Invarianza temporal

La correlación es invariante con el tiempo. Dada la convolución

$$\mathcal{F}(n) ** \mathcal{G}(n) = \mathcal{H}(n) \quad \text{Ecuación 14.}$$

Se tiene que:

$$\mathcal{F}(n - \alpha) ** \mathcal{G}(n) = \mathcal{H}(n - \alpha)$$

$$\mathcal{F}(n) ** \mathcal{G}(n - \beta) = \mathcal{H}(n + \beta)$$

$$\mathcal{F}(n - \alpha) ** \mathcal{G}(n - \beta) = \mathcal{H}(n - \alpha + \beta)$$

- Relación con la convolución

$$\mathcal{F}(n) ** \mathcal{G}(n) = \mathcal{F}(n) * \mathcal{G}(-n) \quad \text{Ecuación 15.}$$

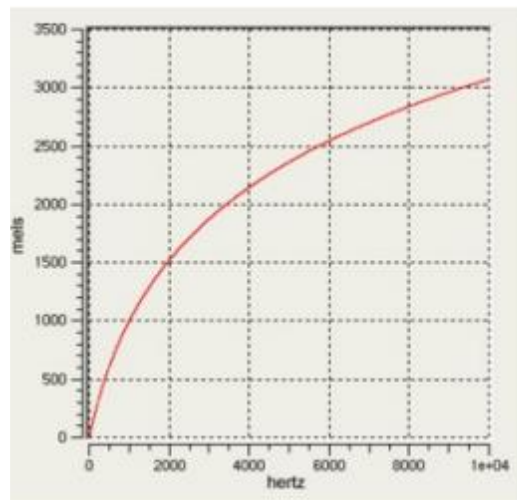
$$\mathcal{F}(n) * \mathcal{G}(n) = \mathcal{F}(n) ** \mathcal{G}(-n)$$

3.4 MFCC (COEFICIENTES CEPTRALES EN LAS FRECUENCIAS DE MEL)

Los MFCC's son esenciales a la hora de extracción de características en diferentes tipos de señales específicamente para la representación del habla con respecto a la perspectiva de la audición humana. En este caso se utilizaron como herramienta para el proceso de identificación de las diferentes señales de emergencia que se están trabajando.

Como conceptos se tiene que ²¹Los MFCC (Coeficientes Cepstrales de las frecuencias de Mel – Mel Frequency Cepstral Coefficients) son coeficientes para la representación del habla basados en la percepción auditiva humana. Los MFCC muestran las características locales de la señal de voz asociadas al tracto vocal (dependiendo del instante de análisis), Los coeficientes cepstrales se derivan de la transformada de Fourier (FT – Fourier Transform) o de la transformada del coseno discreta (DCT – Discrete Cosine Transform), pero la particularidad básica es que en MFCC las bandas de frecuencia están situadas logarítmicamente, según la escala Mel, en la que el punto de referencia se define equiparando un tono de 1000 Hz., 40 dBs por encima del umbral de audición del oyente, con un tono de 1000 mels. Tal y como se muestra en la siguiente figura:

Figura 4. Grafica de Mel/Hz.



Fuente. Rincón Llorente Carmen. Diseño, Implementación y Evaluación de técnicas de identificación de emociones a través de la voz. Abril 2007. Universidad Politécnica de Madrid (Ed).

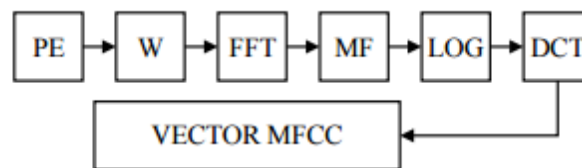
²¹Rincón Llorente Carmen. Diseño, Implementación y Evaluación de técnicas de identificación de emociones a través de la voz. Abril 2007. Universidad Politécnica de Madrid (Ed).

Los MFCCs son una característica ampliamente usada en el reconocimiento automático del discurso o el locutor y fueron introducidos por Davis y Mermelstein en los años 80 y han sido el estado del arte desde entonces.²²MFCCs se calculan comúnmente de la siguiente forma:

1. Separar la señal en pequeños tramos.
2. A cada tramo aplicarle la Transformada de Fourier discreta y obtener la potencia espectral de la señal.
3. Aplicar el banco de filtros correspondientes a la Escala Mel al espectro obtenido en el paso anterior y sumar las energías en cada uno de ellos.
4. Tomar el logaritmo de todas las energías de cada frecuencia mel
5. Aplicarle la transformada de coseno discreta a estos logaritmos.

Aunque los MFCC importan el modelo auditivo del ser humano y mejoran la resolución de la banda de frecuencia, que es sensible a la percepción humana, tiene una desviación orientada entre los formantes y el área sensible de escucha. En la siguiente figura, se muestra el esquema para la obtención de los MFCC.

Figura 5. Esquema para la Obtención de los MFCC



Fuente. Aguilar Torres Gualberto y Martinez Mascorro Guillermo Arturo. Sistema Para identificación de Hablantes robustos a Cambios en la Voz. 2012 Noviembre. Revista Ingeniuis(Ed).

Donde PE es el filtrado de pre énfasis, W es el ventaneo, continúa con la transformada rápida de Fourier (FFT), después se convierte a la escala logarítmica (LOG), y se aplica la Transformada Discreta Coseno (DCT).

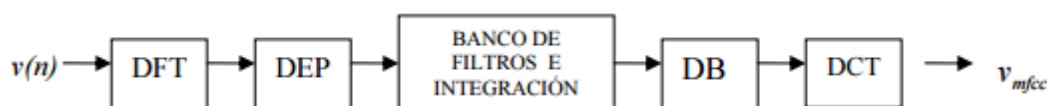
²² Kiyoharu Aizawa, Yuichi Nakamura, Shin'ichi Satoh. Advances in Multimedia Information Processing - .2004: 5th Pacific Rim Conference on Multimedia. Springer

3.4.1 Procedimiento de extracción de los MFCC

El procedimiento es el siguiente: ²³se realiza el enventanado de la señal en bloques de T segundos solapados, mediante ventanas de Hanning. Los pasos a seguir para cada bloque $v(n)$ son los siguientes:

- DFT: cálculo de la transformada discreta de Fourier (DFT)
- DEP: estimación de la Densidad Espectral de Potencia, $2 \text{ DEP}(k) = \text{DFT}(v(n))$
- Vector de coeficientes Mel (VCM): obtenido del filtrado de la DEP mediante un banco de N f filtros triangulares solapados (ver Figura 7), integrando la energía presente en cada banda (se obtienen N f componentes).
- Coeficientes cepstrales MFCC: representación del vector de coeficientes Mel en decibels y aplicación de la transformada coseno (DCT). Esta transformada realiza una compactación de los coeficientes obteniéndose N mfcc elementos.

Figura 6. Diagrama del procedimiento de extracción de características mediante MFCC para cada bloques de muestras $v(n)$



Fuente. Sarachaga Gabriela, Sartori Virginia, Vignoli Laura. Identificación Automática de Resumen en Canciones. Uruguay. Universidad de la Republica (Ed)

3.4.1.1 Bancos de filtros de Mel

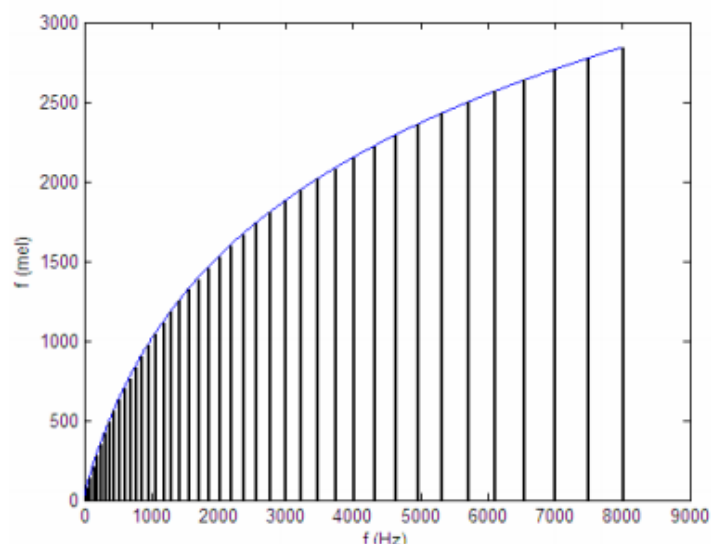
En los diferentes estudios que se realizan ²⁴Para caracterizar la señal de audio, se utiliza un banco de N f filtros triangulares solapados, que se adecuan al comportamiento del oído humano. De esta forma se intenta analizar el contenido energético de la señal, presente en cada una de las bandas críticas del sistema auditivo. Se eligen frecuencias linealmente espaciadas en la escala Mel⁷, y sus correspondientes frecuencias en Hz son los centros de los filtros del banco (ver Figura 7). Por esta razón se le denomina “Banco de filtros Mel”. La base de cada triángulo está comprendida entre las frecuencias centrales de sus filtros adyacentes como se ve en la Figura 8. Las alturas de los filtros se determinan manteniendo el área de cada triángulo unitaria.

La DEP es filtrada utilizando este banco, y se integra la energía de cada una de las N f bandas, obteniéndose así el vector de coeficientes Mel (VCM). Sus componentes representan la información de la DEP presente en cada banda, ponderada por la forma del filtro.

³⁴ Sarachaga Gabriela, Sartori Virginia, Vignoli Laura. Identificación Automática de Resumen en Canciones. Uruguay. Universidad de la Republica (Ed)

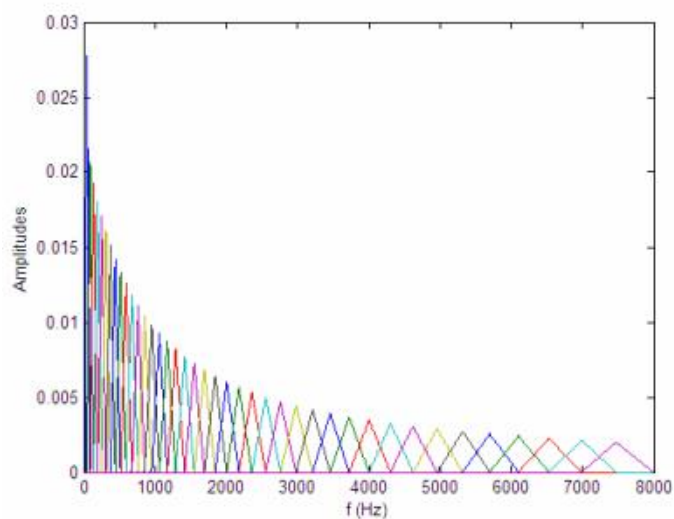
³⁵ Sarachaga Gabriela, Sartori Virginia, Vignoli Laura. Identificación Automática de Resumen en Canciones. Uruguay. Universidad de la Republica (Ed)

Figura 7. Barras que indican las frecuencias de los centros de filtros triangulares



Fuente. Sarachaga Gabriela, Sartori Virginia, Vignoli Laura. Identificación Automática de Resumen en Canciones. Uruguay. Universidad de la Republica (Ed)

Figura 8. Banco de filtros triangulares de MFCC



Fuente. Sarachaga Gabriela, Sartori Virginia, Vignoli Laura. Identificación Automática de Resumen en Canciones. Uruguay. Universidad de la Republica (Ed)

3.4.1.2 Cepstrum

Los coeficientes cepstrales se obtienen a partir del Análisis del Cepstrum real, que ha probado ser una herramienta muy útil en el estudio del reconocimiento automático de voz. El cepstrum se define como la transformada inversa de Fourier (IFT) del logaritmo de la DEP:

$$C(\tau) = IFT(\log(DEP)) \quad \text{Ecuación 16.}$$

Donde τ es una variable en un nuevo dominio del tiempo, llamado dominio cepstral. Éste se mide en unidades de cuelfrecuencias, un anagrama de la palabra frecuencias; de la misma forma se relaciona la palabra cepstrum con espectro (spectrum en inglés). Dado que la DEP es una función real y par, se puede demostrar que aplicar la IFT es equivalente a aplicarle su transformada directa FT, por lo tanto:

$$C(\tau) = FT(\log(DEP)) \quad .$$

Los coeficientes cepstrales de frecuencia Mel son una variante de la ecuación anterior. En primer lugar, el logaritmo no se aplica directamente a la DEP, sino a los coeficientes Mel (VCM)²⁵. La segunda modificación consiste en la utilización de la DCT en lugar de la FT. Esto busca lograr una reducción en el número de coeficientes del cepstrum, resaltando los de más bajas frecuencias

$$C(\tau) = DCT(10 * \log(VCM))$$

El cepstrum es un análisis en frecuencia de la DEP (expresada en decibeles). Sirve para analizar dos tipos de periodicidad que aparecen en la DEP:

- Periodicidades rápidas, debidas a la estructura armónica del espectro, que se repiten en los múltiplos de la frecuencia fundamental F_0 . Corresponden a la información de la fuente excitadora y se sitúan en la parte alta del cepstrum (altas cuelfrecuencias).
- Fluctuaciones mucho más lentas, no periódicas, que dan la envolvente del espectro. Se manifiestan en la parte baja del cepstrum (bajas cuelfrecuencias).

Sea $s(n)$ la señal de muestreada, $u(n)$ la señal de excitación y $h(n)$ la respuesta al impulso del aparato fonador. El modelo de producción de la voz está dado por la expresión siguiente, siendo $*$ la convolución:

$$s(n) = u(n) * h(n) \quad \text{Ecuación 17.}$$

²⁵ . Sarachaga Gabriela, Sartori Virginia, Vignoli Laura. Identificación Automática de Resumen en Canciones. Uruguay. Universidad de la Republica (Ed)

Entonces, Aplicando la transformada de Fourier a la expresión anterior, y recordando que la convolución de dos secuencias en el tiempo, implica su producto en frecuencia ²⁶

$$S(k) = U(k).H(k)$$

Tomando el módulo de S(k) al cuadrado, se obtiene la DEP:

$$DEP(k) = |S(k)|^2 = |U(k)|^2 \cdot |H(k)|^2$$

Ahora calculando el cepstrum

$$Cep(\tau) = IFT(\log(|U(k)|^2 |H(k)|^2)) = IFT(2 \cdot \log(|U(k)|)) + IFT(2 \cdot \log(|H(k)|))$$

Este resultado indica que el cepstrum de una señal de voz está dado por la suma del cepstrum de la excitación y el de la transferencia del aparato fonador. Como se explicó anteriormente, estas señales ocupan partes disjuntas del dominio cepstral, por lo tanto, se pueden obtener sus coeficientes cepstrales separadamente. Las diferentes informaciones se utilizan respectivamente para estimar la frecuencia fundamental (la altura) y para el reconocimiento automático de voz.

²⁶ . Sarachaga Gabriela, Sartori Virginia, Vignoli Laura. Identificación Automática de Resumen en Canciones. Uruguay. Universidad de la Republica (Ed)

4. METODOLOGÍA

El propósito de esta sección es describir los distintos grupos de métodos usados para alcanzar el objetivo del proyecto. Esta sección contiene el alcance del proyecto, creación y análisis de algoritmos, diseño e implementación del algoritmo seleccionado.

4.1 ALCANCE

La meta de este proyecto es diseñar un prototipo de brazalete capaz de identificar sonidos de emergencia tales como Ambulancia, Policía y Bomberos, logrando así discriminar cada uno de estos. Para lograr lo anterior se debe cumplir con los siguientes objetivos los cuales permitirán culminar el desarrollo adecuado del prototipo. Los objetivos de desarrollo a cumplir son:

- Análisis, Caracterización y acondicionamiento de las Señales de Emergencia
- Análisis de métodos de extracción de características de las señales de emergencia usando MATLAB
- Análisis de métodos de Discriminación de las señales de emergencia usando MATLAB
- Selección del Mejor método para la Identificación y discriminación de las señales de emergencia
- Implementación del método de Identificación y Discriminación en un chip procesador Digital de Señales
- Diseño del prototipo de Brazalete.

Como primera etapa de trabajo se debe analizar y determinar cómo son estas señales sonoras a identificar, se buscará conocer sus características en tiempo y frecuencia, de esta manera lograr caracterizar cada una individualmente y definir cuáles son los patrones que se necesitan identificar, además plantear un método de acondicionamiento de estas señales el cual permita conservar los patrones de tiempo y frecuencia identificados. Una vez caracterizadas las señales objetivo, a partir de los patrones obtenidos se buscará identificar las señales mediante métodos de extracción de características. De igual forma las características extraídas al ser distintas para cada señal, se debe disponer de un método de clasificación o discriminación para detectar qué tipo de señal de emergencia se está detectando, para ello se exploraran métodos estadísticos y de inteligencia artificial. A partir de un análisis de resultados se deberá optar por el mejor método de extracción de características y de clasificación el cual ha de ser implementado de forma embebida en un chip procesador de señales. Finalmente resuelto el problema de identificación y clasificación se diseñara el prototipo de brazalete y sus distintas componentes como interfaz mecánica y visual.

4.2 ANÁLISIS, CARACTERIZACIÓN Y ACONDICIONAMIENTO DE LAS SEÑALES DE EMERGENCIA

Una vez definidas las tres señales de Emergencia a identificar (Ambulancia, Policía, Bomberos) se procede a identificar las características de las señales. Para ello se adquirió una Sirena de Perifoneo para vehículos de Emergencia de 100W.

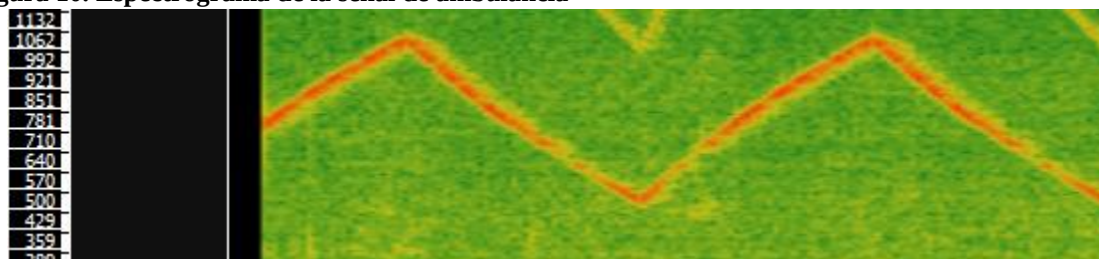
Figura 9. Sistema de Sirena de Perifoneo “ELECTRONIC SIREN SPRINT”



Este sistema de perifoneo de 100W está diseñado para reducción de posibles amenazas y suplementando la movilidad de los vehículos de emergencia. Ambulancias, bomberos, escoltas. Cuenta con 5 tonos de sirenas. Se realizaron Grabaciones con un tiempo de cinco segundos para cada una de las tres señales, con una Frecuencia de Muestreo de 22050 Hz debido a que en un principio se desconocen las características de frecuencia de las señales a tratar, con esta frecuencia de muestreo se garantiza cumplir con el teorema de Nyquist. Posteriormente haciendo uso del Software **Sonic Visualiser** se determinó las características en Frecuencia de cada Señal.

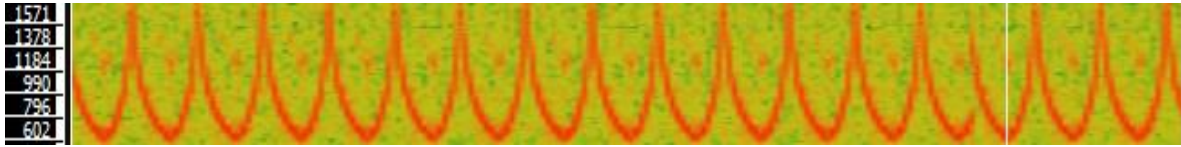
Para la Señal de Ambulancia se obtuvo que la señal se encuentre en un rango de Frecuencias de 530-1100 Hz según el espectrograma obtenido.

Figura 10. Espectrograma de la señal de ambulancia



En cuanto a la señal de Policía mediante el Espectrograma se obtuvo un rango de Frecuencias de 600-1500 Hz.

Figura 11. Espectrograma de la señal de Policía



Por último para la señal correspondiente a Bomberos el rango de frecuencias proporcionado por el Espectrograma es de 700-1300 Hz

Figura 12. Espectrograma de la señal de Bomberos



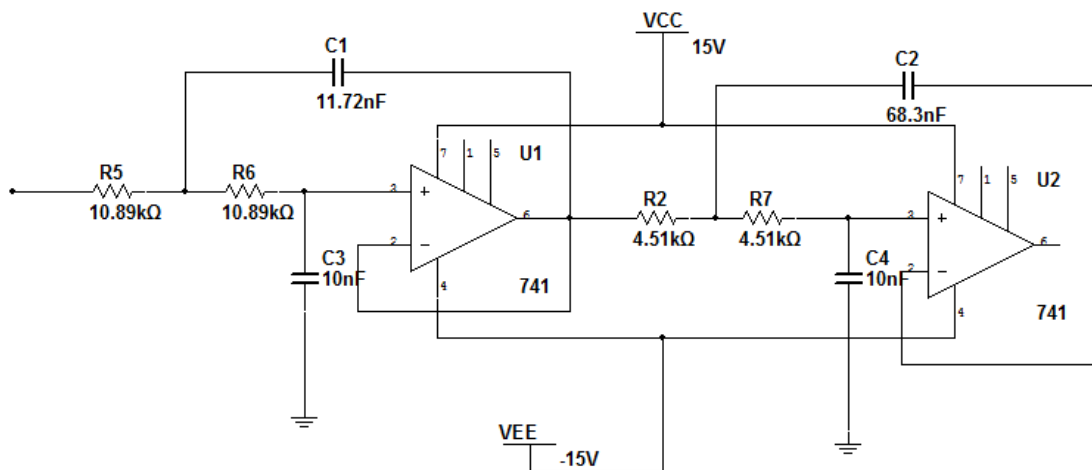
Una vez conocidos los rangos de Frecuencia se realiza el diseño de la etapa de acondicionamiento mediante Filtrado para así conseguir el rango de Frecuencias en el cual se desea trabajar. Se decidió realizar un Filtro Pasa Banda de 500-1350 Hz.

4.2.1 Etapa de Filtrado

El Filtro Pasa Banda es realizado mediante el establecimiento de una etapa de filtrado pasa bajo con una frecuencia de corte de 1350 Hz y una etapa de filtrado pasa alto con una frecuencia de corte de 500 Hz, de esta manera se obtiene el filtro pasa banda requerido para delimitar en frecuencia las señales de emergencia a procesar.

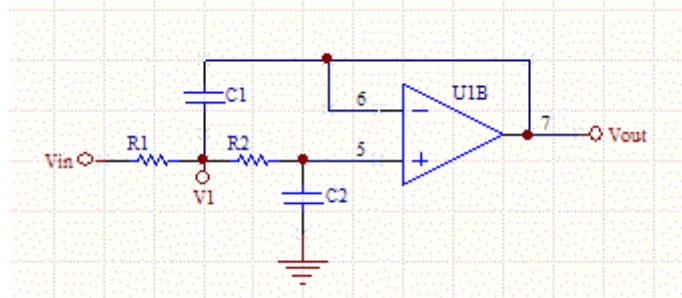
4.2.1.1. Filtro Pasa-Bajo

Figura 13. Filtro Pasa bajo diseñado



A Partir del Software FilterPro de Texas Instruments se genera un Filtro Análogo Pasabajo con Frecuencia de Corte de 1350 Hz . El filtro consta de dos etapas, cada etapa del Filtro Pasa Bajo se puede describir a través de su función de transferencia.

Figura 14. Esquemático de un filtro pasabaja



Fuente. Material extraído de la web (<https://www.unicrom.com>)

$$H(s) = \frac{\frac{1}{R_1 R_2 C_1 C_2}}{s^2 + s \left(\frac{1}{R_2 C_1} + \frac{1}{R_1 C_1} \right) + \frac{1}{R_1 R_2 C_1 C_2}}$$

Ecuación 18

De esta forma para nuestro filtro de tres etapas se tiene:

Etapas 1:

$$H_1(s) = \frac{7.195 \times 10^7}{s^2 + 1.567 \times 10^4 s + 7.195 \times 10^7}$$

Etapas 2:

$$H_2(s) = \frac{7.198 \times 10^7}{s^2 + 6493 s + 7.198 \times 10^7}$$

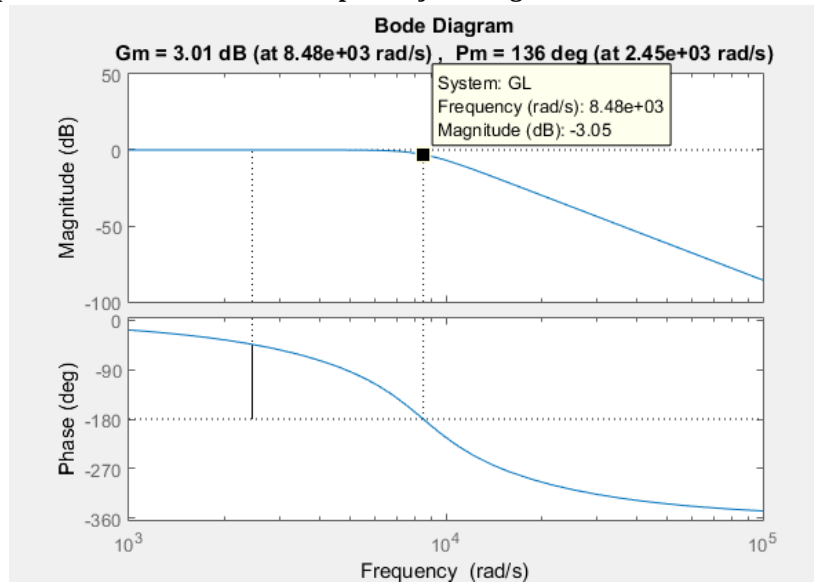
La función de Transferencia que describe el filtro es:

$$H_T(s) = H_1(s) * H_2(s)$$

$$H_T(s) = \frac{5.179 \times 10^{15}}{s^4 + 2.216 \times 10^4 s^3 + 2.457 \times 10^8 s^2 + 1.595 \times 10^{12} s + 5.179 \times 10^{15}}$$

Cuya respuesta en frecuencia es:

Figura 15. Respuesta en frecuencia del filtro pasabaja análogo



Se destaca que:

$$\Omega = 2\pi f$$

$$f = \frac{\Omega}{2\pi} = \frac{8.48 \times 10^3 \text{ rad/s}}{2\pi} \approx 1349 \text{ Hz} \quad \text{Ecuación 19}$$

La respuesta se aproxima bastante al filtro de 1350 Hz deseado.

Una vez obtenido el filtro análogo, se realiza la transformación al Filtro digital por medio de la transformación Bilineal. Esta técnica se usa para el diseño de Filtros IIR, la cual asegura la estabilidad del filtro digital, y mantiene el orden. La transformación es:

$$s = \frac{2}{T} \frac{z-1}{z+1} \quad \text{Ecuación 20}$$

Una particularidad de la transformación bilineal es el mapeo no lineal entre frecuencias analógicas y digitales. La relación entre ambas es:

$$\Omega = \frac{2}{T} * \tan\left(\frac{\omega}{2}\right)$$

$$\omega = 2 * \tan^{-1}\left(\frac{\Omega T}{2}\right) \quad \text{Ecuación 21}$$

Luego si se quiere diseñar un filtro digital de frecuencia de corte ω se debe utilizar un filtro analógico cuya frecuencia de corte sea Ω definida por la **Ecuación 21**. Esta operación de ajuste de frecuencias se denomina “Prewarping”.

Una vez diseñado el filtro análogo y obtenido su correspondiente función de transferencia, se procede a obtener su equivalente Filtro Digital IIR a partir del método de Transformación Bilineal. El correspondiente procedimiento al igual que para el filtro pasa alto se realiza a través de MATLAB.

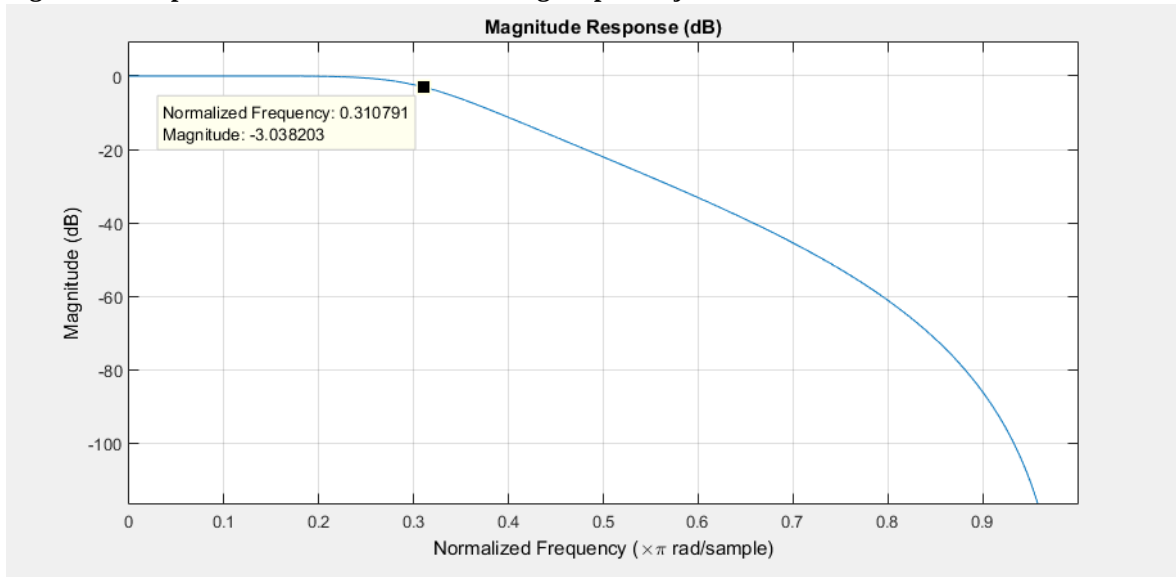
Pero previo a ello es necesario definir una correcta frecuencia de muestreo. A partir de la frecuencia analógica $\Omega=8.48 \times 10^3 \text{ rad/s}$ se busca obtener una frecuencia digital de $\omega = 0.25\pi$, retomando las ecuaciones de transformación lineal:

$$T = \frac{2}{\Omega} * \tan\left(\frac{\omega}{2}\right)$$

$$T = \frac{2}{8.48 \times 10^3 \text{ rad/s}} * \tan\left(\frac{0.3107\pi}{2}\right) = 88.507 \mu s \rightarrow F_s \approx 8000 \text{ Hz} \quad \text{Ecuación 22}$$

La frecuencia de Muestreo seleccionada es de 8000 Hz. La respuesta en frecuencia de este filtro digital será:

Figura 16. Respuesta en frecuencia del filtro digital pasabaja



Se realiza la verificación de la correspondencia entre el dominio digital y el análogo, para ello se retoman las ecuaciones de relación entre las frecuencias analógicas y digitales.

$$\omega = 0.3107\pi$$

$$\Omega = \frac{2}{\frac{1}{8000}} * \tan\left(\frac{0.3107\pi}{2}\right) = 8494.1 \text{ rad/s} \rightarrow f = \frac{8494.1}{2\pi} \approx 1351 \text{ Hz}$$

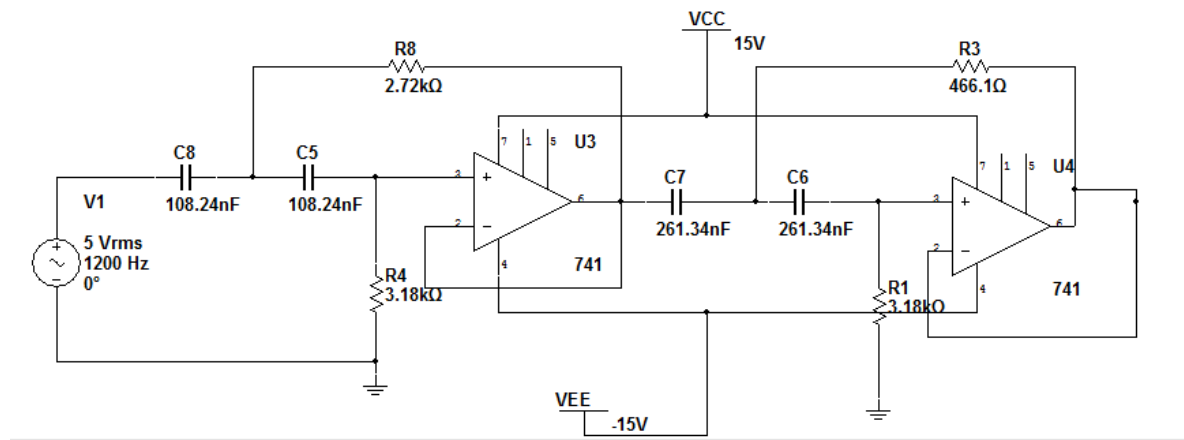
$$\omega = 2 * \tan^{-1}\left(\frac{8494.1 \text{ rad/s} * \frac{1}{8000}}{2}\right) = 0.3107\pi$$

Como se observa la respuesta en frecuencia del Filtro Digital Pasa Bajo es de 0.3107π rad/muestra en su frecuencia de corte, lo que equivale a 1351 Hz en frecuencia analógica, es decir, un valor muy aproximado a la respuesta del filtro analógico de 1350 Hz.

4.2.1.2 Filtro Pasa-Alto

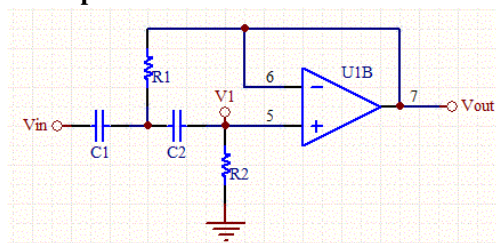
Por medio del Software Filter Pro se obtiene el filtro análogo pasa alto, Chebyshev con topología Sallen-Key.

Figura 17. Esquemático del filtro pasa-alto análogo diseñado



El filtro consta de dos etapas, cada etapa del Filtro Pasa Alto se puede describir a través de su función de transferencia:

Figura 18. Esquemático de un filtro pasa-alto



Fuente. Material extraído de la web (<https://www.unicrom.com>)

$$H(s) = \frac{s^2}{s^2 + s\left(\frac{1}{R_2 C_1} + \frac{1}{R_2 C_2}\right) + \frac{1}{R_1 R_2 C_1 C_2}}$$

Ecuación 23

El filtro consta de tres etapas, cada etapa del Filtro Pasa Bajo se puede describir a través de su función de transferencia:

Etapas 1:

$$H_1(s) = \frac{s^2}{s^2 + 5811 s + 9.868 \times 10^6}$$

Etapas 2:

$$H_2(s) = \frac{s^2}{s^2 + 2407 s + 9.878 \times 10^6}$$

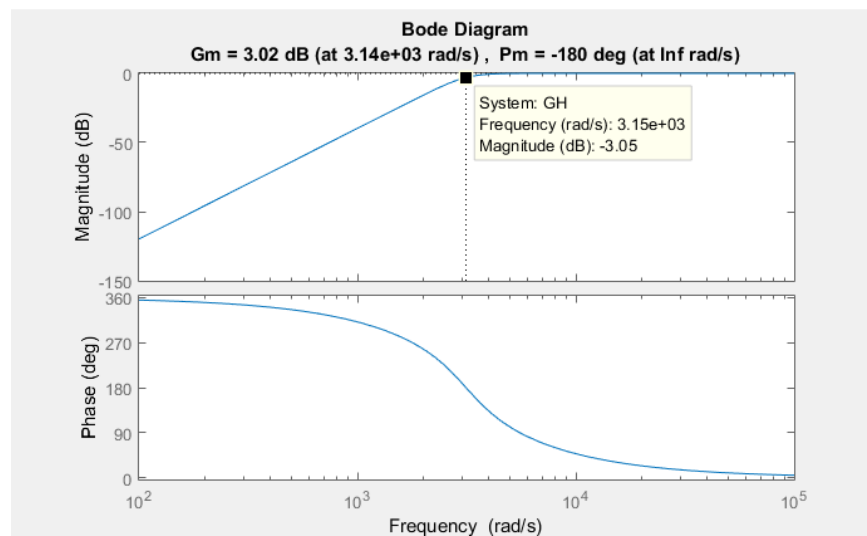
La función de Transferencia que describe el filtro es:

$$H_T(s) = H_1(s) * H_2(s)$$

$$H_T(s) = \frac{s^4}{s^4 + 8217 s^3 + 3.373 \times 10^7 s^2 + 8.115 \times 10^{10} s + 9.748 \times 10^{13}}$$

Cuya respuesta en frecuencia es:

Figura 19. Respuesta en Frecuencia del filtro pasa-alto analógico.



Se infiere que:

$$\Omega = 2\pi f$$

$$f = \frac{\Omega}{2\pi} = \frac{3.15 \times 10^3 \text{ rad/s}}{2\pi} \approx 501 \text{ Hz}$$

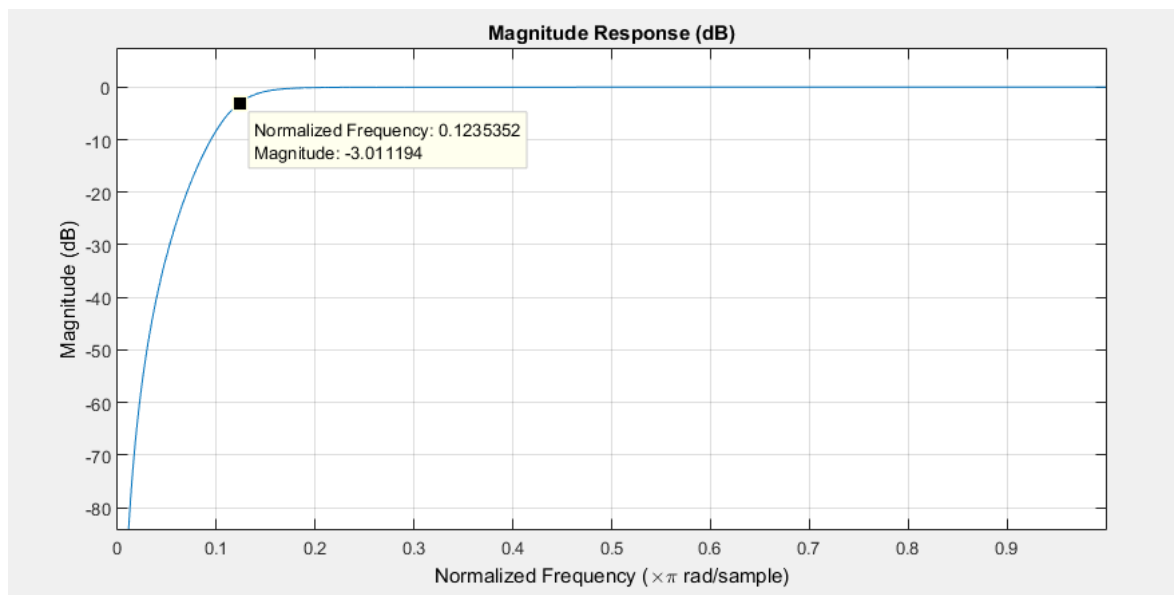
La respuesta se aproxima bastante al filtro de 500 Hz deseado. Para realizar la transformación bilineal se hace uso de MATLAB. En este caso se pretende usar la misma frecuencia de muestreo al igual que el filtro Pasa Bajo ($F_s=8000$ Hz), por tanto se debe encontrar el valor de la Frecuencia digital:

$$\omega = 2 * \tan^{-1} \left(\frac{\Omega T}{2} \right)$$

$$\omega = 2 * \tan^{-1} \left(\frac{3.15 \times 10^3 \frac{\text{rad}}{\text{s}} * \frac{1}{8000}}{2} \right) = 0.12375\pi \text{ rad/muestra}$$

La respuesta en frecuencia del Filtro Digital obtenida es:

Figura 20. Respuesta en frecuencia del filtro digital pasa-alto



Se verifica la correspondencia del Filtro Pasa Alto análogo con el digital:

$$\omega = 0.12375\pi$$

$$\Omega = \frac{2}{\frac{1}{8000}} * \tan \left(\frac{0.12375\pi}{2} \right) = 3149.95 \frac{\text{rad}}{\text{s}} \rightarrow f = \frac{3143.42}{2\pi} \approx 501 \text{ Hz}$$

$$\omega = 2 * \tan^{-1} \left(\frac{3143.42 \frac{rad}{s} * \frac{1}{8000}}{2} \right) = 0.12375 \pi \text{ rad/muestra}$$

Como se observa, la respuesta en frecuencia del Filtro Digital es de 0.12375π rad/muestra en su frecuencia de corte, lo que equivale a 501 Hz en frecuencia analógica. Se obtiene el mismo valor de frecuencia de corte del filtro analógico de 501 Hz.

4.3 ANÁLISIS DE MÉTODOS DE EXTRACCIÓN DE CARACTERÍSTICAS DE LAS SEÑALES DE EMERGENCIA USANDO MATLAB

La extracción de características se refiere al proceso de extraer algunas medidas numéricas de los datos en bruto de los patrones (representación inicial). Por otro lado, se define también como el proceso de formar un conjunto de características (de dimensión n) partiendo de los datos de entrada (de dimensión $m > n$). La extracción de características soluciona problemas de dimensionalidad que suelen ocurrir al implementar un clasificador, además de reducir el gasto computacional que realiza el hardware al momento de procesar los datos, permitiendo así la implementación de sistemas en tiempo real²⁷.

Realizada la etapa de Identificación y caracterización de las señales, se proponen distintos métodos de extracción de características para obtener los patrones señalados en el inciso anterior, se realizará su análisis matemático y su implementación en MATLAB. Estos métodos están principalmente basados en el análisis frecuencial. Los métodos a tratar serán: Transformada rápida de Fourier, Cepstrum, Coeficientes Cepstrales de Mel, Espectrograma. Todos los métodos anteriores tienen en común el uso de la FFT, se busca experimentar si sólo con la FFT es suficiente para obtener resultados eficientes, debido a que aliviaría el gasto computacional a la hora de realizar el embebido del sistema.

4.3.1 Transformada Rápida de Fourier

Una transformada de Fourier es una operación matemática que transforma una señal de dominio de tiempo a dominio de frecuencia y viceversa.

La FFT se origina a partir de la DFT dada por la siguiente ecuación:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-\frac{2\pi kn}{N}} \quad k = 0, 1, 2, \dots, N-1$$

Ecuación 24

²⁷ Romero, Luis Alonso; Calonge Cano, Teodoro. Capítulo 1.- Redes Neuronales y Reconocimiento de Patrones. Dpto. de Informática y Automática. Universidad de Salamanca- España ** Dpto de Informática. Universidad de Valladolid. España (Ed).

Una FFT (Transformada Rápida de Fourier) es una versión más rápida de la DFT que puede ser aplicada cuando el número de muestras de la señal es una potencia de dos. Un cálculo de FFT toma aproximadamente $N * \log_2(N)$ operaciones, mientras que DFT toma aproximadamente N^2 operaciones, así es que la FFT es significativamente más rápida²⁸.

En conclusión la FFT se implementó como un algoritmo que solucionaba la ineficiente forma en que se debía implementar la DTF, de esta manera permitió realizar análisis de frecuencia con un gasto computacional mucho menor y facilitar la realización de sistemas en tiempo real que tengan como base el análisis frecuencial.

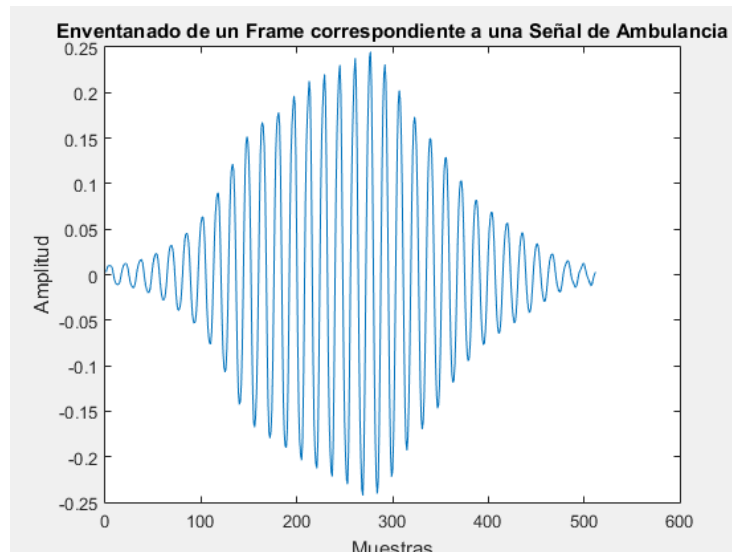
Visto el porqué de su ineficiencia e implementación de la DFT, se presenta su implementación de una FFT como algoritmo de extracción de Características en MATLAB:

Figura 21.Extraccion de Características mediante FFT



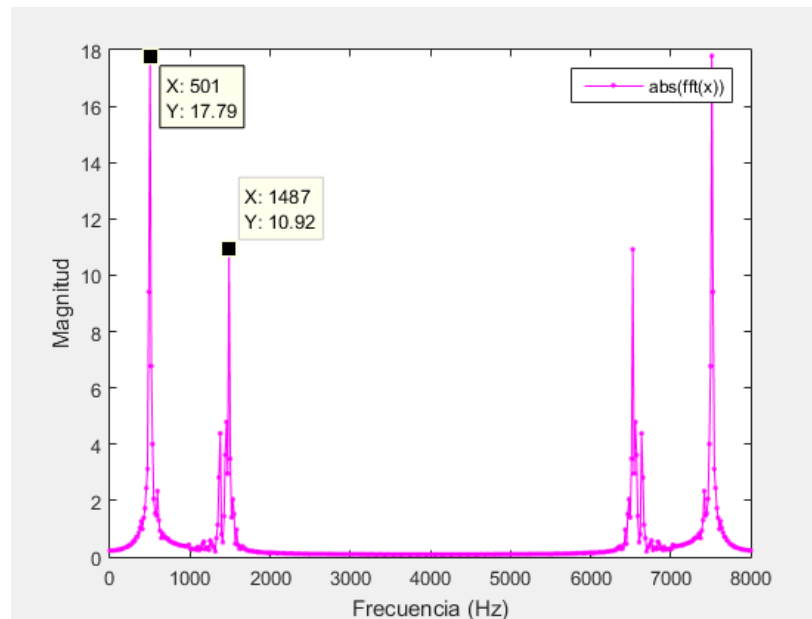
Básicamente se toma un Buffer de 4096 muestras correspondiente a un poco más de 0,5 segundos de señal, este Buffer se compone de 8 Frames de 512 muestras. Cada Frame pasa por un proceso de Filtrado, correspondiente al Filtro Pasa Banda anteriormente diseñado, posteriormente se realiza un proceso de enventanado mediante una ventana de Hamming (Figura 22) y finalmente se aplica la FFT de 512 Puntos. Se realizó uso de las funciones `fft` y `fftshift` en MATLAB para la realización de las pruebas.

Figura 22. Enventado de un Frame correspondiente a la señal de Ambulancia



²⁸ National Instruments.Transformada de Fourier y Analisis de Frecuencias. 2014
<http://www.ni.com/support/esa/cvi/analysis/analy3.html>

Figura 23. FFT de 512 pts de la señal de Ambulancia



4.3.2 Cepstrum

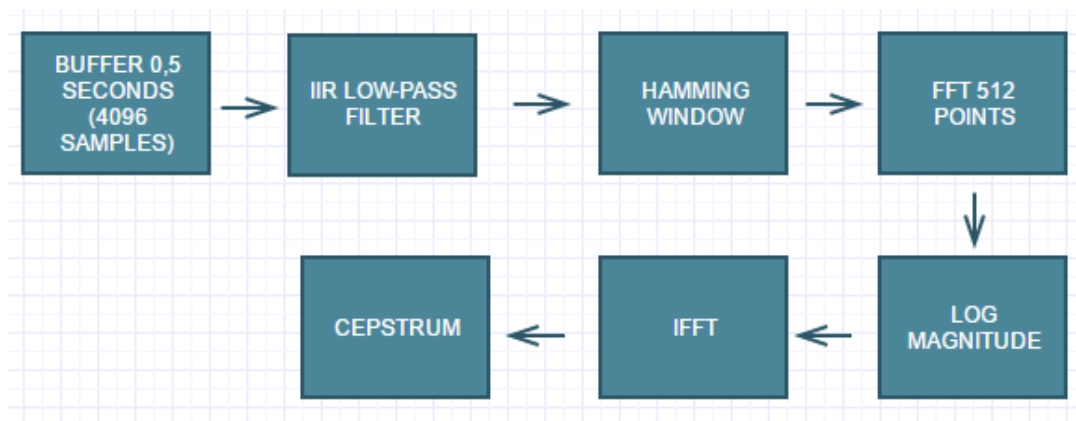
El cepstrum se define como la transformada inversa del logaritmo del módulo de la transformada de Fourier de la señal.

$$c(n) = TF^{-1}[\log|X(w)|], \quad X(w) = TF[x(n)]$$

Ecuación 25

El cepstrum permite distinguir entre sonidos vocalizados y no-vocalizados. El tono se puede determinar a partir del cepstrum. En cuanto a la implementación del Cepstrum como algoritmo de extracción de características en MATLAB se elaboró el siguiente procedimiento:

Figura 24.Extraccion de Características mediante Cepstrum



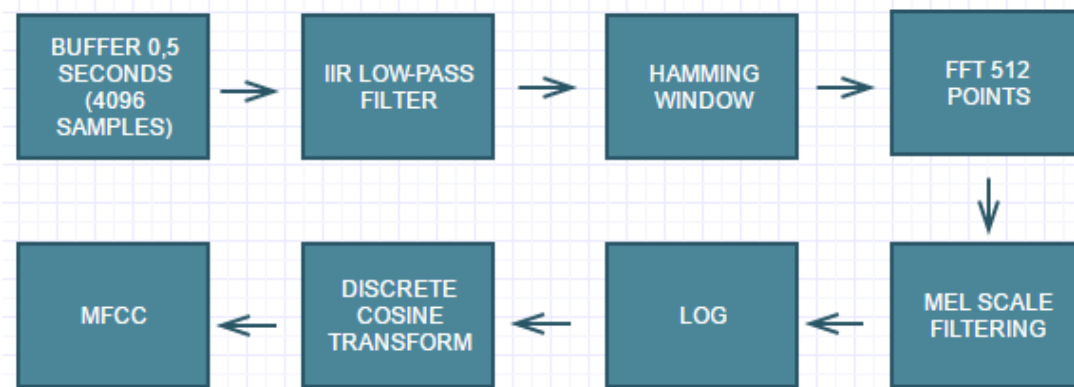
Nuevamente se extrae un Buffer de 4096 muestras, conformado por 8 frames de 512 muestras, cada una de las cuales se han filtrado y enventanado, para luego obtener la FFT de 512 puntos y completar el algoritmo del cepstrum mediante el logaritmo de la transformada inversa del logaritmo de la fft. Para su respectiva implementación se utilizaron las funciones de MATLAB **47**nális, **fft**, **log**, **ifft**.

4.3.3 Coeficientes Ceptrales de Mel (MFCC's)

Al aplicar esta técnica a la señal captada, se trasforma las muestras de la señal a un conjunto de coeficientes que representan eficientemente las propiedades espectrales y concentraciones de Energía, tratando de emular el tipo de procesamiento que realiza nuestro sistema auditivo. Al tener en cuenta las características del oído, se trata de asemejar el sistema al reconocimiento hecho por una persona. Este análisis se basa en el uso de la escala de frecuencia Mel, la cual es un espaciamiento lineal de la frecuencia por debajo de los 1000Hz y un espaciamiento logaritmo por arriba de los 1000Hz²⁹.

El esquema planteado para la implementación de este procedimiento es el siguiente:

Figura 25.Extraccion de características por medio de los MFCC's

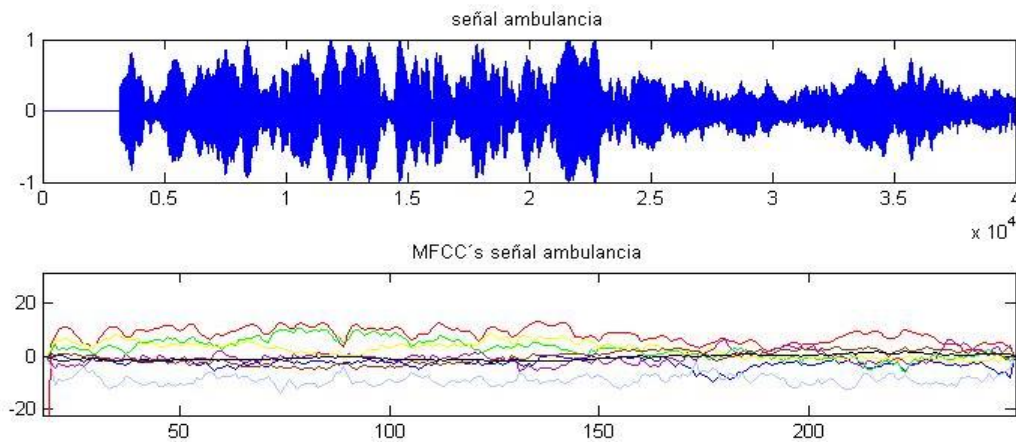


Para trabajar con este procedimiento se han encontrado bastantes proyectos de investigación similares, pero se realizará un enfoque en el trabajo guía realizado en el instituto norteamericano WPI (Worcester Polythecnic Institute) en el condado de Nueva Inglaterra, el proyecto llamado “Bird Call Identifier”, es un identificador de diferentes tipos de especies de aves silvestres a través de sus particulares cantos y silbidos. Se toma como base el algoritmo propuesto por medio de correlación y MFCC's para ser aplicado a señales de emergencia.

²⁹ HERNÁNDEZ PÉREZ, Adriana. Reconocimiento de entornos acústicos para dispositivos móviles. España. Universidad Politécnica de Valencia(Ed).

En primera Instancia se definen las 3 señales que estarán en la base de datos de consulta, se realiza su respectivo análisis, graficando la señal y los MFCC's de cada una de ellas. Las señales son: Sirena Ambulancia, Sirena Policía y Sirena de Bomberos. Utilizando posteriormente las diferentes funciones matemáticas alojadas en el MATLAB para realizar las operaciones necesarias con fin de obtener los MFCC's de cada una de las señales y guardarlos en una Base datos principal que se utilizan como comparación para cuando se ingresen señales de prueba.

Figura 26. Señal y MFCC's obtenidos de la sirena de Ambulancia



4.3.4 Espectrograma

Para el método 3 se implementa un análisis tiempo-frecuencia. Mediante STFT se realiza el análisis de Fourier a cada segmento de señal resultante de cortar la señal en segmentos de tiempo sobrepuestos por medio del uso de una ventana. La reunión del contenido frecuencial (espectros en frecuencia) de cada uno de estos segmentos cambiantes en el tiempo se denomina espectrograma.

$$STFT_x(T, \omega) = \int_{-\infty}^{\infty} x(t) w * (t - T) e^{-j\omega t} dt \quad \text{Ecuación 26}$$

Donde ω denota la frecuencia en una ventana $w(t)$ alrededor de $t = \tau$. El espectrograma es la magnitud al cuadrado de la STFT de la señal $x(t)$.

$$Spectrogram, x(T, \omega) = |STFT_x(T, \omega)|^2$$

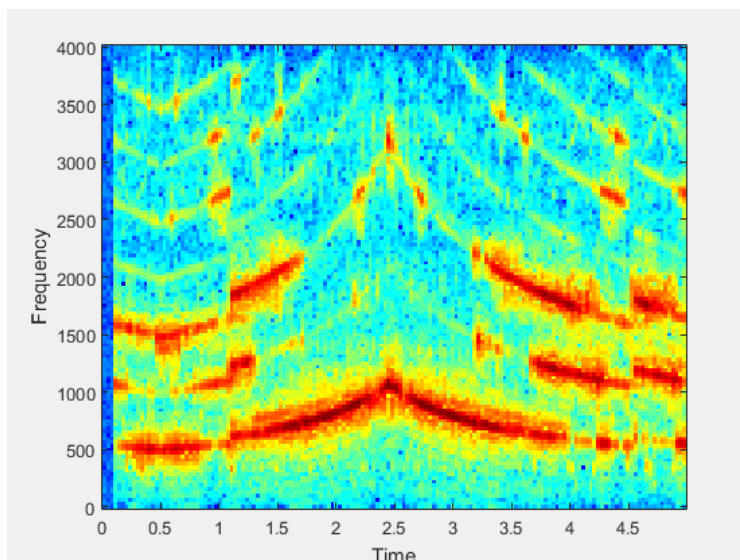
La metodología para el tercer método será la Siguiente:

Figura 27. Extracción de Características por Espectrograma



En la fase inicial de desarrollo todo el procesamiento espectral descrito es realizado mediante la función de MATLAB *spectrogram*, la cual retorna el espectrograma de las señales en consideración, y cuyos parámetros de funcionamiento son la señal a procesar, el número de puntos de la FFT por ventana, la frecuencia de muestreo de la señal, la función de ventana y el número de puntos de traslape entre ventanas.

Figura 28. Espectrograma Señal de Ambulancia



4.4 ANÁLISIS DE MÉTODOS DE DISCRIMINACIÓN DE LAS SEÑALES DE EMERGENCIA USANDO MATLAB

Así como se han planteado distintos métodos de extracción de características, en esta sección se presentan tres métodos distintos a considerar, cuya tarea es definir qué tipo de alarma ha sido detectada. Se presentan dos métodos de Inteligencia artificial, más específicamente para Redes Neuronales Artificiales, de las cuales se ponen a consideración dos tipos: Redes Neuronales para el Reconocimiento de Patrones y Redes Neuronales por mapas auto-organizados. Por otra parte como método estadístico se explorará la correlación cruzada.

4.4.1 Redes Neuronales para el Reconocimiento de Patrones

Un sistema de reconocimiento de patrones tiene uno de los siguientes objetivos³⁰:

³⁰ Romero, Luis Alonso; Calonge Cano, Teodoro. Capítulo 1.- Redes Neuronales y Reconocimiento de Patrones. Dpto. de Informática y Automática. Universidad de Salamanca- España ** Dpto de Informática. Universidad de Valladolid. España (Ed).

- Identificar el patrón como miembro de una clase ya definida (clasificación supervisada).
- Asignar el patrón a una clase todavía no definida (clasificación no supervisada, agrupamiento o clustering).

El diseño de un sistema de reconocimiento de patrones se lleva a cabo normalmente en tres fases:

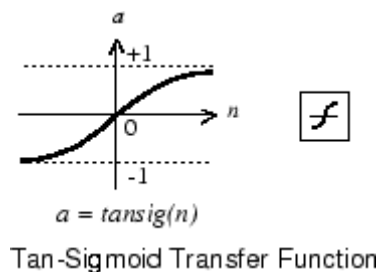
- Adquisición y preproceso de datos.
- Extracción de características.
- Toma de decisiones o agrupamiento.

Para la detección automática en cuanto a la determinación de qué tipo de señal de Emergencia ha sido captada se opta por el uso de Redes Neuronales, para este caso en particular, una red Feedforward en una de sus versiones especializadas conocida como Red de Reconocimiento de Patrones (patternet), la cual es básicamente una red que se usa para clasificar las entradas de acuerdo a los target de distintas clases. Las funciones de activación para este tipo de red son: Tansig para las capas ocultas y softmax para la capa de salida. MATLAB provee la función **patternet** mediante la cual se realizaran los entrenamientos y las pruebas respectivas.

Función Tansig

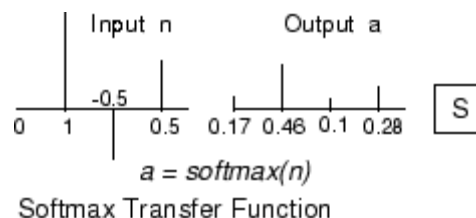
La función de transferencia neuronal tangente hiperbólica sigmoide es usada para calcular la salida de las capas ocultas.

Figura 29.Funcion Tansig



Función Softmax

Figura 30.Funcion Softmax



La función de Transferencia neuronal Softmax es una función usada para las capas de salida de la red patternet.

Para realizar el entrenamiento de la Red Neuronal patternet a partir de los patrones obtenidos por algoritmos de extracción de 51 análisis estadísticos es necesario realizar un proceso de Normalización de datos, ya que es lo óptimo y por lo tanto lo que se pretende con esto, es que tanto las entradas como las salidas se concentren en un rango de longitud aceptable y estén dispersas dentro de ese rango para un tratamiento de los datos más sencillo.

En este caso se aplica una Normalización min-max que busca que los valores de la matriz de entrenamiento estén entre [-1,1]

$$x'_i = (\max_{\text{objetivo}} - \min_{\text{objetivo}}) \times \left[\frac{(x_i - \min_{\text{valor}})}{(\max_{\text{valor}} - \min_{\text{valor}})} \right] + \min_{\text{objetivo}}$$

Ecuación 27

Donde:

x'_i = Entrada normalizada de índice i

x_i = Entrada de índice i

$\max_{\text{objetivo}}$ = valor máximo del rango que quiero aplicar a las entradas

$\min_{\text{objetivo}}$ = valor mínimo del rango que quiero aplicar a las entradas

$\max_{\text{valor}}$ = valor máximo real de las entradas

$\min_{\text{valor}}$ = valor mínimo real de las entradas

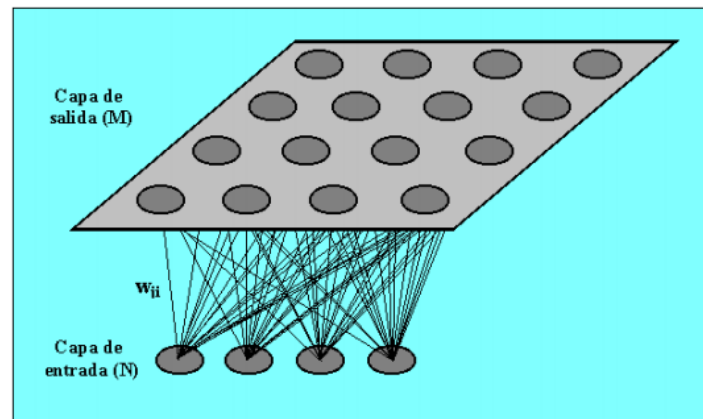
4.4.2 Redes Neuronales por mapas auto-organizados

Los mapas auto-organizados o redes de Kohonen (SOM por sus siglas en inglés, self-organizing map) fueron introducidos por el profesor finlandés Teuvo Kohonen. Un mapa auto-organizado es una herramienta que analiza datos en muchas dimensiones con relaciones complejas entre ellos y los presenta en una visualización sencilla en solo dos dimensiones. La propiedad más importante de una SOM es que preserva las propiedades topológicas de los datos, es decir, que datos próximos aparecen próximos en la visualización³¹.

Un modelo SOM está compuesto por dos capas de neuronas. La capa de entrada (formada por N neuronas, una por cada variable de entrada) se encarga de recibir y transmitir a la capa de salida la información procedente del exterior. La capa de salida (formada por M neuronas) es la encargada de procesar la información y formar el mapa de rasgos. Normalmente, las neuronas de la capa de salida se organizan en forma de mapa bidimensional como se muestra en la figura:

³¹ Ramos Díaz, Antonio. Redes Neuronales No supervisadas con topología Dinámica para la segmentación de imágenes en color. 2010. Universidad de Málaga Escuela técnica Superior de Ingeniería Informática (Ed).

Figura 31. Distribución de neuronas en la capa de salida de una Red SOM



Fuente. Material extraído de la web

(<http://halweb.uc3m.es/esp/Personal/personas/jmmarin/esp/DM/tema5dm.pdf>)

Las redes SOM se caracterizan esencialmente por tener un aprendizaje no supervisado competitivo, es decir, no hay un maestro que indique externamente a la red la forma correcta en que debe operar. La red auto-organizada debe descubrir rasgos comunes, regularidades, correlaciones o categorías en los datos de entrada, e incorporarlos a su estructura interna de conexiones. Se dice, por tanto, que las neuronas deben auto-organizarse en función de los estímulos (datos) procedentes del exterior.

Para la implementación en MATLAB de la red SOM se da uso a la función **newsom** para su respectivo entrenamiento y la función **sim** usada para las respectivas simulaciones.

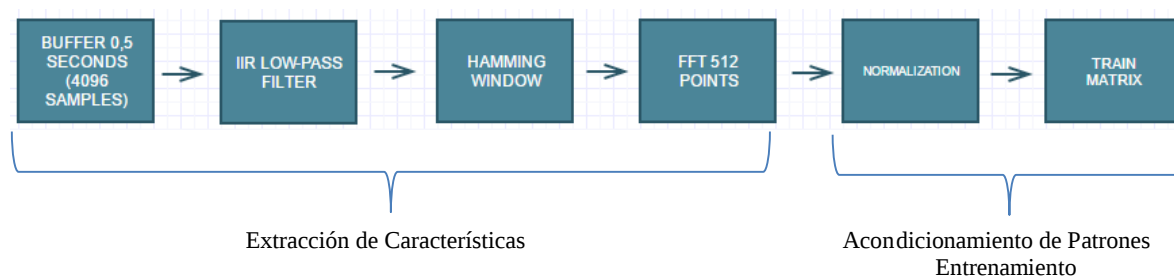
4.5 SELECCIÓN DEL MEJOR MÉTODO PARA LA IDENTIFICACIÓN Y DISCRIMINACIÓN DE LAS SEÑALES DE EMERGENCIA

Una vez descritos los distintos métodos de extracción de características y discriminación, se proponen las siguientes parejas de métodos como solución al problema de identificación y discriminación de señales sonoras de alarma.

4.5.1 Transformada Rápida de Fourier y Red por reconocimiento de Patrones

Retomando el Algoritmo de extracción de características por FFT, se puede resumir la metodología de este método de la siguiente manera:

Figura 32. Metodo extracción de características e identificación DFT y Red Patternet



Una vez obtenido el Buffer resultante de la extracción de características por medio de FFT, esta matriz es Normalizada por el método de Normalización min-max mencionado anteriormente, obteniendo así una matriz de entrenamiento. Como ejemplo de un entrenamiento ilustrativo se propone la siguiente situación en la que se entrenará una red patternet a partir de 3 muestras de 0,5 segundos capturados de cada señal. Se realiza la metodología instruida en el diagrama de bloques para una captura de sonido de cada una de las señales objetivo, es decir, se obtiene tres vectores de entrenamiento, uno para cada señal Objetivo, de esta manera se conforma el vector entrada:

Tabla 2. Vector Entrada 1x12288

Vector Característico 1x4096 elementos Ambulancia	Vector Característico 1x4096 elementos Policia	Vector Característico 1x4096 elementos Bomberos
---	--	---

Ya con el vector de entrada compuesto, se crea la matriz clasificación, esta matriz contiene los targets pertenecientes a cada una de las clases (Ambulancia, Policía, Bomberos). Esta matriz debe ser correspondiente al tamaño del Vector de Entrada, es decir, debe tener un tamaño de 3x12288 para este caso. La composición de esta matriz se puede ilustrar de la siguiente manera:

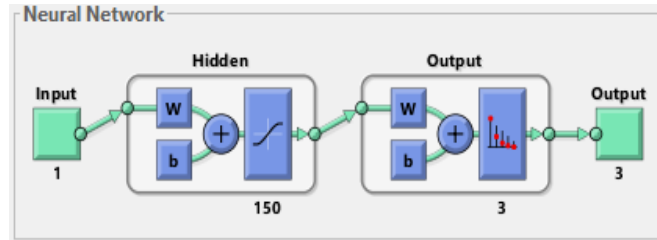
Tabla 3. Matriz Clasificación 3x12288

4096 Ones	4096 Zeros	4096 Zeros
4096 Zeros	4096 Ones	4096 Zeros
4096 Zeros	4096 Zeros	4096 Ones

Cada Fila de la matriz de clasificación posee 1/3 de elementos conformados por unos (1), esta se corresponde con los elementos del Vector de Entrada, de esta manera se adjudican los target a cada clase del Vector de Entrada para la realización del entrenamiento.

La red Neuronal estándar patternet que se ha planteado tiene la siguiente topología:

Figura 33. Topología red Patternet



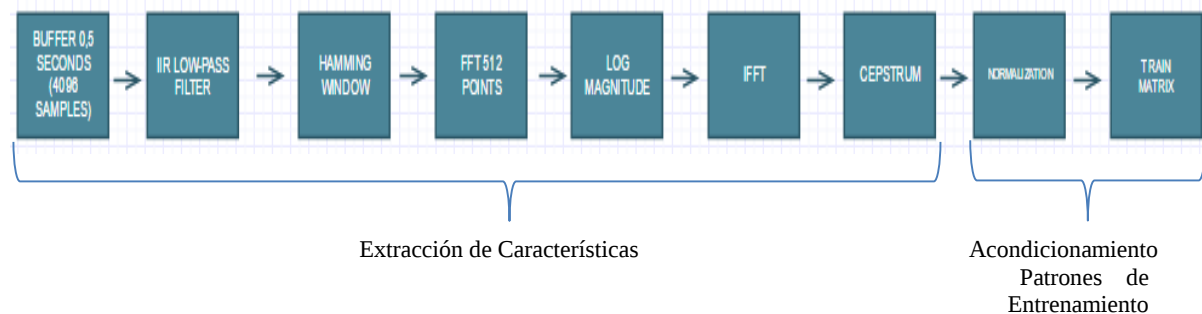
Fuente: Extraído del Software : MATLAB/ NNTOOL

Una matriz de entrada, la cual sería correspondiente a los patrones normalizados obtenidos por el algoritmo de extracción de características. Una capa Oculta compuesta por 150 Neuronas y una capa de salida compuesta por tres neuronas correspondientes a cada señal de emergencia a discriminar.

4.5.2 Cepstrum y Red por reconocimiento de Patrones

La metodología implementada es la siguiente:

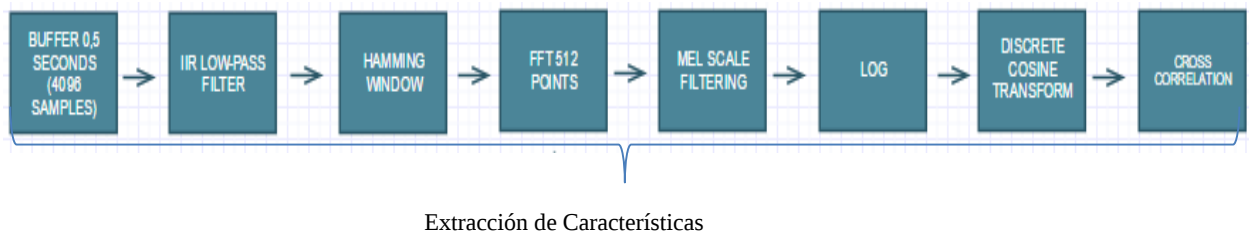
Figura 34. Método extracción de Características e identificación Ceptrsum y Red patternet



Retomando el algoritmo del Cepstrum mencionado en la sección 5.5.1. La conformación de los vectores de Entrada y Clasificación para la red Patternet se realizan de forma idéntica a la mencionada en 3.5.1. La topología de la Red se conserva de igual manera.

4.5.3 Coeficientes Cepstrales de Mel y Correlación Cruzada

Figura 35. Método extracción de Características e identificación MFCC's y Correlación Cruzada



Habiendo ya obtenido las tres (3) señales principales que componen la base de datos, se procede a realizar los respectivos análisis y comparaciones con las diferentes señales de prueba a trabajar para las pruebas del algoritmo.

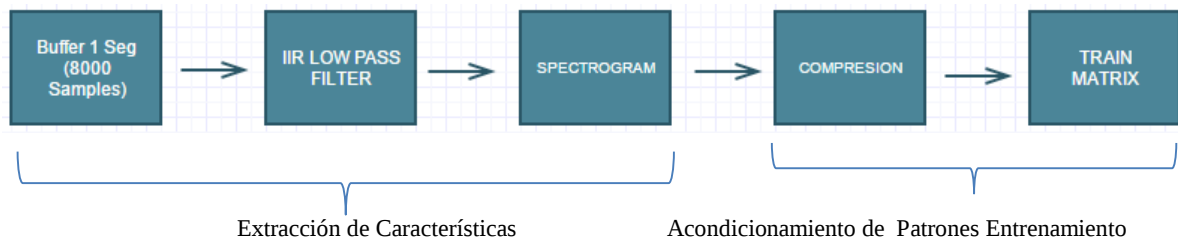
Analizando de las 3 señales de emergencia que alojadas en la base de datos, respectivamente en la representación de los MFCC's de cada una de las señales se trata de identificar cada una de las muestras respectivas con un color diferente, son 8 muestras, pues ese fue el número que se decidió para trabajar, sin embargo según recomendaciones observadas en el trabajo "Bird Call Identifier" similares a la hora de cálculos se trabajó directamente con 7, excluyendo una que en las gráficas se denota con un color azul-aguamarina (Figura 37), ya que maneja valores muy similares por cuestiones a la hora de grabar con el micrófono del dispositivo utilizado (PC), en algunos casos representaban datos idénticos, y esto generaba confusión a la hora que el algoritmo iba a realizar la correlación para tratar de identificar una señal.

Para cada señal principal (Ambulancia, Policía, Bomberos), se realizan 5 pruebas diferentes, es decir, se trabaja con 5 señales de prueba diferentes conformando 15 pruebas, en donde cada señal de prueba tenía alguna particularidad, pero que en esencia eran las mismas señales principales. Un ejemplo ocurre que en algunas señales de prueba, se les agrego un poco de ruido al ambiente donde se tomó la grabación; caso contrario a otras señales de prueba número 3 fueron grabaciones de los mismos tipos de sirenas emergencia pero con la particularidad de que fueron de otras ciudades, sin embargo se aclara que manejan las mismas características originales. En cada escenario de prueba planteado se buscó más complejidad o alguna diferencia con respecto a las señales de emergencia originales. En general los 5 ambientes de prueba resultaron exitoso, sin embargo situaciones de complejidad computacional en su implementación y procesamiento en tiempo real, no llenan las expectativas para poder aplicarlo como método definitivo de discriminación de sonidos.

4.5.4 Espectrograma y Red por mapas auto-organizados (SOM)

El último método utilizado es el de Espectrograma junto a las redes SOM. Este método posee un algoritmo más complejo tanto en extracción de características como en el método de discriminación. La razón de ello será explicada a continuación tras ilustrar la metodología de los algoritmos planteados.

Figura 36. Método extracción de Características e identificación Espectrograma y Redes SOM



Para este método se aplica un algoritmo de compresión de frecuencia, la matriz resultante del espectrograma para un Buffer 1 segundo de señal tiene un tamaño de 128x31 elementos, lo que se busca es reducir el tamaño de esa matriz para optimizar el proceso de entrenamiento. Una vez aplicada la compresión, por medio de eliminación de bandas de frecuencias innecesarias (por debajo de 500 Hz y por encima de 1350 Hz), se obtiene una matriz resultante de 30x31, por cada buffer analizado.

La topología de la Red SOM planteada se compone de 31 redes SOM correspondientes a las 31 ventanas usadas en el Espectrograma, adicionalmente otra Red SOM final la cual se entrenará con las respuestas de las 31 SOM anteriores, para así lograr la discriminación de las señales de alarma. Cada SOM tiene una dimensión 2x3.

Figura 37. Topologías primeras 31 SOM

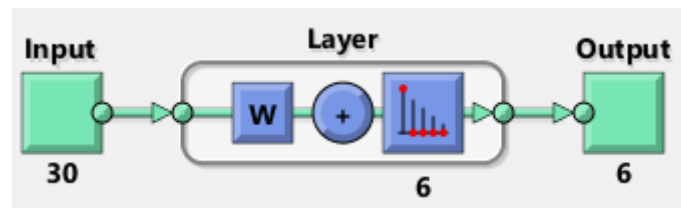
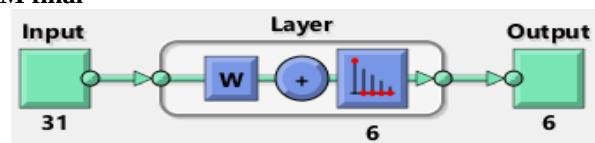
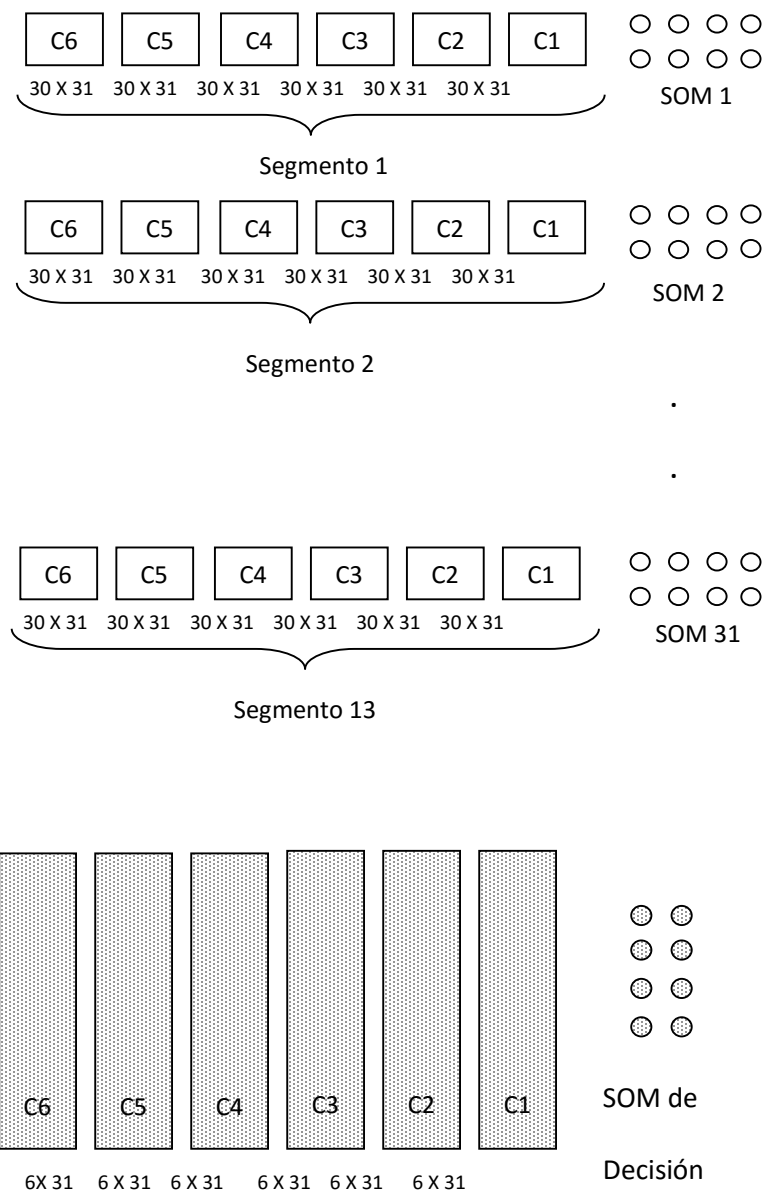


Figura 38. Topología SOM final



De una forma más detalla el sistema Neuronal planteado se puede apreciar de la siguiente forma:

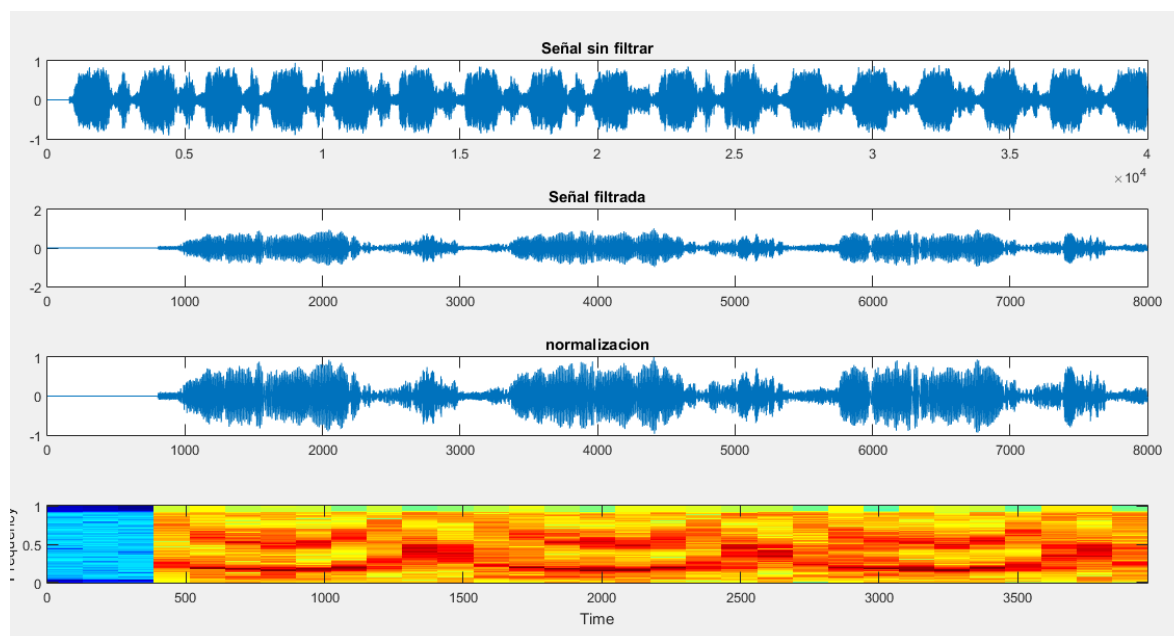
Figura 39. Sistema neuronal de clasificación.



Fuente. Mosquera Vladimir. Control de Aplicaciones mediante comandos orales reconocidos por redes Neuronales. Colombia. Universidad del Cauca(Ed).2001.

La SOM de Decisión recibe las respuestas de las 31 anteriores redes anteriores para determinar la neurona ganadora correspondiente a la clase Ambulancia, Policía o Bomberos de acuerdo a los entrenamientos realizados.

Figura 40. Tratamiento de un Buffer de 0,5 seg de Señal y Espectrograma Resultante



4.5.5 Discusión y Selección del método más Adecuado

Tabla 4 .Resultados de las Pruebas de los Métodos de extracción de Características e identificación

	FFT+ANN (%)				CEPSTRUM +ANN (%)				MFCC'S + CORRELACION CRUZADA (%)				ESPECTROGRAMA+ ANN (%)			
	E1	E2	E3	E4	E1	E2	E3	E4	E1	E2	E3	E4	E1	E2	E3	E4
1	42	45	42	40	59	57	60	62	78	77	77	80	71	73	76	75
2	39	46	40	41	58	62	60	63	78	83	82	80	76	75	77	75
3	40	45	43	41	61	60	57	55	80	80	82	85	75	75	76	76
4	47	43	41	41	61	56	58	63	82	82	76	78	77	76	75	74
5	42	37	32	40	60	61	61	58	80	80	78	76	76	71	77	73
M	42	43	40	41	60	59	59	60	79	80	79	80	75	74	76	75

Nota: La tabla se estructura de la Siguiete Manera, 4 Columnas principales que involucran la combinación de los 4 métodos (extracción de características + identificación) que se utilizaron, posterior a estas, se subdividen en 4 columnas más de cada uno de los métodos en donde se plantearon 4 diferentes escenarios con características particulares (Mayor ruido ambiental, Menor Volumen de la Sirena, etc.).En lo que respecta a filas, la tabla se compone de 6 filas que representan las 5 pruebas realizadas en cada escenario planteado dentro de los 4 métodos propuestos, finalmente la sexta fila es el promedio de cada uno de

los resultados en los diferentes escenarios. Los porcentajes presentados en la tabla hacen referencia a porcentajes de eficiencia de cada método ante cada escenario.

En lo competente al Método 1(FFT+ANN) el cual en vistas de ser implementado en hardware exigiría poco coste computacional desafortunadamente no presenta buen rendimiento. Se realizaron pruebas incluso aumentando el número de puntos de transformada sin obtener mejoras significantes. La principal razón de que este método presente tan poca eficiencia radica en que existen componentes frecuenciales iguales en las tres sirenas objetivo (Ambulancia, Policía, Bomberos) y por consecuencia se presenta conflictos en su discriminación mediante la red de reconocimiento de patrones.

Para el Método 2(Cepstrum+ANN) inicialmente se obtuvo un rendimiento muy bajo a lo esperado, el cual mejoró notablemente aumentando el número de puntos de transformada a 1024, en este método por momentos se obtuvieron resultados de 90% de eficiencia bajo el entrenamiento de un solo escenario, lo cual no sería ideal, ya que la red sólo respondería a ese tipo de escenario y no a una situación distinta. Finalmente se obtuvo un rendimiento de aciertos en la discriminación de solo el 60% lo cual sigue siendo muy bajo. Además de la baja eficiencia la implementación de este sería muy poco viable debido que el implementar una FFT e IFFT de 1024 puntos requeriría un gran costo computacional, adicional a esto el cálculo del Cepstrum requiere el uso de operaciones logarítmicas, las cuales ocupan un considerable número de ciclos de reloj en el microcontrolador lo cual afectaría notablemente la operación en tiempo real.

Para el Método 3 (MFCC's+Correlación) se alcanza un porcentaje de rendimiento no satisfactorio. En cada escenario de prueba planteado se buscó más complejidad o alguna diferencia con respecto a las señales de emergencia originales. En general solo uno de los ambientes de prueba que se trabajó resultó totalmente exitoso y eso no llena las expectativas para poder aplicarlo como método definitivo de discriminación de sonidos, más cuando aún analizado los resultados más generales ninguno de los 3 promedios de correlación supera el umbral del 0.6 (1 es el máximo) que sería un rango “aceptable”; de las 5 pruebas realizadas a cada señal solo la de policía logro pasar el umbral de aceptable.

En lo que respecta al Método 4 (Espectrograma + ANN) se llegó a un rendimiento del 75%, el cual se considera aceptable para dar por cumplido el objetivo de una correcta clasificación. La principal causa de la obtención de un mejor rendimiento es debido a que la topología de la red es mucho más compleja a diferencia de la red por reconocimientos de patrones, esta red cuenta con 32 redes de Mapas auto-organizados compuesto por 6 Neuronas cada uno. El espectrograma aplicado usa una FFT de 256 Puntos junto a una Ventana de Hamming de tamaño igual a los puntos de transformada. Además del uso de una topología de red más compleja se aplica un algoritmo de compresión por bandas de frecuencias, como consecuencia de ello se logró un entrenamiento más eficiente al reducirse el número de entradas de cada red.

4.6 IMPLEMENTACIÓN DEL MÉTODO DE IDENTIFICACIÓN Y DISCRIMINACIÓN EN UN CHIP PROCESADOR DIGITAL DE SEÑALES

El microcontrolador seleccionado para llevar a cabo las labores del procesamiento Digital de las Señales ha sido el DSPIC33EP256GP502 del Fabricante MICROCHIP, posee una

arquitectura de 16-bits, una velocidad de CPU de 70 MIPS, 32Kbytes de Ram, Conversor Analógico Digital 10x12-bit @ 500(kSPS) y Módulo DMA.

Figura 41. Proceso Previo de obtención y tratamiento de la información para su compresión



4.6.1 Conversión analógica a digital

La configuración del ADC se debe hacer en base a la frecuencia de muestreo establecida para el diseño de los filtros digitales y el posterior procesamiento de la señal. Para ello se configura el ADC del DSPIC33EP256GP502 para que trabaje a 40 MIPS (**Millones de instrucciones por segundo**) y luego para que la conversión se haga a 8000 Hz.

Para que el ADC del microcontrolador realice la conversión analógica a digital a 8000 Hz se configura el DMA (**Acceso directo a memoria**). DMA es una característica esencial en todos los ordenadores modernos, ya que permite a dispositivos de diferentes velocidades comunicarse sin someter a la CPU a una carga masiva de interrupciones.

La programación del dispositivo se hace mediante el Software MPLAB X y el compilador MPLAB xc16.

Figura 42 .Proceso de Obtención de la Señal



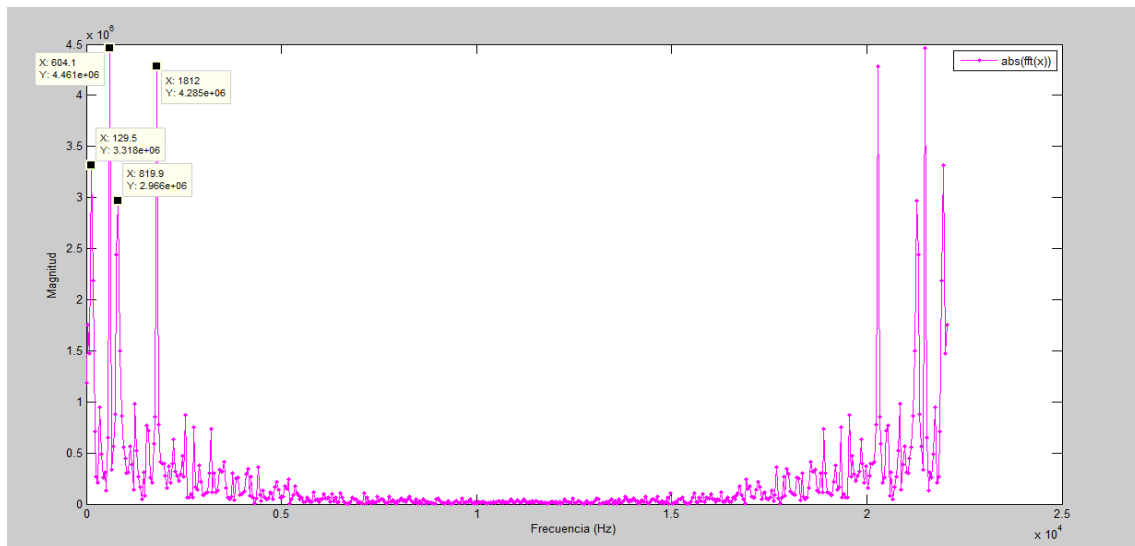
4.6.2 Implementación Filtro Digital

Una vez obtenido el funcionamiento correcto del ADC, se procede a realizar el filtro digital en el microcontrolador, para ello se hace uso de la librería `#include<dsp.h>` y la función **IIRTransposed()** que proporciona el fabricante MICROCHIP. A partir de los Filtros Digitales diseñados y la función de transferencia obtenida se añade el filtro al microcontrolador.

Para el uso de las instrucciones DSP es necesario estar claro con el uso de números fraccionales de 16 bits, también conocido como formato Q15, no es más que numeración de "punto fijo" (en analogía a punto flotante) pues la librería y el motor DSP emplea este formato.

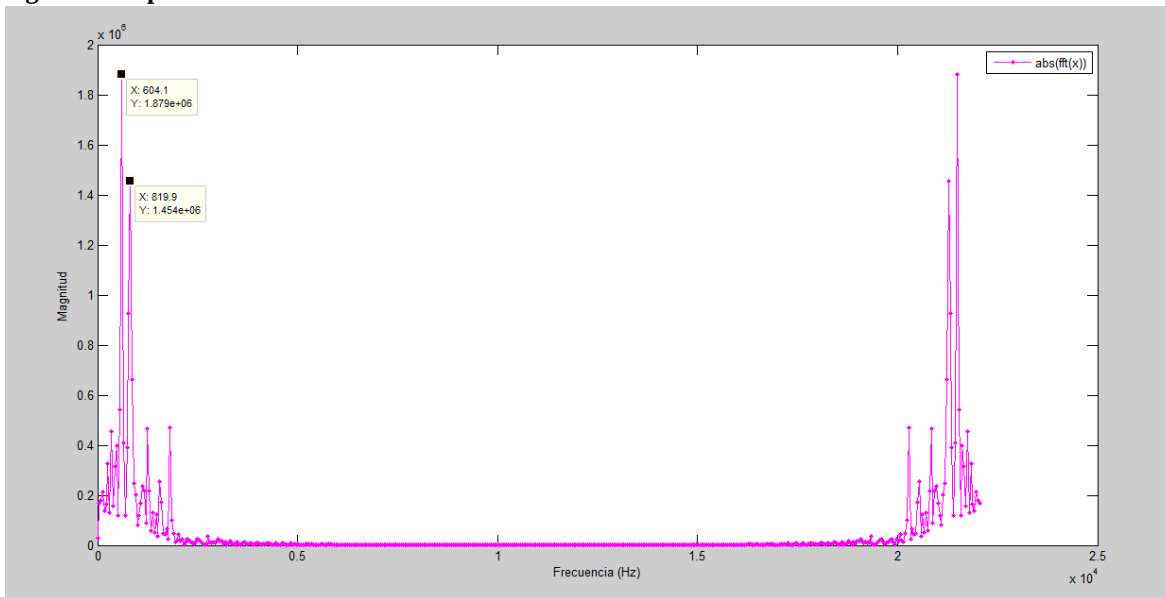
Con el filtro programado se procede a probarlo, para ello se hace uso de **DSPICWORKS**, software del fabricante MICROCHIP, con este se genera la siguiente señal de prueba a partir de una combinación lineal de señales sinusoidales de 130, 600, 800, 1800 Hz. Con la Ayuda de MATLAB se calculará y graficará el espectro de las señales con el fin de comprobar su contenido frecuencial.

Figura 43. Espectro de la Señal de Prueba



Tras ser aplicado el Filtro Pasabanda el resultado obtenido es el siguiente:

Figura 44. Espectro de la Señal de Prueba Filtrada

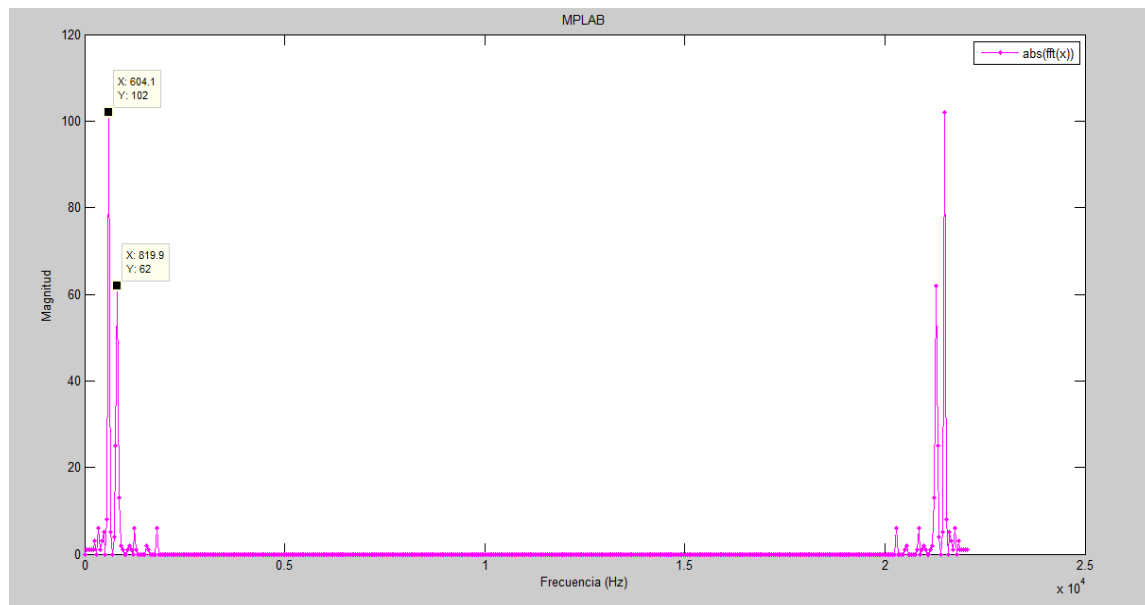


Se observa que las componentes Frecuenciales de 130 y 1800 fueron eliminadas, conservando así solo las componentes de 600 y 800.

4.6.3 Implementación de la Transformada de Fourier

Anteriormente el cálculo de las FFT se hacía por medio de MATLAB, ahora bien se programa la FFT de 512 puntos en el DSPIC, ya que se busca principalmente extraer las características en frecuencia de la señal. El procedimiento a realizar comienza por cargar los twiddlefactors o factores de giro correspondientes a la FFT de 256 puntos que se desea implementar, finalmente mediante las funciones del compilador **FTComplexIP()**, **BitReverseComplex()** y **SquareMagnitudeCplx()** se finaliza el proceso de cálculo de la FFT. Los resultados obtenidos son el mismo resultado que en MATLAB pero lógicamente más escalados como se muestra a continuación:

Figura 45. Espectro Calculado de la Señal de Prueba Filtrada en el DSPIC



4.6.4 Implementación de Ventana de Hamming

La salida espectral depende de la amplitud de la discontinuidad. A mayor discontinuidad mayor salida espectral. Se puede utilizar el windowing para la reducción de las discontinuidades en los límites de cada periodo. El windowing consiste en multiplicar el tiempo de grabación por una ventana de longitud finita cuya amplitud varía suave y gradualmente hacia 0 en los bordes. Al realizar la transformada de Fourier o el análisis espectral del dato finito se puede utilizar la ventana para minimizar los efectos de la transición del muestreo de la forma de onda.

La ventana implementada en este caso es una Hamming, la cual es una modificación de la Hanning, su forma también es similar a la de la onda coseno y se define:

$$w[n] = 0.54 - 0.46 \cos(2 \pi n/N) \text{ donde } n=0, 1, 2, \dots, N-1 \quad \text{Ecuación 28}$$

De igual forma se programó la ventana en MATLAB y en el DSPIC, a continuación se presenta una comparación de los resultados obtenidos para la señal de prueba sin Filtrar aplicando FFT de 512 Puntos y su respectiva ventana de Hamming.

Figura 46. FFT de 512 Puntos para el Espectro Unilateral y Ventana de Hamming para la Señal de Prueba sin Filtrar realizada por el DSPIC.

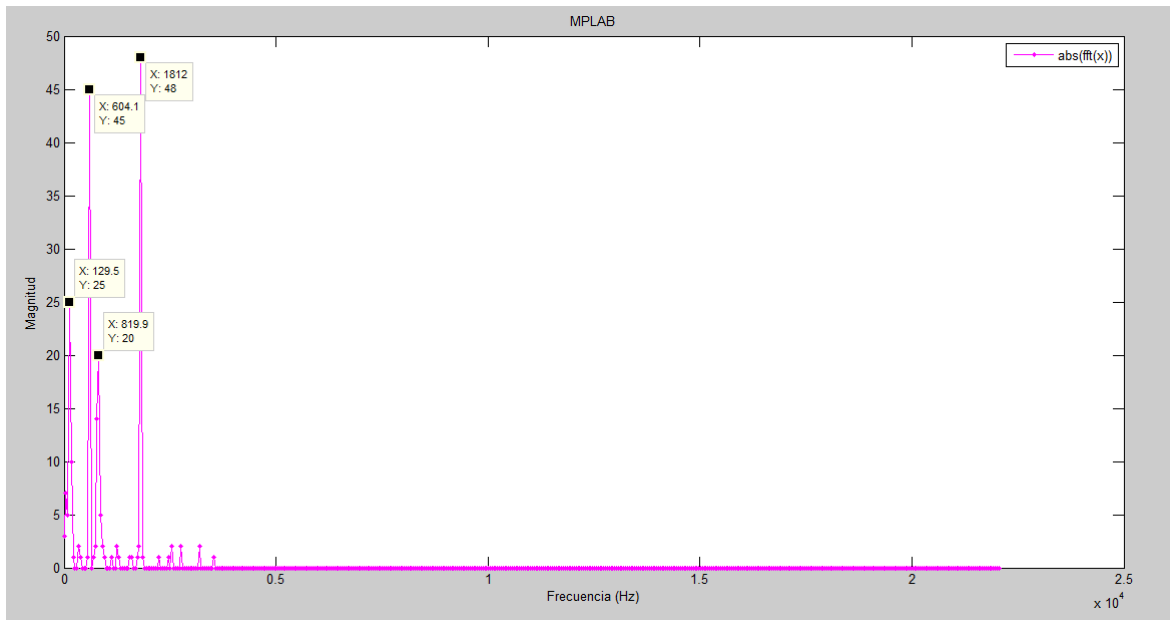
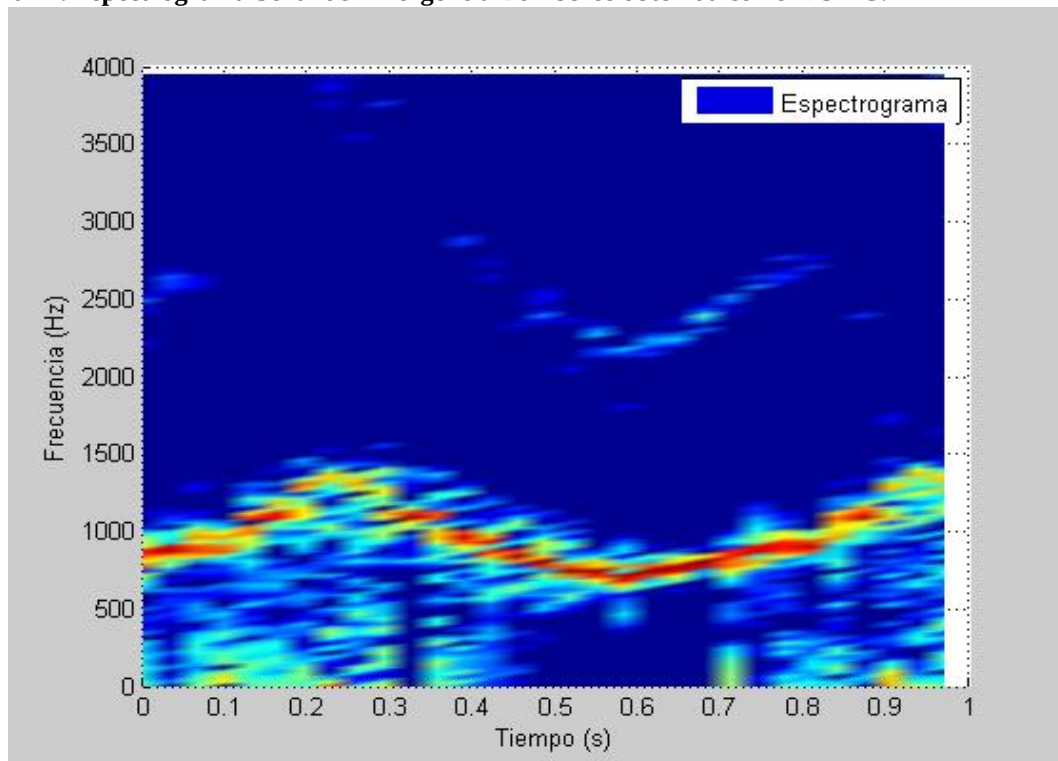


Figura 47. Espectrograma Señal de Emergencia Bomberos obtenida con el DSPIC.



Se obtiene el mismo resultado en ambos casos pero de igual manera el resultado que proporciona el DSPIC es mucho más escalado. Los procedimientos mencionados en los incisos 5.6.2 al 5.6.4 son la base para obtener el espectrograma de cada señal realizando la iteración de estos por 31 ciclos.

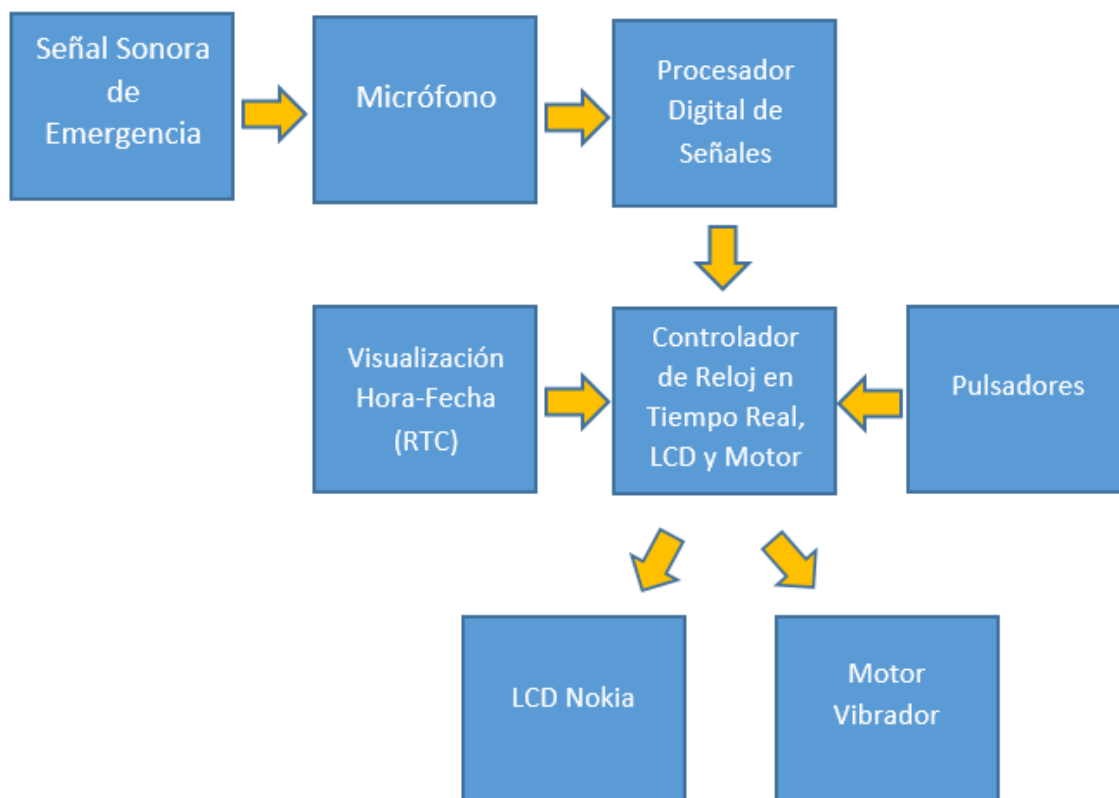
4.6.5 Implementación Red Neuronal

Culminada la implementación de la extracción de características se continua con la ejecución del discriminador de alarmas, en este caso la red neuronal de mapas auto-organizados. Con el algoritmo de extracción de características implementado, se realizan distintas pruebas bajo diferentes condiciones. Con un número suficiente de escenarios se realiza el entrenamiento en MATLAB con las matrices obtenidas del dsPIC. Verificada la buena eficiencia de la red se guardan los pesos de la red en la memoria del microcontrolador y se implementa el algoritmo de su respectiva función de activación.

4.7 DISEÑO DEL PROTOTIPO DE BRAZALETE.

Con el objetivo de ofrecer un dispositivo que satisfaga al usuario en términos de comodidad a la hora de portar y manipular el brazalete, se propone realizar una interfaz agradable, sencilla de manejar y por supuesto permita transmitir al usuario de forma visual y vibratoria las alertas de emergencia que lo rodeen. La figura 48, representa un esquemático de cómo está compuesto el hardware del prototipo.

Figura 48. Diagrama de Bloques del Hardware del dispositivo.



4.7.1 Interfaz Visual

Se selecciona un Display LCD Nokia 6100LCD PCF8833 para dicha labor debido a características de trabajo que permiten un fácil desarrollo y además de gran calidad y accesibilidad de Mercado; estas son:

- Voltaje de Alimentación: 3.3-5.0 V
- 4096 Colores
- 132x132 Píxeles
- Contraste Ajustable
- Dimensiones: 1.7" x 2.2"

Y cuenta con 6 pines de conexión:

- VCC (3.3V a 5.5V)
- GND
- RESET (Activo en Bajo)
- SDATA (Serial Data)
- SCLK (Serial Clock)
- CS (Chip Select, Activo en Bajo)

Figura 49. Nokia 6100LCD PCF8833



Fuente. Material extraído de la web (<https://www.sparkfun.com>)

Para el control de la LCD se seleccionó el microcontrolador de PIC18F4685, el cual recibirá las correspondientes señales provenientes DSPICXXX para indicar el tipo de alerta que esté ocurriendo. En la Tabla 5 expondrá las características del controlador elegido, argumentando además de que trabaja con buen nivel y como eje principal que se encuentran disponibles con facilidad sus encapsulados para montaje superficial y para hacer las respectivas pruebas en la Protoboard de Laboratorio.

Tabla 5. Características del Dispositivo PIC18F4685

Parameter Name	Value
Tipo de Memoria de Programa	Flash
Memoria de Programa (KB)	96
Velocidad CPU (MIPS)	10
RAM Bytes	3,328
Data EEPROM (bytes)	1024
Periféricos de Comunicaciones Digitales	1-UART, 1-A/E/USART, 1-SPI, 1-I2C1-MSSP(SPI/I2C)
Timers	1 x 8-bit, 3 x 16-bit
ADC	11 ch, 10-bit
Comparadores	2
CAN	1 CAN
Temperature Range °C	-40 to 125
Rango de Voltaje de Operación (V)	2 a 5.5
Número de Pines	40

Fuente. Material extraído de la web (<https://www.microchip.com>)

El PIC18F4685 además del control de la LCD tendrá a cargo también el del Motor Vibrador y el RTC DS1302; El motor vibrador será el utilizado como el actuador de las alarmas que tendrá el dispositivo, para el caso del DS1302 este será el encargado de:

- Proveer de la Fecha y Hora Actual
- Liberar tiempo y recursos (Contadores y/o Temporizadores) al microprocesador de la tarea de medir el tiempo
- Proporcionar una fuente suficientemente confiable en cuanto a la medición del tiempo
- Tienen funciones de bajo consumo de energía para cuando se hacen aplicaciones con baterías
- Administrar alarmas, para avisarle al microprocesador sobre algún evento que se tenga que realizar. Estas alarmas se conectan directamente a alguna interrupción del microprocesador

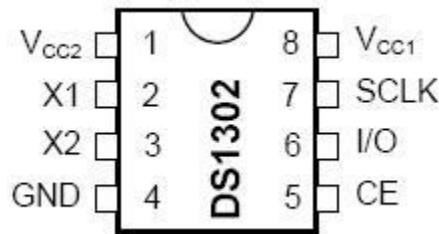
El DS1302 es un circuito integrado el cual tiene un contador de tiempo real, además tiene 31 bytes de memoria RAM para cualquier data que se quiera guardar allí. La comunicación con este integrado es vía Serial Síncrona (3 Pines). Y se puede transmitir o recibir la

información byte por byte o en modo Burst. El DS1302 está diseñado para trabajar en aplicaciones de bajo consumo de potencia, de hecho cuenta con los dispositivos adecuados para ser alimentado por baterías recargables. Este dispositivo cuenta con tan solo 8 pines, de los cuales:

- 2 Pines son para conectar un cristal de cuarzo de 32.768kHz. (X1 y X2 Pines: 2 y 3)
- 1 Pin para voltaje (Vcc2 Pin:1)
- 1 Pin para tierra (Gnd Pin:4)
- 1 Pin de Voltaje de Batería (Vcc1 Pin:8)
- 1 Pin de Inicio/Fin de Comunicación Serial (CE o RST' Pin:5)
- 1 Pin para el Reloj de la Comunicación Serial (SCLK Pin:7)
- 1 Pin para la transmisión de datos seriales (I/O Pin:6)

El dispositivo tiene funciones de bajo consumo de energía para cuando se hacen aplicaciones con baterías como en nuestro aplicación.

Figura 50. Pines DS1302



Fuente. Material extraído de la web (<https://developer.mbed.org/components/DS1302-Timekeeping-Chip/>)

Con el objetivo de brindar una interfaz visual agradable al usuario, basándose en una librería Open Source asentadas en los drivers de James P. Lynch permitió desplegar la respectiva hora y fecha en el Display, además de diseñar un menú desplegable para la configuración del reloj, visualización del estado de la batería, ajuste del contraste de la LCD.

4.7.1.1 Visualización Hora y Fecha

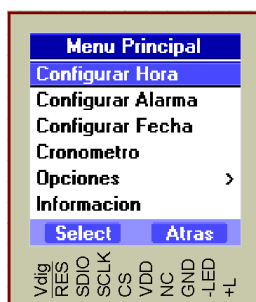
Figura 51. Visualización Hora-Fecha



La hora está disponible en formato Militar. Previamente está precargada en el Microcontrolador PIC18F4685 una hora de inicio a las 23:59:45 con fecha del 13/04/15.

4.7.1.2 Menú y Configuración

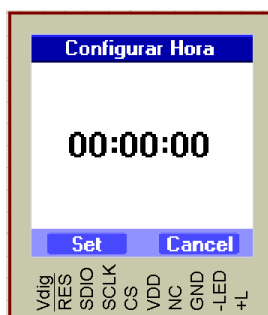
Figura 52. Visualización Menú Principal



En el Menú se cuenta principalmente con tres opciones de configuración:

a. Configuración Hora

Figura 53. Visualización Configuración Hora



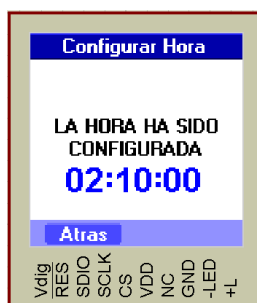
Para la configuración de la hora, es oportuno mencionar ahora bien los controles del Brazaletes, el cual cuenta con 5 pulsadores, los cuales cumplen las siguientes funciones:

- Arriba
- Abajo
- Select /Set
- Atrás
- Menú Principal

Con el botón de Arriba se aumenta la cuenta de la Hora o con el Botón de Abajo disminuye dicha cuenta desde 24 hasta 0, para continuar con el proceso se pulsa el botón Set y se procede a configurar los minutos haciendo uso de los botones Arriba- Abajo repitiendo el

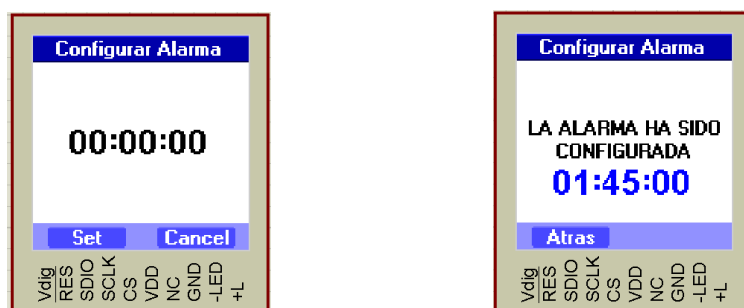
proceso exacto para la hora, Finalmente al presionar Set, se obtiene una notificación de confirmación de la Hora seleccionada

Figura 54. Visualización Notificación de la Hora Configurada



b. Configuración Alarma

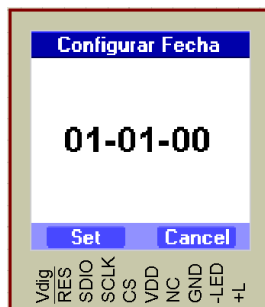
Figura 55. Visualización Configuración Alarma (izq), Visualización Notificación de la Hora Configurada (der)



La configuración de la alarma es exactamente igual a la realizada en la configuración de hora. De igual forma también se adquiere una notificación de la hora de alarma que ha sido configurada.

c. Configuración Fecha

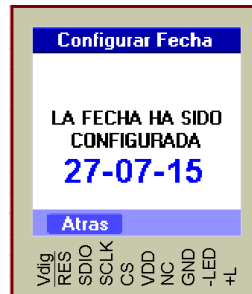
Figura 56. Visualización Configuración Fecha



Para el respectivo proceso de configuración de Fecha es muy similar al que se usa en la configuración de Hora y Alarma, de esta manera para configurar el día se usa el botón de Arriba para aumentar la cuenta de los Años o con el Botón de Abajo disminuirla o iniciar

dicha cuenta desde 99 hasta 1, para continuar con el proceso se pulsa el botón Set y se procede a configurar el Mes haciendo uso de los botones Arriba- Abajo repitiendo el proceso exacto para los días, nuevamente al presionar Set, avanza para configurar el día de igual manera realizado en año y mes. Al finalizar se observa la respectiva notificación de la fecha que se ha configurado.

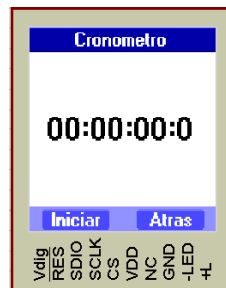
Figura 57. Visualización Notificación de la Fecha Configurada



Después de las Opciones de Configuración relacionadas con las funciones de Reloj del Brazaletes, restan dos opciones del Menú Principal:

d. Cronómetro

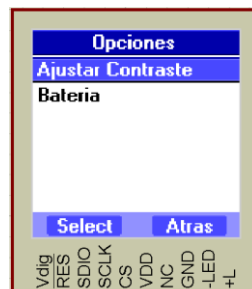
Figura 58. Visualización Cronómetro



Para Iniciar el Cronómetro se hace uso del Botón Set, así mismo con este botón se detiene y reinicia la, con el botón atrás retorna al menú principal.

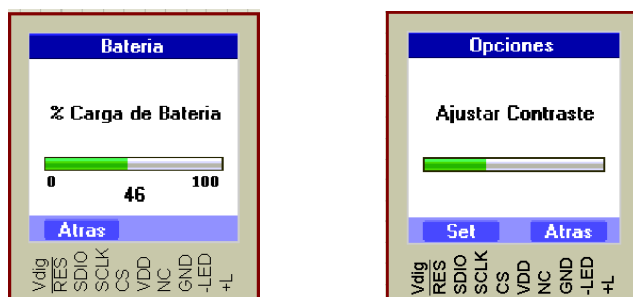
e. Opciones

Figura 59. Visualización Menú de Opciones



En esta sección se localiza el Ajuste del contraste de la LCD y el estado de carga de la Batería. Para la configuración de Ajuste de Contraste se hace uso de los botones de Arriba para el aumento y Abajo para la disminución de dicho parámetro, por último para ajustar el valor seleccionado por el usuario ha de ser pulsado el botón Set.

Figura 60. Visualización Configuración de Ajuste de Contraste y % de carga de la Batería



f. Visualización de las Alertas de Emergencia

Figura 61. Visualización de las tres alertas de Emergencia



Cuando es detectada una señal de emergencia el dispositivo emite una alerta visual acompañada de una alerta vibratoria ilustrando cual es el tipo de alarma detectada.

4.7.2 Interfaz Mecánica

La interfaz mecánica principalmente está compuesta por cuatro (4) elementos, la batería, el módulo cargador, el micrófono y finalmente el motor vibrador.

4.7.2.1 Batería

Se va trabajar con una batería de Lithium portable de 3.7V y 1000mAh, debido a las características eléctricas de los diferentes dispositivos e integrados que se van a utilizar en la implementación del brazalete. Es la de mayor facilidad de encontrar en el mercado no solo para realizar las diferentes pruebas como prototipo sino para cualquier situación de cambio que los que utilicen el dispositivo necesiten, además sus características de trabajo cumplen a cabalidad las necesidades de consumo final del dispositivo.

Figura 62 .Batería de Lithium 3.7 V – 1000Ma

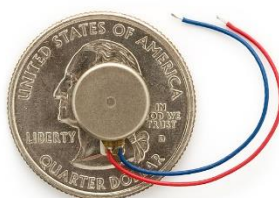


Fuente : Extraído de la Web (<https://www.sparkfun.com/products/13851>)

4.7.2.2 Motor Vibrador

En cuanto al Motor Vibrador se seleccionó uno que funciona a una velocidad de 12000 rpm, con un diámetro de 10mm y una longitud de 3.4mm y 1.2g de peso. Su tamaño fue clave para su selección, además de que se referencio debido a que se utiliza en dispositivos móviles (como celulares) y que se conoce su productividad en respecto a su funcionalidad.

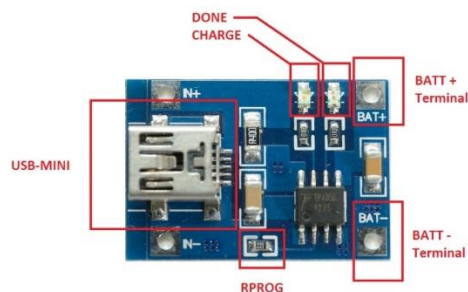
Figura 63. Motor Vibrador



Fuente: Extraído de la web (<https://www.sparkfun.com/products/8449>)

4.7.2.3 Módulo de Carga de la Batería

Figura 64. Modulo cargador de Batería TP4056



Fuente: Extraído de la Web (<http://electrosdr.com/shields-y-sensores/192-modulo-cargador-baterias-de-litio-tp4056-electronica-arduino-prototipado.html>)

Se decidió trabajar con el módulo de carga de Batería TP4056, por sus características útiles, eficaces y de bastante amabilidad en su adecuación al brazaletes, su tamaño es también un

tópico importante que se tomó para escogerlo, además de su facilidad de manejo de ubicación y uso. Se adjuntan algunas características generales del módulo. Un dispositivo que se adapta a las características de carga moderna, pues se carga mediante conexión USB.

Parámetros:

- Tensión de entrada: 4.5-5.5V
- Tensión de carga completa: 4.2V
- Corriente de salida máxima: 1A
- Indicador de encendido: El rojo representa la carga, Verde representa cargada de modo llenado.
- Interfaz de entrada: Interfaz MINI
- Tamaño del artículo: 25 mm * 19 mm.

4.7.2.4 Micrófono

El micrófono es la base principal del proceso de adquisición de datos, es por eso que su elección fue clave y para ella se determina tres (3) ítems importantes para tomar una decisión: Tamaño, accesibilidad comercial y funcionalidad. ¿Su ventaja mas importante?, que es Omnidireccional

Figura 65. Micrófono SPQ1410



Fuente: Extraído de la web (<https://www.jayconsystems.com/spq-1410-mems-microphone.html>)

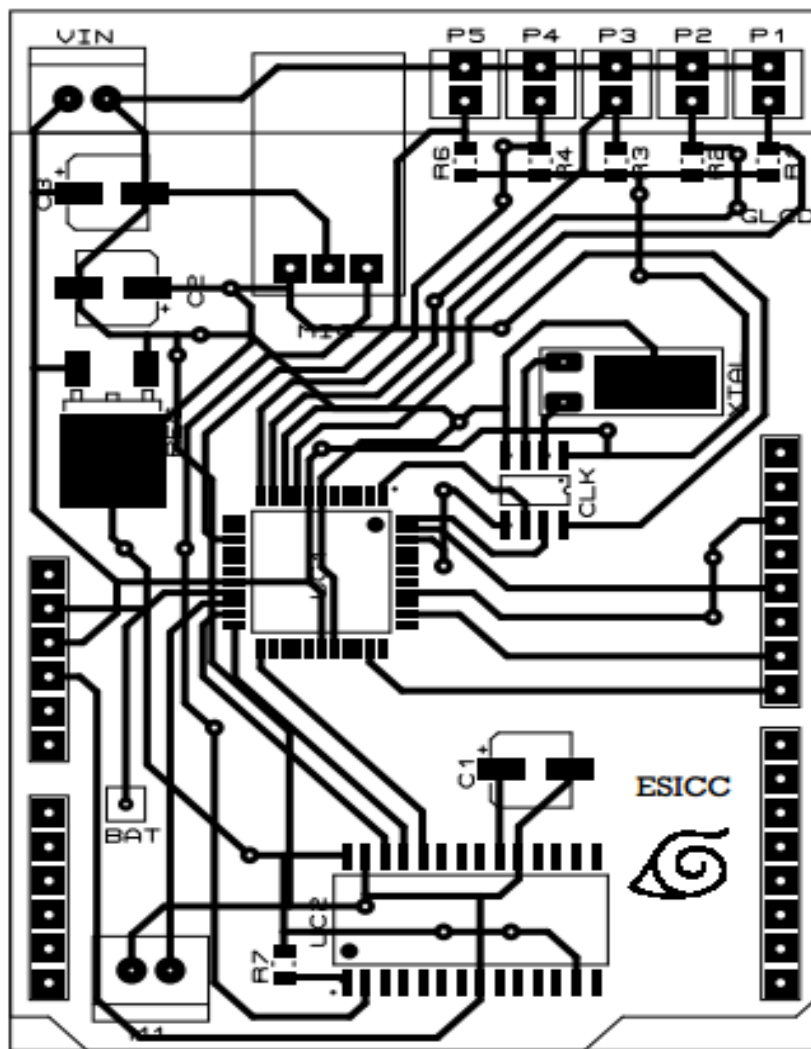
Las medidas de la placa que contiene el micrófono son de 12mm x 8mm, y permiten una ventaja muy grande en la ubicación dentro del dispositivo, además La placa incluye un circuito de amplificación con una ganancia de 67 que coincide con los requisitos de ancho de banda del SPQ1410. La salida de AUD del amplificador flotará a una mitad de VCC cuando no se está capturando ningún sonido y el amplificador produce una salida de pico a pico de aproximadamente 200mV cuando el micrófono se mantiene a la longitud de los brazos y se está hablando a niveles de volumen de conversación normales. El conveniente rango de voltaje de salida de AUD significa que puede conectarse directamente al ADC de un microcontrolador típico. Sus otras principales características de trabajo son:

- 3dB roll off a 100Hz y 15kHz
- Tensión de alimentación de 1,5 a 3,6 V CC
- Debe emitir cómodamente 40mW
- Sensibilidad de -42dBA

4.7.3 DISEÑO DE LA PCB Y EL ENCAPSULADO DEL BRAZALETE

A la hora de realizar el proceso del diseño del brazalete se tomaron consideraciones principales, como estética y comodidad, aunque claro hay que anticipar que la complejidad del circuito implica tener unos sacrificios en estética debido a él volumen que se ocupa. El diseño de la pista del circuito se realizó en el software PROTEUS-ARES , especializado para este tipo de tareas.

Figura 66. Pista del Circuito y Esquemático Final.



En la imagen anterior se observa la interfaz de trabajo para el diseño de la pista del circuito del brazalete, agregando las dimensiones y medidas necesarias para cada uno de los elementos electrónicos que lo componen y las respectivas pistas que transportan la tensión que energiza el circuito.

Como paso siguiente, los temas finales a definir ya teniendo la placa final del circuito completa con elementos y sus demás características, se procedió a escoger los materiales con los cuales se iba a realizar la base protectora del circuito y lo que formaría la representación estética del brazalete. Como se observó este presenta funciones de reloj por ende su diseño también buscaba que tuviera semejanzas a la estética de otros con tales funciones, más claramente como los hoy comúnmente conocidos como Reloj inteligentes. El acrílico fue el material primario escogido para realizar la base del brazalete por su resistencia y maleabilidad para su construcción, este primer diseño se buscó trabajar con colores primarios sencillos de obtener (amarillo, azul o rojo).

Posteriormente se agregó como sujetador del brazalete una correa bastante casual que pueda ser manipulable con facilidad para su comodidad y multiuso. La construcción de toda esta parte se realiza con impresoras 3D, debido a su factibilidad y precisión a la hora de desarrollar todo este tipo de elementos.

Figura 67. Esquemático Final del diseño industrial del prototipo del brazalete

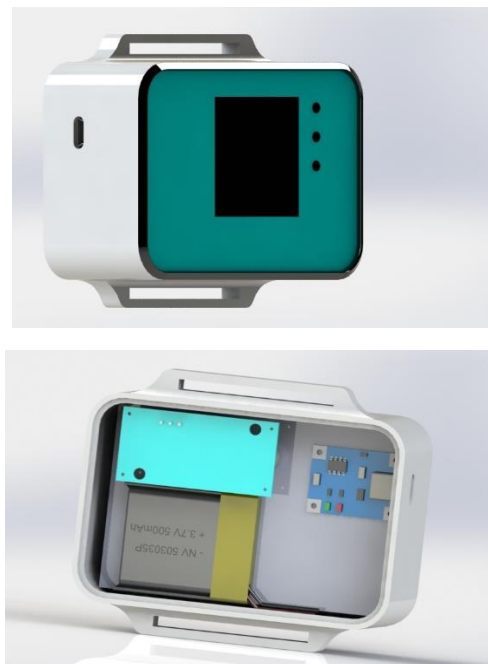


Figura 68. Fotos del dispositivo implementado.



Finalmente se realizaron unas pruebas de campo del dispositivo con personas con discapacidad auditiva, en compañía y asesoría de la fonoaudióloga especialista en el departamento (Elizabeth Moreno Mendez), quienes nos compartió sus conocimientos y practicas respectivas para tener una evaluación aproximada de los alcances del dispositivo. En la página de Anexos (Anexo A.) se observa el certificado de las pruebas realizadas.

5. CONCLUSIONES

En este capítulo se presenta el compendio de conclusiones basadas en los resultados y análisis de la realización del proyecto.

5.1 DISCUSIÓN ACERCA DE LOS RESULTADOS OBTENIDOS EN MATLAB

El uso de MATLAB como herramienta de verificación de ideas resulta sumamente útil y se puede considerar una piedra angular de la realización del proyecto. Los resultados obtenidos en las distintas pruebas condujeron hacia los métodos correctos de extracción de características y de discriminación. Los diferentes paquetes de herramientas de redes neuronales son excelentes aplicaciones para este tipo de simulaciones y pruebas, además de la construcción en sí de la red y su entrenamiento.

5.2 DISCUSIÓN ACERCA DE LOS MÉTODOS DE IDENTIFICACIÓN Y DISCRIMINACIÓN PROPUESTOS

Los Métodos propuestos se basaron principalmente en análisis tiempo-frecuencia, casi todos tenían como base el uso de la FFT, aunque se obtuvieron resultados malos, otros mediocres y por último uno aceptable en los métodos propuestos, se valida la premisa de que el análisis frecuencial permite identificar y discriminar entre señales sonoras de emergencia.

En cuanto a los métodos de discriminación utilizados, en lo que refiere a las Redes Neuronales artificiales sólo se trataron redes de tipo feedforward, las topologías implementadas sólo se variaron en cuanto a la cantidad de neuronas y la cantidad de entradas para las redes Patternet y no se lograron resultados aceptables con este tipo de red, posiblemente la efectividad hubiese mejorado aplicando algoritmos de compresión o una topología de red más robusta.

En lo correspondiente al método de MFCC's no se propuso junto a un sistema de red Neuronal debido a que resultaría un sistema de costo computacional bastante alto, lo cual a la hora de poner en marcha el sistema en tiempo real no sería óptimo al momento de detectar una señal de emergencia. Si se encontrara un sistema Embebido que tuviese optimizadas librerías para el desarrollo de los MFCC's sería ideal para realizar nuevas pruebas en vistas de mejores resultados.

En el método de MFCC's+Correlación cruzada se obtuvo un porcentaje de rendimiento no satisfactorio para consideraciones de aplicación e implementación. En cada escenario de prueba planteado se buscó más complejidad o alguna diferencia con respecto a las señales de emergencia originales. De manera generalizada de las 15 pruebas realizadas solo 7 presentaron acierto y de esas 3 fueron con la ambulancia, esto aclara que aunque el sistema tuvo un acierto al menos en cada tipo de señal de emergencia no es lo suficiente confiable, no presenta la mejor precisión de identificación. Prácticamente solo uno de los ambientes de prueba que se trabajó resultó totalmente exitoso y eso no llena las expectativas para poder aplicarlo como método definitivo de discriminación de sonidos, más cuando aún analizado los resultados más generales ninguno de los 3 promedios de correlación supera el umbral del 0.6 (1 es el máximo) que sería un rango "aceptable"; de las 5 pruebas realizadas a cada señal solo la de policía logró pasar el umbral de aceptable.

Finalmente el método de espectrograma y red SOM resultó ser el método más eficiente, debido a la simplicidad del método de extracción de características que sólo emplea FFT y ventaneo, además el complejo sistema de 32 redes SOM correspondientes a las ventanas del espectrograma y una SOM de decisión, por último sumado a lo anterior el sistema de compresión por frecuencias lo cual permite una reducción de datos significativa, empezando que el espectrograma resultante es de 256x31 pero realmente se hace uso sólo del espectro unilateral, así que el espectrograma resultante en materia de datos se reduce a 128x31, empleando el sistema de compresión por frecuencias el espectrograma resulta siendo una matriz de 30x31. La compresión de datos realizada en este método permitió implementar un sistema que trabajara en tiempo real y que fuese eficiente en materia de la discriminación de sonidos de alarma.

5.3 DISCUSIÓN ACERCA DE LA IMPLEMENTACIÓN DE LOS ALGORITMOS DE IDENTIFICACIÓN Y DETECCIÓN EN EL SISTEMA EMBEBIDO

Se logra ejecutar un dispositivo que funciona en tiempo real, con una buena tasa de acierto de discriminación de las señales sonoras de emergencia. Los tiempos de respuesta del sistema son coherentes a la necesidad de que el usuario detecte el tipo de emergencia con una buena anticipación.

El microcontrolador seleccionado al poseer librerías optimizadas para el procesamiento digital de señales y la capacidad del convertidor analógico digital de trabajar sin consumir recursos de la GPU por medio del DMA resulta sumamente beneficioso en búsqueda de que el sistema trabajase en tiempo real, esas dos principales características del microcontrolador culminan en el éxito de buenos tiempos de ejecución y respuesta de los algoritmos.

El método de Espectrograma seleccionado resulta ser un método de extracción de características eficiente, ya que solo emplea una FFT de 256 puntos lo cual es un costo computacional mínimo a comparación de los otros métodos planteados, así mismo el tamaño de las redes SOM planteadas al ser de una Longitud 6x6 permite un ahorro de memoria de programa sustancial del microcontrolador DSP.

Implementar la Ventana de Hamming por medio de la memoria de programa del sistema microcontrolado permite un ahorro sustancial de RAM en el sistema, debido a que es un proceso constantemente iterativo.

6. RECOMENDACIONES

Para futuros trabajos sería interesante poder coordinar el tipo de dato que trabaja MATLAB y el procesador DSP, ya que en este caso particular, MATLAB trabaja en punto Fijo, mientras el microprocesador lo hace en formato fraccional y punto flotante lo que hace más dispendioso el proceso de simulación y comparación.

En vistas de mejoras del prototipo se recomienda no solo trabajar con algoritmos de extracción de características tiempo frecuencia, ideal sería buscar métodos alternos que permitan aportar nueva semántica en el campo de reconocimiento de señales sonoras de emergencia.

Explorar métodos como la Transformada de Wavelet, el cual no se tuvo en cuenta en este proyecto debido a su complicada implementación en un hardware embebido.

Usar otro tipo de red neuronal como la Backpropagation y diferentes tipos de topologías sería un punto de aporte importante en vistas de explorar nuevos resultados. Se recomienda usar distintos tipos de normalización a la hora de realizar los entrenamientos. También se recomienda implementar algoritmos de discriminación como Modelos Ocultos de Markov y Modelos de Mezclas Gaussianas

En cuanto a los algoritmos de compresión, en el presente trabajo se ha utilizado un algoritmo de compresión por sub-bandas de frecuencias, se recomienda usar algoritmos como PCA y LDA y buscar su implementación en hardware embebido, probablemente el uso de estos métodos permita mejorar la eficiencia de entrenamiento de las redes neuronales.

Las recomendaciones que se puede sumar al área específica del prototipo de brazalete serian buscar tal vez diseñar con otro material base para mirar cómo se alteraría ante el uso diario y los factores como clima u otras circunstancias, pues las pruebas realizadas fueron exclusivamente de funcionamiento del brazalete. También la búsqueda de elementos electrónicos un poco más pequeños, ya que gracias al avance de la tecnología esto es posible sin embargo fueron situaciones que no se planearon porque comúnmente estas consideraciones realizan alzas en los costos de producción

Además se considera útil agregar, que sería interesante realizar otra combinación de métodos de extracción de características e identificación de los propuestos que no se haya realizado pues aún se estima que cada uno tiene un buen rango de aceptación sin embargo los escenarios y las situaciones a tratar son diferentes en cada trabajo o investigación que se vaya realizar y cada uno aporta una importancia crucial diferente.

BIBLIOGRAFIA

LIZONDO, Maximiliano. AGÜERO, Pablo. URIZ, Alejandro. TULLI, Juan C. GONZÁLEZ, Esteban. Verificador de locutores implementado en un dsPIC. Argentina. Universidad Nacional de Mar del Plata

GONZÁLEZ CADENILLAS, Clayder. MURRUGARRRA-LLERNA, Nils. Isolated Words Recognition using a Low Cost Microcontroller. Perú, USA. 2013. Universidad Nacional de Trujillo, Universidad de Pittsburgh

HERNÁNDEZ PÉREZ, Adriana. Reconocimiento de entornos acústicos para dispositivos móviles. España. Universidad Politécnica de Valencia

CARROL TYLER, Colangelo. STROTT, Tom. Bird Call Identifier.USA.2010. Worcester Polytechnic Institute

MIYAZUKI, Takuya, KITAZONO, Yukhi. SHIMAKAWA, Manabu. Ambulance Siren Detector using FFT on dsPIC. Japón. 2013. Kitakyushu National College of Technology, Kumamoto National College of Technology.

PINO, Raúl. GÓMEZ,Alberto. DE ABAJO, Nicolás. Introducción a la Inteligencia Artificial: Sistemas Experto, Redes Neuronales Artificiales y Computación Evolutiva. Oviedo, España. 2001. Servicios de publicaciones Universidad de Oviedo(Ed).

BASOGAIN, Xabier Olabe. Redes Neuronales y sus Aplicaciones. Bilbao, España. 2006. Escuela Superior de Ingenieria de Bilbao(Ed).

POSE GESTAL, Marcos. Introducción a las redes de neuronas artificiales. España. 2008. Universidade da Coruña (Ed).

MARTIN DEL BRIO, Bonifacio; SERRANO CINCA, Carlos. Fundamentos de las redes neuronales artificiales: Hardware y software. Zaragoza, España. 1995.Universidad de Zaragoza (Ed)

HILERA, Juan. MARTÍNEZ, Victor. “REDES NEURONALES ARTIFICIALES: Fundamentos, modelos y aplicaciones”. .1995. RA-MA Editorial

AGUILAR, Jorge. Introducción a las Redes Neuronales Artificiales. Material extraído de la web

(http://www.flacsoandes.edu.ec/comunicacion/aaa/imagenes/publicaciones/pub_25.pdf)

MATICH, Damián Jorge. Redes Neuronales: Conceptos Básicos y Aplicaciones. Rosario, Argentina. 2001. Universidad Tecnológica Nacional-Facultad Regional Rosario.

Sistemas de adquisición y Procesamiento de datos.2007.Tecnología e información

IBARRA MANZANO, Oscar. Notas Académicas Curso Procesamiento Digital de Señales.Mexico.2007 Universidad de Guanajuato(Ed).

RINCÓN LLORENTE, Carmen. Diseño, Implementación y Evaluación de técnicas de identificación de emociones a través de la voz. Abril 2007. Universidad Politécnica de Madrid (Ed).

KIYOHARU, Aizawa. NAKAMURA, Yuichi. SHINICHI, Satoh. Advances in Multimedia Information Processing - .2004: 5th Pacific Rim Conference on Multimedia. Springer

SARACHAGA, Gabriela. SARTORI, Virginia. VIGNOLI, Laura. Identificación Automática de Resumen en Canciones. Uruguay. Universidad de la Republica (Ed)

ANEXOS

A.

ELIZABETH MORENO MENDEZ
FONOAUDIOLOGA
UNIVERSIDAD MANUELA BELTRAN

NIT: 36176905

HACE CONSTAR:

Que los estudiantes de la Universidad Surcolombiana del programa de Ingeniería Electrónica Nicolás Charry Moreno y David Eduardo Clavijo Osorio que adelantan el proyecto **DISEÑO E IMPLEMENTACIÓN DE UN PROTOTIPO DE BRAZALETE DE DETECCIÓN DE SONIDOS DE ALARMA PARA PERSONAS SORDAS E HIPOACÚSICAS** realizaron pruebas del dispositivo en población con discapacidad auditiva moderada y severa a pacientes que reciben rehabilitación en Fonoaudiología en consulta externa durante el periodo del 26 al 28 de Junio del 2017.

Esta constancia se expide a petición de los interesados, dado el día 6 de Julio del 2017.

Elizabeth Moreno Mendez
Fonaudióloga
Reg. 5402 UMB
C.C. 36176905

ELIZABETH MORENO MENDEZ
CELULAR: 3162201584
PROFESIONAL EN FONOAUDIOLOGÍA
REGISTRO 5402