



**I
N
A
O
E**

Conjuntos Difusos tipo 2 aplicados a la Comparación Difusa de Patrones para Clasificación de llanto de infantes con riesgo neurológico

Por

Karen Santiago Sánchez

Tesis sometida como requisito parcial para obtener el grado de **Maestra en Ciencias** en la especialidad de **Ciencias Computacionales** en el Instituto Nacional de Astrofísica, Óptica y Electrónica.

Supervisada por:

Dr. Carlos Alberto Reyes García

Investigador Titular del INAOE

Dra. Maria del Pilar Gómez Gil

Investigadora Asociada del INAOE

©INAOE 2009

Derechos Reservados

El autor otorga al INAOE el permiso de reproducir y distribuir copias de esta tesis en su totalidad o en partes.



RESUMEN

El interés por analizar el llanto de infantes se ha incrementado en los últimos años; actualmente se utilizan modelos computacionales para automatizar la identificación de patologías en el llanto de infante. Para la clasificación del llanto de infante se han utilizado, entre otros, métodos basados en redes neuronales artificiales. En esta tesis se propone un método de clasificación basado en la comparación difusa de patrones aplicando conjuntos difusos tipo 2. La aplicación de este método utiliza, para realizar la clasificación, información tanto de la pertenencia de un elemento a determinada clase como de su importancia relativa con respecto a otras clases. Se probó la clasificación para llantos etiquetados como asfixia, hiperbilirrubinemia y normales, de los cuales, la clase hiperbilirrubinemia no se había clasificado de forma automática anteriormente. El proceso de clasificación incluye la extracción de características de la señal de llanto, las cuales se utilizan para entrenar el método propuesto en esta tesis, y posteriormente para realizar pruebas. Para extraer las características del llanto de infante se utilizaron cuatro métodos comúnmente aplicados a análisis de habla que son: LPC (por sus siglas en inglés *Linear Predictive Coding*), MFCC (por sus siglas en inglés *Mel Frequency Cepstral Coefficients*), cálculo de intensidad y cálculo de cocleograma. Se realizaron pruebas con distintas combinaciones de características y clases. Los mejores resultados para la clasificación utilizando las tres clases (hiperbilirrubinemia, asfixia y normal) se obtuvieron con la combinación de las características LPC-cocleogramas, donde se obtuvo una precisión del 91.74%. En la clasificación utilizando dos clases (hiperbilirrubinemia y normal) el mejor resultado se obtuvo con la combinación de las características LPC-MFCC con una precisión de 95.56%. Y con la combinación: asfixia y normal, se obtuvieron mejores resultados (99.02% de precisión) que trabajos similares de clasificación de llanto de infantes reportados anteriormente.

ABSTRACT

The interest for analyze the baby's cry has increased in the last years, nowadays computational models are in use to automate the identification of pathologies in the baby's cry. For the classification of baby's cry have been used methods based on artificial neural networks. In this thesis we propose a classification method based on the fuzzy pattern matching by applying type-2 fuzzy sets. This method use the Information of a element about its membership to a particular class and its relative importance with respect to other classes. We tested the classification with samples of baby's crying labeled as asphyxia, hyperbilirubinemia and normal. Class hyperbilirubinaemia has not been classified automatically before. The classification process includes the extraction of characteristics of the signal crying, which are used to train the method proposed in this thesis, and subsequently to realize tests. To extract the characteristics of signal crying we used four methods commonly used for the analysis of speech: LPC (Linear Predictive Coding), MFCC (Mel Frequency Cepstral Coefficients), Intensity and Cochleograms. We performed tests with different combinations of characteristics and classes. The best results for classification using the three clases (hyperbilirubinemia, asphyxia and normal) were obtained with the combination of the characteristics LPC and Cochleograms, where we obtained an accuracy of 91.74 %. In the classification using two classes (hiperbilirrubinemia and normal) the best result was obtained with the combination of the characteristics LPC and MFCC with an accuracy of 95.56%. And with the combination: asphyxia and normal, better results were obtained (99.02% of accuracy) than similar works of classification of baby's cry previously reported.

AGRADECIMIENTOS

Agradezco con la misma importancia a mis Asesores de Tesis: Dr. Carlos Alberto Reyes García y Dra. María del Pilar Gómez Gil por su tiempo y apoyo durante la realización de esta tesis.

A mis Sinodales: Dr. Francisco Martínez Trinidad, Dr. Gustavo Rodríguez Gómez y Dr. Leopoldo Altamirano Robles por su paciencia y aportaciones para mejorar este trabajo de Tesis.

A mi Familia, A mi Esposo e Hijo por su comprensión y a Mi mamá por su apoyo incondicional.

Al Instituto Nacional de Astrofísica Óptica y Electrónica por su apoyo durante la realización de mis estudios de Maestría.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) de México, por el apoyo económico proporcionado para mis estudios de maestría.

Índice

ABSTRACT	III
-----------------	------------

CAPÍTULO 1 INTRODUCCIÓN	1
--------------------------------	----------

1.1. ANTECEDENTES Y ESTADO DEL ARTE.	1
1.2.OBJETIVO	4
1.2.1.OBJETIVOS ESPECÍFICOS	4
1.3. ALCANCES Y LIMITACIONES	4
1.4. METODOLOGÍA	5
1.5. ORGANIZACIÓN DEL DOCUMENTO	6

CAPÍTULO 2 FUNDAMENTOS TEÓRICOS	7
--	----------

2.1. PATOLOGÍAS NEUROLÓGICAS REFLEJADAS EN SEÑALES DE LLANTO	7
2.1.1.HIPERBILIRRUBINEMIA	7
2.1.2.ASFIXIA PERINATAL	10
2.2. PROCESAMIENTO DIGITAL DE LA SEÑAL	11
2.2.1.CODIFICACIÓN DE LA SEÑAL ANÁLOGA A DIGITAL	12
2.3. PRODUCCIÓN DE LA VOZ	13
2.3.1.MODELO DIGITAL DE LA PRODUCCIÓN DE LA SEÑAL DE VOZ.	13
2.4. EL OÍDO	14
2.5. ATRIBUTOS ACÚSTICOS DEL SONIDO	15
2.5.1.CODIFICACIÓN PREDICTIVA LINEAL	15
2.5.2.COEICIENTES CEPSTRALES DE FRECUENCIA MEL (MFCC)	17
2.5.3.INTENSIDAD	21
2.5.4.COCLEOGRAMA	21
2.6. LÓGICA DIFUSA	24

2.7. CONJUNTOS DIFUSOS TIPO 2	25
-------------------------------	----

CAPÍTULO 3 PROCESO DE CLASIFICACIÓN DE LLANTO DE INFANTES	29
--	-----------

3.1. DIAGRAMA DE BLOQUES DEL CLASIFICADOR	29
3.1.1.BASE DE LLANTOS GRABADOS	30
3.1.2.EXTRACCIÓN DE CARACTERÍSTICAS	31
3.1.3.REDUCCIÓN DE DATOS	33
3.1.3.1ANÁLISIS DISCRIMINANTE LINEAL	33
3.1.3.2REDUCCIÓN POR OPERACIONES ESTADÍSTICAS	35
3.1.4. CLASIFICACIÓN	36
•COMPARACIÓN DIFUSA DE PATRONES	36
•ALGORITMO DE COMPARACIÓN DIFUSA DE PATRONES CON CONJUNTOS DIFUSOS TIPO 2.	39

CAPÍTULO 4 EXPERIMENTACIÓN Y RESULTADOS	43
--	-----------

4.1. DESCRIPCIÓN DE LAS PRUEBAS REALIZADAS	43
4.1.1.DESCRIPCIÓN DEL MÉTODO DE VALIDACIÓN CRUZADA <i>10-FOLDS</i>	43
4.1.2.MÉTRICAS DE EVALUACIÓN	43
4.2. PRUEBAS DE FUNCIONALIDAD	46
4.3. PRUEBAS CON TRES CLASES DE LLANTO	48
4.3.1.RESULTADOS USANDO TODAS LAS CARACTERÍSTICAS	48
4.3.2.RESULTADOS OBTENIDOS CON DATOS REDUCIDOS	51
4.4. PRUEBAS CON DOS CLASES DE LLANTO: NORMAL E HIPERBILIRRUBINEMIA	55
4.5. COMPARACIÓN CON OTROS TRABAJOS.	57
4.6. DISCUSIÓN	59

CAPÍTULO 5 CONCLUSIONES Y TRABAJO FUTURO	61
---	-----------

5.1. CONCLUSIONES Y APORTACIONES PRINCIPALES	61
5.2. TRABAJO FUTURO	62

APÉNDICE A.	65
APÉNDICE B	73
LISTA DE ACRÓNIMOS	77
BIBLIOGRAFÍA	79

ÍNDICE DE FIGURAS

Figura 2.1. Partes básicas de un convertidor análogo-digital [Proakis, Manolakis, 1998]	12
Figura 2.2. Modelo digital del habla [Smith S., 1999].....	14
Figura 2.3. Diagrama funcional del oído humano [Smith S., 1999].....	15
Figura 2.4. Diagrama en bloques del algoritmo para obtener los MFCC's [Zbynì y Psutka, 1999].....	17
Figura 2.5. Escala Mel [Wikipedia, 2008].....	19
Figura 2.6. Banco de filtros en escala Mel.....	20
Figura 2.7. Representación gráfica de la relación entre Hertz y escala Bark,.....	23
Figura 2.8. Enmascaramiento de sonidos [Wikipedia, 2008].	23
Figura 2.9. Representación espectral de un cocleograma para el llanto de infante con hiperbilirrubinemia.....	24
Figura 2.10 (a) Función de membresía tipo 1 (b) Función de membresía tipo 1 sombreada [Mendel y Bob, 2002].....	25
Figura 2.11. Ejemplo de una función de membresía tipo 2 [Mendel y Bob, 2002]....	26
Figura 3.1 Diagrama en Bloques de un Clasificador de llanto	29
Figura 3.2. Gráfica del Cocleograma para una muestra de la clase Hiperbilirrubinemia sin reducir.....	34

Figura 3.3. Cocleograma de una muestra de la clase Hiperbilirrubinemia reducido con el método LDA	34
Figura 3.4. Gráfica de los Cocleogramas de la clase Hiperbilirrubinemia reducidos por el método de operaciones estadísticas.	35
Figura 3.5. Gráfica de los Cocleogramas de la clase Asfixia reducidos por el método de operaciones estadísticas.	36
Figura 3.6. Algoritmo de Comparación Difusa de Patrones [Saïd, Sayed y Billaudel, 2007]	37
Figura 4.1. Algoritmo <i>validación cruzada 10-folds</i>	44

ÍNDICE DE TABLAS

Tabla 2.1. Intensidad de algunos sonidos conocidos [Smith S., 1999]	21
Tabla 3.1. Segmentos de las grabaciones de llanto de infante.	30
Tabla 3.2. Parámetros utilizados para la extracción de características	31
Tabla 3.3. Atributos Obtenidos por característica.....	33
Tabla 3.4. Valores obtenidos con la reducción por operaciones estadísticas para el cocleograma obtenido de una muestra de hiperbilirrubinemia.	35
Tabla 4.1. Ejemplo de una matriz de confusión con precisión y recuerdo..	46
Tabla 4.2. Pruebas con los datos Fisher Iris.....	47
Tabla 4.3. Pruebas con los datos Fisher Iris.....	47
Tabla 4.4. Pruebas con datos obtenidos de las muestras de Llanto de infante, sin reducir.	49
Tabla 4.5. Recuerdo de las pruebas con datos obtenidos de las muestras de Llanto de infante, sin reducir.....	50
Tabla 4.6. Resultados de pruebas con combinación de características.	51
Tabla 4.7. Precisión obtenida en la clasificación utilizando datos reducidos con el método LDA.	52
Tabla 4.8. Recuerdo obtenida en la clasificación utilizando datos reducidos con el método LDA.	53
Tabla 4.9. Clasificación utilizando datos reducidos con el método de reducción estadística.	54

Tabla 4.10. Recuerdo obtenido en la clasificación utilizando datos reducidos con el método de reducción estadística.	54
Tabla 4.11. Pruebas realizadas para la clasificación de dos clases: Normal e Hiperbilirrubinemia.	55
Tabla 4.12. Pruebas realizadas con dos clases (Normal e Hiperbilirrubinemia) utilizando la combinación de dos características, LPC y MFCC.	56
Tabla 4.13. Matriz de Confusión de las Pruebas con dos clases utilizando LPC-MFCC	57
Tabla 4.14. Resultados obtenidos por Amaro [Amaro, 2008].....	58
Tabla 4.15. Resultados obtenidos en éste trabajo.....	58
Tabla 4.16 Descripción de los conjuntos de datos Pima y LBC.....	59
Tabla 4.17. Resultados obtenidos de pruebas realizadas con los conjuntos de datos Pima y LBC.	59
Tabla A.1 Matriz de Confusión de las pruebas con la base de datos de las Flores Iris	65
Tabla A.2. Matriz de Confusión de las pruebas con datos de la base de datos de llanto sin aplicar reducción utilizando las cuatro características (LPC, MFCC, Intensidad y Cocleogramas).....	66
Tabla A.3. Matriz de Confusión de las pruebas utilizando datos reducidos con el método LDA utilizando cuatro características (LPC, MFCC, Intensidad y Cocleogramas).....	67

Tabla A.4. Matriz de Confusión de las pruebas realizadas con datos reducidos con el método de Operaciones Estadísticas utilizando cuatro características (LPC, MFCC, Intensidad y Cocleogramas).....	68
Tabla A.5. Matriz de Confusión de las pruebas realizadas con datos sin reducir utilizando la característica LPC.	69
Tabla A.6. Matriz de Confusión de las pruebas realizadas con datos sin reducir utilizando las características LPC y Cocleogramas.	70
Tabla A.7. Matriz de confusión de pruebas con datos sin reducir utilizando la característica LPC para dos clases: Normal e Hiperbilirrubinemia	71
Tabla A.8. Evaluación de las pruebas realizadas con datos sin reducir utilizando la característica LPC para dos clases: Normal e Hiperbilirrubinemia.	71
Tabla A.9. Matriz de confusión de pruebas con datos sin reducir utilizando la característica MFCC para dos clases: Normal e Hiperbilirrubinemia	72
Tabla A.10. Evaluación de las pruebas realizadas con datos sin reducir utilizando la característica LPC para dos clases: Normal e Hiperbilirrubinemia.	72
Tabla B.1. Tamaño de las matrices generadas por característica.....	74

Capítulo 1

INTRODUCCIÓN

En este capítulo se introduce el trabajo de tesis “Conjuntos Difusos Tipo 2 aplicados a la Comparación Difusa de Patrones para Clasificación de llanto de infantes con riesgo neurológico”, sus objetivos y metodología. Asimismo se presentan algunos antecedentes y el estado del arte de esta área de conocimiento.

1.1. Antecedentes y Estado del Arte.

Estudio del llanto de infantes

El estudio del llanto de infantes se ha incrementado en los últimos años. La primera referencia que se tiene es de 1850, cuando Charles Darwin diagnosticó el llanto, a través de una serie de fotografías y dibujos que ilustraban expresiones de emoción o dolor [Petroni, 1995]. Prácticamente el punto de partida para el análisis del llanto de infantes fue en 1968, cuando Wasz-Höckert realizó el primer estudio de llanto de infante basándose en un análisis de espectrogramas [Wasz-Höckert, 1968]. Posteriormente Golub realizó un experimento al cual llamó “Primer Análisis Auditivo” donde entrenó a 93 madres para que identificaran el llanto de su bebé entre el llanto de varios infantes [Golub y Corwin, 1982]. El resultado fue que más del 95% de las madres pudo identificar a su propio bebé.

Gracias a la evolución de la tecnología, actualmente se utilizan modelos computacionales para automatizar la identificación de patologías en el infante. El proceso consta de la extracción de características de la señal de llanto, las cuales se utilizan para entrenar un modelo de clasificación, y posteriormente, para realizar pruebas. Para extraer características se usan, entre otros métodos: Codificación

Predictiva Lineal (LPC por sus siglas en inglés *Linear Predictive Coding*), intensidad, frecuencia fundamental.

En [Cano, 2007] se establece que el llanto es un evento acústico que contiene información sobre el funcionamiento del sistema nervioso central, y que el análisis del llanto de infante puede ser un apoyo en el diagnóstico diferencial en casos como asfixia e hiperbilirrubinemia. También se presentan los resultados obtenidos en la clasificación de estas dos patologías, los autores utilizaron un método basado en mapas auto-organizados de Kohonen y obtuvieron una precisión del 85%.

Un ejemplo de este tipo de clasificación es el trabajo desarrollado por Orozco García [Orozco, 2002] quien clasificó el llanto de infantes normales y sordos. Los resultados que reportó fueron de 98.5% de precisión para la clasificación de una patología y casos normales, utilizando una red neuronal *feed-forward* entrenada con el método de gradiente conjugado escalado. Orozco utilizó cuatro tipos de características: Coeficientes de Predicción Lineal, Coeficientes Cepstrales de Frecuencia Mel (MFCC por sus siglas en inglés *Mel Frequency Cepstral Coefficients*), Frecuencia Fundamental e Intensidad.

Uno de los trabajos más recientes es el de Erika Amaro [Amaro, 2008], donde se propone un sistema de reconocimiento automático del llanto de infantes basado en la combinación de características. Clasificando dos clases, normales y sordos, con una red neuronal *feed forward*, logró resultados del 98.66% de precisión con la combinación de cuatro características que son: LPC, MFCC, Intensidad y Cicleogramas.

En [Verduzco, 2004] se explica cómo es que las patologías que afectan el sistema nervioso y que pueden tener secuelas de tipo neurológicas, producen cambios en el llanto de los infantes. En el caso de la hiperbilirrubinemia el llanto es afectado en la frecuencia fundamental, incrementa el porcentaje de fonación y se incrementa la variabilidad del primer formante (la intensidad más alta en una frecuencia

determinada). En llantos de infantes con asfixia se detecta un incremento en la frecuencia fundamental por encima de los casos de llanto de infantes normales.

Comparación Difusa de Patrones y Lógica Difusa

Los conjuntos difusos tipo 2 y los sistemas basados en lógica difusa se han utilizado en aplicaciones como: control de robots móviles, toma de decisiones, pre-procesamiento de imágenes [Mendel y Bob, 2002]. Asimismo, la teoría de los conjuntos difusos tipo 2 se ha usado para modelar la incertidumbre asociada a las membresías de los clusters [Chung y Rhee, 2007]. El método llamado *Interval type-2 fuzzy C-means* se utilizó para reconocer imágenes teniendo una precisión del 100% en comparación con el método *fuzzy C-means* que obtuvo una precisión del 90%.

En [Sayed, 2002] se describe un método de comparación difusa de patrones (CDP) y se realizan experimentos con clases cuyos datos son sobrepuestos¹, con una precisión promedio de 92.58%.

En [Chen, Wu y Yachida, 1995] utilizan el método CDP para la detección del rostro de una persona dentro de una imagen, Una característica de este tipo de datos es que son sobrepuestos. Después de aplicar el método CDP, concluyen que el método es flexible y robusto, ya que lograron detectar rostros en distintas imágenes basándose en la extracción de mapas de colores.

En esta tesis se presentan los resultados al clasificar llantos de infantes recién nacidos con que tienen enfermedades que pueden provocar problemas neurológicos como son: hiperbilirrubinemia y asfixia perinatal, estas patologías se definen en la sección 2.1. Este tipo de llantos tiene como característica que son clases sobrepuestas, es por eso que se propone un nuevo clasificador de comparación difusa de patrones que utilice los conjuntos difusos tipo 2 (CDP-T2 Comparación Difusa de Patrones Tipo 2), con el cual se pretende establecer una mejor diferenciación entre clases para

¹ Las clases sobrepuestas son las que tiene datos son muy similares [Saïd, Sayed y Billaudel, 2007], es decir, que los valores de los datos no tienen una distancia significativa entre clases.

clasificar mejor este tipo de llanto, ya que este tipo de conjuntos almacena mayor información que describe la pertenencia de un elemento a una clase.

1.2.Objetivo

El objetivo de este trabajo es el definir y probar un método basado en la Comparación Difusa de Patrones (CDP) que utilice conjuntos difusos tipo-2 adecuado para clasificar el llanto de infantes con riesgo neurológico.

1.2.1. Objetivos Específicos

- Incorporar los Conjuntos difusos tipo 2 en el método CDP.
- Probar el método CDP con conjuntos difusos tipo 2 en el reconocimiento de llanto de infantes con patologías neurológicas de asfixia e hiperbilirrubinemia.
- Determinar las características que mejor identifiquen a los llantos de infantes con hiperbilirrubinemia.

1.3.Alcances y limitaciones

En esta tesis se aborda el tema de la clasificación de llanto de infantes con patologías de riesgo neurológico. Se propone un nuevo método de clasificación basado en comparación difusa de patrones aplicando conjuntos difusos tipo 2 (CD-T2).

El nuevo clasificador se basa en el descrito en [Saïd, Sayed y Billaudel, 2007] utilizando la teoría de conjuntos difusos tipo dos como se explica en [Mendel, 2007] y en [Mendel y Bob, 2002].

El clasificador ha sido probado con muestras de llanto de infantes normales y con patologías de hiperbilirrubinemia y asfixia perinatal. Para el análisis de las señales de llanto se utilizaron cuatro técnicas para la obtención de atributos acústicos, dos basadas en modelos perceptuales (MFCC y Cocleogramas) y dos basadas en modelos de producción de voz (LPC e Intensidad).

1.4. Metodología

La metodología utilizada es la siguiente:

1. Extracción y análisis de características acústicas de la señal de llanto de infantes etiquetada con patologías de hiperbilirrubinemia, asfixia o normales.
 - (a) Se realiza un procesamiento de la señal de llanto unificando los parámetros (como el número de coeficientes, el salto de tiempo, entre otros) utilizados para la obtención de los vectores característicos (LPC, MFCC, Intensidad y cocleogramas) de las muestras.
 - (b) Se segmenta las muestras de llanto en fragmentos de un segundo, esto para tener un número mayor de muestras de entrenamiento y para facilitar el manejo y extracción de características y la comparación con trabajos previos como en [Amaro, 2008] y [Orozco, 2002].
 - (c) Se extraen las características acústicas: LPC, MFCC, Cocleogramas e intensidad. Este proceso se realiza con ayuda del software PRAAT [Boersma y Weenink, 2007] que permite la manipulación de señales y extracción de características, entre otras funciones.
 - (d) Limpieza y etiquetado de los vectores característicos obtenidos los cuales posteriormente son utilizados en programas de MatLab que realizan la clasificación.
2. Selección de técnicas de reducción de información.
 - (a) Se utilizan las técnicas de reducción de datos: LDA y el método de reducción estadística propuesto y utilizado en [Amaro, 2008].
3. Análisis e implementación de un método de clasificación basado en comparación difusa de patrones utilizando la teoría de los conjuntos difusos tipo 2.

- (a) Se analiza del método CDP [Saïd, Sayed y Billaudel, 2007] y de los conjuntos difusos tipo 2 [Mendel y Bob, 2002] para diseñar un reconocedor que utilice estas dos metodologías.
 - (b) Se implementa un reconocedor basado en el método CDP con CD-T2.
4. Pruebas realizadas con los datos extraídos de las señales de llanto de infantes con patologías detectadas de hiperbilirrubinemia y asfixia.
- (a) Se realizan pruebas utilizando las características de forma individual y combinadas para la clasificación de las tres clases: asfixia, hiperbilirrubinemia y normal; y también para la clasificación de dos clases: hiperbilirrubinemia y normal.
5. Evaluación y análisis del reconocedor haciendo una comparación de los resultados obtenidos.
- (a) Se evalúa el desempeño del reconocedor, utilizando las métricas de precisión y recuerdo, definidas en la sección 4.1.2. Métricas de evaluación.

1.5.Organización del documento

La organización del presente documento es la siguiente: en el Capítulo 2 se explican los fundamentos teóricos, la descripción de las patologías de llanto que se clasifican, el procesamiento digital de señales de audio, las características que se obtienen de las muestras de llanto, la lógica difusa, en el Capítulo 3 se describe el proceso de clasificación del llanto de infantes y el algoritmo que se utilizó para clasificar, en el Capítulo 4 se muestran las pruebas realizadas y los resultados obtenidos con base en las métricas de evaluación, por último en el Capítulo 5 se presentan las conclusiones y trabajo futuro.

Capítulo 2

FUNDAMENTOS TEÓRICOS

2.1. Patologías neurológicas reflejadas en señales de llanto

Se sabe que el llanto en un recién nacido, además de ser su único medio de expresión, contiene información sobre la salud del infante. Cano [Cano, 2007] y Verduzco [Verduzco, 2004] hacen mención de las patologías del sistema nervioso que se pueden detectar en el llanto de infantes. En [Jara, 2006] se realiza un estudio de llanto de recién nacidos en término (es decir que nacieron en tiempo adecuado) y que padecían hiperbilirrubinemia o asfixia, donde encontraron características especiales en estos tipo de llanto. De acuerdo a ellos, cada patología tiene distintas características que pueden ser detectadas en el llanto. A continuación se describen dos de las patologías que se clasifican en este trabajo y las características acústicas que las diferencian.

2.1.1. Hiperbilirrubinemia

Barkin [Barkin, 2000] describe a la hiperbilirrubinemia como un trastorno causado por una cantidad excesiva de bilirrubina en la sangre. La bilirrubina es el resultado de la descomposición de la hemoglobina, la cual es una proteína en los glóbulos rojos que transporta oxígeno. Para los bebés es difícil deshacerse de la bilirrubina, así que es posible que ésta se acumule en su sangre, sus tejidos y fluidos corporales, a lo que se le llama hiperbilirrubinemia. Dado que la bilirrubina tiene un pigmento o coloración amarilla, la piel y los tejidos del bebé se tornan amarillentos, por lo que también se le llama ictericia.

Durante el embarazo la placenta excreta la bilirrubina, pero cuando nace el bebé, su hígado debe cumplir con esa función. Existen diversas condiciones que causan la hiperbilirrubinemia, algunas se describen a continuación:

- **Ictericia fisiológica:** La ictericia fisiológica se presenta como una respuesta "normal" a la capacidad limitada del bebé para expulsar bilirrubina durante los primeros días de vida. Esta ictericia normalmente no causa problema alguno.
- **Ictericia por leche materna:** Aproximadamente del 5% al 10 % de los bebés alimentados con leche materna desarrollan ictericia después de la primera semana y ésta puede mantenerse en niveles bajos durante un mes o más. Normalmente esta ictericia no causa ningún problema, aunque en casos especiales es necesario retirar la leche materna mientras desaparece la ictericia.
- **Ictericia por hemólisis:** La ictericia puede presentarse como resultado de la destrucción de una excesiva cantidad de glóbulos rojos debida a una enfermedad hemolítica del recién nacido (como la enfermedad de Rh por incompatibilidad entre los grupos sanguíneos de la madre y el bebé), o también por algún sangrado.
- **Ictericia relacionada a la deficiencia hepática:** La ictericia puede estar relacionada a una deficiencia hepática (daño en el hígado) ocasionada por una infección u otros factores.

Aproximadamente el 60% de los recién nacidos a término y el 80% de los prematuros desarrollan ictericia. Los bebés de madres diabéticas y de madres con enfermedad Rh tienen mayores probabilidades de desarrollar hiperbilirrubinemia [Hayes, 2000].

La importancia de diagnosticar la hiperbilirrubinemia

Generalmente los niveles bajos de bilirrubina no son problema, pero en niveles altos la bilirrubina afecta los tejidos del cerebro y puede provocar sordera, convulsiones, parálisis cerebral y hasta kernícterus (daño cerebral a causa de altos niveles de bilirrubina). Estos riesgos neurológicos se pueden prevenir con un diagnóstico y tratamiento a tiempo [Barkin, 2000].

Un factor importante en el diagnóstico de la hiperbilirrubinemia es el momento en que se presenta la ictericia. Si se presenta en las primeras 24 horas de nacido, el mal es muy grave y requiere tratamiento inmediato. Si se presenta durante el segundo o tercer día de nacido, generalmente es un caso de ictericia fisiológica. Si se presenta entre el tercer día y la primera semana, puede ser una infección. Si se presenta durante la segunda semana de nacido, probablemente está asociada a la lactancia [Hayes, 2000]. En cualquier caso puede agravarse el trastorno. Entre los procedimientos para el diagnóstico de la hiperbilirrubinemia se pueden incluir:

- Medición de niveles de bilirrubina directa e indirecta. Reflejan si la bilirrubina está ligada a otras sustancias producidas por el hígado y puede excretarse (directa), o es parte de la circulación sanguínea (indirecta).
- Recuentos de glóbulos rojos.
- Determinación del grupo sanguíneo y exámenes para determinar la incompatibilidad de Rh (prueba de Coomb).

Características Acústicas del llanto de infantes con hiperbilirrubinemia

Verduzco [Verduzco, 2004] menciona que el llanto se produce debido a la interacción entre las estructuras anatómicas y los mecanismos fisiológicos, incluyendo el control respiratorio, laríngeo y sistema supralaríngeo. Los músculos de la laringe, faringe, tórax y parte superior del cuello están controlados por los nervios craneales número IX (glossofaríngeo), X (vago), XI (accesorio) y XII (hipogloso). Un daño a cualquiera de estos nervios craneales o a su núcleo provoca cambios en su función y por lo tanto en las características del llanto. Por otra parte la hiperbilirrubinemia es un factor de riesgo de los núcleos del tallo cerebral. Debido a esto se han realizado investigaciones sobre la afectación que causa sobre el llanto de infantes, tales como el incremento de la frecuencia fundamental².

² La Frecuencia Fundamental es la frecuencia más baja del espectro de frecuencias de un sonido de tal forma que las frecuencias dominantes pueden expresarse como múltiplos de esta frecuencia.

Jara [Jara, 2006] realiza un análisis espectrofonográfico sobre el llanto del infantes que presentaban patologías de riesgo neurológico, reportó que la frecuencia fundamental de llantos afectados por hiperbilirrubinemia, incrementó hasta 2120 Hz. y como mínimo tuvo 960 Hz. comparando con un máximo de 670 Hz. y mínimo de 390 Hz, en llanto de infantes sanos.

2.1.2. Asfixia Perinatal

La asfixia perinatal está definida como la insuficiencia de oxígeno en el sistema circulatorio del feto y del recién nacido [Dworkin, 1995]. Esta patología puede ocurrir antes del nacimiento, durante el embarazo, en el trabajo de parto, en el parto y después del nacimiento. La disminución del oxígeno durante períodos prolongados provoca disminución de la presión arterial de manera que la sangre deja de irrigar a los órganos vitales. Las complicaciones inmediatas son las lesiones orgánicas y entre las secuelas principales están las lesiones cerebrales.

La importancia de diagnosticar la Asfixia Perinatal

El diagnóstico de la Asfixia Perinatal está ligada con la valoración del *Apgar* [Barkin, 2000]. El *Apgar* es una valoración sobre el estado del recién nacido realizada al minuto y a los cinco minutos de nacido. Esta valoración se hace en base a cinco parámetros que son: frecuencia cardíaca, irritabilidad, coloración, esfuerzo respiratorio y tono muscular. Se considera que las condiciones son adecuadas cuando el recién nacido obtiene un *Apgar* de 8-10. La asfixia perinatal se considera grave cuando el *Apgar* del recién nacido es menor a 5 [Dworkin, 1995]. La detección y el control prenatal adecuados junto a la asistencia oportuna del parto y del neonato disminuyen las complicaciones de esta patología.

En México, las principales causas de muerte de recién nacidos son: asfixia perinatal, prematuridad, malformaciones congénitas relacionadas con la falta de control prenatal y diagnóstico tardío de embarazo de alto riesgo. En el 2004 la tasa nacional de mortalidad infantil fue de 19.4 por cada 1,000 nacidos vivos, la tasa nacional de

mortalidad por asfixia del recién nacido fue de 4.59 de cada 10 mil nacidos vivos, [SNVE, 2006].

La mayoría de las causas de asfixia perinatal son de origen intrauterino, aproximadamente el 5% ocurre antes del inicio de trabajo de parto, 85% durante el parto y el periodo expulsivo y el 10% durante el periodo neonatal [SNVE, 2006]. Es importante diagnosticar y atender la asfixia perinatal, sobre todo en lugares de nuestro país donde no cuentan con un servicio médico apropiado, para así disminuir la tasa de mortalidad y evitar consecuencias graves.

Características Acústicas del llanto de infantes con asfixia perinatal

Verduzco [Verduzco, 2004] sostiene que en el llanto de infantes asfixiados la frecuencia fundamental incrementa sobre la del llanto de infantes sanos. Por otro lado Jara [Jara, 2006] no encontró diferencias significativas entre la frecuencia fundamental de llanto de infantes asfixiados y normales. Sin embargo documentó altas intensidades en las señales de llanto de infantes con asfixia.

2.2.Procesamiento Digital de la Señal

El Procesamiento Digital de Señales (DSP por sus siglas en inglés *Digital Signal Processing*) es el área especializada en el análisis y procesamiento de señales digitales, como son señales de audio, voz, imágenes, entre otras. Las señales que se encuentran de forma analógica se convierten a datos digitales por distintas razones: es más fáciles de procesar, transmitir y/o almacenar [Smith, 1999].

Entre las ventajas del procesamiento de señales en forma digital sobre las señales analógicas, está que es más fácil de implementar y utiliza menos recursos (como por ejemplo la memoria). Una de las desventajas es que puede haber pérdida de información al convertir la información continua a información discreta; sin embargo, esta pérdida puede no ser significativa si se realiza correctamente [Proakis y Manolakis, 1998].

2.2.1. Codificación de la Señal Análoga a Digital

La señal de audio es una señal análoga la cual debe convertirse en señal digital; esto es necesario para tener una secuencia de números que tienen una precisión finita. La conversión de análogo a digital es un proceso que se compone principalmente de tres pasos [Proakis y Manolakis, 1998] como se muestra en la Figura 2.1 y que se describen de forma general a continuación:

- **Muestreo:** Es la conversión de una señal en tiempo continuo a una señal en tiempo discreto la cual se obtiene tomando “muestras” de la señal en tiempo continuo a instantes de tiempo discreto. De esta forma si $x_a(t)$ es la entrada de una muestra, la salida es $x_a(nT) \equiv x(n)$, donde T es el *intervalo de muestreo*. En otras palabras, se divide la señal continua en pequeños intervalos de tiempo discretos.
- **Cuantificación:** Es la conversión de una señal de tiempo discreto con valores continuos en una señal de tiempo discreto con valores discretos (señal digital). El valor de cada muestra de la señal es representado por un valor seleccionado de un grupo finito de posibles valores. La diferencia entre las muestras no cuantificadas $x(n)$ y la salida cuantificada $x_q(n)$ se conoce como error de cuantificación.
- **Codificación:** En este proceso cada valor $x_q(n)$ se representa por una secuencia de bit's.

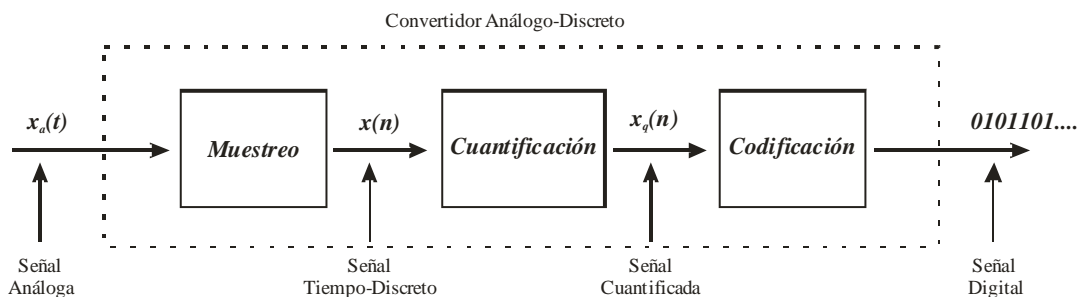


Figura 2.1. Partes básicas de un convertidor análogo-digital [Proakis, Manolakis, 1998]

2.3.Producción de la voz

El sonido que generamos al hablar es originado por el aire que sale de los pulmones y que llega al exterior atravesando la laringe y el tracto vocal, provocando dos tipos de sonidos, el *sonoro* y el *sordo*. En la laringe tenemos las cuerdas vocales (dos pliegues musculares) que al abrirse y cerrarse provocan la vibración de las moléculas de aire produciendo un sonido que sale por la boca. Este tipo de sonido se conoce como *sonoro*. Las vocales son un ejemplo del sonido *sonoro*. El tipo de sonido *sordo*, sin intervención de las cuerdas vocales, es originado como ruido aleatorio, no de las vibraciones, cuando el flujo de aire es casi bloqueado por el tracto vocal (la lengua, labios y/o los dientes); las consonantes *s*, *f*, *z*, son ejemplos de este tipo de sonido. Las cuerdas vocales vibran a frecuencias de entre 50 y 1000 Hz. [Smith S., 1999].

2.3.1. Modelo digital de la producción de la señal de voz.

El habla puede ser modelada de forma digital, en intervalos de tiempo de entre 2 a 40 milisegundos, por tres parámetros: el tipo del sonido (sonoro o sordo), la frecuencia fundamental y los coeficientes de un filtro digital [Smith S., 1999]. El modelo de producción digital de la señal de voz se muestra en la Figura 2.2, donde se observa un bloque llamado generador de ruido, el cual genera el ruido correspondiente al ruido sordo, el bloque llamado Generador de Pulsos recibe el parámetro de frecuencia fundamental y genera el ruido sonoro, estos dos tipos de sonidos son elegidos uno a la vez para pasarlo por el bloque de filtro digital, en el cual se aplican los coeficientes generados artificialmente y de esta forma genera el habla artificial.

El filtro digital representa las modificaciones que se hacen a los sonidos por el tracto vocal, estas modificaciones están representadas por los coeficientes obtenidos del sonido original. Una señal de voz puede ser reconstruida con los datos de este modelo.

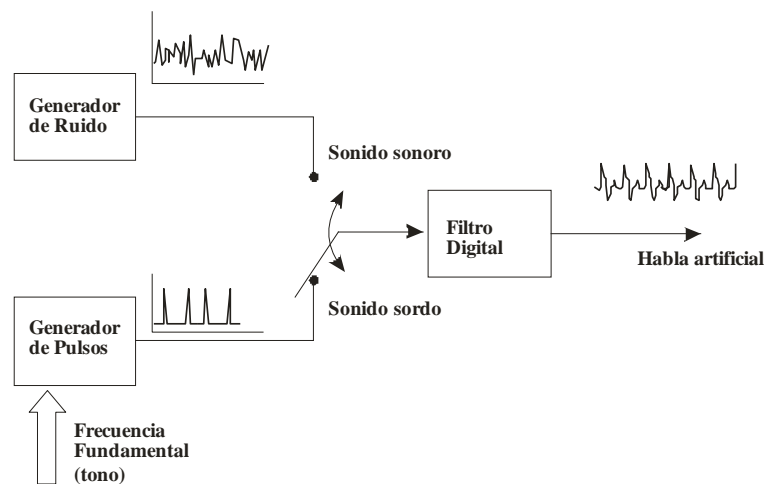


Figura 2.2. Modelo digital del habla [Smith S., 1999].

2.4.El oído

El oído humano se encarga de convertir las ondas sonoras en señales eléctricas que se transmiten por el nervio acústico hasta el cerebro, donde se interpreta el sonido. Los cocleogramas y los MFCC son dos características utilizadas para reconocimiento de llanto, inspirada en la percepción auditiva del humano. El oído se divide en tres partes principales: oído externo, oído medio y oído interno [Smith S., 1999]. En la Figura 2.3 se muestra el diagrama funcional del oído humano.

Oído externo: Es la parte externa del oído, la oreja, el canal auditivo y el tímpano. Se encarga de recoger las ondas sonoras y conducirlas a través del canal auditivo hacia el tímpano, el cual vibra en sincronización con las ondas recibidas.

Oído Medio: En esta parte del oído se encuentran tres huesos pequeños que son: el martillo, el yunque y el estribo. La cabeza del martillo se apoya sobre el tímpano y transmite las vibraciones a través del yunque al estribo, que a su vez se apoya en la ventana oval.

Oído Interno: Consta principalmente de tres elementos: los canales semicirculares, el vestíbulo y la cóclea. Los canales semicirculares tienen que ver con el equilibrio. La ventana oval del vestíbulo recibe las vibraciones por parte del estribo y son

transformadas en la cóclea. Dentro de la cóclea se encuentra la membrana basilar, la cual tiene los nervios auditivos y aproximadamente 12,000 células nerviosas. Debido a la rigidez de la membrana basilar, cada célula nerviosa solo responde a un pequeño rango de frecuencias auditivas, por lo que el oído funciona como un analizador de frecuencia.

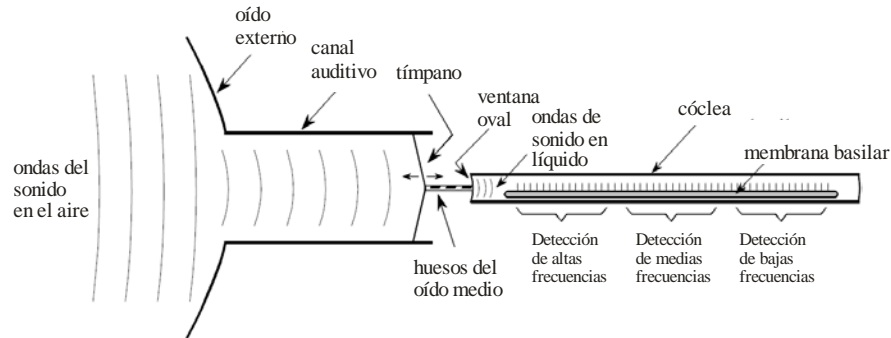


Figura 2.3. Diagrama funcional del oído humano [Smith S., 1999].

2.5. Atributos acústicos del sonido

En este trabajo, las señales de llanto de infante son tratadas como habla. La extracción de características utilizada en este trabajo, es un proceso ya probado en trabajos anteriores, como en [Amaro, 2008], [Orozco, 2002]. Existe una gran variedad de métodos que representan una señal de audio en forma digital, para este trabajo se utilizaron cuatro métodos: Coeficientes de Predicción Lineal, Coeficientes Cepstrales de Frecuencia Mel, Intensidad y Cocleogramas, los cuales ya han sido utilizados en la clasificación de llanto de infantes en trabajos como [Orozco, 2002] y [Amaro, 2008] y se ha comprobado su eficacia en la clasificación. Estos métodos se describen en las siguientes secciones.

2.5.1. Codificación Predictiva Lineal

El método de codificación predictiva lineal es utilizado principalmente para la compresión de datos y el análisis de una señal digital de habla. Este método está basado en el hecho de que una señal que transporta mensajes no es completamente

aleatoria, es decir, existe una correlación entre muestras sucesivas, debido a que la voz mantiene sus propiedades prácticamente invariantes durante determinados intervalos de tiempo. LPC utiliza esta característica para reducir la cantidad de datos cuando se guarda la información de la señal [Smith S., 1999].

El método de predicción lineal pretende extrapolar el valor de la muestra de voz $s(n)$ como la suma ponderada de muestras pasadas $s(n-1)$, $s(n-2)$, ..., $s(n-k)$. De ésta forma es posible determinar el valor de cualquier muestra a partir de las últimas p muestras. Entonces podemos definir una muestra $s(n)$ por medio de la siguiente función de transferencia:

$$s(n) = \sum_{k=1}^p a_k s(n-k) + \delta(n) \quad (2.1)$$

donde $\delta(n)$ es la excitación que llega al filtro digital (puede ser sonido sonoro o sordo) y puede consistir en impulsos periódicos o en ruido aleatorio, los a_k son los parámetros del filtro digital (los que utiliza el bloque de filtro digital de la Figura 2.2). Si se realiza una aproximación de (2.1) por medio de un polinomio predictor lineal se obtiene la siguiente ecuación:

$$s'(n) = \sum_{k=1}^p \alpha_k s(n-k) \quad (2.2)$$

donde los α_k son los coeficientes del predictor a calcular. Una forma de hallar estos coeficientes es minimizando el error de predicción, es decir, minimizar la media de los cuadrados de la diferencia entre las muestras reales y las que se generaron con la predicción lineal durante un intervalo de tiempo. Dadas p muestras de la señal, para calcular α_k , se debe minimizar alguna función de error E , concretamente de mínimos cuadrados, sobre una ventana de tamaño N .

El error en cualquier momento, e_n está dado por:

$$e_n = s(n) - s'(n) \quad (2.3)$$

$$= s(n) - \sum_{k=1}^p \alpha_k s(n-k) \quad (2.4)$$

donde $s(n)$ es la señal real, α_k son los coeficientes que se están calculando, p es el número de muestras de la señal. Entonces la suma del error cuadrático E a lo largo de una ventana finita de longitud N , se calcula con la ecuación:

$$E = \sum_{n=0}^{N-1} e_n^2 = \sum_{n=0}^{N-1} (s(n) - \sum_{k=1}^p \alpha_k s(n-k))^2 \quad (2.5)$$

Donde N es la longitud de la ventana, $s(n)$ es el valor real de la señal, p es el número de muestras de la señal, α_k son los coeficientes que se están calculando y se van ajustando de forma iterativa de tal forma que se minimice E .

2.5.2. Coeficientes Cepstrales de Frecuencia Mel (MFCC)

Los Coeficientes Cepstrales de Frecuencia Mel (MFCC, por sus siglas en inglés), son coeficientes que representan el habla basándose en la percepción auditiva del ser humano. El proceso de obtención de MFCC se muestra en la Figura 2.4 y la descripción de cada bloque se realiza a continuación.



Figura 2.4. Diagrama en bloques del algoritmo para obtener los MFCC's [Zbrynì y Psutka, 1999]

Pre-énfasis y Enventanado

El pre-énfasis consiste en enfatizar las frecuencias que tienen mayor información de la señal de voz. Las secciones de voz generalmente caen en frecuencias altas, debido a las características físicas del sistema de producción del habla. En las frecuencias

bajas se encuentra la mayoría del ruido, entonces se deben eliminar y enfatizar las frecuencias altas. Esto se hace por medio de un filtro “pasa altas”. Al eliminar las frecuencias bajas se elimina el ruido de fondo de una señal y aunque también se pueden eliminar frecuencias que hayan sido generadas por voz, esto no es un problema tan grave ya que el sistema auditivo humano puede compensar las frecuencias bajas que faltan en la señal.

El enventanado consiste en seleccionar una trama con la que se va a trabajar. Existen varios tipos de enventanado; en esta tesis se utiliza la ventana rectangular, para facilitar la comparación con trabajos previos como en [Amaro, 2008] y [Orozco, 2002] que utilizaron este tipo de ventana y que se define como:

$$w_n = \begin{cases} 1 & 0 \leq n \leq N \\ 0 & \text{resto} \end{cases} \quad (2.6)$$

Donde w_n es el elemento n de la ventana y N es el tamaño de la ventana.

Transformada Discreta de Fourier

La transformada discreta de Fourier o DFT (por sus siglas en inglés, *Discrete Foureier Transform*) se utiliza en el procesamiento de señales de audio para analizar las frecuencias presentes en una señal muestreada. Esta transformada se debe aplicar sobre la trama de muestras enventanadas. En [Wikipedia, 2008] definen la DFT como:

$$f_j = \sum_{k=0}^{N-1} x_k e^{-j \left(\frac{2\pi i k}{N} \right)} \quad j=0,1, \dots, N-1 \quad (2.7)$$

Donde N es el número de muestras de la ventana que se va a analizar, e es la base de los logaritmos naturales, i es la unidad imaginaria ($i^2 = -1$). Esta transformada descompone la señal de voz en sus componentes de frecuencia f_j .

Aplicación de Filtros Mel

Antes de analizar la aplicación de los filtros Mel se describe la escala Mel. La escala Mel fue propuesta por Smith Steven, John Volkman y Edwin Newman, y es una escala perceptual de frecuencias. En la Figura 2.5 se muestra en una gráfica la correspondencia de Hertz a su escala Mel [Wikipedia, 2008].

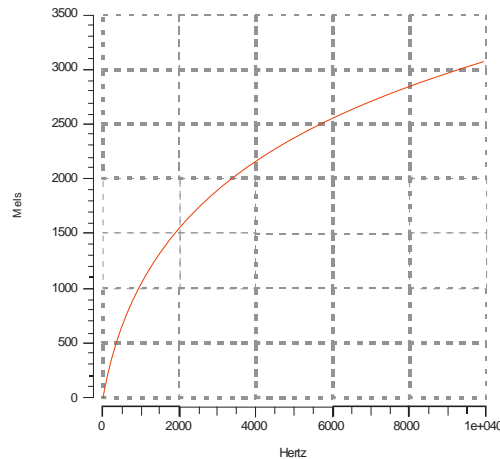


Figura 2.5. Escala Mel [Wikipedia, 2008].

La escala Mel resulta de dividir el espectro de frecuencias en un banco de filtros, en los cuales para las frecuencias bajas son muy estrechos y separados linealmente y en frecuencias altas son muy amplios y espaciados de forma logarítmica, es decir, están espaciados uniformemente por debajo de 1kHz. y logarítmicamente sobre 1 kHz. Con estos filtros se da mayor importancia a la información contenida en las bajas frecuencias, simulando la percepción de frecuencias del oído humano, y para capturar mejor las características fonéticas del habla [Zbynì y Psutka, 1999].

Para aplicar los filtros Mel, se calcula la energía de la señal en cada una de las bandas de frecuencias. Para esto, se suman los módulos al cuadrado de la DFT del fragmento que se encuentra contenido en cada una de las bandas de los filtros, como se muestra en la Figura 2.6.

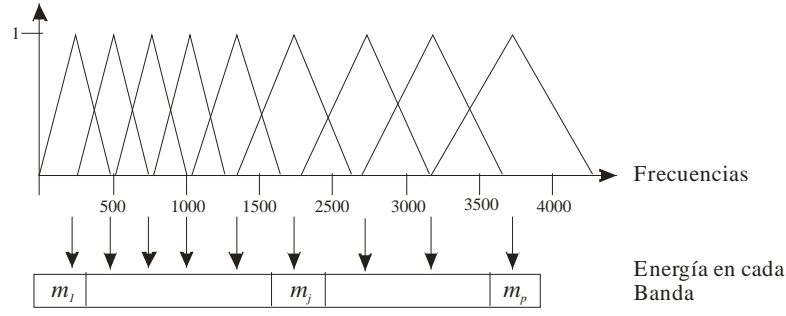


Figura 2.6. Banco de filtros en escala Mel [Zbyni y Psutka, 1999]

Posteriormente cada valor f_i obtenido del DFT es multiplicado por las ganancias correspondientes a los filtros relacionados respectivamente.

Aplicación de log

Se aplica el \log_{10} a los valores m obtenidos. Con esto se logra comprimir el rango dinámico de valores, similar a lo que hace el sistema auditivo, y también consigue una extracción de características menos sensitiva a variaciones de la entrada.

Aplicación de la transformada de coseno discreta

Para terminar, y obtener los coeficientes cepstrales, se tiene que aplicar la transformada de coseno discreta o DCT (por sus siglas en inglés, *discrete cosine transform*), que hace la función de transformada inversa de las energías logarítmicas obtenidas con anterioridad [Zbyni y Psutka, 1999]. El cálculo de los MFCC's se realiza aplicando la siguiente ecuación.

$$MFCC_j^i = \sum_{k=1}^L e(j, k) \cos \left[i \left(k - \frac{1}{2} \right) \frac{\pi}{L} \right] \quad (2.8)$$

Donde $i=1, 2, \dots, M$, k representa a la banda de frecuencias, j es la trama en curso, $e(j, k)$ es el logaritmo de la suma de los módulos al cuadrado de la DFT en la banda k de la trama j , L es el número de bandas o filtros, M es el número total de coeficientes MFCC.

2.5.3. Intensidad

La intensidad de un sonido es una medida de su poder sobre un área dada. Su fórmula es:

$$I = P/A \quad (2.9)$$

Donde I es la intensidad del sonido, P es la potencia acústica y A es el área donde se escucha. Por ejemplo, el sonido que genera una conversación normal tiene 10^{-10} Watts/cm² de intensidad [Smith S., 1999].

La intensidad del sonido se expresa en una escala logarítmica llamada decibel SPL (por sus siglas en inglés *Sound Power Level*) y se obtiene con la fórmula:

$$I_{dB} = 10 \log (I / I_0) \quad (2.10)$$

Donde I_{dB} es la intensidad en decibelios, I es la escala acústica en la escala lineal (Watts/cm²) y I_0 es el umbral de audición (10^{-16} Watts/cm²) [Smith S., 1999]. En la Tabla 2.1 se muestran los decibeles de algunos sonidos conocidos.

Tabla 2.1. Intensidad de algunos sonidos conocidos [Smith S., 1999]

Decibeles	Ejemplo de Sonido
10 Decibeles	La respiración de una persona
60 Decibeles	Conversación Normal
70 Decibeles	Límite para el ruido de una industria
80 Decibeles	El ruido ocasionado por una aspiradora
110 Decibeles	Concierto de Rock
120 Decibeles	Motor de avión en marcha (Incómodo)
140 Decibeles	Avión a reacción (El umbral de dolor)

2.5.4. Cocleograma

Un cocleograma representa la excitación de los filamentos del nervio auditivo de la membrana basilar, la cual se encuentra en la cóclea, en el oído interno. Esta excitación se representa como una función en el tiempo (en segundos) y frecuencia Bark.

Escala Bark

La escala Bark es una escala psicoacústica, es decir, que estudia la percepción subjetiva de las características del sonido. La escala Bark tiene un rango de 1 a 24 y corresponde a las primeras 24 bandas críticas del oído. En la Figura 2.7 se muestra la correspondencia de Hertz a Bark. Una de las características de la Escala de Bark es que tiene una escala de frecuencia altamente lineal para frecuencias menores a 1000 Hz, y una escala de frecuencia altamente logarítmica sobre los 1000 Hz.

El rango de frecuencia de percepción del ser humano es de 20 a 20,000 Hz que corresponden con un rango de frecuencia de 0.22 a 28.84 Bark. En el software de procesamiento de señales de habla PRAAT [Boersma P. y Weenink D. 2007], la frecuencia Bark correspondiente a una frecuencia en Hertz es calculada con la siguiente fórmula:

$$Bark = 7 \times \ln \left(\frac{f}{650} + \sqrt{1 + \left(\frac{f}{650} \right)^2} \right) \quad (2.11)$$

Donde f es la frecuencia en Hertz que se desea convertir a Bark.

Un cocleograma también modela el volumen del sonido y el enmascaramiento de las frecuencias. El enmascaramiento de frecuencias en el oído, se produce cuando se escuchan dos sonidos de diferente intensidad al mismo tiempo, el sonido más débil no se distinguirá ya que el cerebro solo procesa el sonido enmascarador, como se aprecia en la Figura 2.8.

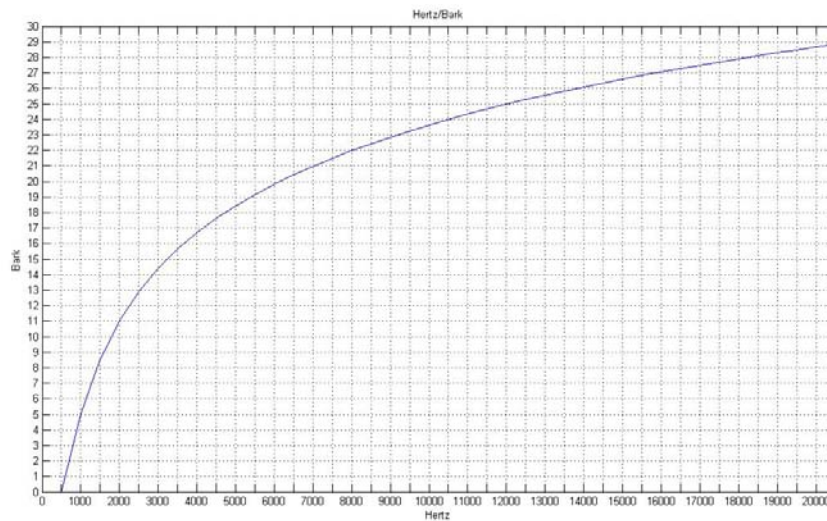


Figura 2.7. Representación gráfica de la relación entre Hertz y escala Bark, generada con la fórmula (2.11)

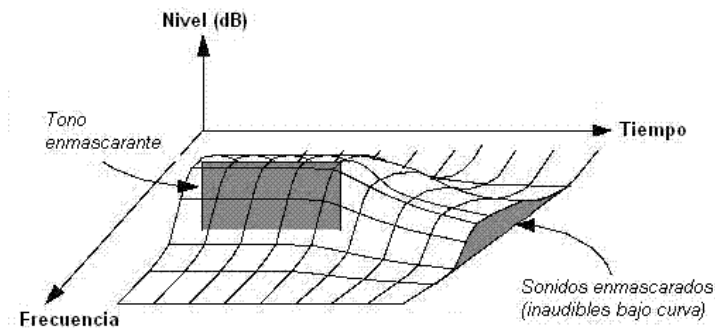


Figura 2.8. Enmascaramiento de sonidos [Wikipedia, 2008].

En la Figura 2.9 se muestra un cocleograma en su representación espectral para el llanto de infante con hiperbilirrubinemia, generado con el software PRAAT [Boersma y Weenink, 2007].

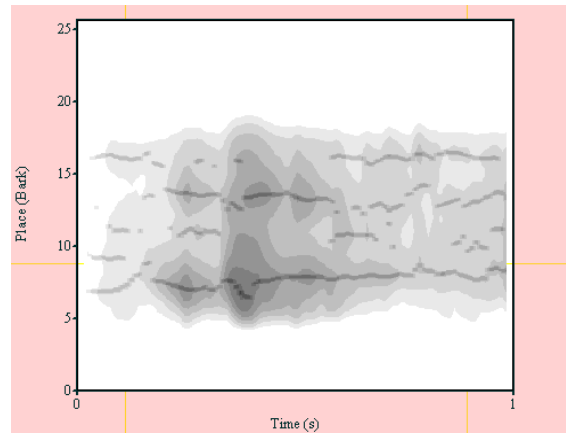


Figura 2.9. Representación espectral de un cocleograma para el llanto de infante con hiperbilirrubinemia.

2.6.Lógica Difusa

En 1965 Lofti A. Zadeh describió la teoría de la lógica difusa, donde presenta a los conjuntos difusos como aquellos a los cuales pertenece un elemento en cierto grado. El grado de pertenencia de un elemento x a un conjuntos difuso A es determinado por una función de membresía; estos grados de pertenencia son llamados *grados de membresía* o simplemente *membresía* y son valores en el intervalo $[0,1]$ [Zadeh, 1965]

Formalmente, un conjunto difuso A se define como una función de membresía que asigna a los elementos de un dominio o universo de discurso X un elemento en el intervalo $[0,1]$:

$$A: X \rightarrow [0,1] \quad (2.12)$$

Cuanto más cerca esté $A(x)$ del valor 1, mayor será la membresía del objeto x al conjunto A . Entonces los valores posibles de membresía varían entre 0 (no pertenece en absoluto) y 1 (membresía total). Este tipo de funciones de membresía se les conoce como tipo 1. Un ejemplo se muestra en la Figura 2.10 (a).

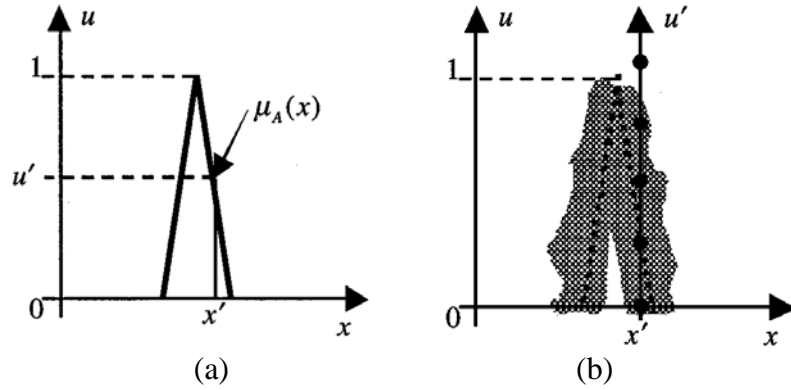


Figura 2.10 (a) Función de membresía tipo 1 (b) Función de membresía tipo 1 sombreada

[Mendel y Bob, 2002]

2.7. Conjuntos Difusos tipo 2

En 1975 se presentaron otro tipo de conjuntos difusos, los tipo 2, como unos conjuntos mas expresivos, debido a que pueden modelar mas información con respecto a la pertenencia de un punto a un conjunto.

Para describir estos conjuntos tomemos como base Figura 2.10 (a), que describe una función de membresía tipo 1, ahora imaginemos que se distorsiona la función de membresía, si movemos los puntos del triangulo hacia la izquierda y hacia la derecha de forma irregular, obtendríamos lo que se muestra en la Figura 2.10 (b).

Entonces para un valor específico x' no hay un valor específico de membresía u' , sino que existen varios valores u' (llamados membresía primaria). Sin embargo, estos valores no tienen la misma importancia, así que se asignan pesos para cada uno de ellos (membresía secundaria), por lo que para cada uno de los valores de x' , se generaría una función de membresía tridimensional. La Figura 2.11 muestra un ejemplo de las funciones de membresía de cada valor de x [Mendel y Bob, 2002].

Un conjunto difuso tipo 2, denotado como \tilde{A} se define de la siguiente forma:

$$\tilde{A} = \{((x, u), \mu_{\tilde{A}}(x, u)) | \forall x \in X, \forall u \in J_x \subseteq [0, 1]\} \quad (2.13)$$

\tilde{A} es caracterizado por una función de membresía tipo dos $\mu_{\tilde{A}}(x,u)$ donde $x \in X$ y $u' \in J_x$ que es el conjunto de membresías de x dentro de $[0,1]$, y $0 \leq \mu_{\tilde{A}}(x,u) \leq 1$.

\tilde{A} puede ser expresado utilizando el símbolo \int , como en la ecuación (2.14) donde \int denota unión sobre todos los x y sus u' calculados, recordemos que la unión de dos conjuntos A y B es por definición otro conjunto que contiene los elementos que se encuentran en A o en B , y el símbolo $/$ no es una división, se utiliza para separar la membresía secundaria de la primaria: [membresía secundaria]/[membresía primaria].

Si viéramos a cada elemento de un conjunto difuso tipo 2 como un subconjunto, entonces las uniones en (2.14) cumplen con la definición de unión, ya que cada elemento de ese conjunto es distinto. Para un valor específico de x y u solo se hace una unión [Mendel y Bob, 2002].

$$\tilde{A} = \int_{x \in X} \int_{u \in J_x} \mu_{\tilde{A}}(x,u)/(x,u) \quad J_x \subseteq [0,1] \quad (2.14)$$

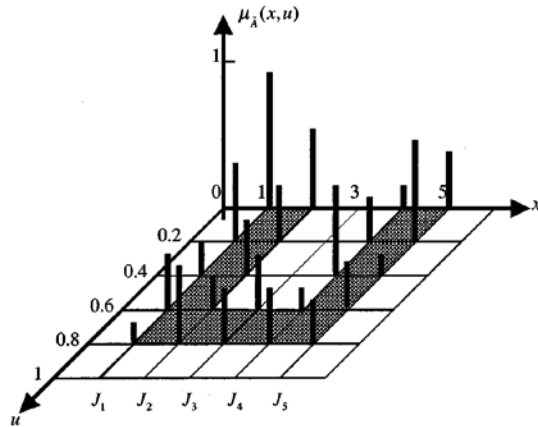


Figura 2.11. Ejemplo de una función de membresía tipo 2 [Mendel y Bob, 2002].

Para explicar como se obtiene $\mu_{\tilde{A}}(x,u)$ analicemos el ejemplo de la Figura 2.11, que describe una función de membresía tipo 2. Los valores de x y u son discretos, en particular, $X = \{1,2,3,4,5\}$ y $U = \{0,0.2,0.4,0.6,0.8\}$. Para cada valor de x , digamos

$x=x'$, el eje del segundo plano representa los valores de u y $\mu_{\tilde{A}}(x',u)$ está representado por la parte vertical. Una función de membresía secundaria es la parte vertical de $\mu_{\tilde{A}}(x,u)$. Esto es $\mu_{\tilde{A}}(x=x',u)$ para $x \in X$ y $\forall u \in J_{x'} \subseteq [0,1]$ es decir:

$$\mu_{\tilde{A}}(x=x',u) \equiv \mu_{\tilde{A}}(x') = \int_{u \in J_{x'}} f_{x'}(u)/u \quad J_{x'} \subseteq [0,1] \quad (2.15)$$

donde $0 \leq f_{x'}(u) \leq 1$ y \int representa la operación unión [Mendel y Bob, 2002].

Continuando con el ejemplo, la función de membresía tipo 2 que describe la Figura 2.11, para $x=2$, tiene la función de membresía:

$$\mu_{\tilde{A}}(2) = 0.5/0 + 0.35/0.2 + 0.35/0.4 + 0.2/0.6 + 0.5/0.8.$$

donde el símbolo $+$ no significa una suma aritmética, solo es la unión de las membresías, cada membresía primaria tiene asignada una membresía secundaria. Nótese que la obtención de cada valor de $\mu_{\tilde{A}}(x')$ es a través de funciones de membresía primaria y secundaria, que en el caso de éste ejemplo, estas funciones se representan con la figura 2.11.

Ya teniendo las membresías de un determinado punto, se puede hacer una operación para obtener la membresía total del elemento al conjunto, esto aplicando alguna operación como por ejemplo promedio ponderado, esta elección es dependiente de la aplicación de los conjuntos difusos tipo 2.

Capítulo 3

PROCESO DE CLASIFICACIÓN DE LLANTO DE INFANTES

3.1. Diagrama de Bloques del Clasificador

La Figura 3.1 muestra el diagrama completo del proceso de clasificación de llanto de infantes. El bloque llamado comparación difusa de patrones es el correspondiente al algoritmo de clasificación propuesto en esta tesis, el cual se describe a detalle más adelante en este capítulo.

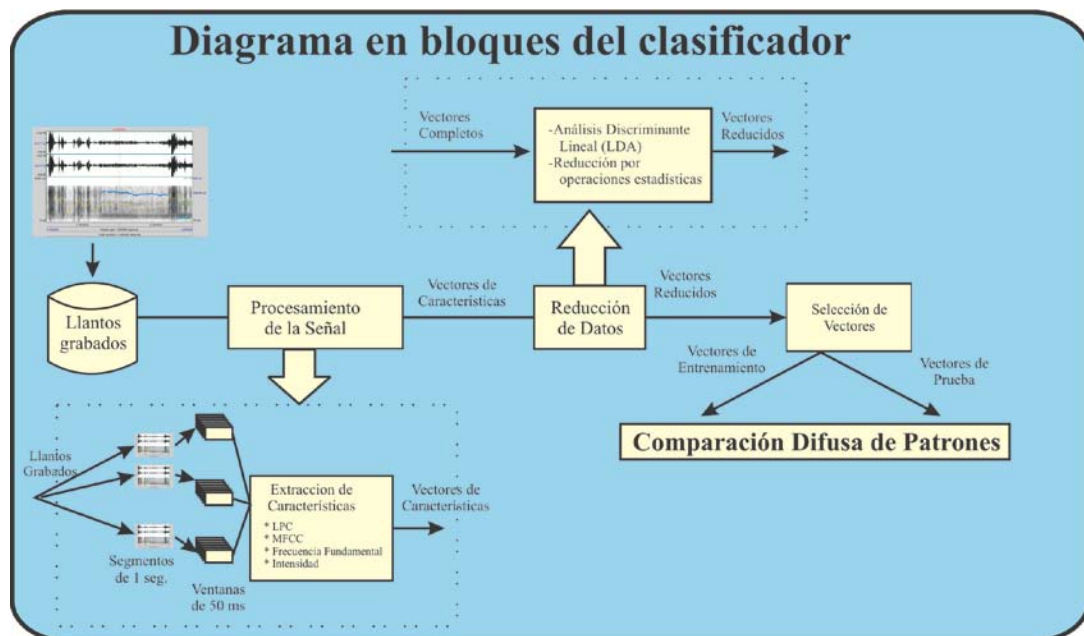


Figura 3.1 Diagrama en Bloques de un Clasificador de llanto

El proceso de clasificación de llanto de infante se inicia con el procesamiento de la señal, a fin de obtener las características de los llantos y formar los vectores de datos de cada clase. Posteriormente puede ejecutarse el módulo de reducción de datos, en el cual se utilizan métodos como “Análisis Discriminante lineal” ó “Reducción por Operaciones Estadísticas”. En el módulo de selección de vectores, se dividen las muestras para entrenamiento y clasificación. Finalmente, se ejecuta el proceso de entrenamiento del clasificador a través de la comparación difusa de patrones. Una vez entrenado el clasificador puede utilizarse para reconocer señales de llanto.

3.1.1. Base de Llantos Grabados

Las muestras de llanto utilizadas en este trabajo fueron recolectadas en el Instituto Nacional de la Comunicación Humana (INCH), en la ciudad de México y en el Instituto Mexicano del Seguro Social (IMSS) en Puebla. Se utilizaron muestras de infantes sanos y con padecimientos de asfixia e hiperbilirrubinemia. En la Tabla 3.1 se describe los tipos de muestras seleccionadas para esta tesis.

Tabla 3.1. Segmentos de las grabaciones de llanto de infante.

Tipo	Segmentos de 1 seg.	Observaciones
Hiperbilirrubinemia	418 segmentos	
Asfixia	340 segmentos	
Normales	400 segmentos	Contiene muestras de llanto originados por hambre, dolor, baño e incomodidad.
Total de segmentos	1,158 segmentos	

Cabe mencionar que de la clase normales se tienen 1,136 segmentos de 1 segundo, sin embargo, se seleccionaron solo 400 para no desbalancear las clases. Los segmentos de 1 segundo de cada clase se extrajeron de los archivos grabados de llanto de distintos infantes a quienes se les detectaron las patologías de hiperbilirrubinemia y asfixia y de infantes sin estas patologías. Estos archivos se

procesaron anteriormente limpiándolos de espacios en silencio y normalizándolos a la frecuencia de 8,000 Hertz debido a que la mayoría de los archivos tenían esa frecuencia.

3.1.2. Extracción de características

El módulo de extracción se realizó utilizando el software PRAAT [Boersma y Weenink, 2007]. Cada segmento de 1 segundo se dividió en ventanas de 50 ms. Este tamaño de ventana es mayor al usado comúnmente en el reconocimiento de habla, el cual es de 40 ms, debido a que en el llanto de infante las variaciones de características son más lentas, en comparación con las del habla. Para la extracción de las características, se utilizaron los parámetros que describe la Tabla 3.2.

Tabla 3.2. Parámetros utilizados para la extracción de características

Característica	Parámetros de PRAAT	Valores
MFCC	No. Coeficientes	16
	Tamaño de la ventana	0.05 seg.
	No. de filtros Mel	100
	Salto en tiempo	0.05 seg.
LPC	No. Coeficientes	16
	Tamaño de la ventana	0.05 seg.
Intensidad	Pitch mínimo	100 Hz
	Tamaño de la ventana	0.05 seg
	Salto en tiempo	0.05 seg
Cocleograma	Salto en tiempo	0.1
	Resolución Frecuencia Bark	0.5

El *salto en tiempo* es el tiempo que va a avanzar el algoritmo para obtener la siguiente muestra. Por ejemplo, en el caso de intensidad se definió la ventana en 0.05 seg. así que el algoritmo se inicia en el segundo 0 y extrae la primera ventana de 0.05 seg. a la cual se calcula su intensidad, posteriormente avanza lo correspondiente al *salto en*

tiempo, en este caso es 0.05 seg., y toma la segunda ventana de 0.05 seg. para obtener su intensidad.

Para obtener los coeficientes MFCC y LPC con el software PRAAT, definiendo el *salto en tiempo* en 0.05 seg. y el tamaño de la ventana en 0.05 seg., se obtienen 16 coeficientes por cada ventana tomada de la muestra del tamaño definido en 0.05 seg. Entonces el número de coeficientes por cada segmento de 1 segundo es $19 \times 16 = 304$ coeficientes MFCC o LPC para cada muestra. En [Orozco, 2002] se demuestra que para el llanto de infantes se obtienen mejores resultados utilizando 16 coeficientes en MFCC y LPC, por esta razón se decidió utilizar 16 coeficientes.

En el caso de la intensidad, se definió el mismo *salto en tiempo* y se obtiene un valor de intensidad por cada segmento de 0.05 seg. Así, el número de valores de intensidad para un segmento de un segundo es de 19. Este valor va ligado al tamaño de ventana y se optó por definir el *salto en tiempo* igual al tamaño de la ventana para que el salto fuera igual y no existiera traslape entre ventanas.

Para obtener los parámetros de excitación de cocleogramas utilizando PRAAT se tiene que definir un valor de *salto en tiempo* y de resolución de Frecuencia Bark. El procedimiento es: se divide el segmento entre el valor del *salto en tiempo* (1 segundo / *salto en tiempo*), se toma el valor de la frecuencia más alta en Bark que maneja el software PRAAT (25.6 Bark), y se divide entre el valor de la resolución de la frecuencia que se asigna. Para el *salto en tiempo* y para la resolución de frecuencia Bark, se optó por asignarles los valores 0.1 y 0.5 respectivamente, de esta forma se tienen $25.6 \text{ Bark} / 0.5$ que redondeando es igual a 51 bandas de frecuencia. El número de pasos es $1 \text{ segundo} / 0.1 = 10$. Entonces por cada segundo se obtienen $51 \times 10 = 510$ parámetros que modelan la excitación de la membrana bacilar. PRAAT maneja una resolución por omisión de 0.01 con lo que se obtendrían 25,600 parámetros, que resultan ser vectores demasiado grandes, por lo tanto se optó por utilizar la resolución de 0.5 para obtener 510 parámetros, que de acuerdo a las pruebas realizadas en la

sección 4.3.1 Resultados usando todas las características, mantienen su capacidad descriptiva.

3.1.3. Reducción de Datos

En la Tabla 3.3 se resume el número de atributos obtenidos por cada característica.

Tabla 3.3. Atributos Obtenidos por característica

Característica	No. de atributos
LPC	304
MFCC	304
Intensidad	19
Cocleograma	510

Para probar el clasificador propuesto en esta tesis se utilizan las matrices de datos sin reducir, así como matrices generadas utilizando dos métodos de reducción: “Análisis Discriminante Lineal” [Duchene y Leclerq, 1988] y “Reducción por Operaciones Estadísticas” [Amaro, 2008].

3.1.3.1. Análisis Discriminante Lineal

El método del Análisis Discriminante Lineal (LDA por sus siglas en inglés *Linear Discriminand Analysis*) es utilizado para la clasificación y reducción de datos. Este método se utiliza en los casos donde los elementos de una misma clase son distintos entre ellos y su comportamiento es similar al de un generador de datos aleatorio.

El objetivo de este método es encontrar el mejor conjunto de vectores discriminantes para separar las clases. El método maximiza el radio de la varianza entre-las-clases y la varianza dentro-de-las-clases en cualquier conjunto de datos particular de tal modo que garantiza la separación máxima. El algoritmo LDA utilizado para la reducción de datos es el propuesto por Duchene y Leclerq [Duchene y Leclerq, 1988].

En la Figura 3.2 se observa la representación gráfica del cocleograma obtenido para una muestra de la clase hiperbilirrubinemia sin reducir y en la Figura 3.3 su reducción con LDA. Cabe mencionar que el método LDA toma en cuenta la participación de cada elemento dentro de las distintas clases.

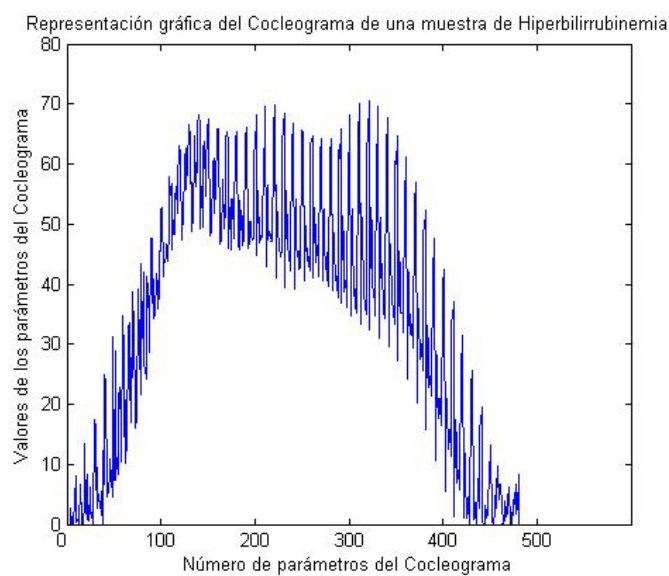


Figura 3.2. Gráfica del Cocleograma para una muestra de la clase Hiperbilirrubinemia sin reducir.

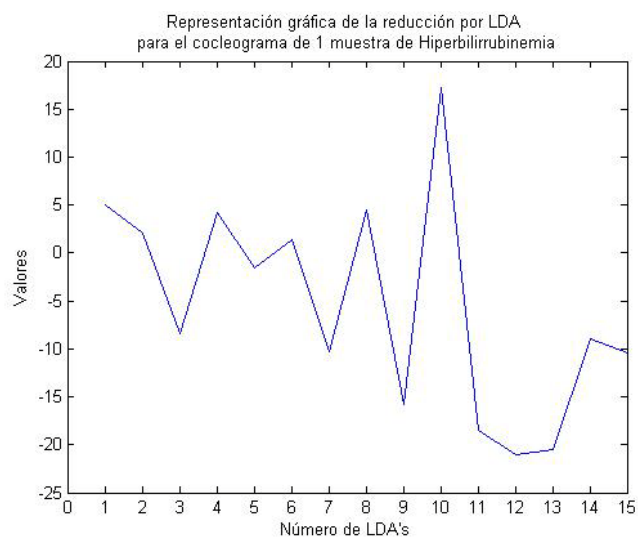


Figura 3.3. Cocleograma de una muestra de la clase Hiperbilirrubinemia reducido con el método LDA

3.1.3.2. Reducción por Operaciones Estadísticas

Este método de reducción de datos fue propuesto y utilizado en [Amaro, 2008], y consiste en obtener cinco parámetros del vector de características a reducir, los cuales son: valor mínimo, valor máximo, media, desviación estándar y varianza. Este método se aplica a los datos de cada clase de manera local sin considerar la clase a la que pertenece. En la Tabla 3.4 se muestra la reducción por operaciones estadísticas para los cocleogramas de la muestra de hiperbilirrubinemia de la Figura 3.2.

Tabla 3.4. Valores obtenidos con la reducción por operaciones estadísticas para el cocleograma obtenido de una muestra de hiperbilirrubinemia.

Mínimo	Máximo	Media	Desviación Estándar	Varianza
0.0001	70.5325	33.7771	22.6483	512.9436

En las figuras Figura 3.4 y Figura 3.5 se muestra la representación gráfica de los cocleogramas obtenidos de las 418 muestras de Hiperbilirrubinemia y de las 340 muestras de Asfixia, cuyos valores fueron reducidos por operaciones estadísticas.

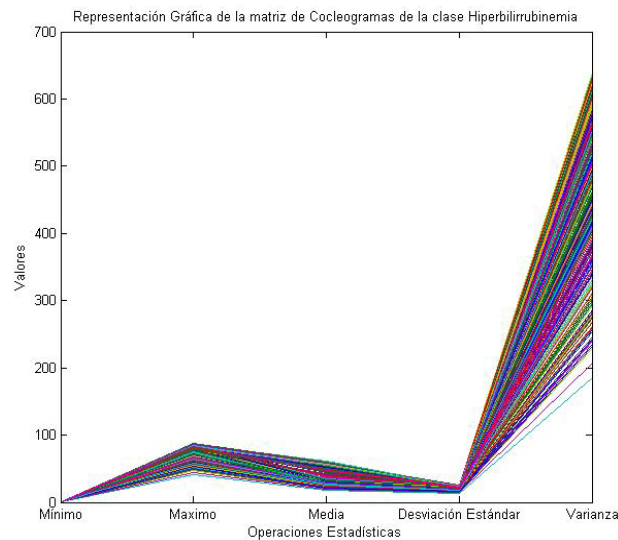


Figura 3.4. Gráfica de los Cocleogramas de la clase Hiperbilirrubinemia reducidos por el método de operaciones estadísticas.

Puede notarse visualmente que las representaciones de las clases son muy similares, debido a que las clases son sobrepuestas, los datos obtenidos con las operaciones no tienen mucha diferencia entre las dos clases.

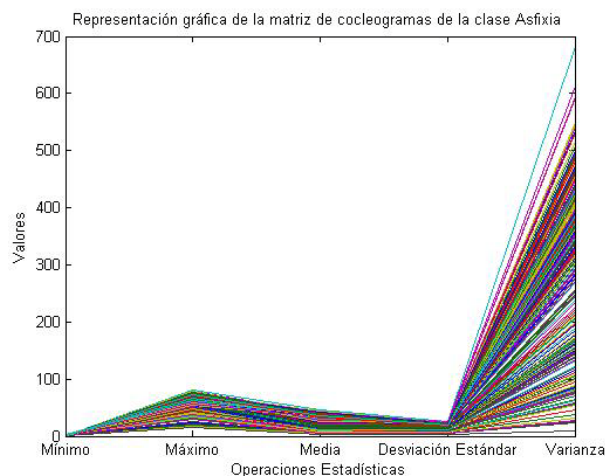


Figura 3.5. Gráfica de los Cocleogramas de la clase Asfixia reducidos por el método de operaciones estadísticas.

3.1.4. Clasificación

Teniendo los vectores característicos de cada clase, ya sea reducidos o sin reducir, se procede a la clasificación del llanto de infantes. Para esto utilizamos el método de comparación difusa de patrones modificado con conjuntos difusos tipo 2 (CDP-T2). Antes de explicar el método CDP-T2, explicaremos el método de comparación difusa de patrones utilizado como base para el presente trabajo [Sayed, 2002], descrito en la Figura 3.6.

- **Comparación Difusa de Patrones**

La comparación difusa de patrones toma en cuenta la representación semántica asociada a los componentes de los datos con lo que se genera una base de conocimiento de cada clase y con la cual se puede realizar la evaluación, en el intervalo $[0,1]$, de la pertenencia de un elemento a cada clase [Saïd, Sayed y Billaudel, 2007]. El método de comparación difusa de patrones es un método de

clasificación supervisada basada en el uso de histogramas de probabilidad e involucra dos fases:

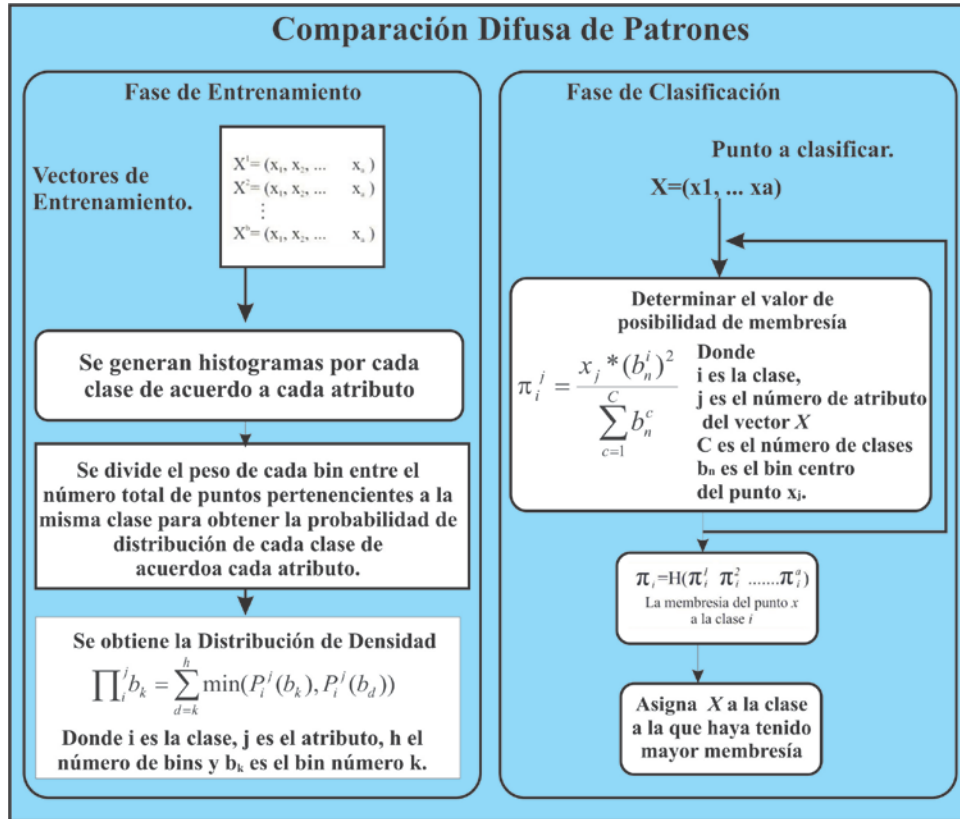


Figura 3.6. Algoritmo de Comparación Difusa de Patrones [Saïd, Sayed y Billaudel, 2007]

I. Fase de Aprendizaje.

Dadas las clases C_1, C_2, \dots, C_c , descritas por atributos, se construyen los histogramas de datos para cada clase de acuerdo a cada atributo. El histograma se compone por cierto número de bins³, el número de bins para un histograma, denotado por h , se determina de forma experimental.

³ El número de bins en un histograma es el número de intervalos en que se divide un histograma. Un bin es un intervalo.

El peso de cada bin b_k es el número de puntos de aprendizaje localizados en ese intervalo. La probabilidad de distribución P_i^j de la clase C_i de acuerdo al atributo j se calcula dividiendo el peso de cada bin b_k entre el número total de muestras de aprendizaje pertenecientes a la misma clase N_i . Por último se obtiene la distribución de densidad utilizando la fórmula:

$$\prod_i^j b_k = \sum_{d=k}^h \min(P_i^j(b_k), P_i^j(b_d)) \quad (3.1)$$

II. Fase de Clasificación

Dado un punto $X(x_1, \dots, x_n)$ se determina el valor de posibilidad de membresía para cada $x_i \in X$ utilizando la fórmula:

$$\pi_i^j = \frac{x_j * (b_n^i)^2}{\sum_{c=1}^C b_n^c} \quad (3.2)$$

Donde b_n es el bin que corresponde al atributo j . Se obtiene la membresía total de X a cada clase C_c :

$$\pi(X^c) = H(\pi_1^c, \dots, \pi_n^c)$$

donde H puede ser una operación difusa o el promedio de las membresías de X para la clase c , en [Sayed, 2002] utilizaron la operación promedio.

Por último se asigna X a la clase donde su promedio de membresías fue el mayor.

- **Algoritmo de Comparación Difusa de Patrones con conjuntos difusos tipo 2.**

El algoritmo propuesto se basa en la Comparación Difusa de Patrones y se complementa con la utilización de conjuntos difusos tipo 2. Se divide en dos fases, la fase de aprendizaje, donde se recopila la información de las clases y se calculan las membresías primaria y secundaria, posteriormente en la fase de clasificación, se calcula la membresía de cada elemento a cada una de las clases, y se decide a cual pertenece. El algoritmo es el siguiente:

Fase de Aprendizaje

En la fase de aprendizaje se reciben los vectores que contienen los datos de las características de cada clase, y se forma una matriz de entrenamiento para cada clase, denotada por $MatrizEntrenamiento_c$ donde $c \in (c_1, c_2, \dots, c_m)$ donde m es el número de clases. Se obtiene una matriz llamada $HisT2_c$ por cada clase, la cual es un histograma que concentra información de la matriz de entrenamiento de la clase c y se obtiene de la siguiente manera:

Se determina el valor mayor y el valor menor que se tienen dentro de las matrices de entrenamiento de las clases para las que se va a entrenar el clasificador. Se calcula un vector b , cuyo primer valor es el dato menor y el último valor es el dato mayor y los valores intermedios son los datos obtenidos de calcular intervalos iguales entre el dato menor y el dato mayor. El número de intervalos se determina de manera experimental y representan los bins de un histograma. Posteriormente se determina el valor d , que es el número de intervalos en que se dividirá cada intervalo dado en b con lo que se construye una matriz llamada *MatrizGuía* la cual contiene los valores determinados por los intervalos y sirve como guía para determinar el bin centro de un punto x .

Para cada atributo $x \in X=(x_1, x_2, \dots, x_n)$ que se encuentra en la matriz de entrenamiento de la clase c :

Se obtienen los índices ij tal que MatrizGuía_{ij} es el valor más cercano a x y se actualiza el HisT2_{ij}^c

$$\text{HisT2}_{ij}^c = \text{HisT2}_{ij}^c + 1 \quad (3.3)$$

Posteriormente se calculan las membresías primarias y secundarias correspondientes a los atributos que se contaron en cada HisT2_{ij}^c de la siguiente forma:

$$MsP_{ij}^c = \frac{\text{HisT2}_{ij}^c}{\text{HisT2}_{\max}^c} \quad (3.4)$$

Donde MsP_{ij}^c es la membresía primaria del elemento HisT2_{ij}^c , c es la clase a la que pertenece y HisT2_{\max}^c es el valor máximo contenido en el HisT2 de la clase c . Esta membresía primaria denota el grado de pertenencia de un elemento x a la clase c .

La membresía secundaria se calcula:

$$MsS_{ij}^c = \frac{\text{HisT2}_{ij}^c}{\sum_{c=1}^C \text{HisT2}_{ij}^c} \quad (3.5)$$

Donde MsS_{ij}^c es la membresía secundaria del elemento HisT2_{ij}^c , este valor se divide entre la suma de los valores en HisT2_{ij} de cada clase. Esta membresía secundaria denota la importancia que tiene un elemento x a la clase c tomando en cuenta las otras clases. Otro punto a tomar en cuenta es que la sumatoria de las membresías secundarias de cada clase de ij es igual a 1. Es decir.

$$\sum_{c=1}^C MsS_{ij}^c = 1$$

Por ejemplo si un elemento x tiene las membresías 0.3/0.8 para la clase c (membresía secundaria / membresía primaria) se puede decir que x es importante para la clase c , pero para las otras clases también es importante, por lo tanto es uno de los elementos sobrepuestos y no nos proporcionará información suficiente para la clasificación. Por

otro lado si el elemento x tuviera las membresías 0.8 / 0.6 para la clase c , se puede decir que este elemento tiene mayor importancia para la clase c que para otras clases.

Fase de Clasificación

En la fase de clasificación se recibe un elemento $X=(x_1, x_2, \dots, x_n)$ del cual se obtiene su membresía a cada clase para ello se obtiene los índices ij tal que $MatrizGuía_{ij}$ es el valor más cercano a x_k , donde $x_k \in X$, y si $MsS^c_{ij} > P_M$ se calcula la membresía con la ecuación (8) de lo contrario se calcula con la ecuación (9). P_M es un parámetro que se puede modificar de forma experimental, y puede tomar valores de $[0,1]$, y que sirve para evaluar la membresía secundaria del elemento x_k , debido a que dependiendo del valor de la membresía secundaria podemos saber si el elemento es importante para la clase, entonces si la membresía secundaria es mayor a P_M se maximiza su membresía final de lo contrario se minimiza.

$$\pi^c(x_k) = MsP^c_{ij} + MsS^c_{ij} \quad (3.6)$$

$$\pi^c(x_k) = (MsP^c_{ij} + MsS^c_{ij}) / 2 \quad (3.7)$$

donde $\pi(x^c_k)$ representa la membresía del punto x_i a la clase c , si la membresía secundaria que se tiene para el punto ij en la clase c es mayor a P_M significa que ese punto es importante para esa clase así que se suman las membresías primaria y secundaria, de lo contrario, el punto no es tan importante para la clase c así que se divide la suma de las membresías primaria y secundaria entre dos.

Posteriormente se obtiene la membresía total de X para la clase c

$$\pi^c(X) = H[\pi(x^c_1), \pi(x^c_2) \dots \pi(x^c_n)] \quad (3.8)$$

donde $\pi^c(X)$ es la membresía de la muestra X a la clase c y la función H es el promedio de las membresías obtenidas por cada punto x_k .

Se asigna el elemento X a la clase a la que haya obtenido mayor membresía. Es importante hacer mención que para las pruebas realizadas en este trabajo se utilizaron

los siguientes parámetros: $b=20$ y $d=10$, es decir, que los *HisT2* generados en cada prueba fueron del tamaño 20×10 al igual que las matrices de las membresías primaria y secundaria. Nótese que sin importar el tamaño de las muestras ni de los atributos, cada clase se representaba las matrices del mismo tamaño, lo que hace a la fase de clasificación ser un algoritmo simple y rápido.

Capítulo 4

EXPERIMENTACIÓN Y RESULTADOS

4.1. Descripción de las pruebas realizadas

Para probar el clasificador propuesto en esta tesis, se realizaron pruebas de clasificación a dos conjuntos de datos. El primero es un conjunto de datos de flores iris [Fisher, 1936], comúnmente utilizado para realizar pruebas de clasificadores, el cual se usó para validar experimentalmente la implementación del algoritmo. El segundo conjunto utilizado contiene los llantos de infantes con hiperbilirrubinemia, asfixia y normales. Para cada conjunto de datos se realizaron las pruebas basándonos en el método de validación cruzada *10-folds*, que se describe a continuación.

4.1.1. Descripción del método de validación cruzada *10-folds*

Para realizar las pruebas del clasificador se utilizó el método de validación cruzada *10-folds* que consiste básicamente en dividir el conjunto total de las muestras en 10 subconjuntos y probar con cada uno. Se toman 9 subconjuntos como entrenamiento y un subconjunto como prueba, se guarda el resultado de la evaluación de la prueba y se vuelve a repetir el procedimiento tomando otro subconjunto para prueba y entrenando con el resto. Este proceso se repite 10 veces reportándose resultados para cada prueba la media de todo el experimento.

4.1.2. Métricas de evaluación

Para evaluar el clasificador por cada corrida obtenida con el método de validación cruzada *10-folds* se obtuvo su matriz de confusión, la precisión y el recuerdo del clasificador para cada clase.

Una matriz de confusión M de tamaño $n \times n$, donde n es el número de clases, es una herramienta de visualización donde cada columna de la matriz representa el número de veces que se clasificó un elemento de prueba en cada clase y cada fila representa las instancias en que se clasificaron en la clase real. Una ventaja de la matriz de confusión es que se puede ver si el clasificador está confundiendo dos clases.

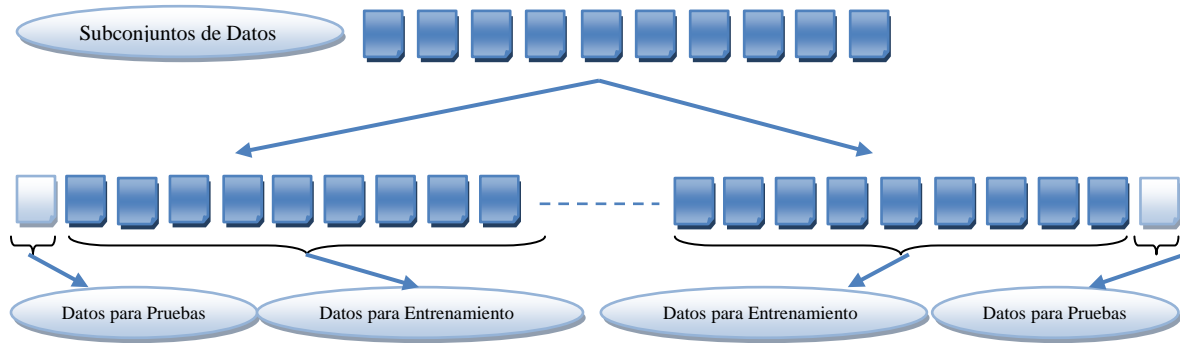


Figura 4.1. Algoritmo validación cruzada 10-folds.

En la Tabla 4.1 se muestra un ejemplo de la matriz de confusión de una corrida de la clasificación utilizando el conjunto de datos de llanto de infantes. Se tienen tres clases: normal, asfixia e hiperbilirrubinemia. La primera fila de datos muestra que de los 30 elementos de la clase normal que se usaron para la prueba de clasificación, 27 fueron clasificados correctamente (en la clase normal) y 3 fueron clasificados erróneamente, dos en la clase asfixia y 1 en la clase hiperbilirrubinemia.

La precisión de un clasificador para la clase c es el porcentaje de elementos clasificados de forma correcta en la clase c . En [Montes y Villaseñor, 2007] la definen como la probabilidad de que si una instancia x es clasificada en la clase c , la instancia realmente pertenezca a esa clase. La fórmula es la siguiente:

$$precisión_c = \left(\frac{E_c}{E_t} \right) * 100 \quad (4.1)$$

donde E_c es el total de elementos clasificados en la clase c y E_t es el total de elementos que pertenecían a la clase c .

El recuerdo con respecto a una clase se define como la probabilidad de que si una instancia x pertenece a una clase c , el clasificador la clasifica correctamente. Se calcula de la siguiente manera [Montes y Villaseñor, 2007]:

$$recuerdo_c = \left(\frac{E_c}{E_{tc}} \right) * 100 \quad (4.2)$$

donde E_c es el total de elementos clasificados en la clase c y E_{tc} es el total de elementos que se clasificaron en la clase c , incluyendo los que pertenecían a otras clases.

Para el ejemplo mostrado en la Tabla 4.1, la precisión de cada clase se calcula como:

$$precisión_{normal} = \left(\frac{27}{27 + 2 + 1} \right) * 100 = 90\%$$

$$precisión_{asfixia} = \left(\frac{34}{0 + 34 + 0} \right) * 100 = 100\%$$

$$precisión_{hiperbilirrubinemia} = \left(\frac{42}{0 + 0 + 42} \right) * 100 = 100\%$$

La última columna de la Tabla 4.1 es el recuerdo, el cual es otro método de evaluación de un clasificador. El recuerdo para las tres clases del ejemplo de la Tabla 4.1 se calcula de la siguiente manera:

$$recuerdo_{normal} = \left(\frac{27}{27 + 0 + 0} \right) * 100 = 100\%$$

$$recuerdo_{asfixia} = \left(\frac{34}{2 + 34 + 0} \right) * 100 = 94.44\%$$

$$recuerdo_{hiperbilirrubinemia} = \left(\frac{42}{1 + 0 + 42} \right) * 100 = 97.67\%$$

Tabla 4.1. Ejemplo de una matriz de confusión con precisión y recuerdo..

	Normal	Asfixia	Hiperbilirrubinemia	Precisión	Recuerdo
Normal	27	2	1	90.00%	100.00%
Asfixia	0	34	0	100.00%	94.44%
Hiperbilirrubinemia	0	0	42	100.00%	97.67%

En todas las pruebas que se realizaron se calcularon las matrices de confusión, su precisión y recuerdo correspondientes. En este capítulo se presentan los resultados principales y en el Apéndice A se detallan los resultados de todas las pruebas.

4.1. Pruebas de Funcionalidad

Se realizó una prueba de funcionalidad del algoritmo implementado utilizando el conjunto de datos de las flores iris, recopilado por Sir Ronald Aylmer Fisher y que se acostumbra utilizar como medio de comparación de eficacia entre clasificadores. La base de datos consta de los datos de 50 muestras de cada una de las tres especies de flores iris (versicolor, virginica y setosa) de las cuales se obtuvieron cuatro características: el ancho y largo del sépalo y pétalo de las flores. El conjunto de datos se obtuvo de [Fisher, 1936].

En la Tabla 4.2 se muestra el resumen de los resultados obtenidos de las pruebas de funcionalidad con el conjunto de datos de las flores Iris. En esta tabla podemos observar que en cuatro corridas se obtuvo la media de precisión del **100 %**, en dos corridas se obtuvo la media de precisión del **87.67 %** y en cuatro corridas la media de precisión del **93%**.

Tabla 4.2. Pruebas con los datos Fisher Iris

Corrida (validación cruzada 10-folds)	Precisión			Media
	Iris-setosa	Iris-versicolor	Iris-virginica	
1	100.00%	80.00%	80.00%	86.67%
2	100.00%	100.00%	100.00%	100.00%
3	100.00%	80.00%	100.00%	93.33%
4	100.00%	80.00%	80.00%	86.67%
5	100.00%	100.00%	100.00%	100.00%
6	100.00%	80.00%	100.00%	93.33%
7	100.00%	80.00%	100.00%	93.33%
8	100.00%	100.00%	100.00%	100.00%
9	100.00%	100.00%	100.00%	100.00%
10	100.00%	80.00%	100.00%	93.33%
Promedios Totales	100.00%	88.00%	96.00%	94.67%

En la Tabla 4.3 se muestra el recuerdo del clasificador. Se puede observar que el promedio total del recuerdo es de **95.22%** lo que significa que la confusión entre las clases fue pequeña.

Tabla 4.3. Pruebas con los datos Fisher Iris

Corrida (validación cruzada 10-folds)	Recuerdo			Media
	Iris-setosa	Iris-versicolor	Iris-virginica	
1	83.33%	100.00%	80.00%	87.78%
2	100.00%	100.00%	100.00%	100.00%
3	83.33%	100.00%	100.00%	94.44%
4	100.00%	80.00%	80.00%	86.67%
5	100.00%	100.00%	100.00%	100.00%
6	100.00%	100.00%	83.33%	94.44%
7	83.33%	100.00%	100.00%	94.44%
8	100.00%	100.00%	100.00%	100.00%
9	100.00%	100.00%	100.00%	100.00%
10	100.00%	100.00%	83.33%	94.44%
Promedios Totales	95.00%	98.00%	92.67%	95.22%

4.2.Pruebas con tres clases de llanto

4.2.1. Resultados usando todas las características

El reconocedor fue probado con los vectores de características combinando las cuatro características. Por cada clase se tiene el siguiente número de muestras: 400 para la clase normal, 340 para la clase asfixia y 418 para la clase hiperbilirrubinemia. Para cada muestra se obtuvieron 4 vectores de características: LPC, MFCC, Intensidad y Cocleogramas de tamaño 304, 304, 19 y 510 respectivamente.

Según las características que se utilizaron para cada prueba, se formaron vectores por cada clase, en el caso de las pruebas con la combinación de las cuatro características por cada muestra se tienen $304+304+19+510=1137$ atributos. Para la clase hiperbilirrubinemia se tiene una matriz de 418×1137 ya que son 418 muestras por 1137 atributos, para la clase normal la matriz es de 400×1137 y para la clase asfixia el tamaño de la matriz es de 340×1137 .

Para aplicar el método validación cruzada 10-folds se dividieron las matrices, para el caso de la clase normal se dividió la matriz en 10 subconjuntos de 40 muestras, para la clase asfixia se obtuvieron 10 subconjuntos de 34 muestras y para la clase hiperbilirrubinemia se obtuvieron 9 subconjuntos de 42 muestras y uno de 41 muestras.

Los resultados de las pruebas se muestran en la tabla 4.4 , donde se obtuvo una precisión media del 89.85% para el total de la prueba, sin embargo, también podemos observar que para la clase hiperbilirrubinemia se obtuvo mayor porcentaje de precisión 99.29 % en promedio de las 10 corridas, mientras que para la clase normal su promedio fue del 79.67% en promedio de las 10 corridas, para la clase Asfixia se obtuvo en promedio 90.59%. Esto nos indica que la clase que mejor se clasifica es la de Hiperbilirrubinemia.

Tabla 4.4. Pruebas con datos obtenidos de las muestras de Llanto de infante, sin reducir.

Características: Cocleogramas, LPC, MFCC e Intensidad				
Corrida (validación cruzada 10- folds)	Precisión			Media
	Normal	Asfixia	Hiperbilirrubinemia	
1	90.00%	100.00%	100.00%	96.67%
2	53.33%	100.00%	97.62%	83.65%
3	100.00%	76.47%	100.00%	92.16%
4	100.00%	91.18%	100.00%	97.06%
5	96.67%	100.00%	95.24%	97.30%
6	76.67%	100.00%	100.00%	92.22%
7	83.33%	100.00%	100.00%	94.44%
8	53.33%	64.71%	100.00%	72.68%
9	70.00%	88.24%	100.00%	86.08%
10	73.33%	85.29%	100.00%	86.21%
Promedios Totales	79.67%	90.59%	99.29%	89.85%

La media de precisión más alta por corrida fue la número 5 con **97.30 %** donde la precisión por clase fue **96.67%, 100.00% y 95.24%** para las clases Normal, Asfixia e Hiperbilirrubinemia, respectivamente.

La tabla 4.5 muestra el recuerdo obtenido en cada clase y la media total del recuerdo que es de 93.92%, nótese que la clase normal tiene como recuerdo el 100% siendo que su precisión fue la más baja para estas pruebas, de 79.67%, y el recuerdo de la clase hiperbilirrubinemia fue el más bajo, de 86.49%, mientras que el recuerdo de la clase normal fue del 100%, esto significa que ningún elemento fue erróneamente clasificado como normal y que existieron elementos que se clasificaron erróneamente en la clase hiperbilirrubinemia así que existió confusión de las clases normal y asfixia con la clase hiperbilirrubinemia.

Tabla 4.5. Recuerdo de las pruebas con datos obtenidos de las muestras de Llanto de infante, sin reducir.

Características: Cocleogramas, LPC, MFCC e Intensidad				
Corrida (validación cruzada 10-folds)	Recuerdo			Media
	Normal	Asfixia	Hiperbilirrubinemia	
1	100.00%	94.44%	97.67%	97.37%
2	100.00%	69.39%	100.00%	89.80%
3	100.00%	100.00%	84.00%	94.67%
4	100.00%	100.00%	93.33%	97.78%
5	100.00%	94.44%	97.56%	97.34%
6	100.00%	94.44%	89.36%	94.60%
7	100.00%	100.00%	89.36%	96.45%
8	100.00%	100.00%	61.76%	87.25%
9	100.00%	100.00%	76.36%	92.12%
10	100.00%	100.00%	75.47%	91.98%
Promedios Totales	100.00%	95.27%	86.49%	93.92%

También se realizaron pruebas combinando características para determinar cuales identifican a las clases hiperbilirrubinemia y asfixia. En la Tabla 4.6 se muestra la media obtenida para la precisión y el alcance de cada clase. Se observa que la combinación de características con mejores resultados es la de LPC-Cocleogramas superando incluso los resultados que obtuvieron todas las características juntas. En cuanto a los resultados de las características individuales el mejor resultado fue el obtenido con la característica LPC y el peor resultado fue el obtenido con la característica intensidad.

Es importante mencionar que aunque la matriz de Cocleogramas es la que tiene mayor número de atributos, su desempeño fue superado por LPC, por lo que podemos decir que los LPC hacen una mejor representación de las clases asfixia e hiperbilirrubinemia.

Tabla 4.6. Resultados de pruebas con combinación de características.

Características Utilizadas	Clases						Media	
	Normal		Asfixia		Hiperbilirrubinemia			
	Precisión	Recuerdo	Precisión	Recuerdo	Precisión	Recuerdo	Precisión	Recuerdo
Cocleogramas	85.33%	86.05%	85.59%	87.71%	84.08%	83.43%	85.00%	85.73%
Intensidad	51.67%	55.59%	40.59%	47.69%	91.67%	69.61%	61.31%	57.63%
LPC	72.75%	96.42	94.71	100.00%	100.00%	81.69	89.15	92.70
MFCC	66.19%	93.08%	70.00%	90.71%	100.00%	73.00%	78.73%	85.60%
LPC-Cocleogramas	86.75%	96.75	94.41	92.80	94.05%	88.05	91.74%	92.53%
LPC-Cocleogramas-MFCC	78.33%	99.64%	78.53%	71.75%	97.14%	77.10%	84.67%	82.83%
LPC-Cocleogramas-MFCC-Intensidad	79.67%	100.00%	90.59%	95.27%	99.29%	86.49%	89.85%	93.92%

4.2.2. Resultados obtenidos con datos reducidos

También se realizaron pruebas con datos reducidos con el método LDA (ver sección 3.1.3.1) y con el método de reducción estadística (ver sección 3.1.3.2), para tener otro punto de comparación para los resultados obtenidos del clasificador.

Resultados con reducción LDA

En el caso de la clasificación con las matrices reducidas por el método LDA a 15 elementos, se realizaron pruebas combinando todas las características, obteniendo 3 matrices con $15+15+19+15=64$ características por clase. El tamaño de los vectores reducidos por LDA es 15 para LPC, 15 para MFCC y 15 para Cocleogramas, la característica intensidad no fue reducida debido a que su tamaño es 19 y no es muy grande. La división de las muestras para el método validación cruzada 10-folds es la misma que en las matrices sin reducir, ya que son el mismo número de muestras, solo cambia el número de atributos por muestra.

Los resultados de la clasificación se pueden observar en la Tabla 4.7, donde la precisión más alta fue para la clase hiperbilirrubinemia en la mayoría de las corridas, y la media de precisión más alta fue para la corrida 1 donde se alcanzó un **91.99 %** y la precisión más baja fue para la corrida 8 donde solo se obtuvo el **71.08 %**.

Tabla 4.7. Precisión obtenida en la clasificación utilizando datos reducidos con el método LDA.

Características: Cocleogramas, LPC, MFCC e Intensidad				
Corrida (validación cruzada 10-folds)	Precisión			Media
	Normal	Asfixia	Hiperbilirrubinemia	
1	92.50%	88.24%	95.24%	91.99%
2	90.00%	76.47%	92.86%	86.44%
3	97.50%	44.12%	100.00%	80.54%
4	97.50%	44.12%	100.00%	80.54%
5	82.50%	76.47%	100.00%	86.32%
6	82.50%	97.06%	80.95%	86.84%
7	90.00%	70.59%	97.62%	86.07%
8	67.50%	55.88%	92.86%	71.08%
9	67.50%	88.24%	92.86%	82.86%
10	47.50%	100.00%	100.00%	82.50%
Promedios Totales	81.50%	74.12%	95.24%	83.62%

En la Tabla 4.8 se muestra el cálculo del recuerdo para cada corrida de la clasificación con datos reducidos con el método LDA. Cabe destacar que la clase que obtuvo menor recuerdo fue Hiperbilirrubinemia, lo que significa que se confundieron elementos de la clase Normal y Asfixia con la clase de Hiperbilirrubinemia.

Tabla 4.8. Recuerdo obtenida en la clasificación utilizando datos reducidos con el método LDA.

Características: Cocleogramas, LPC, MFCC e Intensidad				
Corrida (validación cruzada 10-folds)	Recuerdo			Media
	Normal	Asfixia	Hiperbilirrubinemia	
1	94.87%	100.00%	85.11%	93.33%
2	90.00%	92.86%	81.25%	88.04%
3	76.47%	93.75%	85.71%	85.31%
4	86.67%	100.00%	75.00%	87.22%
5	91.67%	100.00%	77.78%	89.81%
6	97.06%	80.49%	82.93%	86.82%
7	100.00%	96.00%	74.55%	90.18%
8	84.38%	86.36%	62.90%	77.88%
9	100.00%	90.91%	69.64%	86.85%
10	100.00%	100.00%	66.13%	88.71%
Promedios Totales	92.11%	94.04%	76.10%	87.42%

Resultados Obtenidos con reducción por operaciones estadísticas

Se obtuvieron matrices de características reducidas utilizando el método de reducción estadística propuesto por Amaro (2008). Las características de cada matriz se redujeron a 5 valores de características, así se combinaron las matrices de cuatro características (cocleogramas, intensidad, LPC y MFCC) obteniendo al final $5+5+5+5=20$ características para cada clase. Para la clase normal la matriz obtenida es de 400 muestras x 20 características. Para la clase asfixia la matriz es de 340 muestras x 20 características y para la clase hiperbilirrubinemia su matriz es de 418 muestras x 20 características. La división de las muestras para el método validación cruzada *10-folds* es la misma que en las matrices sin reducir, ya que son el mismo número de muestras, solo cambia el número de características por muestra.

Los resultados se muestran en la Tabla 4.9, donde se observa que la única clase con precisión alta fue la de hiperbilirrubinemia mientras que las clases normal y asfixia tienen su precisión muy baja en todas las corridas.

Tabla 4.9. Clasificación utilizando datos reducidos con el método de reducción estadística.

Características: Cocleogramas, LPC, MFCC e Intensidad				
Corrida (validación cruzada 10-folds)	Precisión			Media
	Normal	Asfixia	Hiperbilirrubinemia	
1	2.56%	26.47%	100.00%	43.01%
2	7.50%	55.88%	97.62%	53.67%
3	15.00%	26.47%	100.00%	47.16%
4	20.00%	14.71%	100.00%	44.90%
5	12.50%	14.71%	100.00%	42.40%
6	20.00%	0.00%	100.00%	40.00%
7	12.50%	0.00%	83.33%	31.94%
8	5.00%	0.00%	100.00%	35.00%
9	5.00%	0.00%	100.00%	35.00%
10	5.00%	2.86%	100.00%	35.95%
Promedios Totales	10.51%	14.11%	98.10%	40.90%

Tabla 4.10. Recuerdo obtenido en la clasificación utilizando datos reducidos con el método de reducción estadística.

Características: Cocleogramas, LPC, MFCC e Intensidad				
Corrida (validación cruzada 10-folds)	Recuerdo			Media
	Normal	Asfixia	Hiperbilirrubinemia	
1	50.00%	60.00%	42.86%	50.95%
2	60.00%	54.29%	53.95%	56.08%
3	100.00%	75.00%	42.86%	72.62%
4	72.73%	100.00%	42.00%	71.58%
5	100.00%	55.56%	41.18%	65.58%
6	80.00%	0.00%	39.62%	39.87%
7	83.33%	0.00%	35.35%	39.56%
8	100.00%	0.00%	36.84%	45.61%
9	100.00%	0.00%	36.84%	45.61%
10	50.00%	100.00%	39.05%	63.02%
Promedios Totales	79.61%	44.48%	41.05%	55.05%

En la Tabla 4.10 se muestra el recuerdo obtenido para cada clase, donde observamos que la clase con menor alcance es hiperbilirrubinemia a pesar de que su precisión fue la más alta de 98.10%, esto debido a que se confundieron las clases Asfixia y Normal con Hiperbilirrubinemia así que, aunque se clasificaron bien la mayoría de elementos

etiquetados como hiperbilirrubinemia, se clasificaron mal las clases Asfixia y Normal ya que se clasificaron como hiperbilirrubinemia.

Cabe mencionar que el recuerdo más bajo fue para la clase hiperbilirrubinemia, que es un resultado opuesto a la precisión que obtuvo, lo que significa que hubo mucha confusión de las clases normal y asfixia con la de hiperbilirrubinemia. Debido a que las clases son sobrepuestas, los datos reducidos por operaciones estadísticas resultan ser muy similares entre las clases.

4.3. Pruebas con dos Clases de llanto: Normal e Hiperbilirrubinemia

Con la finalidad de detectar qué características son las que mejor describen a la clase hiperbilirrubinemia se realizaron una serie de pruebas solo con dos clases, utilizando las cuatro características de forma individual y haciendo combinaciones (Intensidad, LPC, MFCC y Cocleogramas). Para estas pruebas se utilizaron matrices sin reducir.

Tabla 4.11. Pruebas realizadas para la clasificación de dos clases: Normal e Hiperbilirrubinemia.

Resumen de Resultados						
Características utilizadas	Clases				Media	
	Normal		Hiperbilirrubinemia			
	Precisión	Recuerdo	Precisión	Recuerdo	Precisión	Recuerdo
Intensidad	58.50%	91.88%	92.14%	74.75%	75.32%	83.32%
LPC	85.50%	87.75%	88.33%	84.49%	86.92%	86.12%
MFCC	86.75%	94.68%	95.00%	88.45%	90.88%	91.57%
Cocleogramas	84.25%	75.41%	63.46%	87.18%	73.86%	81.30%
LPC-MFCC	93.50%	97.49%	97.62%	94.05%	95.56%	95.77%
LPC-Intensidad	50.25%	92.24%	94.76%	68.90%	72.51%	80.57%
LPC-Cocleogramas	75.25%	97.14%	98.33%	82.19%	86.79%	89.67%
MFCC-Cocleogramas	73.75%	85.77%	88.10%	82.97%	80.93%	84.37%

Como se observa en la Tabla 4.11 los mejores resultados fueron los obtenidos con la característica LPC y el segundo mejor fue con MFCC, así que se realizaron pruebas

con la combinación de dos características y se obtuvo el mejor resultado con la combinación de las características LPC y MFCC.

En la Tabla 4.12 se muestran los resultados obtenidos de las pruebas realizadas para la clasificación de dos clases utilizando las características LPC y MFCC. Estas pruebas fueron las que obtuvieron mejores resultados.

Tabla 4.12. Pruebas realizadas con dos clases (Normal e Hiperbilirrubinemia) utilizando la combinación de dos características, LPC y MFCC.

Característica: LPC-MFCC						
Corrida validación cruzada <i>10 folds</i>	Normal		Hiperbilirrubinemia		Media	
	Precisión	Recuerdo	Precisión	Recuerdo	Precisión	Recuerdo
1	95.00%	100.00%	100.00%	95.45%	97.50%	97.73%
2	92.50%	88.10%	88.10%	92.50%	90.30%	90.30%
3	90.00%	97.30%	97.62%	91.11%	93.81%	94.20%
4	95.00%	97.44%	97.62%	95.35%	96.31%	96.39%
5	92.50%	97.37%	97.62%	93.18%	95.06%	95.28%
6	95.00%	100.00%	100.00%	95.45%	97.50%	97.73%
7	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
8	97.50%	100.00%	100.00%	97.67%	98.75%	98.84%
9	90.00%	94.74%	95.24%	90.91%	92.62%	92.82%
10	87.50%	100.00%	100.00%	88.89%	93.75%	94.44%
Promedios Totales	93.50%	97.49%	97.62%	94.05%	95.56%	95.77%

Como se puede observar el recuerdo de la clase normal e hiperbilirrubinemia tiene como media 95.77% lo que significa que no existió mucha confusión.

En la Tabla 4.13 se observan las matrices de confusión obtenidas por cada corrida del método de validación cruzada *10-folds*. En esta tabla se puede observar que existía mayor confusión de elementos que pertenecían a la clase normal con elementos de hiperbilirrubinemia.

Tabla 4.13. Matriz de Confusión de las Pruebas con dos clases utilizando LPC-MFCC

Matriz de confusion de pruebas con LPC-MFCC			
No. De corrida validación cruzada <i>10-folds</i>	Clases	Normal	Hiperbilirrubinemia
1	Normal	38	2
	Hiperbilirrubinemia	0	42
2	Normal	37	3
	Hiperbilirrubinemia	5	37
3	Normal	36	4
	Hiperbilirrubinemia	1	41
4	Normal	38	2
	Hiperbilirrubinemia	1	41
5	Normal	37	3
	Hiperbilirrubinemia	1	41
6	Normal	38	2
	Hiperbilirrubinemia	0	42
7	Normal	40	0
	Hiperbilirrubinemia	0	42
8	Normal	39	1
	Hiperbilirrubinemia	0	42
9	Normal	36	4
	Hiperbilirrubinemia	2	40
10	Normal	35	5
	Hiperbilirrubinemia	0	40

4.4. Comparación con otros trabajos.

Se hace una comparación con algunos resultados obtenidos por Amaro (2008), ya que realizó pruebas con llantos de infantes de las clases normal, sordos y asfixia utilizando las mismas características que en este trabajo: LPC, MFCC, Cocleogramas e Intensidad. Realizó la clasificación con distintos algoritmos entre los que se encuentra: la red neuronal *Feed Forward* y *Naive Bayes*. En la Tabla 4.14 se muestra una tabla con algunos resultados obtenidos por Amaro.

Cabe mencionar que Amaro [Amaro, 2008] utilizó diferentes parámetros para la obtención de las características, ya que obtuvo 2500 valores por cocleograma, los

cuales redujo a 500 eligiendo valores de forma aleatoria y posteriormente aplicando la reducción LDA o reducción estadística. También difieren los datos LPC y MFCC ya que incluyo un valor llamado ganancia, que describe el error de los coeficientes LPC, calculada por el software PRAAT [Boersma y Weenink, 2007] para cada conjunto de 16 coeficientes; esta ganancia no fue considerada en este trabajo.

Tabla 4.14. Resultados obtenidos por Amaro [Amaro, 2008]

Características de la Clasificación	Clases	Reducción	Precisión
LPC-Cocleogramas	Normal-Sordos-Asfixia	LDA 15 atributos	87.41%
LPC-Cocleogramas-Intensidad-MFCC	Normal-Sordos-Asfixia	LDA 20 atributos	91.30%
LPC	Normal-Asfixia	LDA 9 atributos	91.87%
MFCC	Normal-Asfixia	LDA 9 atributos	89.5%

Tabla 4.15. Resultados obtenidos en éste trabajo.

Características de la Clasificación	Clases	Reducción	Precisión
LPC-Cocleogramas	Normal-Asfixia-Hiperbilirrubinemia	LDA 15 atributos	91.49%
LPC-Cocleogramas-Intensidad-MFCC	Normal-Asfixia-Hiperbilirrubinemia	LDA 15 atributos	89.85%
LPC-MFCC	Normal- Hiperbilirrubinemia	Sin reducción	95.56 %
LPC-MFCC	Normal-Asfixia	Sin reducción	99.02 %

En la tabla 4.15 se muestran algunos resultados obtenidos en este trabajo, considerando características similares a las que utilizó Amaro para poder hacer una comparación. En el caso de la clasificación de dos clases con las características LPC y MFCC se mejoró el desempeño con el algoritmo de Comparación Difusa de

Patrones con conjuntos Difusos tipo 2. En el caso de la clasificación de tres clases utilizando cuatro características no se mejoró el desempeño.

Otras comparaciones

También se realizaron pruebas con los conjuntos de datos de *Ljubljana Breast Cancer* (LBC) y el de *Pima Indians Diabetes* [Newman, Hettich, Blake y Merz, 2008], la descripción de estos conjuntos de datos se muestra en la tabla 4.16.

Tabla 4.16 Descripción de los conjuntos de datos Pima y LBC.

Conjunto de Datos	Clases	Atributos por elemento	Elementos por clase
Pima	2, Con diabetes y sin diabetes	8	268 de la clase con diabetes, 500 de la clase sin diabetes
LBC	2, Con Cancer de seno y sin Cancer.	9	218 de la clase sin cáncer, 68 de la clase con cáncer

Se utilizó el algoritmo CDP-T2 para clasificar también las bases de datos Pima y LBC para poder comparar con los resultados reportados en [Saïd, Sayed y Billaudel, 2007], cuyos resultados se muestran en la tabla 4.17.

Tabla 4.17. Resultados obtenidos de pruebas realizadas con los conjuntos de datos Pima y LBC.

Criterio de evaluación		Precisión
Metodo	CDP-T2	CDP [Saïd, Sayed y Billaudel, 2007]
Pima	75.22%	69.27%
LBC	79.65%	74.13%

Como se puede observar el método CDP-T2 mejoro los resultados reportado por [Saïd, Sayed y Billaudel, 2007].

4.5. Discusión

Los mejores resultados, para la clasificación de tres clases de llanto, se obtuvieron utilizando las características LPC-Cocleogramas, que fueron las que mejor describieron a las clases. También se obtuvo un buen resultado con la combinación de

las cuatro características (LPC, MFCC, Cocleogramas e Intensidad), aunque no mejoro lo obtenido con las dos características LPC-Cocleogramas, probablemente porque al utilizar en la combinación la característica intensidad, provoco que se confundieran las clases, ya que las pruebas que se realizaron con la característica intensidad demostró que con esta característica existe mayor confusión entre las clases.

Los resultados obtenidos en la clasificación con los datos reducidos con LDA no mejoraron los resultados en comparación con las pruebas realizadas con los datos sin reducir, aunque no fueron , ya que con esta reducción se obtienen los datos más representativos de cada clase, debido a que el algoritmo toma en cuenta los datos característicos de todas las clases.

El bajo desempeño que presentó la clasificación con datos reducidos por operaciones estadísticas es debido a que la reducción solo obtiene datos de cada clase sin considerar la importancia de los datos entre clases. Debido a que las clases con las que se realizaron las pruebas en esta tesis son clases sobrepuestas, las matrices obtenidas con esta reducción fueron muy similares, así que el clasificador no obtuvo suficiente información de cada clase generándose confusión.

Por otro lado con las pruebas realizadas con dos clases de llanto, normal e hiperbilirrubinemia, se puede ver que las características que mejor identifican a la clase hiperbilirrubinemia son LPC y MFCC.

Capítulo 5

CONCLUSIONES Y TRABAJO FUTURO

5.1. Conclusiones y aportaciones principales

Se cumplió el objetivo principal de este trabajo que fue la implementación del algoritmo de Comparación Difusa de Patrones utilizando Conjuntos Difusos tipo 2 para clasificar llanto de infantes diagnosticados con patologías de riesgo neurológico, que en este caso fueron solo hiperbilirrubinemia y asfixia, y que además tienen como característica que son clases sobrepuestas.

Cabe hacer notar que hasta el momento de la preparación de este documento esta clase patológica no se había clasificado de forma automática en trabajos anteriores.

Dentro de las conclusiones importantes obtenidas está que los mejores resultados son los que se obtienen clasificando solo dos clases simultáneamente, utilizando las características LPC y MFCC.

A partir de los resultados obtenidos se determina que las clases asfixia e hiperbilirrubinemia resultaron ser las más sobrepuestas entre sí, debido a que los valores de recuerdo para cada clase fueron regularmente, los más bajos de las pruebas.

También se demostró que las características LPC y Cocleogramas identificaron mejor el llanto de infantes que tienen patologías de riesgo neurológico. Los mejores resultados al clasificar las tres clases simultáneamente (normal, asfixia e hiperbilirrubinemia) se obtuvieron utilizando juntas las características LPC y Cocleogramas. Asimismo se observó que las características que mejor definen a la clase hiperbilirrubinemia son LPC y MFCC.

El método de Comparación Difusa de Patrones con Conjuntos Difusos tipo 2 es un método simple de orden polinomial, que obtiene buenos resultados sin utilizar demasiados recursos del ordenador.

El uso de los conjuntos difusos tipo 2 nos da una mejor representación de los elementos que pertenecen a cada clase, y en el caso de nuestra implementación, nos dio información de la importancia de cada elemento a cada clase dando así mejores resultados en la clasificación de elementos de clases sobrepuestas.

También pudimos notar que los mejores resultados fueron con el conjunto de datos que no tuvieron ningún tipo de reducción, por lo que podríamos omitir el bloque de reducción de vectores en el caso de la clasificación con el método CDP-T2.

5.2. Trabajo Futuro

Como trabajo futuro se propone agregar nuevas clases de llanto de infante para ser clasificadas. Asimismo, es conveniente realizar pruebas utilizando otras características que puedan representar a la clase de hiperbilirrubinemia, como la frecuencia fundamental, que no se utilizó en este trabajo.

Es necesario obtener más muestras de llanto de cada clase para mejorar el desempeño de entrenamiento ya que se cuentan con pocas muestras para las clases asfixia e hiperbilirrubinemia.

También se puede incursionar probando el algoritmo propuesto con otras operaciones, como el promedio ponderado o operaciones difusas, para la obtención de la membresía.

APÉNDICE A.

En éste apéndice se muestran algunas matrices de confusión obtenidas al realizar las pruebas, así como los mejores resultados obtenidos para la realización de las pruebas con 2 clases.

Tabla A.1 Matriz de Confusión de las pruebas con la base de datos de las Flores Iris

No. De corrida	Clase	Iris-setosa	Iris-versicolor	Iris-virginica
1	Iris-setosa	5	0	0
	Iris-versicolor	0	4	1
	Iris-virginica	1	0	4
2	Iris-setosa	5	0	0
	Iris-versicolor	0	5	0
	Iris-virginica	0	0	5
3	Iris-setosa	5	0	0
	Iris-versicolor	1	4	0
	Iris-virginica	0	1	4
4	Iris-setosa	5	0	0
	Iris-versicolor	0	4	1
	Iris-virginica	0	0	5
5	Iris-setosa	5	0	0
	Iris-versicolor	0	5	0
	Iris-virginica	0	0	5
6	Iris-setosa	5	0	0
	Iris-versicolor	0	4	1
	Iris-virginica	0	0	5
7	Iris-setosa	5	0	0
	Iris-versicolor	1	4	0
	Iris-virginica	0	0	5
8	Iris-setosa	5	0	0
	Iris-versicolor	0	5	0
	Iris-virginica	0	0	5
9	Iris-setosa	5	0	0
	Iris-versicolor	0	5	0
	Iris-virginica	0	0	5
10	Iris-setosa	5	0	0
	Iris-versicolor	0	4	1
	Iris-virginica	0	0	5

Tabla A.2. Matriz de Confusión de las pruebas con datos de la base de datos de llanto sin aplicar reducción utilizando las cuatro características (LPC, MFCC, Intensidad y Cocleogramas).

No. De corrida	Clase	Normal	Asfixia	Hiperbilirrubinemia
1	Normal	36	2	2
	Asfixia	0	34	0
	Hiperbilirrubinemia	0	0	42
2	Normal	21	19	0
	Asfixia	0	34	0
	Hiperbilirrubinemia	0	1	41
3	Normal	40	0	0
	Asfixia	0	26	8
	Hiperbilirrubinemia	0	0	42
4	Normal	40	0	0
	Asfixia	0	31	3
	Hiperbilirrubinemia	0	0	42
5	Normal	39	0	1
	Asfixia	0	34	0
	Hiperbilirrubinemia	0	2	40
6	Normal	31	3	6
	Asfixia	0	34	0
	Hiperbilirrubinemia	0	0	42
7	Normal	33	0	7
	Asfixia	0	34	0
	Hiperbilirrubinemia	0	0	42
8	Normal	21	0	19
	Asfixia	0	22	12
	Hiperbilirrubinemia	0	0	42
9	Normal	28	0	12
	Asfixia	0	30	4
	Hiperbilirrubinemia	0	0	42
10	Normal	29	0	11
	Asfixia	0	29	5
	Hiperbilirrubinemia	0	0	40

Tabla A.3. Matriz de Confusión de las pruebas utilizando datos reducidos con el método LDA utilizando cuatro características (LPC, MFCC, Intensidad y Cocleogramas).

No. De corrida	Clase	Normal	Asfixia	Hiperbilirrubinemia
1	Normal	37	0	3
	Asfixia	0	30	4
	Hiperbilirrubinemia	2	0	40
2	Normal	36	2	2
	Asfixia	1	26	7
	Hiperbilirrubinemia	3	0	39
3	Normal	39	1	0
	Asfixia	12	15	7
	Hiperbilirrubinemia	0	0	42
4	Normal	39	0	1
	Asfixia	6	15	13
	Hiperbilirrubinemia	0	0	42
5	Normal	33	0	7
	Asfixia	3	26	5
	Hiperbilirrubinemia	0	0	42
6	Normal	33	0	7
	Asfixia	1	33	0
	Hiperbilirrubinemia	0	8	34
7	Normal	36	0	4
	Asfixia	0	24	10
	Hiperbilirrubinemia	0	1	41
8	Normal	27	0	13
	Asfixia	5	19	10
	Hiperbilirrubinemia	0	3	39
9	Normal	27	0	13
	Asfixia	0	30	4
	Hiperbilirrubinemia	0	3	39
10	Normal	19	0	21
	Asfixia	0	34	0
	Hiperbilirrubinemia	0	0	41

Tabla A.4. Matriz de Confusión de las pruebas realizadas con datos reducidos con el método de Operaciones Estadísticas utilizando cuatro características (LPC, MFCC, Intensidad y Cocleogramas)

No. De corrida	Clase	Normal	Asfixia	Hiperbilirrubinemia
1	Normal	2	6	32
	Asfixia	1	9	24
	Hiperbilirrubinemia	0	0	42
2	Normal	3	15	22
	Asfixia	2	19	13
	Hiperbilirrubinemia	0	1	41
3	Normal	6	3	31
	Asfixia	0	9	25
	Hiperbilirrubinemia	0	0	42
4	Normal	8	0	32
	Asfixia	3	5	26
	Hiperbilirrubinemia	0	0	42
5	Normal	5	4	31
	Asfixia	0	5	29
	Hiperbilirrubinemia	0	0	42
6	Normal	8	0	32
	Asfixia	2	0	32
	Hiperbilirrubinemia	0	0	42
7	Normal	5	4	31
	Asfixia	1	0	33
	Hiperbilirrubinemia	0	7	35
8	Normal	2	0	38
	Asfixia	0	0	34
	Hiperbilirrubinemia	0	0	42
9	Normal	2	0	38
	Asfixia	0	0	34
	Hiperbilirrubinemia	0	0	42
10	Normal	2	6	32
	Asfixia	2	1	32
	Hiperbilirrubinemia	0	0	41

Tabla A.5. Matriz de Confusión de las pruebas realizadas con datos sin reducir utilizando la característica LPC.

No. De corrida	Clase	Normal	Asfixia	Hiperbilirrubinemia
1	Normal	36	0	4
	Asfixia	1	33	0
	Hiperbilirrubinemia	0	0	42
2	Normal	11	0	29
	Asfixia	0	30	4
	Hiperbilirrubinemia	0	0	42
3	Normal	40	0	0
	Asfixia	2	31	1
	Hiperbilirrubinemia	0	0	42
4	Normal	40	0	0
	Asfixia	3	31	0
	Hiperbilirrubinemia	0	0	42
5	Normal	21	0	19
	Asfixia	2	32	0
	Hiperbilirrubinemia	0	0	42
6	Normal	40	0	0
	Asfixia	2	32	0
	Hiperbilirrubinemia	0	0	42
7	Normal	15	0	25
	Asfixia	0	34	0
	Hiperbilirrubinemia	0	0	42
8	Normal	25	0	15
	Asfixia	0	34	0
	Hiperbilirrubinemia	0	0	42
9	Normal	28	0	12
	Asfixia	0	34	0
	Hiperbilirrubinemia	0	0	42
10	Normal	35	0	5
	Asfixia	3	31	0
	Hiperbilirrubinemia	0	0	40

Tabla A.6. Matriz de Confusión de las pruebas realizadas con datos sin reducir utilizando las características LPC y Cocleogramas.

No. De corrida	Clase	Normal	Asfixia	Hiperbilirrubinemia
1	Normal	33	0	7
	Asfixia	0	32	2
	Hiperbilirrubinemia	0	0	42
2	Normal	33	7	0
	Asfixia	0	32	2
	Hiperbilirrubinemia	0	3	39
3	Normal	31	7	2
	Asfixia	0	32	2
	Hiperbilirrubinemia	0	0	42
4	Normal	40	0	0
	Asfixia	0	34	0
	Hiperbilirrubinemia	0	5	37
5	Normal	33	0	7
	Asfixia	0	34	0
	Hiperbilirrubinemia	2	3	37
6	Normal	33	0	7
	Asfixia	0	30	4
	Hiperbilirrubinemia	1	1	40
7	Normal	35	1	4
	Asfixia	0	32	2
	Hiperbilirrubinemia	0	0	42
8	Normal	37	0	3
	Asfixia	0	32	2
	Hiperbilirrubinemia	5	0	37
9	Normal	37	1	2
	Asfixia	0	32	2
	Hiperbilirrubinemia	5	0	37
10	Normal	35	0	5
	Asfixia	0	31	3
	Hiperbilirrubinemia	0	0	40

Tabla A.7. Matriz de confusión de pruebas con datos sin reducir utilizando la característica LPC para dos clases: Normal e Hiperbilirrubinemia

Matriz de confusion de pruebas con LPC			
No. De corrida del 10 folds cross validation	Clases	Normal	Hiperbilirrubinemia
1	Normal	28	12
	Hiperbilirrubinemia	4	38
2	Normal	36	4
	Hiperbilirrubinemia	0	42
3	Normal	40	0
	Hiperbilirrubinemia	7	35
4	Normal	28	12
	Hiperbilirrubinemia	8	34
5	Normal	32	8
	Hiperbilirrubinemia	11	31
6	Normal	31	9
	Hiperbilirrubinemia	4	38
7	Normal	38	2
	Hiperbilirrubinemia	1	41
8	Normal	33	7
	Hiperbilirrubinemia	11	31
9	Normal	33	7
	Hiperbilirrubinemia	3	39
10	Normal	31	9
	Hiperbilirrubinemia	0	40

Tabla A.8. Evaluación de las pruebas realizadas con datos sin reducir utilizando la característica LPC para dos clases: Normal e Hiperbilirrubinemia.

Característica: LPC						
Corrida (10 fold cross validation)	Normal		Hiperbilirrubinemia		Media	
	Precisión	Recuerdo	Precisión	Recuerdo	Precisión	Recuerdo
1	70.00%	87.50%	90.48%	76.00%	80.24%	81.75%
2	90.00%	100.00%	100.00%	91.30%	95.00%	95.65%
3	100.00%	85.11%	83.33%	100.00%	91.67%	92.55%
4	70.00%	77.78%	80.95%	73.91%	75.48%	75.85%
5	80.00%	74.42%	73.81%	79.49%	76.90%	76.95%
6	77.50%	88.57%	90.48%	80.85%	83.99%	84.71%
7	95.00%	97.44%	97.62%	95.35%	96.31%	96.39%
8	82.50%	75.00%	73.81%	81.58%	78.15%	78.29%
9	82.50%	91.67%	92.86%	84.78%	87.68%	88.22%
10	77.50%	100.00%	100.00%	81.63%	88.75%	90.82%
Promedio	82.50%	87.75%	88.33%	84.49%	85.42%	86.12%

Tabla A.9. Matriz de confusión de pruebas con datos sin reducir utilizando la característica MFCC para dos clases: Normal e Hiperbilirrubinemia

Matriz de confusión de pruebas con MFCC			
No. de corrida Validación Cruzada 10 folds.	Clases	Normal	Hiperbilirrubinemia
1	Normal	38	2
	Hiperbilirrubinemia	1	41
2	Normal	32	8
	Hiperbilirrubinemia	1	41
3	Normal	30	10
	Hiperbilirrubinemia	0	42
4	Normal	36	4
	Hiperbilirrubinemia	5	37
5	Normal	32	8
	Hiperbilirrubinemia	2	40
6	Normal	33	7
	Hiperbilirrubinemia	9	33
7	Normal	36	4
	Hiperbilirrubinemia	1	41
8	Normal	35	5
	Hiperbilirrubinemia	1	41
9	Normal	37	3
	Hiperbilirrubinemia	1	41
10	Normal	38	2
	Hiperbilirrubinemia	0	40

Tabla A.10. Evaluación de las pruebas realizadas con datos sin reducir utilizando la característica LPC para dos clases: Normal e Hiperbilirrubinemia.

Característica: MFCC						
Corrida (10 fold cross validation)	Normal		Hiperbilirrubinemia		Media	
	Precisión	Recuerdo	Precisión	Recuerdo	Precisión	Recuerdo
1	95.00%	97.44%	97.62%	95.35%	96.31%	96.39%
2	80.00%	96.97%	97.62%	83.67%	88.81%	90.32%
3	75.00%	100.00%	100.00%	80.77%	87.50%	90.38%
4	90.00%	87.80%	88.10%	90.24%	89.05%	89.02%
5	80.00%	94.12%	95.24%	83.33%	87.62%	88.73%
6	82.50%	78.57%	78.57%	82.50%	80.54%	80.54%
7	90.00%	97.30%	97.62%	91.11%	93.81%	94.20%
8	87.50%	97.22%	97.62%	89.13%	92.56%	93.18%
9	92.50%	97.37%	97.62%	93.18%	95.06%	95.28%
10	95.00%	100.00%	100.00%	95.24%	97.50%	97.62%
Promedio	86.75%	94.68%	95.00%	88.45%	90.88%	91.57%

APÉNDICE B

En éste apéndice se describe la base de datos de llantos utilizada en las pruebas que se realizaron en ésta tesis.

La base de datos utilizada fue generada a partir de las grabaciones de llanto de bebés con las patologías de hiperbilirrubinemia, asfixia y sin estas patologías que son etiquetadas como “normales.” Estas grabaciones se segmentaron en pequeños archivos de 1 segundo y se tienen 418 segmentos de la clase hiperbilirrubinemia, 340 de la clase asfixia y 1136 de la clase normales. Cabe mencionar que para las pruebas realizadas solo se utilizaron 400 muestras de la clase normales para tener el entrenamiento con 3 clases balanceadas.

Por cada muestra de 1 segundo se obtuvieron 4 vectores de características (LPC, MFCC, Intensidad y Cocleogramas). Cada vector de distinto tamaño según los parámetros que se utilizaron, como se explica en la sección 3.1.2. *Extracción de características*. Para obtener cada vector característico se siguieron los siguientes pasos:

- Se obtuvo las características con el software PRAAT [Boersma P. y Weenink D. 2007]
- Cada archivo obtenido en PRAAT tenía información adicional innecesaria para éste trabajo, por ejemplo, encabezados para cada sección de coeficientes en el caso de los LPC, entre otros datos. Así que se limpio cada archivo con ayuda de un programa que realizaba esta tarea automáticamente.
- Finalmente por cada muestra de un segundo obtuvimos 1 vector de 304 coeficientes LPC, 1 vector de 304 coeficientes MFCC, 1 vector con 19 valores de Intensidad y 1 vector con 510 valores de Cocleogramas.

Posteriormente por cada clase se unieron los vectores para generar 1 matriz por cada característica, en archivos de texto. Para la Clase hiperbilirrubinemia se tiene 1 matriz de la característica LPC de tamaño 418x304, 418 muestras por 304 coeficientes LPC. En la Tabla B.1 se observa el tamaño de las matrices generadas por cada clase y característica. Se generaron 12 matrices, en formato .mat para ser utilizadas en el programa MatLab, cada fila de la matriz pertenece a una muestra de 1 segundo, y al final de la fila se coloco la clase a la que pertenece.

Tabla B.1. Tamaño de las matrices generadas por característica.

Tamaño de las matrices generadas				
Clase	Característica			
	LPC	MFCC	Intensidad	Cocleogramas
Hiperbilirrubinemia	418x305	418x305	418x20	418x511
Asfixia	340x305	340x305	340x20	340x511
Normal	1136x305	1136x305	1136x20	1136x511

También se generaron las matrices con datos reducidos, en el caso de la matriz reducida por el método de reducción estadística [Amaro, 2008], cada fila tiene los siguientes datos:

- Mínimo
- Máximo
- Promedio
- Desviación estándar
- Varianza
- Clase

El tamaño de las muestras es el mismo para cada clase y también se generaron 12 matrices, que corresponden a las 4 características utilizadas (LPC, MFCC, Intensidad y Cocleogramas) por cada clase: hiperbilirrubinemia, asfixia y normal.

En el caso de la reducción por el método LDA, solo se redujeron las características LPC, MFCC, Cocleogramas, así que se tienen 9 matrices que corresponden a tres clases: hiperbilirrubinemia, asfixia y normal, por tres características.

Cada fila de la matriz corresponde a una muestra y contiene 16 columnas que son los 15 LDA's obtenidos y la clase a la que pertenecen.

Para los tres tipos de matrices, el valor de las clases es un número y corresponde a: 1 para la clase normal, 2 para la clase asfixia y 3 para la clase hiperbilirrubinemia.

LISTA DE ACRÓNIMOS

CDP	Comparación Difusa de Patrones
CD-T2	Conjuntos Difusos Tipo 2
DCT	De las siglas en inglés Discrete Cosine Transform
DSP	De las siglas en inglés Digital Signal Processing.
LBC	Conjunto de datos de <i>Ljubljana Breast Cancer</i>
LD-T2	Lógica Difusa Tipo 2
LDA	De las siglas en inglés Linear Discriminant Analysis
LPC	De las siglas en ingles Linear Predictive Coding
MFCC	De las siglas en ingles Mel Frequency Cepstral Coefficients
SPL	De las siglas en ingles Sound Power Level

BIBLIOGRAFÍA

- [Amaro, 2008] Amaro Camargo E., Caracterización Automática del Llanto de Bebé para su Estudio con Modelos de Clasificación, Tesis de Maestría en Ciencias, Instituto Nacional de Astrofísica Óptica y Electrónica de la Coordinación de Computación, 2008.
- [Barkin, 2000] Barkin Roger, Urgencias Pediátricas, Ed. Elsevier España, Capítulo 11, pag. 107 – 115, 2000
- [Boersma y Weenink, 2007] Boersma, P. & Weenink D; *PRAAT: Doing phonetics by computer (Versión 5.0.1)*, programa computacional. Descargado en Enero 2007, última version disponible en <http://fonsg3.hum.uva.nl/praat/> .
- [Cano, 2007] Cano S. D., et al. “15 Años del Cry Analysis en Cuba: Resultados y Perspectivas”, VI Congreso Internacional de Informática en Salud, Santiago de Cuba, 2007.
- [Chen, Wu y Yachida, 1995] Qian Chen, Haiyuan Wu y Masahiko Yachida, “Face Detection by Fuzzy Pattern Matching” Department of Systems Engineering, Osaka University, IEEE 0-8186-7042-8/95 pag. 591-596, 1995.
- [Chung y Rhee, 2007] Chung Frank y Rhee Hoon, “Uncertain Fuzzy Clustering: Insights and Recommendations”, Hanyang University, Korea, IEEE Computational Intelligence Magazine, February 2007.
- [Duchene y Leclerq, 1988] Duchene, L. y Leclerq, S., “An Optimal Transformation for Discriminant and Principal Component Analysis”, IEEE Transactions

on Pattern Analysis and Machine Intelligence, Vol. 10, No 6, pag. 978-983, 1988

[Dworkin, 1995] Dworkin Paul H., Pediatría- Serie médica nacional para el estudio independiente, ed. McGraw Hill Interamericana, 4ª edición, pag. 151-157 y 170-173, 1995.

[Fisher, 1936] R.A. Fisher, “The use of multiple measurements in taxonomic problems”, Cambridge University Press, publicada por Annals of Eugenics, vol. 7, pag. 179-188, 1936.

[Golub y Corwin, 1982] Golub, H., Corwin, M. “Infant Cry: a Clue to Diagnostic”. Pediatrics, vol. 69. pp. 197-201. 1982.

[Hayes, 2000] Hayes Gloria, “Ictericia del recién nacido: Cómo disminuir los riesgos”. Revista: Atención Médica, vol. 13, No. 10 Octubre 2006, pag. 54-67

[Jara, 2006] Jara Hernández N. Análisis Espectrofonográfico del llanto en recién nacidos de término con riesgo neurológico, Tesis de especialidad en Pediatría Médica, Dirección de enseñanza en Neonatología del Instituto Nacional de Perinatología, México, 2006

[Mendel, 2007] Mendel J.M., “Type-2 Fuzzy Sets and Systems: an overview”, University of Southern California, USA.

[Mendel y Bob, 2002] Mendel J.M. y Bob John R.I., “Type-2 Fuzzy Sets Made Simple, IEEE” Transactions on Fuzzy Systems, Vol. 10, No. 2, 2002.

[Montes y Villaseñor, 2007] M. Montes y L. Villaseñor, “Notas del curso de clasificación de texto”, Laboratorio de Tecnologías del Lenguaje, INAOE, 2007

- [Newman, Hettich, Blake y Merz, 2008] Newman D.J., Hettich, S., Blake C.L. and Merz, C.J, UCI Repository of Machine learning databases, <http://www.ics.uci.edu/~mlearn/MLRepository.html>, Departamento de información y ciencias computacionales, Universidad de California, 1998.
- [Orozco, 2002] Orozco García J., Clasificación de Llanto de Bebé utilizando una Red Neural de Gradiente Conjugado Escalado, Tesis de Maestría, INAOE, Puebla, Mex., 2002.
- [Petroni, 1995] Petroni, M., Malowany, et al. “Identification of Pain from Infant Cry Vocalizations Using Artificial Neural Networks (ANNs)”. The International Infant Cry Research Group. Applications and Science of Artificial Neural Networks. The International Society for Optical Engineering. Volumen 2492. 1995. pp.729-738.
- [Proakis y Manolakis, 1998] Proakis J.G., Manolakis D.G.; Tratamiento Digital de Señales, Principios, algoritmos y aplicaciones; 3ra Edición, Prentice Hall, Cap. 1. pp. 1-38, 1998
- [Sayed, 2002] Sayed Mouchaweh, M., A. Devillez, G.V. Lecolier y P Billaudel, “Incremental learning in real time using fuzzy pattern matching”, Fuzzy Sets and Systems, Vol. 132/1, pag. 49-62, 2007.
- [Saïd, Sayed y Billaudel, 2007] Saïd Bouguelid M., Sayed Mouchaweh M. y Billaudel P., “New Results on Diagnosis by Fuzzy Pattern Recognition”, CReSTIC, Université de Reims-Champagne-Ardennes Moulin de la Housse, France, 2007.
- [Smith, 1999] Steven W. Smith, The Scientist and Engineer’s Guide to Digital Signal Processing, Ed. California Technical Publising, 2da. Edición. 1999
- [SNVE, 2006] “Epidemiología, Sistema único de información, Mortalidad Infantil, Perfil Epidemiológico de las defunciones por Asfixia del Recién nacido”,

Sistema Nacional de Vigilancia Epidemiológica SNVE Num. 45, Vol. 23,
Semana 45, del 5 al 11 de Noviembre del 2006, ISSN 1405-2636

[Verduzco, 2004] Verduzco A. et al. “Características del llanto de infantes de alto riesgo”, Ciencias Clínicas, Vol. 5, Num. 1, pp. 21-29, 2004.

[Wasz-Höckert, 1968] Wasz-Höckert, O., et al. “The Infant Cry: A Spectrographic and Auditory Analysis”. Spastics International Medical Publication, Lavenham, U.K., 1968.

[Zadeh, 1965] Zadeh L.A., “Fuzzy Sets”, Departament of Electrical Engineering and Electrics Research Laboratory, University of California, Berkeley, California, 1965.

[Wikipedia, 2008] De la enciclopedia en línea Wikipedia, <http://en.wikipedia.org/>

[Zbynì y Psutka, 1999] Zbynì k Tychtl y Josef Psutka, “Speech Production Based on the Mel-Frequency Cepstral Coefficients”, Departament of Cybernetics, University of West Bohemia, Czech Republic, 1999.