



UNIVERSIDAD AUTÓNOMA METROPOLITANA
UNIDAD IZTAPALAPA

DIVISIÓN DE CIENCIAS BÁSICAS E INGENIERÍA

**“MAPA DE SONIDOS
RESPIRATORIOS ADVENTICIOS
DISCONTINUOS”**

Tesis para obtener el título de
Maestra en Ciencias (Ingeniería Biomédica)
presenta:

GUADALUPE DORANTES MÉNDEZ

Asesores:

Dr. A. Tomás Aljama Corrales

Dr. Ramón González Camarena

Sinodales:

Dr. J. Guillermo Carrillo Rodríguez (Presidente)

Dr. A. Tomás Aljama Corrales (Secretario)

M. en I. Miguel Ángel Peña Castillo (Vocal)

Esta Tesis está dedicada a:

*Mis padres Guadalupe y José
Manuel por brindarme apoyo,
fortaleza y amor
incondicionalmente.*

*A mis hermanos Lolita y
Manuelito, por hacer mi vida más
alegre e interesante y llenarme de
cariño.*

AGRADECIMIENTOS

A la Dra. Sonia Charleston Villalobos y al Dr. A. Tomás Aljama Corrales por sus enseñanzas y su apoyo a lo largo de este tiempo. Además de contagiarme el gusto por la investigación.

A mis asesores el Dr. Tomás Aljama y el Dr. Ramón González por su colaboración y en la realización de esta tesis.

Al M. en I. Miguel Ángel Peña Castillo y al Dr. Carrillo por sus valiosas sugerencias sobre el trabajo de tesis.

A mis profesores de la Maestría por compartir parte de sus conocimientos durante mi trayecto en el posgrado.

A Aldo Mejía por su cariño y su apoyo en todo momento.

A mis amigos Omar Piña, Alejandra Torres, Eduardo Estrada, Lidia Albuerno y Bersain Reyes, por su amistad y su apoyo.

A mis compañeros del posgrado y a todas las personas que me brindaron apoyo durante esta maestría.

RECONOCIMIENTOS

Al Laboratorio de Imágenes y Procesamiento de Señales en Ingeniería Biomédica (LIPSIB) de la Universidad Autónoma Metropolitana (UAM).

A CONACYT por otorgarme una beca para la realización de esta maestría.

Al Instituto Nacional de Enfermedades Respiratorias (INER).

ÍNDICE

RESUMEN	1
CAPITULO I	
INTRODUCCIÓN	3
1.1 LA AUSCULTACIÓN PULMONAR.....	3
1.2 LOS SONIDOS RESPIRATORIOS Y SU IMPORTANCIA CLÍNICA	4
1.3 TIPOS DE SONIDOS RESPIRATORIOS Y SUS PRINCIPALES CARACTERÍSTICAS	5
1.3.1 Sonidos Respiratorios Normales	5
1.3.2 Sonidos Traqueales Normales	6
1.4 SONIDOS ADVENTICIOS	7
1.4.1 Sibilancias.....	8
1.4.2 Ronquidos	8
1.4.3 Estertores Crepitantes	8
Referencias.....	12
CAPITULO II	
2.1 ANTECEDENTES EN EL ANÁLISIS DE CREPITANCIAS.....	13
2.1.1 Análisis Unicanal	13
2.1.2 Análisis Multicanal	20
Referencias.....	22
CAPITULO III	
3.1 PLANTEAMIENTO DEL PROBLEMA.....	24
3.2 HIPÓTESIS.....	25
3.3 OBJETIVO.....	25
3.3.1 Objetivo General.....	25
3.3.2 Objetivos Específicos	26
CAPITULO IV	
ELEMENTOS TEÓRICOS EN EL ANÁLISIS DE LOS SONIDOS	
CREPITANTES	27
4.1 ANÁLISIS DINÁMICO NO LINEAL DE UNA SERIE DE TIEMPO: LA DIMENSIÓN	
FRACTAL	28
4.2 MODELOS ESTOCÁSTICOS Y PARAMETRIZACIÓN DE UNA SERIE DE TIEMPO POR	
MEDIO DE UN MODELO AR.....	32

4.2.1	Modelo Estocástico Autorregresivo	34
4.2.1.1	<i>Selección del orden del Modelo AR</i>	39
4.2.2	Modelo Estocástico Autorregresivo Variante en el Tiempo y su Estimación por Mínimos Cuadrados Recursivos Estándar (RLS)	40
4.3	REDES NEURONALES ARTIFICIALES.....	46
4.3.1	Algoritmo de Retropropagación	50
4.3.1.1	<i>Criterios para detener el algoritmo de retropropagación</i>	54
4.3.2	Modo de Entrenamiento	54
4.3.3	Curvas de Aprendizaje y la Técnica de Validación Cruzada.....	55
	Referencias.....	57
CAPITULO V		
METODOLOGÍA.....		58
5.1	ADQUISICIÓN MULTICANAL DE LOS SONIDOS RESPIRATORIOS.....	59
5.2	PREPROCESAMIENTO DE LA INFORMACIÓN ACÚSTICA MULTICANAL	61
5.3	SIMULACIÓN DE SONIDOS ADVENTICIOS DISCONTINUOS (CREPITANTES).....	62
5.4	GENERACIÓN DE LA BASE DE DATOS.....	67
5.5	TÉCNICAS DE PROCESAMIENTO PARA GENERAR ÍNDICES DE CREPITANCIAS.....	68
5.5.1	Dimensión Fractal	68
5.5.2	Modelo AR y Redes Neuronales Artificiales	68
5.5.3	Modelo AR Variante en el Tiempo	77
5.6	EVALUACIÓN DEL DESEMPEÑO DE LOS ALGORITMOS PROPUESTOS	79
5.7	IMAGENOLOGÍA DE SONIDOS CREPITANTES.....	81
	Referencias.....	82
CAPITULO VI		
RESULTADOS		83
6.1	INFLUENCIA DE LA RELACIÓN DE AMPLITUD ENTRE LA SEÑAL (CREPITANCIA) Y EL RUIDO (SONIDO PULMONAR DE FONDO) EN LA MORFOLOGÍA DEL CREPITANTE .	84
6.2	DETECCIÓN DE CREPITANCIAS POR MEDIO DE DIMENSIÓN FRACTAL.....	86
6.3	DETECCIÓN DE CREPITANCIAS MEDIANTE PARAMETRIZACIÓN CON UN MODELO AR INVARIANTE EN EL TIEMPO Y CLASIFICACIÓN CON REDES NEURONALES	89
6.3.1	Resultados del Entrenamiento por Regiones	91
6.3.2	Resultados del Entrenamiento por Canales	94

6.4	DETECCIÓN Y CONTEO DE CREPITANCIAS CON UN MODELO AR VARIANTE EN EL TIEMPO	104
6.5	DETECCIÓN DE CREPITANCIAS COMBINADAS FINAS Y GRUESAS VÍA EL MODELO AR VARIANTE EN EL TIEMPO	112
6.6	MODELO AR VARIANTE EN EL TIEMPO VS. EXPERTO MÉDICO.....	112
6.7	IMAGEN DE SONIDOS ADVENTICIOS DISCONTINUOS EN CASOS REALES	116
CAPITULO VII		
DISCUSIÓN DE RESULTADOS		118
7.1	ANÁLISIS NO LINEAL MEDIANTE DIMENSIÓN FRACTAL	119
7.2	DETECCIÓN DE CREPITANCIAS MEDIANTE PARAMETRIZACIÓN CON UN MODELO AR INVARIANTE EN EL TIEMPO Y CLASIFICACIÓN CON REDES NEURONALES	120
7.3	DETECCIÓN Y CONTEO DE CREPITANCIAS CON UN MODELO AR VARIANTE EN EL TIEMPO	123
7.4	COMPARACIÓN ENTRE UNA DETECCIÓN COMPUTARIZADA Y UNA VISUAL REALIZADA POR EXPERTOS	124
	Referencias.....	125
CAPITULO VIII		
CONCLUSIONES		126
ANEXO		128

RESUMEN

El análisis de los sonidos respiratorios (SR) normales y anormales representa una alternativa en el apoyo al diagnóstico de diversas enfermedades pulmonares. Esta tesis de maestría se enfoca en la identificación automatizada en registros puntuales de los sonidos adventicios discontinuos (estertores crepitantes), con el objetivo de generar una imagen que refleje la localización y distribución espacial de los sonidos crepitantes. La propuesta de la *Imagenología de Sonidos Discontinuos* es novedosa en el campo de los SR y la información espacio-temporal presente en la imagen generada tiene la ventaja que se encuentra asociada a la función pulmonar.

El procedimiento de análisis de sonidos discontinuos conocido como *Time Expanded Waveform Analysis* (TEWA) es ampliamente utilizado para detectar visualmente las crepitancias y extraer parámetros en el dominio del tiempo que definen su morfología. Sin embargo, la técnica requiere de criterios que a la fecha son poco precisos para identificar a las crepitancias por ejemplo, que la amplitud de las crepitancias sea del doble de la amplitud del sonido respiratorio de base. La técnica presenta limitaciones, ya que en situaciones reales las crepitancias pueden estar traslapadas temporalmente, tener una SNR baja y poseer una forma de onda alterada debido al sonido respiratorio pulmonar de base. Por lo tanto TEWA, desde nuestra experiencia, no es adecuada para producir la imagen de sonidos adventicios, esto es entonces motivo de la presente investigación.

La metodología propuesta para la detección de las crepitancias incluyó 3 esquemas de procesamiento con base en conceptos de técnicas no lineales como el cálculo de la Dimensión Fractal (DF), conceptos de técnicas lineales como la parametrización de la información acústica multicanal con un modelo AR invariante en el tiempo y su clasificación por redes neuronales artificiales (AR-RNA) y el concepto de modelo AR variante en el tiempo (ARVT).

Se han realizado varios esfuerzos para detectar estertores crepitantes, sin embargo, los resultados basan sus medidas de desempeño en información obtenida por un experto médico mediante el procedimiento TEWA. En esta tesis, en una primera fase, se prescindió de la detección por un experto y se generó una base de datos con crepitancias simuladas donde se tiene el control sobre el número, tipo, SNR y distribución espacial de las mismas al ser insertadas en información acústica multicanal proveniente de sujetos sanos.

Los resultados demuestran que en casos simulados y con datos adquiridos de sujetos enfermos, la metodología basada en la técnica de ARVT proporciona los mejores resultados en términos del número de crepitancias y su distribución espacial. En consecuencia, las imágenes de sonidos adventicios generadas por el esquema ARVT fueron las más

cercanas a las imágenes patrón en diferentes condiciones de simulación.

Utilizando casos simulados se logró demostrar el desempeño y las limitaciones que tiene un experto médico en la detección de crepitancias comparado con el método de ARVT. Además, el resultado generado con el esquema ARVT, con información acústica proveniente de un sujeto con Neumonía Intersticial Difusa, correlaciona con el número de estertores que el médico especialista, aplicando TEWA y su experiencia clínica, es capaz de detectar.

Los resultados de la tesis se presentaron en el congreso internacional de la IEEE-EMBS que se llevó a cabo en la ciudad de Vancouver en el 2008 mediante dos participaciones, una presentación oral y otra en sesión de posters. Se anexan copias de ambos trabajos al final de la presente tesis.

CAPITULO I

INTRODUCCIÓN

1.1 LA AUSCULTACIÓN PULMONAR

La auscultación de los sonidos respiratorios constituye uno de los procedimientos clínicos de relevancia para la evaluación del estado de salud de los pacientes. Los sonidos producidos por la actividad respiratoria humana se consideran importantes desde que Laënnec, en 1819, inventó el estetoscopio como una herramienta para mejorar la auscultación pulmonar [1].

Durante la práctica clínica, la auscultación se lleva a cabo moviendo el estetoscopio sobre la superficie torácica en una secuencia que permite la comparación acústica entre el hemitórax derecho y el izquierdo. Al mismo tiempo, se le pide al sujeto que realice una serie de inspiraciones o espiraciones, o se le solicita pronunciar algunas palabras sencillas. Sin embargo, la interpretación de los sonidos respiratorios ha sido subjetiva, dependiendo de la experiencia del médico y de su sistema auditivo en conjunto con la respuesta en frecuencia del estetoscopio.

La auscultación pulmonar presenta grandes ventajas como procedimiento de diagnóstico clínico, ya que por su facilidad de realización y ausencia de riesgos es factible de practicar en cualquier paciente, independientemente de su edad y estado de salud. Sin embargo, en la práctica clínica tradicional, la descripción de los hallazgos durante la auscultación depende en gran medida de la experiencia y aptitudes del médico, lo que hace difícil tanto el seguimiento de la evolución de los pacientes como las comparaciones entre el mismo o diferentes médicos, debido al carácter subjetivo de dichas descripciones [2].

Se ha reportado que la auscultación clínica involucra tres limitaciones técnicas bien conocidas. La primera limitación, es que la interpretación es esencialmente subjetiva; la segunda es que los sonidos originales y la señal de flujo asociada no se almacenan, limitando la evaluación médica actual así como las posteriores; y la tercera, es que la auscultación espacial torácica se lleva a cabo por comparación de la información acústica que se escucha en diferentes instantes de tiempo [3].

1.2 LOS SONIDOS RESPIRATORIOS Y SU IMPORTANCIA CLÍNICA

Por los elementos anteriormente expuestos con relación a la auscultación, ha surgido el análisis computarizado de los sonidos respiratorios que básicamente se puede considerar un campo multidisciplinario dados los aspectos fisiológicos, patológicos y de procesamiento de información que demanda. A continuación se describen algunos términos utilizados en el desarrollo de esta tesis, basándose en los estándares recomendados por la Sociedad Europea Respiratoria (ERS Task Force Report, Computerized Respiratory Sounds Analysis, CORSA) [4].

Los sonidos generados en pulmones y vías aéreas sanas a través de la respiración normal, difieren en amplitud de acuerdo a la localización donde son registrados y varían con el ciclo respiratorio. El origen de los sonidos pulmonares no es completamente claro y probablemente, están involucrados múltiples mecanismos. El pulmón por si sólo no puede generar sonido respiratorio si no existe flujo de aire; las diferencias de presión entre estructuras dentro del tórax o diferentes niveles de volúmenes pulmonares no pueden por si mismos inducir sonidos en ausencia de flujo de aire. Los sonidos respiratorios son inducidos probablemente por la turbulencia del aire a nivel de lóbulos o segmentos bronquiales. En bronquios pequeños, la velocidad del gas disminuye y llega a ser menor que la velocidad crítica necesaria para inducir turbulencia. Por lo tanto, se cree que el flujo de aire en vías aéreas pequeñas es laminar y silencioso [5].

Los sonidos respiratorios pueden clasificarse en 2 categorías, los sonidos adventicios y los sonidos respiratorios normales. Los sonidos adventicios, son sonidos ruidosos o musicales que no están presentes normalmente en los sonidos respiratorios de un sujeto sano.

En 1985 en la décima convención de la Asociación Internacional de Sonidos Pulmonares (ILSA) se generó un esquema de clasificación que incluía estertores crepitantes finos y gruesos, sibilancias y ronquidos. Cada uno de estos tipos de sonido puede ser descrito acústicamente y no se establece un mecanismo de generación o localización [6].

1.3 TIPOS DE SONIDOS RESPIRATORIOS Y SUS PRINCIPALES CARACTERÍSTICAS

1.3.1 SONIDOS RESPIRATORIOS NORMALES

Los sonidos asociados con la respiración y escuchados en el tórax de una persona sana se denominan sonidos respiratorios normales [6]. Se consideran dos tipos básicos de sonidos respiratorios, los sonidos traqueo-bronquiales, escuchados sobre la pared torácica en la proyección de dichas estructuras y los sonidos pulmonares sanos que se perciben sobre la pared torácica en sitios periféricos a las vías aéreas primarias, principalmente sobre las bases pulmonares [2].

Los sonidos respiratorios normales poseen frecuencias por debajo de los 100 Hz, donde se mezclan con sonidos musculares y sonidos cardiovasculares y son difíciles de distinguir. Además, la energía de los sonidos respiratorios decae rápidamente entre 100 y 200 Hz y con micrófonos especiales y en un cuarto silencioso aún se puede detectar información por encima de 1000 Hz. El espectro de los sonidos respiratorios normales no tiene picos discretos y en consecuencia no es musical. El componente inspiratorio se genera primariamente dentro de lóbulos y segmentos de vías aéreas, mientras que el componente espiratorio proviene de localizaciones proximales [6].

La amplitud de los sonidos respiratorios difiere entre personas y del sitio de auscultación en la superficie del tórax, pero principalmente varía con el cuadrado del flujo de aire. Sin embargo, el efecto del volumen pulmonar en la amplitud de los sonidos se ha estudiado poco [6]. La fase inspiratoria presenta una duración mayor que la fase espiratoria con una relación inspiración/espiración de 2/1 durante la respiración a volumen corriente. Los sonidos respiratorios no presentan características uniformes a lo largo de la superficie pulmonar: en el ápice pulmonar, el sonido es menos intenso durante una inspiración realizada a partir del volumen residual. En contraste, en la base pulmonar el sonido es menos intenso al inicio de la fase inspiratoria y, gradualmente, se incrementa hasta alcanzar su máximo alrededor del 50% de la capacidad vital [5].

El componente inspiratorio del sonido es generado principalmente dentro de vías aéreas lobulares y segmentarias, mientras que el componente espiratorio proviene de vías

aéreas proximales. El sonido es vesicular en las bases pulmonares, ya sea anteriormente, lateralmente o posteriormente. Conforme se avanza hacia los ápices, especialmente en la parte anterior del tórax, el sonido gradualmente toma un carácter bronquial hasta llegar a la tráquea, donde el sonido es claramente bronquial o tubular [6].

Se asume que la turbulencia del aire genera sonidos respiratorios normales, sin embargo, la turbulencia es un fenómeno que depende de la densidad del gas y, el comportamiento de los sonidos respiratorios en respuesta a la respiración de un gas de baja densidad es peculiar. Austrheim y Kraman encontraron que la amplitud del sonido traqueal disminuye en un 45% al respirar He-O₂ mientras que los registros simultáneos sobre varias localizaciones en el tórax disminuyeron del 13 al 16% [6].

1.3.2 SONIDOS TRAQUEALES NORMALES

Los sonidos traqueales normales son sonidos escuchados en la hendidura supra-esternal. La señal de sonido traqueal tiene una intensidad considerable, cubre un amplio rango de frecuencias mayor al rango de los sonidos pulmonares en la pared torácica, con fases respiratorias claramente separadas y una cercana relación al flujo de aire [6]. Un interés muy particular en la clínica por los sonidos traqueales es que representan indicadores de la obstrucción del flujo de vías aéreas superiores y como fuente para la evaluación cualitativa y potencialmente cuantitativa del flujo de aire.

La generación de sonidos traqueales se relaciona principalmente a un flujo turbulento en vías aéreas superiores, incluyendo faringe, glotis y regiones subglóticas. La turbulencia del flujo y la formación de corrientes en la glotis causan fluctuaciones de presión dentro del lumen de las vías aéreas. Las ondas de presión de sonido dentro de las vías aéreas y el movimiento de sus paredes contribuyen a las vibraciones que llegan a la superficie del cuello y son registradas como sonidos traqueales. Los sonidos traqueales se consideran como sonidos respiratorios puros debido a que sufren un filtrado mínimo dado que las fuentes sonoras que lo producen en las vías aéreas superiores se encuentran ubicadas cerca de los sensores en el cuello y al nulo tejido pulmonar que existe en esa región torácica [6].

1.4 SONIDOS ADVENTICIOS

Los sonidos adventicios se asumen como sonidos agregados a los sonidos normales, por lo que el sonido característico traqueobronquial o vesicular se encuentra subyacente. Estos sonidos agregados pueden generar dos tipos de patrones acústicos, los continuos (ronquidos y sibilancias) y los discontinuos (estertores crepitantes).

1.4.1 SIBILANCIAS

Las sibilancias son sonidos pulmonares adventicios del tipo musical, llamados también sonidos *continuos* debido a que su duración es mayor que la de los sonidos *discontinuos* como los estertores crepitantes [3],[6]. De acuerdo a la ATS (American Thoracic Society) la duración de una sibilancia es mayor a 250 ms. [5]. Su contenido en frecuencia está en el rango de 100 Hz y 1000 Hz [6]. Además, la forma de onda de las sibilancias se asemeja a una señal sinusoidal.

Los mecanismos fisiopatológicos que generan las sibilancias no han sido aclarados por completo. El movimiento de las secreciones en las vías aéreas puede ser importante pero probablemente son más significativas las vibraciones de las paredes de las vías aéreas. Grotberg y Gavriely [7] propusieron un modelo en el cual las sibilancias son producidas por oscilaciones en las vías aéreas. Las oscilaciones comienzan cuando la velocidad del flujo de aire alcanza un valor crítico. Este modelo muestra que las sibilancias siempre están acompañadas por limitaciones de flujo, sin embargo una limitación de flujo no es necesariamente acompañada por sibilancias. El sonido musical de las sibilancias es fácilmente reconocible por el oído debido a que sobrepasa en intensidad a los sonidos pulmonares normales [6]. Las sibilancias son frecuentemente audibles en pacientes con la boca abierta o por auscultación de la laringe. Estas pueden ser monofónicas, cuando se escucha un solo tono o polifónicas, cuando se perciben múltiples frecuencias en forma simultánea [5].

La transmisión de sibilancias a través de vías aéreas es mejor que la transmisión a través del pulmón a la superficie de la pared torácica. Los sonidos de alta frecuencia se detectan mejor sobre la tráquea que en el tórax. Las sibilancias de intensidad media a alta son fáciles de identificar en el espectro de potencia de los sonidos respiratorios como picos

[6]. La detección computarizada de sibilancias es posible con algoritmos que relacionan la amplitud de los picos espectrales con la amplitud promedio de sonido pulmonar. Sin embargo, la detección automática por computadora se vuelve difícil cuando la sibilancia es débil pero, aún audible. La cuantificación de sibilancias ofrece la posibilidad del monitoreo no invasivo en asma nocturno [6].

1.4.2 RONQUIDOS

Los ronquidos son sonidos comúnmente producidos y escuchados durante el sueño. Estos sonidos son producidos por vibraciones en las paredes de la orofaringe y se asocian frecuentemente con el síndrome de apnea obstructiva del sueño y con enfermedades cardiovasculares [5]. Los ronquidos están asociados con una limitación en el flujo inspiratorio y con un aumento en la resistencia de vías aéreas [5].

El ronquido es un sonido inspiratorio, aunque pueden aparecer componentes espiratorios en apnea obstructiva del sueño. Los ronquidos son sonidos ruidosos con una intensidad mayor de 50 dB(A), y contienen componentes periódicos con una frecuencia fundamental entre 30 y 250 Hz.

1.4.3 ESTERTORES CREPITANTES

Los estertores crepitantes son sonidos adventicios discontinuos, explosivos y de naturaleza transiente; ocurren frecuentemente en enfermedades cardiorespiratorias. Su duración es menor a 20 ms. y su contenido en frecuencia se encuentra entre 100 y 2000 Hz [4]. Con base en parámetros en el dominio del tiempo las crepitancias se pueden clasificar en finas y gruesas.

Anteriormente se consideraba que las crepitancias eran generadas por burbujas de aire en las vías aéreas o por el movimiento de secreciones intraluminales. La persistencia de crepitancias después de toser y su naturaleza repetitiva en inspiraciones consecutivas, contradice ésta teoría. En la actualidad se piensa que las crepitancias son causadas por una repentina apertura de vías aéreas cerradas anormalmente. De acuerdo a la teoría de Forgas

se desarrolla un gradiente de presión de gas a través de las vías aéreas en la inspiración, que son colapsadas durante la espiración. El sonido del estertor crepitante se genera cuando una vía aérea cerrada repentinamente se abre durante la inspiración o se cierra durante la espiración, induciendo una rápida ecualización de presión de gas y una onda de presión. Cada estertor puede representar una apertura o cerradura abrupta de una simple vía aérea [8]. Nath and Capel encontraron que estertores individuales ocurrían a volúmenes inspiratorios y presiones transpulmonares similares durante inspiraciones consecutivas [9].

Los estertores crepitantes ocurren usualmente cuando aumenta la presión retráctil elástica de los pulmones. El carácter del sonido de los estertores crepitantes posiblemente depende del diámetro de las vías aéreas que se cierran y se abren, mecanismo que se afecta por la fisiopatología del tejido circundante. Además, se ha mostrado experimentalmente que las características de los estertores crepitantes varían con el tamaño de las vías aéreas, donde vías aéreas pequeñas producen estertores de corta duración. Los estertores crepitantes presentes al comienzo de la inspiración probablemente se generan en vías aéreas más proximales que los estertores crepitantes que aparecen al final de la inspiración [8].

Los alvéolos basales de un pulmón normal se desinflan a un volumen residual y se inflan al final de la inspiración; las vías aéreas basales son las primeras en cerrar hacia el final de la espiración. Por lo tanto, los estertores aparecerían más frecuentemente en la etapa temprana de una enfermedad en áreas basales de los pulmones. Con el envejecimiento, cuando la presión retráctil elástica de los pulmones disminuye la incidencia de estertores crepitantes aumenta [8]. Los estertores crepitantes aparecen en enfermedades como la fibrosis pulmonar donde regularmente son finos, repetitivos y aparecen al final de la inspiración. La obstrucción de vías aéreas presenta estertores crepitantes gruesos, menos repetibles y aparecen al principio de la fase de inspiración. En situaciones relacionadas con fallas cardíacas, los estertores crepitantes se presentan de la mitad al final de la inspiración y principalmente, son estertores crepitantes gruesos. En la asbestosis los estertores aparecen primero en zonas pulmonares basales y después, cuando la enfermedad progresa aparecen en zonas pulmonares superiores.

El número de estertores por respiración está asociado con la severidad de la enfermedad en pacientes con desórdenes pulmonares intersticiales. Además, la forma de onda y el tiempo de aparición de los estertores crepitantes pueden ser significativos

clínicamente en diferentes enfermedades cardiorespiratorias. En la figura 1.1 se muestra una señal de sonido respiratorio con estertores crepitantes.

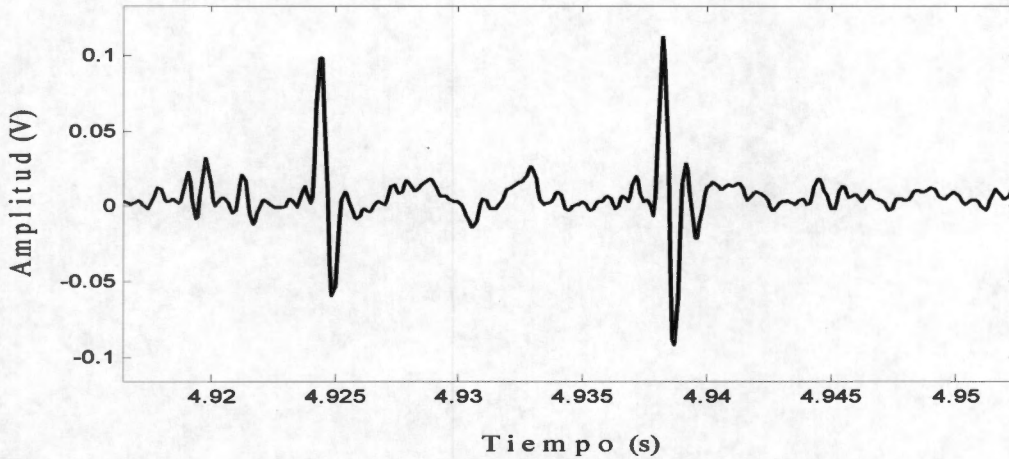


Figura 1.1. Ejemplo de una señal con estertores, tomada de un paciente con alveolitis alérgica extrínseca

Se han detectado estertores crepitantes en sonidos respiratorios de sujetos normales; Thacker y Kraman [10] encontraron crepitancias en sonidos pulmonares de voluntarios normales registrados sobre la cara anterior del tórax durante la inspiración comenzando a un nivel de volumen residual, pero no se encontraron cuando comenzaron a un nivel de capacidad residual funcional. Los estertores crepitantes variaron considerablemente en tono y duración. Al parecer, los estertores crepitantes no siempre son indicadores de procesos pulmonares patológicos, además, su ocurrencia depende de los volúmenes espiratorios durante la auscultación.

Un resumen de la clasificación y de las características de los diferentes sonidos pulmonares se presenta en la Tabla 1.1.

La información presentada en este capítulo es un panorama general sobre los sonidos respiratorios y sus características, remarcando su importancia clínica. El trabajo de la presente tesis se enfoca en los estertores crepitantes, tratando de generar una imagen acústica torácica que refleje su localización. Por esta razón es importante mencionar algunos aspectos relevantes de los estertores crepitantes, como su origen, su contenido en frecuencia, su morfología, etc., con el objetivo de entender su comportamiento y así poder analizarlos.

Tabla 1.1

Clasificación y características de los diferentes sonidos pulmonares [6].

Sonidos Respiratorios	Mecanismos de Producción	Origen	Características Acústicas	Relevancia clínica
<i>Sonidos Básicos</i>				
Sonido normal pulmonar	Vórtices de flujo turbulento, Mecanismos desconocidos	Espiración: vías aéreas centrales, Inspiración: vías aéreas lobulares a segmentales	Sonido sometido a un filtro pasa bajas (< 100 a > 1000 Hz)	Ventilación regional, calibre de vías aéreas
Sonido normal traqueal	Flujo turbulento	Faringe, laringe, tráquea, vías aéreas superiores	Sonido con resonancias (<100 a >3000 Hz)	Configuración de vías aéreas superiores
<i>Sonidos Adventicios Continuos</i>				
Sibilancias	Oscilaciones de la pared de vías aéreas, dispersión de un vórtice	Vías aéreas centrales y periféricas	Senoidal (~100 a > 1000 Hz) Duración típica > 80 ms	Obstrucción de vías aéreas, Flujo limitado.
Ronquidos	Ruptura de capas de fluido, vibraciones en las paredes de vías aéreas	Vías aéreas superiores	Serie de senoidales rápidamente amortiguadas (< 300 Hz y duración > 100 ms)	Secreciones, colapso de vías aéreas anormales.
<i>Sonidos Adventicios Discontinuos</i>				
Estertores	Relajación – tensión de paredes de vías aéreas	Vías aéreas periféricas y centrales	Deflexión de ondas rápidamente amortiguadas (duración < 20 ms)	Cierre de vías aéreas, secreciones.

REFERENCIAS

-
- [1] R.T.H. Laënnec, *De l'auscultation mediate ou traité du diagnostic de maladies des poumons et du Coeur, fondé principalement sur ce nouveau moyen d'exploration*. Paris: Brosson et Chaudé, 1819.
- [2] González Camarena R, Carrasco Sosa S, Charleston Villalobos S, Gaitán González JM, Ramos Martínez GR, Chi Lem G, Pérez Padilla JR, "Sistema para el Análisis Cuantitativo de los sonidos respiratorios," *Revista Mexicana de Ingeniería Biomédica*, vol. 18, no. 2, pp. 9-22, 1997.
- [3] Charleston Villalobos S, Cortés Rubiano S, González Camarena R, Chi-Lem G, AljamaCorrales T, "Respiratory acoustic thoracic imaging (RATHI): assesing deterministic interpolation techniques," *Medical & Biological Engineering & Computing*, vol. 42, no. 5, pp. 618-626, 2004.
- [4] A. R. A. Sovijärvi, L. P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmaso, C. Sacco, M. Rossi, J. E. Earis, "Definition of terms for applications of respiratory sounds." *European Respiratory Review*, vol. 10, no. 77, pp. 597-610, 2000.
- [5] A. R. A. Sovijärvi, L. P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmaso, C. Sacco, M. Rossi, J. E. Earis, "Characteristics of breath sounds and adventitious respiratory sounds," *European Respiratory Review*, vol. 10, no. 77, pp. 591-596, 2000.
- [6] Pasterkamp H, Kraman S, Wodicka G, "Respiratory Sounds. Advances beyond the stethoscope," *American Journal of Respiratory and Critical Care Medicine*, vol. 156, pp. 974-987, 1997.
- [7] Grotberg JB, Gavriely N., "Flutter in collapsible tubes: a theoretical model of wheezes," *J. Appl. Physiol.*, vol. 66, pp. 2262-2273, 1989.
- [8] Piirilä P., Sovijärvi A.R.A., "Crackles: recording, analysis and clinical significance," *European Respiratory Journal*, vol. 8, pp. 2139-2148, 1995.
- [9] Nath AR, Capel LH. "Inspiratory crackles and mechanical events of breathing." *Thorax*, vol. 29, pp. 695-698, 1974.
- [10] Thacker, R. E., S. S. Kraman. "The prevalence of auscultatory crackles in subjects without lung disease," *Chest* vol. 81, pp. 672-674, 1982.

CAPITULO II

2.1 ANTECEDENTES EN EL ANÁLISIS DE CREPITANCIAS

La detección automática de sonidos adventicios discontinuos (crepitancias) y su extracción del sonido de base pulmonar, anteriormente denominado sonido vesicular, representan una tarea compleja, la cual ha sido abordada por diferentes investigadores. En esta sección se describirán varios de los esfuerzos realizados en este sentido y se mencionarán algunas de sus limitaciones para establecer el contexto de la propuesta en esta tesis.

2.1.1 ANÁLISIS UNICANAL

Uno de los primeros métodos para detectar crepitancias que se ha convertido en un método clásico de referencia, fue el propuesto por Murphy *et al* [1] con base en el análisis visual de la señal de sonido respiratorio a una escala de tiempo fija en 800 mm/seg. Con esta expansión temporal de la señal los autores describieron morfológicamente a los sonidos discontinuos como complejos y de corta duración con una deflexión inicial de poca duración seguida de deflexiones con atenuación de la amplitud. De acuerdo con este trabajo, los estertores crepitantes presentan una deflexión inicial con una amplitud de dos veces la amplitud del sonido pulmonar o vesicular. Además, los parámetros en el dominio del tiempo que utilizaron para caracterizar a las crepitancias fueron: el ancho de la deflexión inicial (IDW), la duración de dos ciclos (2CD) y el ancho de la deflexión final (LDW); con base en estos parámetros las crepitancias fueron clasificadas como finas o gruesas.

Con relación a la separación automatizada de crepitancias exclusivamente del tipo fino del sonido pulmonar de base o vesicular, Ono *et al.* propusieron considerar las características estadísticas de las señales involucradas [2]. Por lo tanto, en el trabajo se asume que la señal a procesar es la suma de dos tipos de señales: una señal estacionaria que puede representarse por un modelo autorregresivo y una señal no estacionaria, compuesta

de ondas impulsivas aleatorias que los autores asumen con una frecuencia de ocurrencia baja. El método propuesto es un un filtro digital que combina una función no lineal en paralelo con un filtro de predicción. El filtro de predicción utiliza el algoritmo LMS para adaptar los coeficientes del modelo autorregresivo. El desempeño de la metodología propuesta se evaluó mediante cuatro casos de sujetos enfermos con una relación señal a ruido favorable para las crepitancias. La medida de desempeño fue la correspondencia temporal entre la posición estimada de la crepitancia con la posición determinada en forma visual. Una limitación del procesamiento propuesto es que se requiere de la definición empírica de dos parámetros que establecen la forma matemática de la función no lineal. Además, los parámetros no son del tipo adaptable y permanecen constantes al procesar la señal adquirida.

El concepto de estacionaridad espectral se utiliza en otra investigación para la detección automática de crepitancias con base en el análisis del sonograma de sonidos pulmonares [3]. El estudio fue desarrollado considerando la información acústica de 10 pacientes con alveolitis fibrosante (FA) y 10 con bronquiectasia (BE). Los padecimientos FA y BE presentan crepitancias finas y gruesas, respectivamente. La idea fundamental es que los picos transientes, es decir, bordes en la imagen tiempo-frecuencia del sonograma de sonidos pulmonares indican crepitancias. El espectrograma vía la transformada de Fourier de tiempo corto se calculó dividiendo la señal de tiempo en segmentos de 128 puntos y considerando un traslape entre ellos del 50%. Además, el tamaño de los segmentos usados en el cálculo de la FFT fue aproximadamente igual a la duración de las crepitancias para obtener una mejor resolución en el tiempo. Para calcular el número de crepitancias se aplicó una máscara espacial a la imagen, donde los bordes verticales con mayor intensidad se contabilizaron siempre y cuando éstos sobrepasen un umbral predeterminado. El número de crepitancias detectadas se comparó con el obtenido por dos observadores que utilizaron el método de TEWA de Murphy. Los resultados del método con respecto a un observador fueron 486 crepitancias detectadas, 75 crepitancias no detectadas y 105 crepitancias detectadas erróneamente. La sensibilidad promedio del método automático con respecto a los 2 observadores fue de 89% para pacientes con FA y 80% para pacientes con BE. Sin embargo, el espectrograma o sonograma presenta un compromiso entre la resolución en tiempo y en frecuencia de los eventos que se quieren detectar, es decir, es necesario

determinar de antemano el ancho de la ventana de análisis y este ancho no es adaptable a lo largo de la señal a analizar.

Hadjileontiadis *et al.* [4] propusieron una combinación del filtro no-lineal propuesto por Ono *et al.* [2] con estadística de alto orden. Los autores mencionan que el uso de estadística de alto orden en la estimación del orden del modelo AR y sus coeficientes provee una estimación más confiable de la parte estacionaria de la señal de entrada. Las motivaciones primordiales del uso de estadística de tercer orden fueron: a) la eliminación de ruido Gausiano, dado que su estadística de tercer orden es igual a cero y b) la preservación de la fase verdadera de la señal modelada, en contraste con la estadística de segundo orden que si elimina la fase. Los autores aplicaron la metodología propuesta en registros de sujetos enfermos en los que previamente se determinó el número y posición de los estertores crepitantes por expertos vía TEWA. Sólo se presentan dos casos en forma gráfica donde nuevamente los estertores crepitantes presentan una gran amplitud con relación al sonido pulmonar de base o vesicular, es decir, la SNR para las crepitancias es buena. Para examinar la contribución de su propuesta probaron el filtro no lineal en señales adquiridas que contaminaron con ruido Gausiano de color de media cero y varianza de 0.03. Encontraron que el ruido añadido no influyó en la estimación del orden del modelo AR, y además, que la separación de las partes estacionarias y no estacionarias de las señales con ruido mostraron resultados similares a los obtenidos con las señales sin ruido. El análisis de ruido sólo se presenta en una señal adquirida que contiene squawks o graznidos. Hay que hacer notar que los autores no mencionan en que condiciones se esperaría que las señales respiratorias adquiridas incluyeran ruido Gausiano de color. Además, no presentan comparaciones de sus resultados con la propuesta original de Ono *et al.* [2] para corroborar la hipótesis de una estimación más confiable de la parte estacionaria de la señal de entrada.

En otro trabajo Hadjileontiadis *et al.* emplearon un algoritmo que combina el análisis multiresolución vía transformada *wavelet*, utilizando una función ortonormal de Daubechies, con una umbralización *dura* para crear un filtro que separa señales estacionarias de no estacionarias (WTST-NST) [5]. La idea es que los picos correspondientes a las crepitancias presentan componentes con gran amplitud en varias escalas *wavelets* y que la magnitud del *ruido de fondo*, es decir, del sonido vesicular se disminuye lentamente conforme se incrementa la escala. La metodología incluye la

definición de un umbral en cada escala, si los coeficientes *wavelets* son mayores al umbral, los coeficiente pertenecen a un estertor crepitante, de lo contrario pertenecen al sonido de base. El umbral se establece como 3 veces la varianza de los coeficientes en cada una de las escalas. Los autores justifican con una serie de pruebas la selección del factor igual a 3 indicando que es el valor que proporciona la mejor separación de ambos componentes. Los autores comparan el desempeño de su algoritmo con el método de Ono *et al.* [2] y reportan que fue mejor el desempeño del filtro WTST-NST; esto podría explicarse debido a que los parámetros del filtro de Ono se definen empíricamente, dependiendo de las características de la señal de entrada. Nuevamente sus casos de estudio presentan señales con crepitancias de amplitud considerable con respecto al sonido vesicular y su patrón de referencia son los resultados obtenidos por un médico experto utilizando el método de TEWA. Además, llama la atención que los autores no realizaron comparaciones con el método que ellos mismos formularon en [4].

Tratando de realizar una separación en tiempo real de los sonidos adventicios discontinuos y el sonido vesicular, Toliás *et al.* [6] propusieron un filtro con base en reglas difusas para mejorar la separación de señales estacionarias y no estacionarias. El filtro consiste de 2 sistemas de inferencia difusos que operan en paralelo, entrenados con un algoritmo adaptable propuesto por otros autores. Los autores mencionan que su esfuerzo anterior [5] con base en la transformada *wavelet* obtiene buenos resultados de separación de la señal pero no se puede implementar fácilmente en tiempo real. Sus resultados indican que su metodología detecta adecuadamente las crepitancias finas, gruesas y los denominados *squawks*. Sin embargo, su patrón de oro es el médico especialista que aplicando el método visual TEWA a una base de datos extrae los parámetros de IDW, 2CD y LDW para preclasificar los sonidos crepitantes. Sin embargo, para que un médico especialista logre detectar adecuadamente un crepitante es necesario que éste guarde una relación de amplitud con respecto al sonido vesicular de fondo, por lo menos de 2 veces, y conserve una morfología adecuada, de otra forma los crepitantes no se clasifican adecuadamente.

Kiyokawa *et al.* [7] presentan un estudio donde se simuló matemáticamente crepitancias finas y gruesas, y se agregaron al sonido respiratorio normal registrado a 1 L/s y 2 L/s. Una de las intenciones del trabajo consistió en definir la capacidad de la audición

humana para detectar la presencia de crepitancias. Los resultados indicaron que la detección de crepitancias fue más difícil con sonido respiratorio con intensidad alta; es decir, se detectaron menos crepitancias con un sonido a un flujo de 2 L/s que a flujos menores. Además, los autores reportan que el tipo de crepitancias tiene un efecto en el índice de detección siendo las crepitancias gruesas las más difíciles de detectar. Otro factor importante fue la amplitud de las crepitancias, ya que las crepitancias con amplitud baja no fueron detectadas. Este trabajo es importante dado que señala la dificultad auditiva y visual para la detección de estertores crepitantes en condiciones de flujo respiratorio alto y cuando existe traslape temporal de información de los sonidos crepitantes subsecuentes.

Hadjileontiadis *et al.* [8] utilizan la dimensión fractal (DF), calculada vía el algoritmo de Katz, como una medida de la complejidad de la señal en el dominio del tiempo. Su algoritmo identifica la ubicación y duración de los sonidos adventicios pulmonares e intestinales mediante los picos de la DF, es decir, mediante el cambio de la complejidad de la señal debido a no-estacionariedades. Los autores mencionan que el atractivo de su propuesta es su simplicidad, dado que para el cálculo de la DF se opera directamente en el dominio original de la señal. Una vez que se obtiene la DF vía el algoritmo de Katz, se intenta mejorar la relación señal a ruido en la señal de DF obtenida mediante un algoritmo adicional que realiza el rasurado o limpieza de la señal. Sus resultados demuestran un 100% de detección tomando como base el número de crepitantes detectados por los expertos vía TEWA utilizando la misma base de datos de los esfuerzos previos de los mismo autores. Posteriormente, Hadjileontiadis *et al.* [9],[10] combinan la dimensión fractal con la descomposición por transformada wavelet y una umbralización de la DF; a pesar de que habían indicado que la DF por sí sola es adecuada. Para el cálculo de la DF, en cada una de las escalas wavelets, utilizan dos algoritmos, el algoritmo clásico de Katz y el de Sevcik. Los autores comparan una serie de algoritmos y llegan a la conclusión de que el algoritmo con base en la descomposición *wavelets* y el cálculo de la DF es el mejor, lo anterior en términos de una base de datos preclasificada auditiva y visualmente por expertos médicos.

En un trabajo relacionado con la clasificación de sonidos respiratorios, Xiaoguang Lu *et al.* [11] desarrollaron un sistema automático para la detección y clasificación de crepitancias. Utilizaron un filtro de separación de señales estacionarias y no estacionarias con base en la transformada *wavelet*, el cual aísla las crepitancias del sonido vesicular de

fondo. El análisis consiste en 3 pasos: 1) se aplica un filtro para suprimir el ruido residual estacionario en la señal no estacionaria a través de un procedimiento de eliminación de ruido; 2) se detectan las crepitancias con base en un umbral utilizando dimensión fractal, calculada vía el algoritmo de Sevcik, y 3) las crepitancias extraídas se clasifican como finas o gruesas. Uno de los aspectos relevantes de la propuesta es la definición del umbral en términos de la amplitud del crepitante; según los autores el crepitante presenta de 4 a 10 deflexiones positivas y negativas, que se incrementan gradualmente en magnitud con el tiempo hasta alcanzar un máximo para después disminuir. Los autores generaron un algoritmo que elimina pequeñas oscilaciones en la señal de DF y que enfatiza los cambios de pendientes al involucrar el cálculo de su derivada; posteriormente, generan una señal con valores de 1, 0 ó -1 de acuerdo a los cambios de pendiente de la DF. Una vez que las crepitancias se encuentran aisladas, Xiaoguang Lu *et al.* [11] realizan la clasificación de las crepitancias de 3 formas diferentes: 1) En el dominio de la representación tiempo-frecuencia, las crepitancias se clasificaron con base a la duración del ancho de la deflexión inicial (IDW), la duración de 2 ciclos (2CD), el ancho de la deflexión total, el ancho de la deflexión máxima, la frecuencia pico y el ancho de banda Gaussiano, 2) con la función matemática de Prony, que corresponde a una senoide amortiguada y 3) con base a la transformada *wavelet*; definiendo una *wavelet* madre con una morfología igual al crepitante. Los resultados con la transformada *wavelet* presentaron el mejor desempeño y la función matemática de Prony presentó la clasificación más pobre. Nuevamente se utilizan señales adquiridas que han sido previamente analizadas por expertos médicos para señalar, vía el uso de TEWA, la posición y el carácter de las crepitancias encontradas.

Otros investigadores, con el objetivo de separar las crepitancias del sonido vesicular, han trabajado con la transformada *wavelet packet*, utilizando la *wavelet* 8 de Daubechies y con 5 niveles de descomposición [12]. Nuevamente, la idea es separar el componente estacionario del no estacionario. En este esfuerzo se definieron dos umbrales en términos de las varianzas de los coeficientes en cada escala y dos factores definidos experimentalmente en 0.75 y 2. Los autores compararon sus resultados con los obtenidos por el método de Hadjileontiadis *et al.* [5]. Los resultados con señales de sujetos enfermos, que presentan estertores crepitantes de magnitud importante, son cercanos a los obtenidos por [5]. Sin embargo, la prueba más atractiva la realizan sobre señales provenientes de sujetos normales

donde se observa que el método de [5] produce en la señal no estacionaria estimada muchas transiciones en comparación con la señal no estacionaria estimada por los autores, figura 2.1.

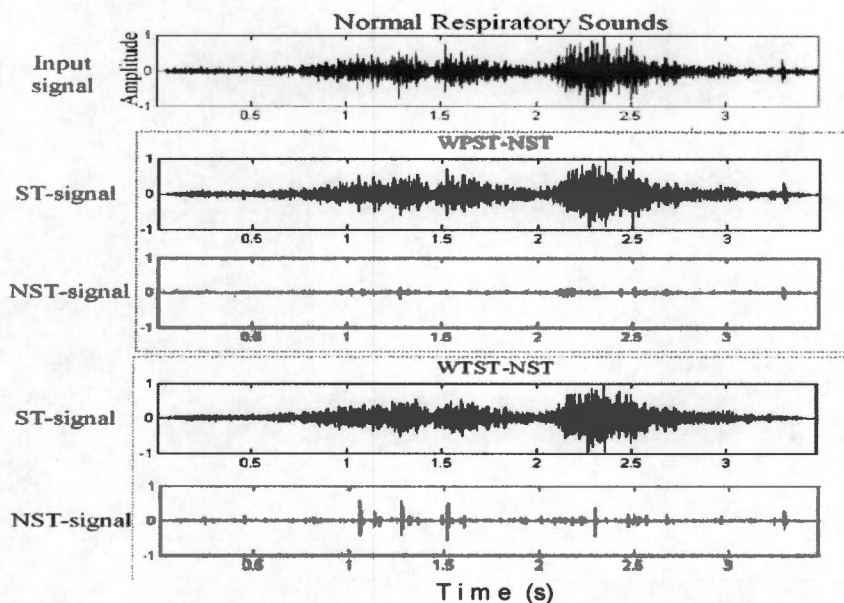


Figura 2.1. Figura tomada de [13]. Comparación entre el desempeño de WPST-NST y WTST-NST para sonidos normales.

En la Universidad Bogazici en Turquía, Kahya *et al.* [13] utilizaron diferentes conjuntos de características para enfrentar el problema de clasificación de señales de sonido respiratorio, proveniente de un sujeto sano o enfermo. Los sonidos fueron adquiridos a través de un micrófono electret en la zona posterior basal del sujeto, la posición del micrófono fue seleccionada por el médico en términos de una zona de mejor auscultación. Se utilizaron 3 conjuntos diferentes de características a) Los coeficientes de un modelo AR, b) Los coeficientes AR mas parámetros temporales de las crepitancias (IDW, 2CD), volumen de ocurrencia (VO) y un parámetro para indicar la presencia o ausencia de crepitancias, 1 ó -1 respectivamente, y c) Parámetros derivados de la transformada *wavelet* de la señal de sonido respiratorio, la cual es descompuesta diádicamente en subbandas de frecuencia usando una *wavelet* Daubechies de orden 8. En este último caso, una vez obtenidas las señales en las diferentes sub-bandas, se efectúa la reconstrucción al dominio original de la señal y se aplica un modelo AR de orden cuarto a cada una de ellas.

Posteriormente, los autores emplearon 2 tipos de clasificadores y analizaron su desempeño, el del vecino más cercano (Nearest Neighbour K-NN) y una red neuronal artificial de tres capas con algoritmo de aprendizaje por retropropagación. Los resultados indican que ambos clasificadores mejoran su desempeño al considerar los coeficientes AR en combinación con parámetros temporales del sonido crepitante.

2.1.2 ANÁLISIS MULTICANAL

Desde el punto de vista de análisis multicanal, el grupo de investigación integrado por Martínez, Charleston, Aljama y González [14], realizó un trabajo de clasificación de sonidos respiratorios, que involucra la presencia de estertores crepitantes, en el cual se lleva a cabo la evaluación de la información acústica multicanal a través de procesamiento digital de señales en enfermedades intersticiales difusas. En éste trabajo se realiza un esfuerzo por analizar y ponderar la ventaja de información acústica respiratoria multicanal sobre la monocal. En el trabajo se analizaron y compararon diferentes técnicas de parametrización de la información acústica como: medidas clásicas en el dominio de la frecuencia, eigenvalores de la matriz de covarianza, modelos autorregresivos monocanales y multicanales. Después, a través de un clasificador del tipo supervisado, se diferenciaron sonidos normales y anormales. Los resultados indican que el modelo autorregresivo monocal logra que exista una mejor clasificación de los datos en la etapa de validación de la red neuronal artificial (85.43 %).

Otra propuesta para el análisis de los sonidos respiratorios es la descomposición de modos empírica (EMD), Charleston *et al.* [15] exploraron el uso de ésta técnica para enfatizar la información de las crepitancias, investigando su desempeño en datos simulados y reales. Los autores adquirieron sonidos respiratorios con un arreglo de sensores de 5x5 y además, simularon crepitancias con la función matemática propuesta por Kiyokawa *et al.* [7]. Sus simulaciones incluyeron diferentes escenarios donde se presentaban crepitancias aisladas múltiples, crepitancias temporalmente traslapadas y combinación de crepitancias finas y gruesas. Las simulaciones permiten manipular el número de crepitancias, relación crepitancia a sonido vesicular de fondo, entre otras características. Una de las principales

observaciones de los autores establece que la información de las crepitancias, relacionada con su número y posición, se encontró principalmente distribuida en los modos de más alta oscilación, IMF1 a IMF4. Además, los autores indican que, dependiendo de la relación señal a ruido (RSR), las crepitancias finas son más fáciles de discriminar del sonido de base respiratorio que las crepitancias gruesas. Una de las condiciones más difíciles es cuando las crepitancias se traslapan temporalmente y en este caso la información aparece desde el IMF1, pero sus componentes individuales no pueden ser diferenciados.

Como conclusión, en esta revisión de antecedentes, que no intenta ser completamente exhaustiva, se observa que se han propuesto diferentes algoritmos para detectar y separar los estertores crepitantes del sonido conocido como de base o vesicular. Es interesante observar que la mayoría de los trabajos utilizan como referencia al experto clínico, aún y cuando se ha demostrado que le es difícil detectar auditiva y visualmente las crepitancias, sean estas finas o gruesas. Considerando los trabajos de investigación arriba citados y de los realizados por el grupo de trabajo del Laboratorio de Imágenes y Procesamiento de Señales en Ingeniería Biomédica (LIPSIB) se establecieron las técnicas de procesamiento digital a utilizar en el desarrollo del presente trabajo.

REFERENCIAS

-
- [1] R.L. Murphy, S. K. Holford, W. C. Knowler, "Visual lung sound characterization by time-expanded waveform analysis," *N. England. J. Med.*, vol. 296, no. 17, pp. 968-971, 1977.
- [2] M. Ono, K. Arakawa, M. Mori, T. Sugimoto, H. Harashima, "Separation of Fine Crackles from Vesicular Sounds by a Nonlinear Digital Filter," *IEEE Trans. Biomed. Eng.*, vol. 36, no. 2, pp. 286-291, 1989.
- [3] T.K. Kaisla, A. Sovijärvi, P. Piirilä, H.M. Rajala, S. Haltsonen, and T. Rosqvist, "Validated method for automatic detection of lung sounds crackles," *Med. & Biol. Eng. & Comput.*, vol. 29, pp. 517-521, 1991.
- [4] L. J. Hadjileontiadis and S. M. Panas, "Nonlinear separation of crackles and squawks from vesicular sounds using third-order statistics," *Proc. IEEE-EMBS 18th Conf.*, vol. 5, pp. 2217-2219, Amsterdam, The Netherlands, 1996.
- [5] L. J. Hadjileontiadis and S. M. Panas, "Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter," *IEEE T. Biomed. Eng.*, vol. 44, no. 12, pp. 1269-1281, 1997.
- [6] Y. A. Tolia, L. J. Hadjileontiadis, and S. M. Panas, "Real-time separation of discontinuous adventitious sounds from vesicular sounds using a fuzzy rule-based filter," *IEEE T. Inform. Technol. Biomed.*, vol. 2, no. 3, pp. 204-215, 1998.
- [7] H. Kiyokawa, M. Greenberg, K. Shirota, and H. Pasterkamp, "Auditory detection of simulated crackles in breath sounds," *Chest*, vol. 119, no. 6, pp. 1886-1892, 2001
- [8] L. J. Hadjileontiadis, I. T. Rekanos, "Detection of explosive lung and bowel sounds by means of fractal dimension," *IEEE Signal Proc. L.*, vol. 10, no. 10, pp. 311-314, 2003.
- [9] L. J. Hadjileontiadis, "Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding Part I: Methodology," *IEEE T. Biomed. Eng.*, vol.52, no. 6, pp. 1143-1148, June 2005.
- [10] L. J. Hadjileontiadis, "Wavelet-based enhancement of lung and bowel sounds using fractal dimension thresholding Part II: Application results," *IEEE T. Biomed. Eng.*, vol.52, no. 6, pp. 1143-1148, June 2005.
- [11] X. Lu, M. Bahoura, "An automatic system for crackles detection and classification," *IEEE CCECE/CCGEI*, pp. 725-729, Ottawa, Mayo 2006.
- [12] M. Bahoura, X. Lu, "Separation of crackles from vesicular sounds using wavelet packet transform," *IEEE ICASSP*, pp. II-1076 - II-1079, 2006.

-
- [13] Y. Kahya, M. Yeginer, B. Bilgic, "Classifying respiratory sounds with different feature sets," *Proceedings of the 28th IEEE*, pp. 2856-2859, New York City, USA, Aug 30-Sept 3, 2006.
- [14] H.G. Martinez-Hernandez, T. Aljama-Corrales, R. Gonzalez-Camarena, S. Charleston-Villalobos, G. Chi-Lem, "Computerized classification of normal and abnormal lung sounds by multivariate linear autoregressive model," *Proc. 27th Ann .Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1464-1467, Shanghai, China 2005.
- [15] S. Charleston, R. González, G. Chi-Lem, T. Aljama, "Crackle sounds analysis by empirical mode decomposition," *Proc. Int. Conf. IEEE-EMBS*, pp. 40-47, 2007.

CAPITULO III

3.1 PLANTEAMIENTO DEL PROBLEMA

El análisis de los sonidos respiratorios ha sido de gran interés por varias décadas con la idea de proporcionar una herramienta de diagnóstico no invasivo y de carácter objetivo de las enfermedades pulmonares. En particular, los sonidos respiratorios anormales como los estertores crepitantes se han logrado asociar con diversas enfermedades pulmonares y cardíacas, y su utilidad clínica se ha evidenciado a través de diversos trabajos. Entre las características más relevantes de los estertores crepitantes que permiten diferenciar patologías y su severidad se encuentran: la forma de onda, el contenido espectral, su distribución en las diferentes zonas pulmonares, el número y su aparición durante el ciclo respiratorio, ya sea en la fase inspiratoria o espiratoria.

Con el propósito de efectuar una clasificación de los estertores crepitantes se han realizado varios esfuerzos para detectarlos y contarlos automáticamente. Sin embargo, por diversas razones, ninguno de los métodos descritos en la sección de antecedentes es de uso común, ya sea en el campo de la investigación o en la clínica. Al evaluar el desempeño de diversas propuestas de procesamiento, el método que se ha considerado como un estándar de oro es el método desarrollado por Murphy *et al.* denominado TEWA (*Time expanded waveform analysis*). Sin embargo, la aplicación de TEWA es tediosa y requiere de experiencia por parte del médico para la extracción en el dominio del tiempo de los parámetros conocidos como 2CD, IDW y LDW. Adicionalmente, el método de TEWA depende de diversos factores entre los que se incluyen: a) el carácter no-estacionario del sonido respiratorio de base o pulmonar y de los estertores crepitantes, b) la relación señal a ruido entre los estertores crepitantes y el sonido respiratorio pulmonar (SNR), c) el traslape temporal de los estertores crepitantes y d) la modificación de la forma de onda de los estertores crepitantes debido al sonido respiratorio pulmonar.

Además, en forma reciente se ha propuesto el uso de sistemas de adquisición multicanal, en un esfuerzo por contar con información espacio-temporal de los sonidos respiratorios normales y anormales. El propósito es generar un procedimiento similar al que se realiza en la práctica clínica, donde el médico especialista solicita al paciente

realizar una serie de maniobras para auscultar diversas zonas pulmonares y generar un probable diagnóstico con base en la información acústica escuchada. Sin embargo, la cantidad de información adquirida por los sistemas multicanales propicia que la aplicación de métodos como TEWA no sean fáciles de aplicar.

Por otro lado, la obtención de imágenes anatómicas o funcionales es de gran interés para diferentes aplicaciones médicas, y en el estudio de la actividad pulmonar comienzan a surgir fuertemente. Por ejemplo, se han desarrollado mapas de contorno, usando señales obtenidas con 2 micrófonos en lugares simétricos y con adquisición simultánea. Otro avance en este campo es la generación de imágenes acústicas que permite profundizar en los mecanismos subyacentes de la actividad respiratoria normal y anormal. Por las razones anteriormente expuestas, en el grupo de investigación en Acústica Respiratoria de la UAM-I surge la iniciativa de explorar en mayor detalle la detección multicanal de estertores crepitantes y se propone la generación de una imagen de sonidos adventicios discontinuos para apoyar el diagnóstico médico.

3.2 HIPÓTESIS

Con base en el procesamiento digital de la información acústica multicanal de sonidos pulmonares es posible formar una imagen del tipo *funcional* acústica que refleje la presencia y relevancia de los sonidos adventicios discontinuos en sujetos con NID (Neumopatías Intersticiales Difusas).

3.3 OBJETIVO

3.3.1 OBJETIVO GENERAL

Proporcionar al médico especialista una imagen de sonidos pulmonares que refleje la localización o distribución espacial de sonidos adventicios discontinuos (sonidos estertores crepitantes), con base en un índice que considere: a) el número estimado de estertores crepitantes o b) el porcentaje de ventanas temporales de información acústica clasificadas como pertenecientes a la clase *enferma*.

3.3.1 OBJETIVOS ESPECÍFICOS

1. Establecer esquemas de procesamiento digital que permitan generar un índice que muestre la localización y distribución torácica de estertores crepitantes.
2. Generar imágenes de sonidos adventicios discontinuos en escenarios simulados y reales de sonidos pulmonares.
3. Evaluar el desempeño de los esquemas de procesamiento digital planteados.

CAPITULO IV

ELEMENTOS TEÓRICOS EN LA GENERACIÓN DE IMÁGENES DE SONIDOS CREPITANTES

El problema que se aborda en esta tesis requiere detectar los estertores crepitantes y definir algún índice que refleje su número o importancia, así como su distribución espacial sobre la superficie torácica de un sujeto. Las características de los crepitantes, como ya se mencionó, incluyen un comportamiento no estacionario por lo que es necesario contender con esta característica. Además, y de acuerdo con la experiencia que se tiene en el grupo de investigación de Acústica Torácica de la UAMI, se conoce que el modelo AR monocanal proporciona buenos resultados en la clasificación de señales que contiene estertores crepitantes. Lo anterior, aunado con las investigaciones reportadas en el capítulo de antecedentes hacen del modelo AR una herramienta prometedora para la detección de información relacionada con el crepitante. Es necesario hacer énfasis en que la mayoría de los esfuerzos de procesamiento digital para la detección de crepitantes han utilizado un número limitado de sensores acústicos. En contraste, en esta tesis se explora el procesamiento multicanal de información acústica, colocando los sensores en posiciones seleccionadas de acuerdo a elementos anatómicos, para la detección de sonidos crepitantes y la generación de una imagen acústica relacionada con su número o importancia.

En esta sección se establecen los conceptos teóricos básicos de la técnica de dimensión fractal que se ha referido como una técnica, de acuerdo con los antecedentes, que podría proporcionar información valiosa relacionada con el estertor crepitante. Adicionalmente, se realiza una revisión de los aspectos relevantes del modelo AR, ya sea invariante o variante en el tiempo, que proporciona una forma de parametrización de la información acústica multicanal. Con los vectores de características obtenidos por las técnicas anteriores, una red neuronal permitirá la clasificación de la información en dos clases, estertor crepitante y sonido de base. Por lo tanto, la teoría de redes neuronales artificiales se expone en este capítulo en conjunto con los elementos necesarios para entender los procesos de entrenamiento y validación de una arquitectura seleccionada. Posteriormente, en la sección de metodología se indicará la forma en que los diferentes elementos teóricos aquí expuestos

se combinaron y adecuaron para obtener la imagen correspondiente a la importancia de los estertores crepitantes en sujetos enfermos.

4.1 ANÁLISIS DINÁMICO NOLINEAL DE UNA SERIE DE TIEMPO: LA DIMENSIÓN FRACTAL

En la ingeniería biomédica las series de tiempo asociadas a eventos fisiológicos se han analizado mediante técnicas de procesamiento derivadas de la teoría de los sistemas lineales. Sin embargo, se acepta que los diferentes sistemas biológicos generan información que tiene un comportamiento no periódico y sus variaciones son del tipo aleatorio. En este sentido, existe un avance significativo en el análisis de las series de tiempo efectuando su procesamiento desde un punto de vista estocástico. Recientemente, ha surgido la idea de que existen situaciones donde las series de tiempo no tienen un comportamiento determinista o estocástico y en consecuencia, se han generado técnicas de procesamiento denominadas técnicas no-lineales para analizar y entender mejor a los sistemas biológicos. Es decir, la salida de sistemas no lineales determinísticos pueden ser del tipo aleatorio. Entre las técnicas no-lineales se encuentra la *Dimensión Fractal* que se utiliza en la presente tesis para detectar la presencia de sonidos adventicios discontinuos.

Las series de tiempo asociadas con fenómenos fisiológicos se pueden concebir como un conjunto ordenado de pares (x_i, y_i) , donde y_i se asocia con la amplitud de la serie en la muestra *i-ésima*, mientras que x_i representa al tiempo discreto. Una característica fundamental de las series de tiempo es que no retroceden en el tiempo y no existen cruzamientos sobre su gráfica, es decir, el tiempo discreto crece monotónicamente. Una forma de clasificar a las series de tiempo es mediante la evaluación de su cercanía a las características de un patrón en específico. Por ejemplo, el análisis de Fourier aproxima a las señales como una serie de senos y cosenos, es decir, se asume un comportamiento de tipo periódico. Existen técnicas de procesamiento de señales que evalúan las características de patrones que no asumen periodicidad en la serie de tiempo. Ejemplo de estas técnicas es la dimensión fractal de un patrón. La dimensión fractal es una medida de la complejidad del patrón- que refleja su extensión espacial, su grado de intrincamiento o su tendencia a llenar un espacio- y que esta fuertemente correlacionado con su forma y dimensionalidad. En

consecuencia, la dimensión fractal se utiliza para cuantificar que tan compleja es una forma de onda o señal.

La dimensión fractal de una serie de tiempo se define de la forma siguiente [1]:

$$DF = \frac{\log(L)}{\log(d)} \quad (1)$$

donde L corresponde a la *longitud total* de la curva mientras que el parámetro d se asocia con el *diámetro* (extensión en el plano) de la curva. Para series de tiempo que cumplen con la propiedad de que son pares ordenados en el plano (x, y) , el tamaño total de la curva es simplemente la suma de las distancias individuales de todos los puntos sucesivos. En la figura 4.1, se ejemplifica el concepto de longitud total de una curva, para esta serie de tiempo constituida de 10 muestras, $L = l_1 + l_2 + \dots + l_9$. En consecuencia, el parámetro L se expresa matemáticamente de la forma siguiente:

$$L = \sum_{i=1}^{N-1} l_i \quad (2)$$

donde $l_i = \text{dist}(i, i+1)$, es decir, la distancia entre dos puntos consecutivos y N es la longitud de la señal (número de muestras).

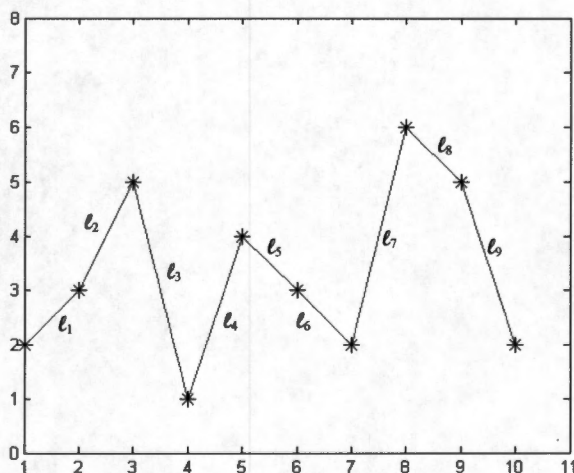


Figura 4.1. Serie de tiempo con dimensión fractal 1.71.

Además, la extensión en el plano o diámetro de una señal que no tienen regresos, señal que progresa monótonicamente, se determina como la máxima distancia entre el conjunto de distancias calculadas entre el primer punto y los puntos restantes, es decir:

$$d = \max(\text{dis}(1, i)) \quad (3)$$

En el ejemplo de la serie de tiempo mostrada en la figura 4.1, la dimensión fractal es:

$$DF = \frac{\log(l_1 + l_2 + l_3 + \dots + l_9)}{\log(\max(\text{dist}(1, i)))} \quad (4)$$

donde

$$l_1 + l_2 + \dots + l_9 = 1.41 + 2.23 + 4.12 + 3.16 + 1.41 + 1.41 + 4.12 + 1.41 + 3.16 = 22.43$$

$$\max(\text{dis}(1, i)) = \max(1.41 \ 3.60 \ 3.16 \ 4.47 \ 5.09 \ 6.00 \ 8.06 \ 8.54 \ 9.00) = 9.0$$

En consecuencia, la dimensión fractal de la serie de tiempo de la figura 4.1 es 1.4163.

La formulación de la dimensión fractal arriba mencionada tiene la desventaja de que su resultado depende de las unidades en que se mida la serie de tiempo, es decir, si asumimos que los valores para la longitud total y el diámetro están en volts, entonces el valor de la dimensión fractal es 1.4163, mientras que si asumimos que los valores se encuentran en microvolts, entonces la dimensión fractal es 0.9213. Con el propósito de evitar la dependencia de las unidades utilizadas y en consecuencia, el poder comparar resultados al obtener la dimensión fractal de varias realizaciones de un experimento, es necesario efectuar una normalización; en el cálculo de la dimensión fractal de series de tiempo se normaliza dividiendo la longitud total y el diámetro por el promedio de las distancias entre puntos sucesivos, es decir:

$$DF = \frac{\log\left(\frac{L}{\text{prom}(d)}\right)}{\log\left(\frac{d}{\text{prom}(d)}\right)} \quad (5)$$

En consecuencia:

$$DF = \frac{\log \left(\frac{\sum_{i=1}^n \ell_i}{\frac{1}{n} \sum_{i=1}^n \ell_i} \right)}{\log \left(\frac{d}{\frac{1}{n} \sum_{i=1}^n \ell_i} \right)} = \frac{\log(n)}{\log \left(\frac{d}{\frac{1}{n} \sum_{i=1}^n \ell_i} \right)} \quad (6)$$

donde $n = N-1$, además:

$$\frac{L}{prom(d)} = n \Rightarrow prom(d) = \frac{L}{n}. \quad (7)$$

Finalmente, la dimensión fractal se establece como:

$$DF = \frac{\log(n)}{\log \left(\frac{d}{\frac{L}{n}} \right)} = \frac{\log(n)}{\log \left(\frac{nd}{L} \right)} = \frac{\log(n)}{\log(n) + \log \left(\frac{d}{L} \right)} \quad (8)$$

En nuestro ejemplo, la dimensión fractal normalizada tiene un valor de 1.71. Para señales que no tienen retrocesos, señales que progresan monotónicamente, la extensión en el plano, normalizada entre el promedio de las distancias individuales, es igual a $N-1$, en nuestro ejemplo tiene un valor de 9.

El rango de la dimensión fractal de una señal se puede establecer si partimos de la siguiente situación: consideremos que las distancias ℓ_i son iguales, figura 4.2, es decir se trata de una recta, entonces los parámetros de longitud total y la extensión en el plano tendrán el valor de $n \cdot \ell$. Es decir, la dimensión fractal de una señal sin variaciones es igual a uno. Es evidente que si deformamos a la recta, entonces la suma de distancias ℓ_i aumentará y en consecuencia, el valor de la dimensión fractal también aumentará. En el caso extremo, se tiene una serie de tiempo con una gran cantidad de variaciones, como es el caso de una señal de ruido. Es por esta razón que la dimensión fractal se ha utilizado para medir la irregularidad de una señal y se ha asociado con el término de complejidad. Es importante

notar que al aplicar la dimensión fractal a una señal no se asume, entre otras cosas, estacionaridad de la señal. Es decir, la señal se trata desde un punto de vista diferente al tratamiento estocástico clásico de una señal [1,2].

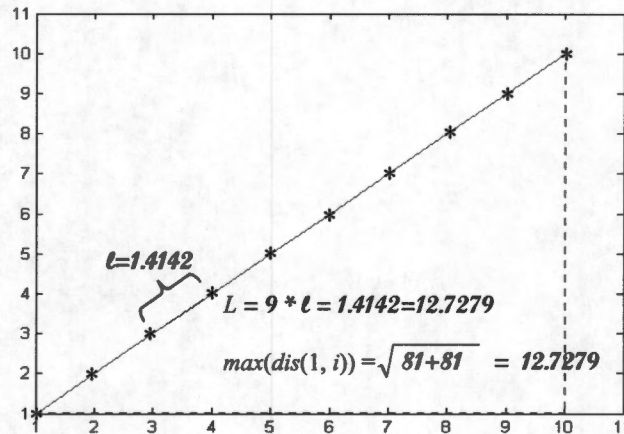


Figura 4.2. Procedimiento para el cálculo de la dimensión fractal.

El valor superior de la dimensión fractal, para series de tiempo relacionadas con eventos fisiológicos, lo podemos establecer con base al hecho de que jamás serán lo suficientemente irregulares (complejas) para llenar el espacio en R^2 y en consecuencia, su dimensión fractal tendrá un valor menor que dos.

4.2 MODELOS ESTOCÁSTICOS Y PARAMETRIZACIÓN DE UNA SERIE DE TIEMPO POR MEDIO DE UN MODELO AR

Una serie de tiempo, $\{u(n)\}$, consiste de observaciones altamente correlacionadas temporalmente que pueden ser generadas teóricamente aplicando *muestras* estadísticamente independientes, $\{v(n)\}$, a un filtro lineal, como se muestra en la figura 4.3. Las *muestras* son variables aleatorias extraídas de una distribución que usualmente se asume Gaussiana con media cero y varianza constante [3]. El conjunto serie de variables aleatorias constituye un proceso aleatorio que comúnmente se conoce como ruido blanco Gaussiano. Específicamente, se puede describir la entrada $\{v(n)\}$, en la figura 4.3, en términos estadísticos como sigue:

$$E[v(n)] = 0 \quad \text{para todo } n \quad (9)$$

$$E[v(n)v^*(k)] = \begin{cases} \sigma_v^2, & k = n \\ 0 & \end{cases} \quad (10)$$

donde $E[\cdot]$ representa el operador valor esperado y σ_v^2 es la varianza de $\{v(n)\}$, el ruido. La ecuación 9 es el resultado de asumir media cero, mientras que la ecuación 10 es el resultado de la suposición de ruido blanco.

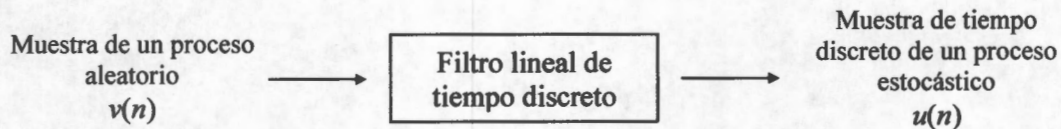


Figura 4.3. Modelo estocástico de una serie de tiempo.

En general, la descripción en el dominio del tiempo de la relación entrada – salida para el modelo estocástico de la figura 4.3 puede ser escrita como sigue:

$$\begin{array}{l} \text{Valor presente de} \\ \text{la salida del} \\ \text{modelo} \end{array} + \begin{array}{l} \text{Combinación lineal} \\ \text{de valores pasados} \\ \text{de la salida del} \\ \text{modelo} \end{array} = \begin{array}{l} \text{Combinación lineal de} \\ \text{valores presentes y} \\ \text{pasados de la} \\ \text{entrada del modelo} \end{array} \quad (11)$$

La estructura del filtro lineal, figura 4.3, se determina por la manera en la que se definen las 2 combinaciones lineales indicadas en la ecuación 11. Se pueden identificar 3 tipos de modelos estocásticos:

1. Modelos autorregresivos, en los cuales se utilizan muestras pasadas de la señal de salida y el valor presente de la señal de entrada del modelo.
2. Modelos de promedio móvil, en los cuales no se usan valores pasados de la salida del modelo y en consecuencia, la salida únicamente depende de la información de la señal de entrada.
3. Modelos de promedio móvil y autorregresivos combinados, en los cuales la descripción de la señal de salida del sistema cumple con la ecuación 11.

4.2.1 MODELO ESTOCÁSTICO AUTORREGRESIVO

La serie de tiempo $u(n)$, $u(n-1)$,... $u(n-M)$ representa la realización de un proceso autorregresivo (AR) de orden M si satisface la ecuación de diferencias [3,4]:

$$u(n) + a_1^* u(n-1) + a_M^* u(n-M) = v(n) \quad (12)$$

donde a_1, a_2, \dots, a_M son constantes llamadas parámetros del modelo AR y $\{v(n)\}$ es un proceso estocástico de ruido blanco. El término $a_k^* u(n-k)$ es un producto interno de a_k y $u(n-k)$, donde $k = 1, \dots, M$.

Para explicar la razón del nombre *autorregresivo*, se puede reescribir la ecuación 12 en la forma:

$$u(n) = w_1^* u(n-1) + w_2^* u(n-2) + \dots + w_M^* u(n-M) + v(n), \quad (13)$$

donde $w_k = -a_k$. Se puede observar que el valor presente del proceso, $u(n)$ es igual a una combinación lineal finita de valores pasados del proceso, $u(n-1), \dots, u(n-M)$, más un término de error $v(n)$, lo anterior es la razón para el nombre *autorregresivo*. Específicamente un modelo lineal de la forma:

$$y = \sum_{k=1}^M w_k^* x_k + v,$$

que relaciona una variable *dependiente* y a un conjunto de variables *independientes* x_1, x_2, \dots, x_M más un término de error v que frecuentemente se le refiere como un modelo de regresión. En la ecuación 13 la variable $u(n)$ se obtiene mediante una regresión sobre valores previos de sí misma.

Por otro lado, el lado izquierdo de la ecuación 12 representa la *convolución* de la secuencia de entrada $\{u(n)\}$ y la secuencia de parámetros $\{a_n^*\}$. Para resaltar este punto, reescribimos la ecuación 12 en forma de una suma de convolución:

$$\sum_{k=0}^M a_k^* u(n-k) = v(n), \quad (14)$$

donde $a_0 = 1$. Tomando la transformada z de ambos lados de la ecuación 14 y definiendo a a_n^* , $H_A(z)$ denota la transformada z de la secuencia $\{a_n^*\}$ de la forma:

$$H_A(z) = \sum_{n=0}^M a_n^* z^{-n}, \quad (15)$$

y a $U(z)$ la transformada z de la secuencia de entrada $\{u(n)\}$:

$$U(z) = \sum_{n=0}^{\infty} u(n) z^{-n}, \quad (16)$$

donde z es una variable compleja, la forma equivalente de la ecuación 12 es:

$$H_A(z)U(z) = V(z) \quad (17)$$

donde

$$V(z) = \sum_{n=0}^{\infty} v(n) z^{-n}. \quad (18)$$

La transformada z de la ecuación 17 ofrece 2 interpretaciones, dependiendo si el proceso AR $\{u(n)\}$ es visto como la entrada o salida de interés:

1. Dado un proceso AR $\{u(n)\}$, podemos usar el filtro mostrado en la figura 4.4 (a) para producir el proceso de ruido blanco $\{v(n)\}$ como salida. Los parámetros de este filtro corresponden uno a uno con los coeficientes del proceso AR $\{u(n)\}$. Este filtro representa un proceso de análisis con función de transferencia discreta $H_A(z) = V(z)/U(z)$. La respuesta impulso correspondiente, que es la transformada inversa z de $H_A(z)$ tiene duración finita.
2. Con el ruido blanco $\{v(n)\}$ actuando como entrada, se puede usar el filtro mostrado en la figura 4.4 (b) para producir el proceso AR $\{u(n)\}$ como salida. Este segundo filtro representa un proceso de generación, cuya función de transferencia es igual a:

$$\begin{aligned} H_G(z) &= \frac{U(z)}{V(z)} \\ &= \frac{1}{H_A(z)} \\ &= \frac{1}{\sum_{n=0}^M a_n^* z^{-n}} \end{aligned} \quad (19)$$

En este modelo la respuesta impulso correspondiente a $H_G(z)$ tiene duración infinita.

La función de transferencia $H_A(z)$ de la figura 4.4 (a) es un filtro de todos ceros, dado que $H_A(z)$ está completamente definida por las localizaciones de sus ceros y por lo tanto, el filtro es inherentemente estable. En contraste, la función de transferencia $H_G(z)$, de la figura 4.4 (b), es un filtro de únicamente polos, es decir:

$$H_G(z) = \frac{1}{(1-p_1z^{-1})(1-p_2z^{-1})\dots(1-p_Mz^{-1})} \quad (20)$$

donde los parámetros p_1, p_2, \dots, p_M son los polos de $H_G(z)$ que se definen por las raíces de la ecuación característica:

$$1 + a_1^*z^{-1} + a_2^*z^{-2} + \dots + a_M^*z^{-M} = 0. \quad (21)$$

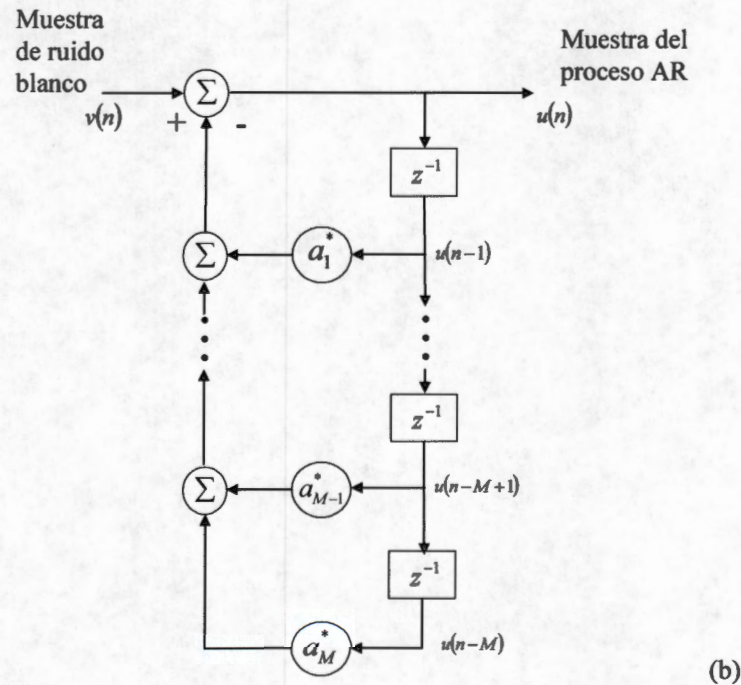
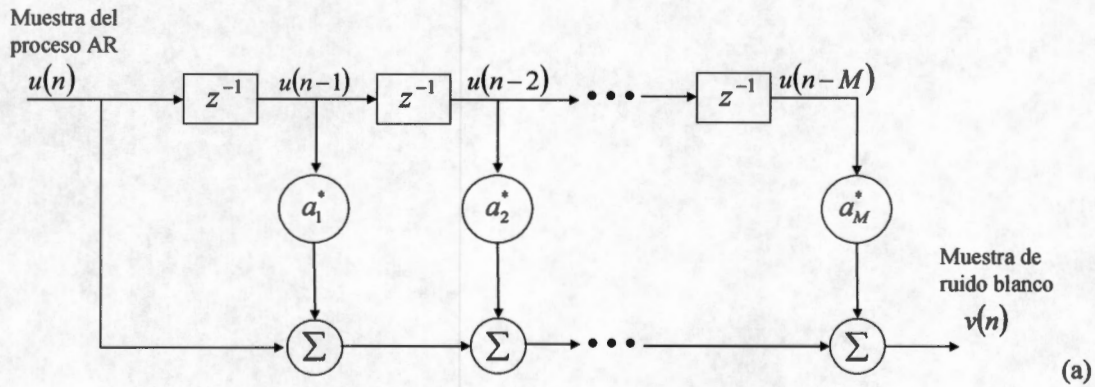


Figura 4.4. (a) Modelo AR como un proceso de análisis AR; (b) Modelo AR como un proceso generador.

Para que $H_G(z)$ sea estable, las raíces de la ecuación característica deben estar dentro del círculo unitario, en el plano z . Lo anterior, es necesario y condición suficiente para que

el proceso producido por el modelo AR de la figura 4.4 (b) sea estacionario en el sentido amplio.

En el caso de que el modelo estocástico AR de la figura 4.4 (b) cumpla la condición de estacionaridad asintótica, se puede demostrar que la función de autocorrelación del proceso $\{u(n)\}$ decae a cero conforme el retraso o *lag* aumenta. Para corroborar este comportamiento es posible derivar una relación recursiva para la función de autocorrelación del proceso AR $u(n)$. Primero se multiplican ambos lados de la ecuación 12 por $u^*(n-l)$ y se aplica el operador valor esperado:

$$E\left[\sum_{k=0}^M a_k^* u(n-k) u^*(n-l)\right] = E[v(n) u^*(n-l)] \quad (22)$$

Se puede simplificar el lado izquierdo de la ecuación 22 intercambiando el operador valor esperado y la sumatoria además, $E[u(n-k) u^*(n-l)]$ es igual a la autocorrelación del proceso AR para un retraso $l-k$. En el lado derecho de la ecuación 22 se logra la simplificación al observar que el valor esperado de $E[v(n) u^*(n-l)]$ es cero para $l > 0$, ya que $u(n-l)$ sólo involucra muestras del proceso de ruido blanco hasta la muestra $n-l$. Por lo tanto, la ecuación 22 se puede reescribir de la siguiente manera:

$$\sum_{k=0}^M a_k^* r(l-k) = 0, \quad l > 0 \quad (23)$$

donde $a_0=1$. Por lo tanto, la función de autocorrelación de un proceso AR satisface la ecuación de diferencias:

$$r(l) = w_1^* r(l-1) + w_2^* r(l-2) + \dots + w_M^* r(l-M), \quad l > 0 \quad (24)$$

donde $w_k = -a_k$, $k=1,2,\dots,M$. Nótese que la ecuación 24 es análoga a la ecuación de diferencias para el proceso AR $\{u(n)\}$. Ahora bien, la solución de la ecuación de diferencias en la ecuación 24 es:

$$r(m) = \sum_{k=1}^M C_k p_k^m,$$

donde C_k , $k=1, \dots, M$ son constantes y p_k , $k=1, \dots, M$ son los polos mencionados. Si $|p_k| < 1 \forall k$ entonces $r(m)$ tiende a cero conforme m tiende al infinito.

Existen varias técnicas para estimar los coeficientes AR de la ecuación 14. Las principales son mínimos cuadrados (Least Squares LS) y el método de Burg. El método

más común de LS está basado en las ecuaciones de Yule Walker, por lo que a continuación se describirá esta técnica.

Si se escribe la ecuación 24 para todos los valores $l=1, 2, \dots, M$, entonces se obtiene un conjunto de M ecuaciones simultáneas con valores $r(0), \dots, r(M)$ de la función de autocorrelación del proceso AR como cantidades conocidas y los parámetros a_1, a_2, \dots, a_M como desconocidos. Este conjunto de ecuaciones puede expresarse en términos matriciales como:

$$\begin{bmatrix} r(0) & r(1) & \dots & r(M-1) \\ r^*(1) & r(0) & \dots & r(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ r^*(M-1) & r^*(M-2) & \dots & r(0) \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{bmatrix} = \begin{bmatrix} r^*(1) \\ r^*(2) \\ \vdots \\ r^*(M) \end{bmatrix}, \quad (25)$$

donde $w_k = -a_k$. El conjunto de ecuaciones (25) es llamado Ecuaciones de Yule-Walker. Se pueden expresar las ecuaciones de Yule - Walker en forma compacta de la siguiente forma:

$$\mathbf{R}\mathbf{w} = \mathbf{r} \quad (26)$$

Asumiendo que la matriz de correlación \mathbf{R} es no singular la solución de la ecuación 26 es:

$$\mathbf{w} = \mathbf{R}^{-1}\mathbf{r} \quad (27)$$

donde \mathbf{R}^{-1} es la inversa de la matriz \mathbf{R} , con \mathbf{R} definida por:

$$\mathbf{R} = \begin{bmatrix} r(0) & r(1) & \dots & r^*(M-1) \\ r^*(1) & r(0) & \dots & r^*(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ r^*(M-1) & r^*(M-2) & \dots & r(0) \end{bmatrix} \quad (28)$$

el vector \mathbf{w} está definido por: $\mathbf{w}^T = [w_1, w_2, \dots, w_M]$, y el vector \mathbf{r} por:

$$\mathbf{r}^H = [r(1), r(2), \dots, r(M)] \quad (29)$$

Por otro lado, el método de Burg puede ser considerado como un método de mínimos cuadrados de orden recursivo, basado en la minimización de los errores hacia adelante y hacia atrás en filtros de predicción lineal, con la restricción de que los parámetros AR satisfacen la recursión de Levinson - Durbin [5].

4.2.1.1 Selección del orden del modelo AR

Como regla general, si se selecciona un modelo AR con un orden p muy bajo, se obtiene un espectro de potencia altamente suavizado. Por otro lado, si p se selecciona con un valor muy alto se corre el riesgo de introducir información espuria y en consecuencia un espectro con una cantidad considerable de picos de baja amplitud. Un índice del desempeño del modelo AR es el valor cuadrático medio del error residual. La característica de este error residual es que disminuye conforme el orden del modelo AR aumenta. Se puede monitorear el comportamiento del ruido residual para decidir el orden del modelo AR [4, 5]. Dos de los mejores criterios para seleccionar el orden del modelo han sido propuestos por Akaike:

- a) *Error de Predicción Final (FPE)*.- Se selecciona el orden para minimizar el índice:

$$\text{FPE}(p) = \hat{\sigma}_{wp}^2 \left(\frac{N+p+1}{N-p-1} \right) \quad (30)$$

donde p es el orden del modelo, N es la longitud de la serie de tiempo y $\hat{\sigma}_{wp}^2$ es la varianza estimada del error de predicción lineal. Este índice de desempeño está basado en minimizar el error cuadrático medio para un predictor de un paso hacia adelante en el tiempo discreto.

- b) *Criterio de Información de Akaike (AIC)*.- Está basado en seleccionar el orden que minimice:

$$\text{AIC}(p) = \ln \hat{\sigma}_{wp}^2 + \frac{2p}{N} \quad (31)$$

En este criterio, el término $\hat{\sigma}_{wp}^2$ disminuye y por lo tanto $\ln \hat{\sigma}_{wp}^2$ también disminuye conforme el orden del modelo AR aumenta. Sin embargo, se tiene la desventaja que $\frac{2p}{N}$ aumenta con un aumento en p .

Para evaluar el ajuste del modelo AR, es necesario calcular el error residual dado el orden p y graficar la función de autocorrelación, y comprobar el comportamiento del error como ruido blanco.

Nótese que una vez estimados los coeficientes AR de la serie de tiempo $\{u(n)\}$ se logra parametrizarla, dado que ahora estos coeficientes representan su estadística en forma óptima, de acuerdo a un criterio de selección de orden. Por lo tanto, para propósitos de clasificación o estimación de información, es posible utilizar a los coeficientes AR en lugar de la serie completa de tiempo $\{u(n)\}$. Sin embargo, la condición para utilizar los modelos AR es que la serie $\{u(n)\}$ debe ser estacionaria en el sentido amplio.

4.2.2 MODELO ESTOCÁSTICO AUTORREGRESIVO VARIANTE EN EL TIEMPO Y SU ESTIMACIÓN POR MÍNIMOS CUADRADOS RECURSIVOS ESTÁNDAR (RLS)

La ecuación 14 representa el modelo AR de una serie de tiempo $\{u(n)\}$. Sin embargo, si $\{u(n)\}$ es una serie de tiempo del tipo no estacionaria, es necesario proponer un modelo AR variante en el tiempo que pueda contender con los cambios en $\{u(n)\}$. Este modelo representa el siguiente nivel de complejidad en la parametrización de la serie $\{u(n)\}$. El modelo AR variante en el tiempo (ARVT) está dado por la ecuación:

$$\sum_{k=0}^M a_k^*(n)u(n-k) = v(n),$$

donde los coeficientes ARVT dependen ahora del tiempo discreto y se van modificando muestra a muestra por lo que es posible generar una serie de tiempo correspondiente a cada coeficiente. Para estimar los coeficientes ARVT, se han desarrollado algoritmos recursivos como el *least mean squares* (LMS) y el *recursive least squares* (RLS), con ventajas del segundo sobre el primero para seguir el comportamiento no estacionario de $\{u(n)\}$. A continuación se exponen los elementos teóricos que dan origen al algoritmo RLS y la forma de calcular los coeficientes $a_k^*(n)$, $k=1, \dots, M$.

Algoritmo RLS

La sección se inicia con algunos elementos relevantes del método de mínimos cuadrados para después utilizar el lema de inversión de matrices que proporciona la ecuación recursiva necesaria para estimar los coeficientes AR.

En la implementación recursiva del método de mínimos cuadrados, el cálculo comienza con condiciones iniciales conocidas y se usa la información contenida en muestras de datos nuevos para actualizar los parámetros estimados pasados. La función de costo a minimizarse $\xi(n)$ depende del tamaño de la ventana de datos observables, es decir n . Además, en la definición de $\xi(n)$ es costumbre introducir un factor de peso o factor de olvido [3], por lo tanto:

$$\xi(n) = \sum_{i=1}^n \beta(n,i) e(i)^2, \quad (32)$$

donde $e(i)$ es la diferencia entre la *respuesta deseada* $d(i)$ y la *salida* $y(i)$ producida por un filtro transversal con entrada (al tiempo i) igual a $u(i), u(i-1), \dots, u(i-M+1)$, tal como se muestra en la figura 4.5. Esto es:

$$\begin{aligned} e(i) &= d(i) - y(i) \\ &= d(i) - \mathbf{w}^H(n) \mathbf{u}(i), \end{aligned} \quad (33)$$

donde $\mathbf{u}(i)$ es el vector de entrada al tiempo i , definido por :

$$\mathbf{u}(i) = [u(i), u(i-1), \dots, u(i-M+1)]^T \quad (34)$$

y $\mathbf{w}(n)$ es el vector de pesos al tiempo n , definido por

$$\mathbf{w}(n) = [\omega_0(n), \omega_1(n), \dots, \omega_{M-1}(n)]^T. \quad (35)$$

Los pesos del filtro transversal permanecen fijos durante el intervalo $1 \leq i \leq n$, ventana temporal donde se define la función de costo $\xi(n)$. El factor de ponderación $\beta(n,i)$ en la ecuación (32) tiene la propiedad de:

$$0 < \beta(n,i) \leq 1, \quad i = 1, 2, \dots, n \quad (36)$$

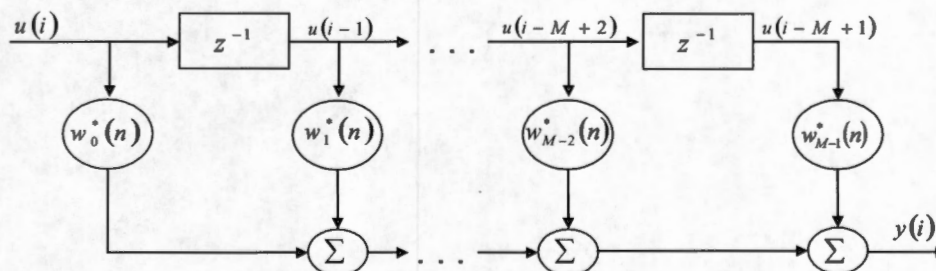


Figura 4.5. Filtro Transversal.

Posteriormente, se indicará a que variables corresponden $y(i)$ y $d(i)$ en el caso de la estimación de los coeficientes AR variantes en el tiempo.

En general, el uso del factor de ponderación $\beta(n, i)$, tiene la intención de asegurar que los datos en un pasado lejano sean “olvidados” para permitir la posibilidad de seguir las variaciones estadísticas de los datos observables, cuando el filtro opera en un ambiente no estacionario. Una forma especial de factor de ponderación que se usa comúnmente es el factor definido por:

$$\beta(n-i) = \lambda^{n-i}, \quad i=1, 2, \dots, n. \quad (37)$$

donde λ es una constante positiva cercana, pero menor que la unidad. Cuando $\lambda=1$, se tiene el método ordinario de mínimos cuadrados. El inverso de $1-\lambda$ es una medida de la *memoria* del algoritmo. El caso especial $\lambda=1$ corresponde a una *memoria infinita*.

Así en el método de mínimos cuadrados ponderados exponencialmente, se minimiza la función de costo:

$$\xi(n) = \sum_{i=1}^n \lambda^{n-i} |e(i)|^2. \quad (38)$$

Para determinar el valor óptimo del vector de pesos, $\mathbf{w}(n)$, para el cual la función de costo $\xi(n)$ alcanza el valor mínimo se tienen dos procedimientos: a) la forma clásica de obtener $\mathbf{w}(n)$ es derivar la expresión en la ecuación 38 con respecto al vector de pesos $\mathbf{w}(n)$ e igualar a cero la expresión resultante y b) utilizar el principio de ortogonalidad que se aplica entre la señal de error $e(i)$ al momento de ser mínimo y la señal de entrada al filtro transversal $\mathbf{u}(i) = [u(i), u(i-1), \dots, u(i-M+1)]^T$. A partir de estas derivaciones se obtienen las ecuaciones normales en forma matricial como:

$$\Phi(n)\mathbf{w}(n) = \theta(n), \quad (39)$$

donde la matriz de correlación $\Phi(n)$, de tamaño $M \times M$, se define como:

$$\Phi(n) = \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i)\mathbf{u}^H(i), \quad (40)$$

y el vector de correlación cruzada $\theta(n)$, de tamaño $M \times 1$, entre las entradas del filtro transversal y la respuesta deseada, se define como:

$$\theta(n) = \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i)d^*(i), \quad (41)$$

donde el asterisco denota la operación complejo conjugado. Es importante observar que las ecuaciones 39 a 41 se encuentran en términos directamente de los valores de tiempo de la serie $\{u(n)\}$ y de la respuesta que se desea del filtro $d(i)$ y no de promedios del ensamble.

Aislando el término correspondiente a $i=n$ del resto de la sumatoria en el lado derecho de la ecuación 40, se puede reescribir como:

$$\Phi(n) = \lambda \left[\sum_{i=1}^{n-1} \lambda^{n-1-i} \mathbf{u}(i) \mathbf{u}^H(i) \right] + \mathbf{u}(n) \mathbf{u}^H(n). \quad (42)$$

Sin embargo, por definición, la expresión dentro de los paréntesis cuadrados en la ecuación 42 es igual a la matriz de correlación $\Phi(n-1)$, entonces se tiene la siguiente ecuación recursiva para actualizar el valor de la matriz de correlación del vector de entrada:

$$\Phi(n) = \lambda \Phi(n-1) + \mathbf{u}(n) \mathbf{u}^H(n) \quad (43)$$

donde $\Phi(n-1)$ es el valor pasado de la matriz de correlación, y el producto $\mathbf{u}(n) \mathbf{u}^H(n)$ es un término de corrección en la actualización.

En forma análoga, se puede usar la ecuación 41 para derivar la siguiente ecuación recursiva para actualizar el vector de correlación cruzada entre la entrada y la respuesta deseada:

$$\theta(n) = \lambda \theta(n-1) + \mathbf{u}(n) d^*(n). \quad (44)$$

Para estimar el vector de pesos $\mathbf{w}(n)$ por mínimos cuadrados con base en la ecuación 39, es necesario calcular el inverso de la matriz de correlación $\Phi(n)$. En la práctica, se intenta comúnmente evitar tal operación ya que consume mucho tiempo, particularmente si el número de pesos M es alto. También se debe estimar $\mathbf{w}(n)$ recursivamente para $n=1,2,\dots,\infty$. Se pueden alcanzar ambos objetivos, usando un resultado básico en el álgebra de matrices conocido como *lema de inversión de matrices*. Se asume que las condiciones iniciales se seleccionan para asegurar la no singularidad de la matriz de correlación $\Phi(n)$.

Para aplicar el lema de inversión, se hacen primeramente las siguientes correspondencias:

$$\begin{aligned}
 \mathbf{A} &= \Phi(n) \\
 \mathbf{B}^{-1} &= \lambda \Phi(n-1) \\
 \mathbf{C} &= \mathbf{u}(n) \\
 \mathbf{D} &= 1
 \end{aligned}
 \tag{45}$$

Sustituyendo estas definiciones en el lema de inversión de matrices, se obtiene la siguiente ecuación recursiva para el inverso de la matriz de correlación:

$$\Phi^{-1}(n) = \lambda^{-1} \Phi^{-1}(n-1) - \frac{\lambda^{-2} \Phi^{-1}(n-1) \mathbf{u}(n) \mathbf{u}^H(n) \Phi^{-1}(n-1)}{1 + \lambda^{-1} \mathbf{u}^H(n) \Phi^{-1}(n-1) \mathbf{u}(n)}.
 \tag{46}$$

Por conveniencia de cálculo, se definen las siguientes variables:

$$\mathbf{P}(n) = \Phi^{-1}(n)
 \tag{47}$$

$$\mathbf{k}(n) = \frac{\lambda^{-1} \mathbf{P}(n-1) \mathbf{u}(n)}{1 + \lambda^{-1} \mathbf{u}^H \mathbf{P}(n-1) \mathbf{u}(n)}.
 \tag{48}$$

Usando estas definiciones, se puede describir la ecuación 46 como:

$$\mathbf{P}(n) = \lambda^{-1} \mathbf{P}(n-1) - \lambda^{-1} \mathbf{k}(n) \mathbf{u}^H(n) \mathbf{P}(n-1)
 \tag{49}$$

donde $\mathbf{P}(n)$ tiene las dimensiones de $M \times M$, y $\mathbf{k}(n)$ tiene dimensiones $M \times 1$. A partir de la ecuación 48 se puede demostrar que el vector de ganancia $\mathbf{k}(n)$ es igual a:

$$\mathbf{k}(n) = \mathbf{P}(n) \mathbf{u}(n)
 \tag{50}$$

$$\mathbf{k}(n) = \Phi^{-1}(n) \mathbf{u}(n)
 \tag{51}$$

En otras palabras, el vector de ganancia $\mathbf{k}(n)$ se define como el vector de entrada $\mathbf{u}(n)$ transformado por el inverso de la matriz de correlación $\Phi(n)$.

Ahora bien, para desarrollar una ecuación recursiva y actualizar el estimado $\hat{\mathbf{w}}(n)$ por mínimos cuadrados de $\mathbf{w}(n)$ se utilizan las ecuaciones 39, 44 y 48 en la secuencia siguiente:

$$\begin{aligned}
 \hat{\mathbf{w}}(n) &= \Phi^{-1}(n) \boldsymbol{\theta}(n) \\
 &= \mathbf{P}(n) \boldsymbol{\theta}(n) \\
 &= \lambda \mathbf{P}(n) \boldsymbol{\theta}(n-1) + \mathbf{P}(n) \mathbf{u}(n) d^*(n)
 \end{aligned}
 \tag{52}$$

Sustituyendo la ecuación 49 $\mathbf{P}(n)$ en el primer término del lado derecho de la ecuación 52 se obtiene:

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) - \mathbf{k}(n)\mathbf{u}^H(n)\hat{\mathbf{w}}(n-1) + \mathbf{P}(n)\mathbf{u}(n)d^*(n). \quad (53)$$

Usando el hecho de que $\mathbf{P}(n)\mathbf{u}(n)$ es igual al vector de ganancia $\mathbf{k}(n)$, se obtiene la ecuación recursiva deseada para actualizar el vector de pesos:

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mathbf{k}(n)\alpha^*(n), \quad (54)$$

donde $\alpha(n)$ es la innovación definida como:

$$\begin{aligned} \alpha(n) &= d(n) - \mathbf{u}^T(n)\hat{\mathbf{w}}^*(n-1) \\ &= d(n) - \hat{\mathbf{w}}^H(n-1)\mathbf{u}(n) \end{aligned} \quad (55)$$

El producto interno $\hat{\mathbf{w}}^H(n-1)\mathbf{u}(n)$ representa un estimado de la respuesta deseada $d(n)$, basado en el estimado del vector de pesos realizado al tiempo $n-1$. De acuerdo con esto, $\alpha(n)$ se puede referir como el error de estimación *a priori*. En general, la innovación $\alpha(n)$ es diferente de la estimación del error *a posteriori*:

$$e(n) = d(n) - \hat{\mathbf{w}}^H(n)\mathbf{u}(n) \quad (56)$$

El cálculo de $e(n)$, en la ecuación 56, involucra el estimado del vector de pesos disponible al tiempo n . En consecuencia, se puede ver a $\alpha(n)$ como un valor tentativo de $e(n)$ antes de la actualización del vector de pesos. Sin embargo, en la optimización de mínimos cuadrados que conduce al algoritmo recursivo de la ecuación 54, se minimiza la función de costo basándose en $e(n)$ y no en $\alpha(n)$ [3].

A continuación se presenta un resumen del algoritmo RLS:

$$\begin{aligned} \mathbf{k}(n) &= \frac{\lambda^{-1}\mathbf{P}(n-1)\mathbf{u}(n)}{1 + \lambda^{-1}\mathbf{u}^H(n)\mathbf{P}(n-1)\mathbf{u}(n)} \\ \alpha(n) &= d(n) - \hat{\mathbf{w}}^H(n-1)\mathbf{u}(n) \\ \hat{\mathbf{w}}(n) &= \hat{\mathbf{w}}(n-1) + \mathbf{k}(n)\alpha^*(n) \\ \mathbf{P}(n) &= \lambda^{-1}\mathbf{P}(n-1) - \lambda^{-1}\mathbf{k}(n)\mathbf{u}^H(n)\mathbf{P}(n-1). \end{aligned}$$

En particular la ecuación 55 describe la operación de filtrado del algoritmo, el filtro transversal se alimenta con $\mathbf{u}(n)$ para calcular el error de estimación *a priori* $\alpha(n)$. Además, la ecuación 54 describe la operación adaptiva del algoritmo, por lo que el vector de pesos es actualizado aumentando su valor pasado por una cantidad igual al complejo conjugado de la estimación del error *a priori* por el vector de ganancia variante en el tiempo $\mathbf{k}(n)$. Las

ecuaciones 48 y 49 actualizan el vector de ganancia por si mismo. Una característica importante del algoritmo RLS descrita por estas ecuaciones es que la inversión de la matriz de correlación $\Phi(n)$ se remplaza en cada paso por una simple división escalar.

Ahora bien, para el cálculo de los coeficientes AR del modelo variante en el tiempo, la correspondencia de las variables es la siguiente:

$$d(n) = u(n)$$

$$\mathbf{u}(n) = [u(n-1), \dots, u(n-M+1)]^T$$

$$a_k(n) = -w_k(n), \quad k=1, 2, \dots, M$$

$$y(n) = \hat{\mathbf{w}}^H(n-1)\mathbf{u}(n).$$

4.3 REDES NEURONALES ARTIFICIALES

Una red neuronal es un ente diseñado para modelar la forma en que el cerebro desempeña una tarea específica o función de interés. Para lograr un buen desempeño, las redes neuronales emplean una interconexión masiva de células computacionales llamadas neuronas. Una red neuronal puede ser considerada como una máquina adaptable, del tipo procesador distribuido en paralelo, que tiene una tendencia natural para almacenar conocimiento experimental y hacerlo disponible para su uso [6]. La red neuronal se asemeja al cerebro en dos aspectos:

1. El conocimiento se adquiere por la red neuronal artificial a través de un proceso de aprendizaje.
2. Las conexiones interneuronales conocidas como pesos sinápticos se utilizan para almacenar el conocimiento.

El procedimiento llamado algoritmo de aprendizaje permite modificar los pesos sinápticos de las redes neuronales en forma ordenada para lograr un objetivo deseado. Además, las redes neuronales tienen beneficios importantes: a) su estructura distribuida en paralelo; y, b) la habilidad para aprender y por lo tanto generalizar. La generalización se

refiere a que la red neuronal produce salidas razonables para entradas que no fueron usadas durante el entrenamiento (aprendizaje) [6].

Fundamental en la operación de una red neuronal es la neurona que es una unidad de procesamiento de información. Se pueden identificar tres elementos básicos en una neurona:

1. Un conjunto de conexiones o sinapsis, el cual se caracteriza por sus pesos w_{kj} que conectan a la neurona k -ésima con la j -ésima.
2. Un sumador de las señales de entrada ponderadas por la respectiva sinapsis de la neurona, que en realidad constituye una combinación lineal de las entradas a la neurona.
3. Una función de activación para limitar la amplitud de la salida de la neurona.

En términos matemáticos, se puede describir a la neurona k -ésima a través de las siguientes ecuaciones:

$$u_k = \sum_{j=1}^p w_{kj} x_j \quad (57)$$

$$y_k = \varphi(u_k - \theta_k) \quad (58)$$

donde x_1, x_2, \dots, x_p son las señales de entrada; $w_{k1}, w_{k2}, \dots, w_{kp}$ son los pesos sinápticos de la neurona k -ésima; u_k es la combinación lineal de las entradas a la red; θ_k es el umbral; $\varphi(\cdot)$ es la función de activación; y y_k es la salida de la neurona. En la figura 4.6 se observa el modelo no lineal de una neurona.

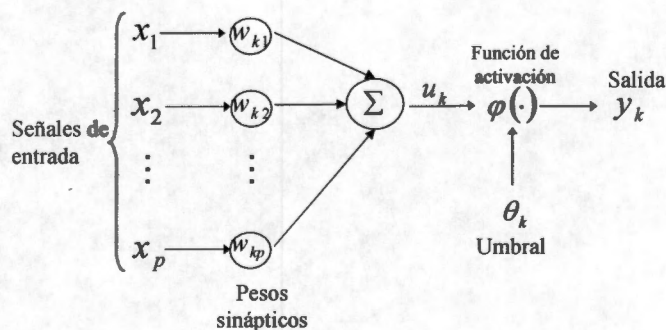


Figura 4.6. Modelo no lineal de una neurona.

El umbral tiene el efecto de disminuir la entrada a la función de activación, además esta entrada puede ser aumentada por un término de sesgo en vez de un umbral. Si se combinan las ecuaciones 57 y 58 se puede escribir:

$$v_k = \sum_{j=0}^p w_{kj} x_j \quad (59)$$

$$y_k = \varphi(v_k) \quad (60)$$

En la ecuación 59 se agrega una nueva entrada considerando:

$$x_0 = -1 \quad (61)$$

$$w_{k0} = \theta_k \quad (62)$$

Alternativamente, se puede fijar $x_0 = +1$ y el peso igual al término de sesgo $w_{k0} = b_k$.

La función de activación $\varphi(\cdot)$ define la salida de una neurona en términos del nivel de actividad a su entrada. Se pueden identificar tres tipos básicos de funciones de activación:

1. La *función umbral* se describe de la siguiente manera:

$$\varphi(v) = \begin{cases} 1 & \text{si } v \geq 0 \\ 0 & \text{si } v < 0 \end{cases} \quad (63)$$

La salida de la neurona k -ésima utilizando una función umbral se expresa como:

$$y_k = \begin{cases} 1 & \text{si } v_k \geq 0 \\ 0 & \text{si } v_k < 0 \end{cases} \quad (64)$$

2. *Función lineal*:

$$\varphi(v) = \begin{cases} 1, & v \geq \frac{1}{2} \\ v, & -\frac{1}{2} < v < \frac{1}{2} \\ 0, & v \leq -\frac{1}{2} \end{cases} \quad (65)$$

3. *Función sigmoidea*: Es una de las funciones más utilizadas en redes neuronales y está definida como una función creciente con propiedades asintóticas [7]. Un ejemplo de la función sigmoidea es la *función logística*, definida como:

$$\varphi(v) = \frac{1}{1 + \exp(-av)}, \quad (66)$$

donde a es la pendiente de la función sigmoidea. Variando el valor de a se pueden obtener funciones sigmoideas con diferentes pendientes. En algunos casos es deseable obtener un rango entre -1 y 1, para esto la función sigmoidea se define de la siguiente forma:

$$\varphi(v) = \frac{2}{1 + \exp(-av)} - 1. \quad (67)$$

La función de la ecuación 67 pertenece a la familia de las funciones tangenciales hiperbólicas de la forma:

$$\varphi(v) = c \frac{1 - \exp(-av)}{1 + \exp(-av)} = c \tanh\left(\frac{av}{2}\right) \quad (68)$$

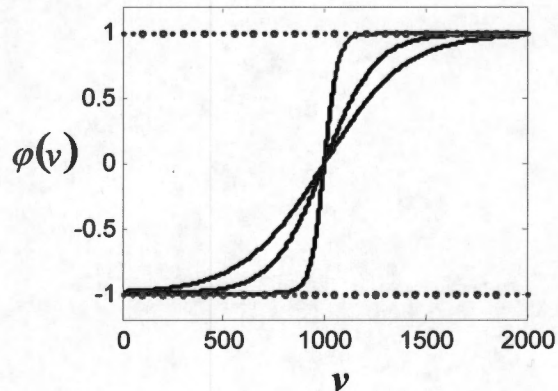


Figura 4.7. Gráfica de una función de activación sigmoidea tangencial.

Una red neuronal puede tener a las neuronas organizadas en capas, en una forma simple sólo se tiene una capa de nodos de datos de entrada que se proyectan a una capa de salida de neuronas, pero no viceversa, es decir, el flujo de información se da sólo desde la entrada hacia la salida; este tipo de redes neuronales se denominan *feedforward*. La segunda clase de red neuronal del tipo *feedforward* se distingue por la presencia de una o más capas ocultas con sus respectivas neuronas, como se muestra en la figura 4.8. Las neuronas de entrada reciben los datos a procesar, las neuronas ocultas reciben las entradas ponderadas por los pesos sinápticos y producen las salidas, mientras que las neuronas de salida realizan la suma ponderada de las salidas de las neuronas ocultas [6].

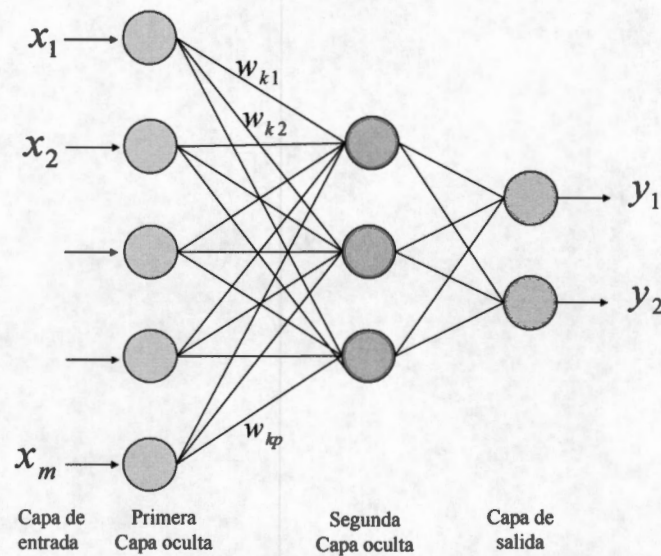


Figura 4.8. Estructura de una red multicapa.

Los perceptrones multicapa tienen tres características distintivas:

1. El modelo de cada neurona en la red incluye una función no-lineal diferenciable. Una forma de no-linealidad que satisface estos requerimientos es la no-linealidad sigmoidea definida por la función logística (ecuación 66).
2. La red contiene una o más capas de neuronas ocultas que no son parte de la entrada o la salida de la red. Las neuronas ocultas habilitan a la red para aprender tareas complejas a través de la extracción progresiva de características principales de los patrones de entrada.
3. La red exhibe un alto grado de conectividad, determinado por la sinapsis de la red. Un cambio en la conectividad de la red requiere un cambio en la población de las conexiones sinápticas o sus pesos [6].

4.3.1 ALGORITMO DE RETROPROPAGACIÓN

El algoritmo de *retropropagación* o regla generalizada Delta se basa en el gradiente descendente de la curva de error; es una extensión natural del algoritmo Least Mean Square (LMS). El algoritmo de *retropropagación* incluye el cálculo del error, proporcional al cuadrado de la diferencia entre la salida actual de la red y la salida deseada, que debe ser evaluado por cada neurona de salida. Además, también es posible calcular un error para

cada neurona oculta, y así incluir una regla de aprendizaje para los pesos sinápticos ocultos [8]. La forma de diseñar un perceptrón multicapa es definir su arquitectura, que se constituye por el número de capas, número de nodos por capa y número de datos de entrada, y calcular los pesos sinápticos para minimizar una función de costo en términos de la salida deseada [7].

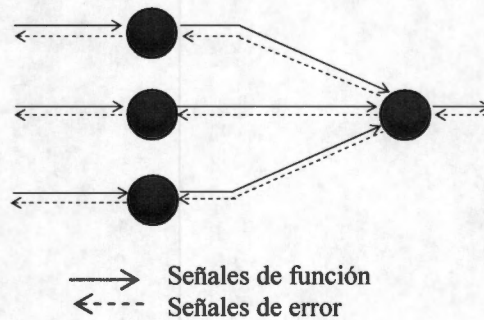


Figura 4.9. Esquema del flujo de las dos señales básicas en un perceptrón multicapa; propagación hacia delante de las señales de función y propagación hacia atrás de las señales de error.

Se pueden identificar dos tipos de señales en un perceptrón multicapa:

1. *Señal función.* Una señal función es una señal de entrada (estímulo) que se propaga hacia adelante neurona por neurona a través de la red, y emerge a la salida de la red como una señal de salida [8], ver figura 4.9.
2. *Señales de error.* Una señal de error se origina en una neurona de salida de la red, y se propaga hacia atrás, capa por capa, a través de la red neuronal.

En el proceso de aprendizaje, se inicia con una red sin entrenar, se presenta un patrón de entrenamiento a la capa de entrada, se propagan las señales a través de la red y se determina la salida en la capa de salida. Después las salidas se comparan con las etiquetas, o salidas deseadas, y cualquier diferencia corresponde a un error. Este error o función de costo es alguna función escalar de los pesos sinápticos y se minimiza cuando las salidas de la red coinciden con las salidas deseadas. Posteriormente, los pesos sinápticos se ajustan para reducir esta medida de error [8].

Se considera el error de entrenamiento sobre un patrón, o vector de características, como la sumatoria de la diferencia al cuadrado entre la salida deseada t_k y la salida actual z_k de la red. A partir de este error, la función de costo a minimizar se define como:

$$J(\mathbf{w}) = \frac{1}{2} \sum_{k=1}^c (t_k - z_k)^2 = \frac{1}{2} (\mathbf{t} - \mathbf{z})^2 \quad (69)$$

donde \mathbf{t} y \mathbf{z} son los vectores correspondientes a las etiquetas o salidas deseadas para todos los patrones de entrada de la red y los vectores de salida de la red, respectivamente. Además, el parámetro c corresponde al número de neuronas en la capa de salida. A partir de la ecuación 69 se puede observar que la función de costo no puede tener un valor negativo, además, la regla de aprendizaje resultante garantiza que el aprendizaje se detendrá, excepto en casos extremos.

La regla de aprendizaje de retropropagación se basa en el algoritmo del gradiente descendente, los pesos sinápticos se inicializan con valores aleatorios y su valor aumenta o disminuye en una dirección que reduce la función de costo en la ecuación 69. Por lo tanto, el cambio de valor de los pesos se genera mediante la siguiente regla de adaptación:

$$\Delta \mathbf{w} = -\eta \frac{\partial J}{\partial \mathbf{w}}, \quad (70)$$

en términos de los elementos del vector de pesos sinápticos:

$$\Delta w_{pq} = -\eta \frac{\partial J}{\partial w_{pq}}, \quad (71)$$

donde η es la *velocidad de aprendizaje*, e indica el tamaño relativo del cambio en los pesos.

El algoritmo iterativo requiere tomar un vector de pesos en la iteración m -ésima y actualizarlo de la siguiente forma:

$$\mathbf{w}(m+1) = \mathbf{w}(m) + \Delta \mathbf{w}(m) \quad (72)$$

La regla de aprendizaje general en la ecuación 70 se modifica dependiendo de la pertenencia de la neurona a una capa de salida o capas ocultas. En el desarrollo del algoritmo de retropropagación que a continuación se presenta se ha eliminado la variable m , sin embargo, cada una de las ecuaciones se realiza por cada iteración m .

Para evaluar la ecuación 71, se consideran primero los pesos sinápticos de la última capa oculta a la capa de salida, w_{kj} y debido a que el error no depende explícitamente de w_{kj} , se utiliza la regla de la cadena para efectuar su diferenciación:

$$\frac{\partial J}{\partial w_{kj}} = \frac{\partial J}{\partial net_k} \frac{\partial net_k}{\partial w_{kj}} = -\delta_k \frac{\partial net_k}{\partial w_{kj}} \quad (73)$$

donde la sensibilidad de la neurona k -ésima se define por:

$$\delta_k = -\frac{\partial J}{\partial net_k}, \quad (74)$$

donde $net_k = \sum_{j=1}^{n_H} y_j w_{kj} + w_{k0}$ y y_j es la salida de la última capa oculta. Asumiendo que la función de activación φ asociada a cada neurona es diferenciable, se deriva la ecuación 74 y se encuentra que para el caso de una neurona de salida, δ_k es simplemente:

$$\delta_k \equiv -\frac{\partial J}{\partial net_k} = -\frac{\partial J}{\partial z_k} \frac{\partial z_k}{\partial net_k} = (t_k - z_k) \varphi'(net_k) \quad (75)$$

La última derivada en la ec. 75 se encuentra utilizando las relaciones:

$$z_k = \varphi(net_k)$$

$$net_k = \sum_{j=1}^{n_H} y_j w_{kj} + w_{k0} = \sum_{j=0}^{n_H} y_j w_{kj} = \mathbf{w}'_k \mathbf{y}, \quad (76)$$

Además, considerando que:

$$\frac{\partial net_k}{\partial w_{kj}} = y_j \quad (77)$$

la regla de aprendizaje para los pesos de la última capa oculta a la capa de salida está dada por:

$$\Delta w_{kj} = \eta \delta_k y_j = \eta (t_k - z_k) \varphi'(net_k) y_j. \quad (78)$$

En el caso en que la neurona de salida incluya una función de activación lineal, es decir, $\varphi(net_k) = net_k$ y $\varphi'(net_k) = 1$, la ec. 78 corresponde al algoritmo LMS.

La regla de aprendizaje para los pesos sinápticos del resto de las capas es más sutil porque no existe una señal deseada para estas neuronas. De la ecuación 71 y usando la regla de la cadena, se puede calcular:

$$\frac{\partial J}{\partial w_{ji}} = \frac{\partial J}{\partial y_j} \frac{\partial y_j}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}}, \quad (79)$$

donde y_j representa la salida de una neurona de la capa más oculta. El primer término del lado derecho requiere sólo un poco de cuidado y puede demostrarse que se define como:

$$\frac{\partial J}{\partial y_j} = -\sum_{k=1}^c (t_k - z_k) \varphi'(net_k) w_{kj} \quad (80)$$

Para el segundo paso, se utiliza otra vez la regla de la cadena. La suma final sobre las neuronas de la capa de salida en la ecuación 80 se expresa como la salida de neuronas ocultas, y_j , afectando al error en cada neurona de salida. Ahora, se utiliza la ecuación 80 para definir la sensibilidad de una neurona de la capa oculta:

$$\delta_j \equiv \varphi'(net_j) \sum_{k=1}^c w_{kj} \delta_k. \quad (81)$$

La sensibilidad en una neurona oculta es simplemente la suma de las sensibilidades individuales de las neuronas de salida afectadas por los pesos de la capa oculta a la capa de salida w_{kj} , multiplicadas por $\varphi'(net_j)$. Entonces la regla de aprendizaje para los pesos de las capas ocultas es:

$$\Delta w_{ji} = \eta y_i \delta_j = \eta y_i \varphi'(net_j) \sum_{k=1}^c w_{kj} \delta_k. \quad (82)$$

Si la neurona j -ésima esta en la primera capa oculta entonces $y_j = x_i$, los datos de entrada a la red a ser clasificados.

4.3.1.1 Criterios para detener el algoritmo de retropropagación.

Se considera que el algoritmo de retropropagación converge cuando la norma euclideana del vector gradiente $\frac{\partial J}{\partial w}$ alcanza un umbral suficientemente pequeño. La desventaja de este criterio puede ser el tiempo de entrenamiento y además de requerir el cálculo del vector gradiente. Otro criterio para detener el algoritmo se basa en el hecho de que la función de costo es estacionaria en el valor óptimo del vector de pesos, por lo tanto, se puede considerar que el algoritmo ha logrado la convergencia cuando la razón de cambio absoluta del error cuadrático promedio por época es muy pequeña. Otro criterio útil es el siguiente: después de cada iteración de aprendizaje se evalúa la generalización que logra la red. El proceso de aprendizaje se detiene cuando el desempeño es adecuado.

4.3.2 MODO DE ENTRENAMIENTO

El entrenamiento de tipo supervisado consiste en presentar un conjunto de vectores de características a la red, determinar la salida de la red y ajustar los pesos para que la salida

actual de la red se aproxime a las etiquetas deseadas. La presentación de los patrones o vectores de características a la red es a través de épocas, una época corresponde a una presentación completa del conjunto de entrenamiento durante el proceso de aprendizaje [6, 8].

Existen dos protocolos de entrenamiento: estocástico y en grupo. En el entrenamiento estocástico, los patrones se seleccionan aleatoriamente del conjunto de entrenamiento y los pesos de la red se actualizan en cada presentación de los patrones. Este método es llamado estocástico porque la búsqueda del mínimo de la superficie formada por la función de costo, en el espacio definido por el vector de pesos, es aleatoria. Lo anterior, reduce la posibilidad de que el algoritmo se quede atrapado en un mínimo local. En el entrenamiento por grupo, los pesos se actualizan después de la presentación de todos los patrones de entrenamiento-etiquetas que constituyen una época. El entrenamiento por grupo es un proceso inherentemente de promediación, esto conduce a una mejor estimación del vector gradiente y a una convergencia con un mejor comportamiento [7].

4.3.3 CURVAS DE APRENDIZAJE Y LA TÉCNICA DE VALIDACIÓN CRUZADA

La esencia del aprendizaje por retropropagación es codificar la relación entrada – salida, representada por un conjunto de patrones de entrenamiento-etiquetas, con un perceptron multicapa bien entrenado, es decir, que haya aprendido lo suficiente del pasado para generalizar al futuro [6]. Para ello, es necesario seleccionar una estructura de una red neuronal multicapas de un conjunto posible, que genere el mejor valor de un criterio dado. En este sentido, la técnica de validación cruzada provee una forma de seleccionar la *mejor* estructura de una red neuronal. En esta técnica, el conjunto de datos se divide de forma aleatoria en un conjunto de entrenamiento y en otro de prueba. El conjunto de entrenamiento posteriormente se divide en dos subconjuntos:

1. Un subconjunto para la estimación del modelo (entrenamiento de la red).

2. Un subconjunto para la evaluación del desempeño del modelo (validación). El subconjunto de validación regularmente es el 10% o 20 % del conjunto de entrenamiento.

La curva de aprendizaje de una red neuronal artificial, figura 4.10, muestra el valor de la función de costo en términos del número de épocas. Antes de que el entrenamiento comience, el error generado por el conjunto de entrenamiento es alto, a través del aprendizaje el error llega a disminuir. El error promedio en un conjunto de prueba siempre es mayor que el error del conjunto de entrenamiento, este error generalmente disminuye, pero también puede aumentar u oscilar. El conjunto de entrenamiento tiene dos usos, primero, establecer el desempeño de la red y para esto se utiliza el conjunto de prueba; segundo, decidir cuando se detiene el entrenamiento, para este propósito se utiliza el conjunto de validación. Algunas veces el entrenamiento se detiene en el primer mínimo del conjunto de validación [8].

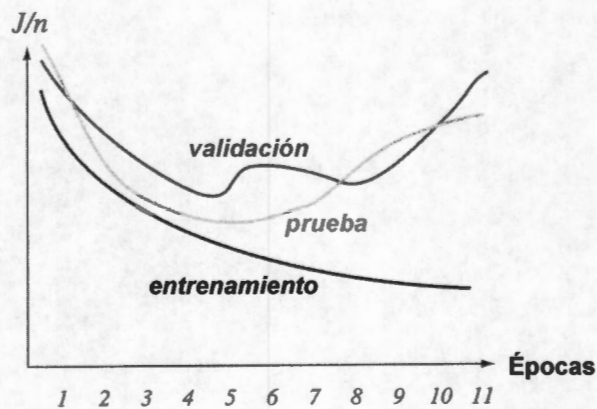


Figura 4.10. Curva de aprendizaje, muestra la función de costo como función de la cantidad de entrenamiento, indicado por el número de épocas. Se grafica el error promedio por patrón: $\frac{1}{n} \sum_{p=1}^n J_p$. La validación del error y la prueba o generalización del error por patrón son más grandes que el error de entrenamiento.

REFERENCIAS

-
- 1 Katz J. Michael, "Fractals and the analysis of waveforms," *Computers in Biology and Medicine*, vol. 18, no. 3, pp. 145-156, 1988.
 - 2 Metin Akay, "Nonlinear Biomedical Signal Processing, Dynamic Analysis and Modeling," vol. 2, IEEE PRESS, pp. 1-11, 2001.
 - 3 Haykin Simon, "Adaptive Filter Theory," Edit. Prentice Hall, 3a. ed., pp. 96 – 130, 1996.
 - 4 Proakis G. John, Manolakis G. Dimitris, "Digital Signal Processing, Principle, Algorithms, and Applications," Edit. Prentice Hall, pp. 876 – 878, 975, 931 – 932, 3a. ed., 1996.
 - 5 Therrien W. Charles, "Discrete Random Signals and Statistical Signal Processing," Edit. Prentice Hall, pp. 411-415, 564-567, 535-550, 1992
 - 6 Haykin Simon, "Neural Networks," Edit. Macmillan College Publishing Company, pp.1-41, 138-220, USA, 1994.
 - 7 Theodoridis Sergios, Koutroumbas Konstantinos, "Pattern Recognition," Edit. Elsevier Academia Press, pp. 104-112, 2ª ed., 2003.
 - 8 Duda O. Richard, hart E. Meter, Store G. David, "Pattern Classification", Edit. Wiley-Interscience Publication, pp. 282-333, 2ª ed., 2001.

CAPITULO V

METODOLOGÍA

La metodología del presente trabajo incluye las 7 etapas siguientes: 1) la adquisición multicanal de sonidos respiratorios (SR), 2) el preprocesamiento de la señal de SR para reducir interferencias cardíacas y musculares, 3) la simulación de sonidos adventicios discontinuos (crepitancias finas y gruesas), 4) la generación de una base de datos de SR, 5) la aplicación de técnicas de procesamiento con la finalidad de generar un índice que refleje el número y distribución espacial de crepitancias, 6) la evaluación del desempeño de los algoritmos propuestos en señales simuladas y en señales adquiridas en sujetos enfermos, y 7) la generación de una imagen funcional que refleje la presencia de crepitancias.

En la figura 5.1 se muestra un diagrama a bloques de la metodología seguida en este trabajo.

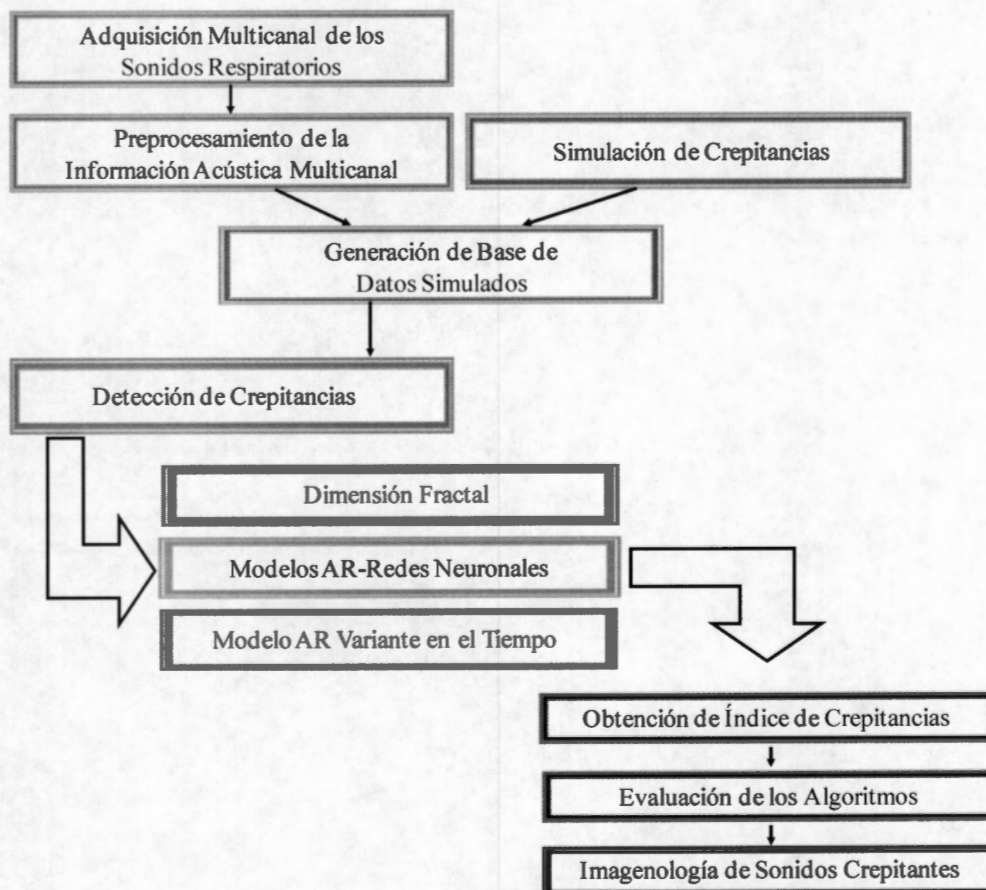


Figura 5.1. Diagrama a bloques de la metodología utilizada.

5.1 ADQUISICIÓN MULTICANAL DE LOS SONIDOS RESPIRATORIOS

Las señales de SR fueron tomadas de una base de datos de información acústica de sujetos sanos, el protocolo de adquisición fue realizado por integrantes del Laboratorio de Imágenes y Procesamiento de Señales en Ingeniería Biomédica (LIPSIB) de la Universidad Autónoma Metropolitana (UAM) en conjunto con personal médico del Laboratorio de Acústica del Instituto Nacional de Enfermedades Respiratorias (INER) y del Laboratorio de Fisiología Humana (UAM).

El equipo y los transductores utilizados para la adquisición de los sonidos respiratorios fue el siguiente:

- ◇ 25 micrófonos subminiatura tipo electret BT-1834 Knowles® con respuesta en frecuencia de 30 a 10 KHz, acoplados cada uno a una campana de acrílico para conformar un estetoscopio electrónico, cada una de ellos se fija a la superficie torácica con un disco de doble adherencia.
- ◇ Una computadora personal con procesador Pentium 4 de 3.4 GHz de velocidad y 512 MB de memoria RAM.
- ◇ Una tarjeta de conversión A/D National Instruments PCI-6071E de 12 bits de resolución. Este sistema tiene una capacidad máxima de 64 canales analógicos con opción de ganancias independientes seleccionadas mediante software. La máxima frecuencia de muestreo por canal es de 1.25×10^6 muestras/segundo. Para los registros se utilizaron 26 canales de entrada a una frecuencia de muestreo de 10 KHz. La tarjeta se configuró en modo de entrada única referenciada *SINGLE ENDED* y el rango de entrada analógico de los canales es de -5 a 5 V.
- ◇ Un neumotacógrafo tipo Fleish, acoplado al transductor de presión Validyne®, para adquirir el flujo en L/seg. La calibración del neumotacógrafo se realizó mediante un rotámetro.

El programa de adquisición fue diseñado en LabView® 6i [1], y permite adquirir 64 canales de forma continua o un número finito de muestras, además se pueden seleccionar ganancias distintas para cada canal. La información adquirida puede almacenarse en archivos binarios o de texto. Tomando en consideración el tiempo de adquisición, el

número de canales y la frecuencia de muestreo la información acústica se guardó en archivos de tipo binario y posteriormente se convirtieron a archivos de texto con un programa diseñado también en LabView®.

El protocolo de adquisición incluyó 25 micrófonos colocados en la pared torácica del sujeto en un arreglo de 5 x 5 micrófonos, con una distancia de 5 cm. entre micrófonos, como se muestra en la figura 5.2 (a). Mediante el arreglo de micrófonos se adquirieron sonidos respiratorios del tipo bronquial, bronquio-vesicular y vesicular, de acuerdo a la ATS. La nomenclatura para los micrófonos utilizada incluye una letra, P (Posterior), L (Left), R (Right), X (Axilar), C (Clavicular), M (Media) y un número del 1 al 5 como se observa en la figura 5.2 (b). Por ejemplo, si un micrófono esta en la posición PLX3, entonces el micrófono esta en la posición posterior izquierda axilar en el renglón 3 del arreglo de micrófonos. Durante el protocolo de adquisición los sujetos permanecieron sentados, usando una pinza nasal, respirando a través del neumatógrafo durante 15 segundos a un flujo de 1.5 L/seg.

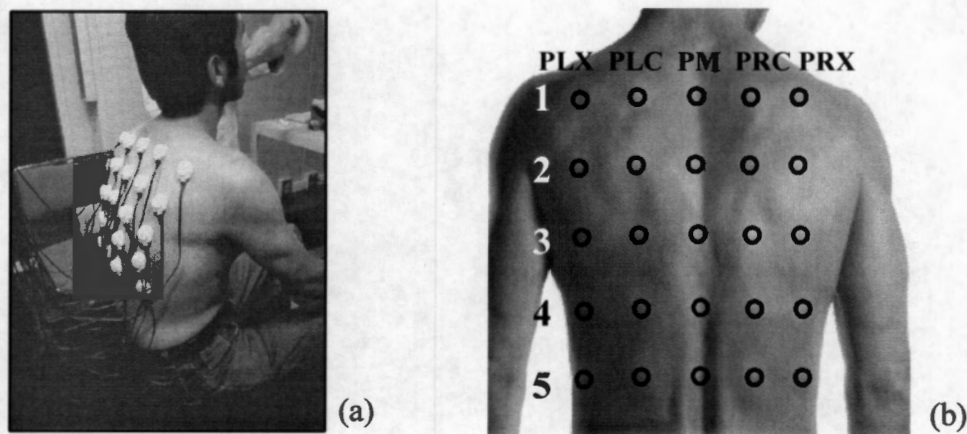


Figura 5.2. (a) Sujeto con 25 micrófonos colocados en la pared torácica, (b) Nomenclatura utilizada en el arreglo de micrófonos.

Además de los 25 canales correspondientes a las señales de sonidos respiratorios se adquirió un canal con la información correspondiente a la señal de flujo. La adquisición se llevó a cabo en un cuarto sub-amortiguado, para evitar ruidos de interferencia del tipo ambiental. Las características promedio de edad, peso y estatura de los 12 sujetos sanos incluidos en el protocolo se mencionan en la tabla 5.1.

Tabla 5.1
Características de los 12 sujetos sanos estudiados.

Edad	Peso	Estatura
24.3 ± 1.5	77.8 ± 11.0	174.8 ± 7.8

5.2 PREPROCESAMIENTO DE LA INFORMACIÓN ACÚSTICA MULTICANAL

Las señales de SR fueron calibradas y filtradas para reducir interferencias cardiacas y musculares, además las señales fueron segmentadas en información correspondiente a la fase inspiratoria o espiratoria.

Calibración

Se utilizaron micrófonos con ganancia y respuesta en frecuencia similar, y con el objetivo de evitar diferencias en los niveles de amplitud de los micrófonos se adquirió una señal de ruido de fondo en todos los canales, además, se calculó un factor de escala y se multiplicó por las señales de sonido respiratorio.

Filtrado de los sonidos respiratorios

Los sonidos respiratorios fueron procesados con un filtro FIR pasa banda de orden 500, con frecuencias de corte de 75 Hz y 1000 Hz, para reducir interferencias cardiacas y musculares. La señal de flujo fue procesada con un filtro pasa bajas tipo FIR de orden 500 y con frecuencia de corte de 1 Hz. La señal de flujo filtrada se requiere para realizar la segmentación del sonido respiratorio en sus fases inspiratoria y espiratoria. En la figura 5.3, a manera de ejemplo, se muestran las 25 señales de sonido respiratorio y la señal de flujo de un sujeto sano con su nomenclatura asociada, calibradas y filtradas.

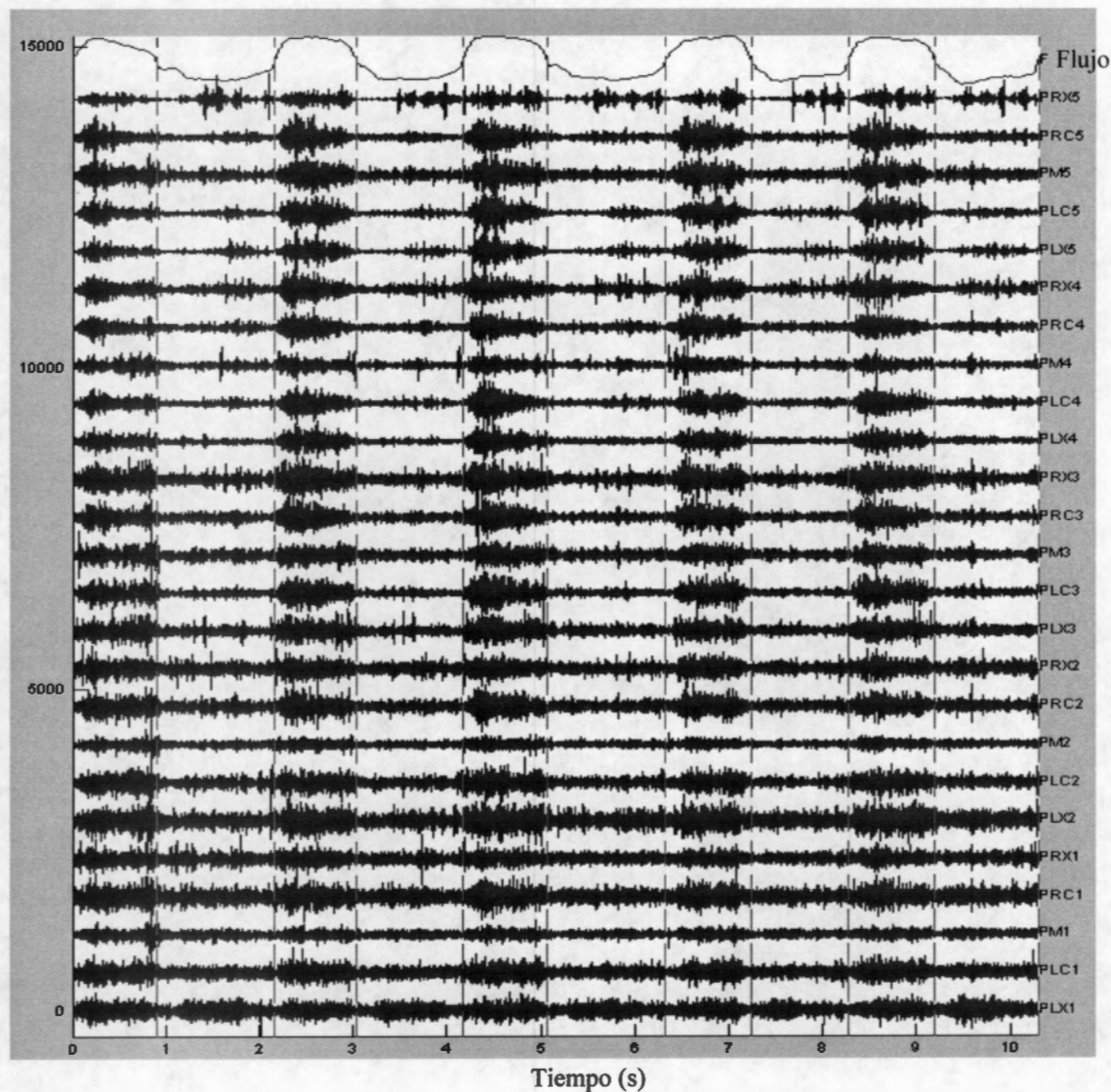


Figura 5.3. Gráfica de 26 señales adquiridas, calibradas y filtradas.

5.3 SIMULACIÓN DE SONIDOS ADVENTICIOS DISCONTINUOS (CREPITANCIAS)

Entre las principales características de las crepitancias se encuentran:

- Carácter explosivo y transitorio.
- Presencia en enfermedades cardio-respiratorias.
- Contenido en frecuencia de 100 – 2000 Hz o mayor.
- Clasificadas como finas o gruesas de acuerdo a características temporales de duración y de amplitud de sus diferente ondas.

Las crepitancias simuladas fueron generadas a través de un modelo matemático propuesto por Kiyokawa *et al.* [2]:

$$y_0(t) = \sin(4\pi^\alpha), \quad \text{con} \quad \alpha = \frac{\log(0.25)}{\log(t_0)}, \quad (1)$$

en donde $y_0(t)$ tiene dos ciclos con la primera intercepción positiva en el eje del tiempo, denominado t_0 . La función de modulación de $y_0(t)$ está dada por:

$$m(t) = 0.5 \{ 1 + \cos[2\pi(t^{0.5} - 0.5)] \} \quad (2)$$

La mayor parte de la potencia de $y(t) = y_0(t)m(t)$ se encuentra concentrada cerca del inicio de la señal. De acuerdo con la terminología de las crepitancias, IDW (ancho de la deflexión inicial) y 2CD (duración de dos ciclos) son parámetros que diferencian a las crepitancias finas y gruesas, de acuerdo con la ATS. En nuestras simulaciones en el caso de las crepitancias finas el IDW fue de 0.5 ms y el 2CD fue de 5 ms, mientras que para los crepitancias gruesas el IDW fue de 1.2 ms y el 2CD fue de 9 ms. En la figura 5.4 se muestra el comportamiento de una crepitancia fina (línea en color azul) y una gruesa (línea en color rojo), señalando los parámetros antes mencionados.

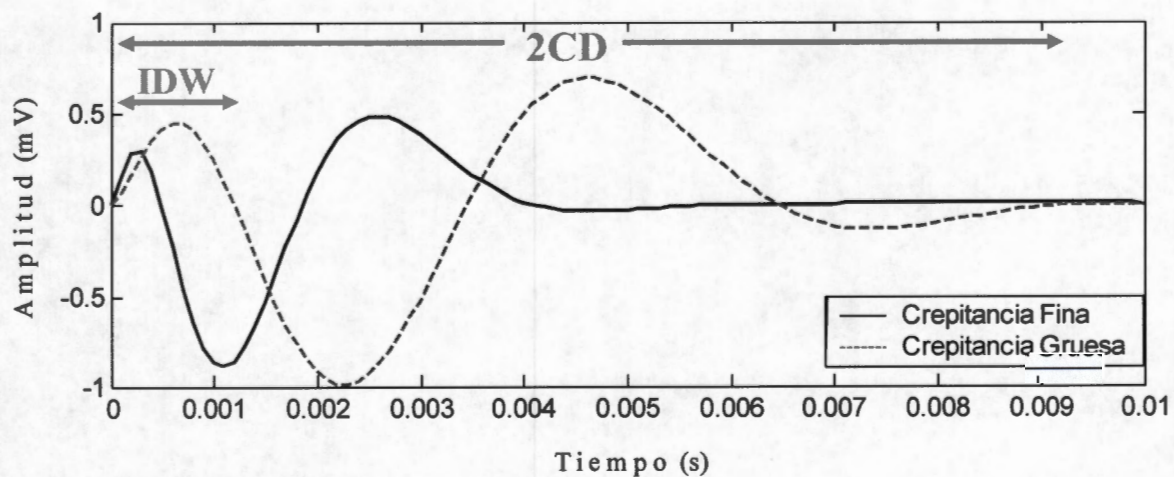


Figura 5.4. Gráfica de crepitancias finas y gruesas simuladas.

El proceso de inserción de crepitancias simuladas, esquematizado en la figura 5.5, en sonido respiratorio de base de sujetos sanos consistió en los siguientes pasos.

- a) En cada fase inspiratoria, con ayuda de la señal de flujo respiratorio, se seleccionó un segmento correspondiente a la segunda mitad de cada inspiración; se ha reportado en la literatura que las crepitancias aparecen frecuentemente en la segunda mitad de la fase inspiratoria [3].
- b) En cada segmento se seleccionó una posición temporal de forma aleatoria.
- c) A partir de la posición seleccionada se consideraron 150 muestras de la información acústica hacia delante y hacia atrás.
- d) Se calculó la desviación estándar del segmento de señal acústica seleccionado.
- e) La amplitud de la crepitancia a insertar, con el propósito de variar la relación crepitancia a sonido de base o relación señal a ruido (SNR), se determinó multiplicando la desviación estándar calculada por un factor. Los factores seleccionados fueron 1.5, 2.5 y 3.5.
- f) La crepitancia simulada se insertó mediante un modelo aditivo al sonido de base respiratorio del inciso (a) en la posición determinada en el inciso (c).

En la figura 5.6(a) se muestra una señal con 4 crepitancias finas insertadas, señaladas por líneas punteadas en color rojo, con un factor de escala de 2.5 y en la figura 5.6(b) se observa una de estas crepitancias, mediante una expansión del eje de tiempo, donde se aprecia de mejor manera el comportamiento morfológico de la crepitancia.

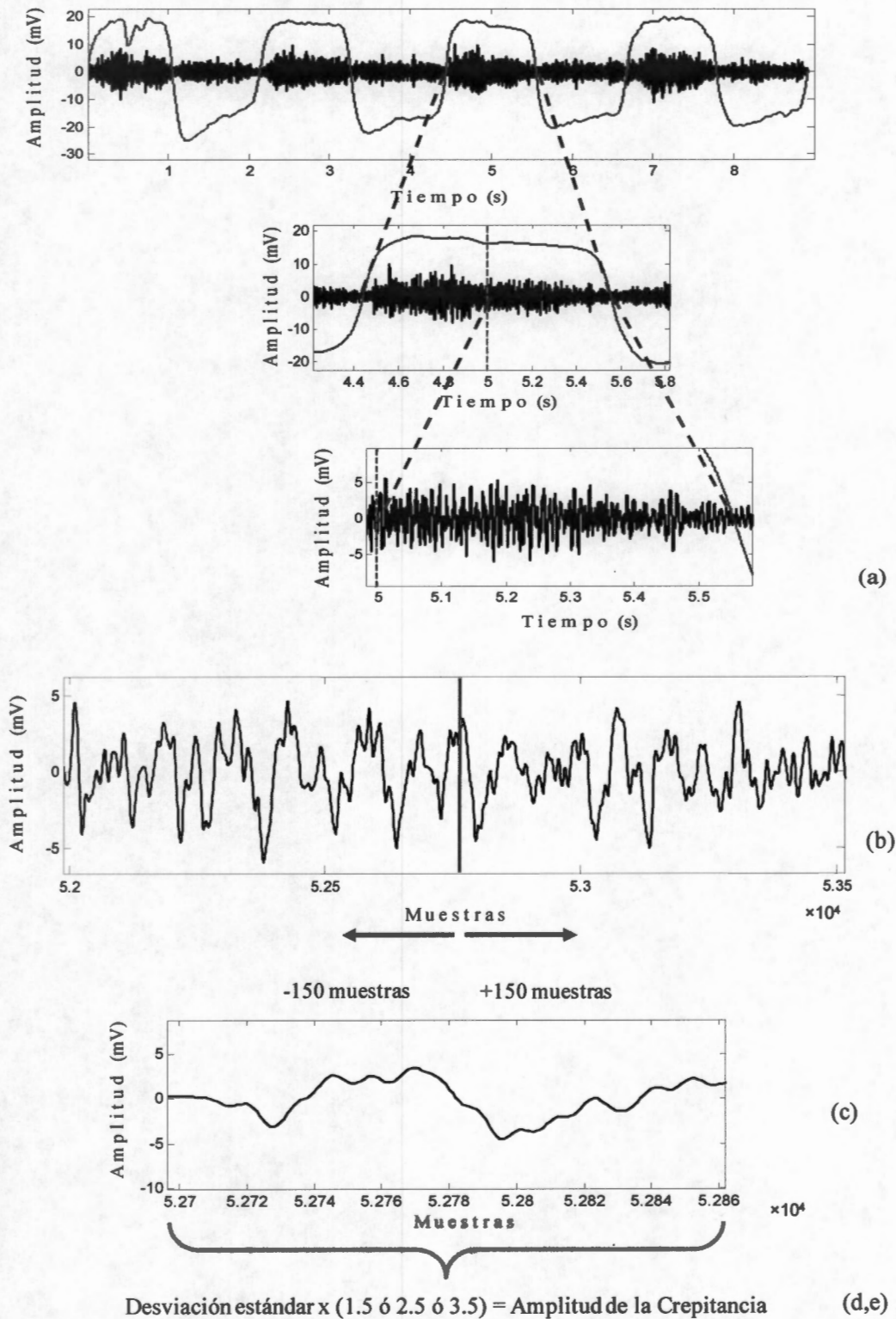


Figura 5.5. Proceso de inserción de crepitancias simuladas en el sonido respiratorio de base de sujetos sanos.

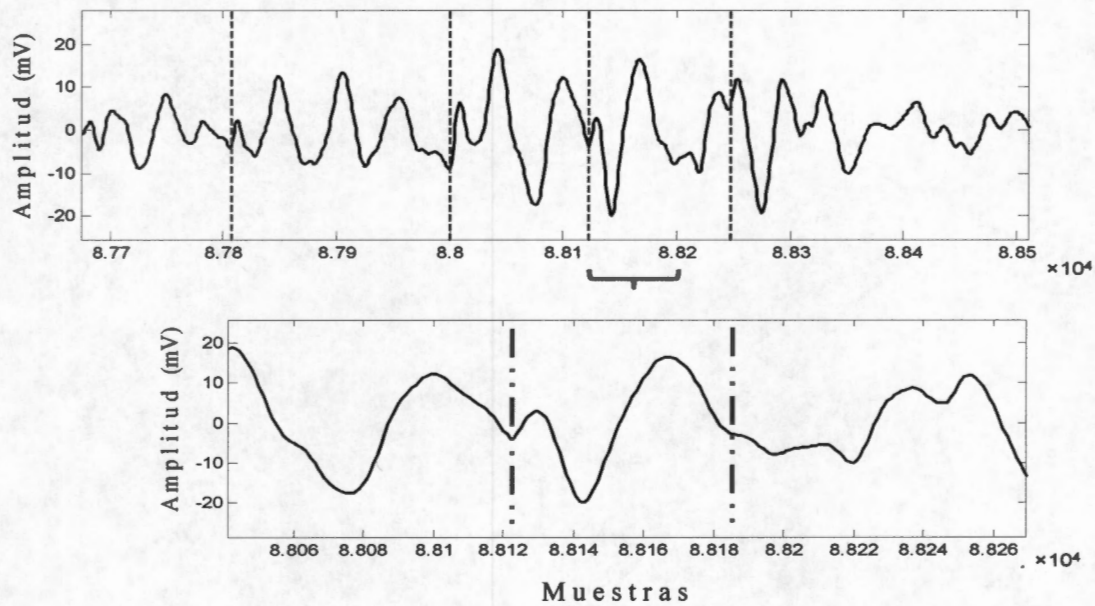


Figura 5.6. Señal de sonido respiratorio con 4 crepitancias insertadas artificialmente, las líneas rojas en (a) indican la posición inicial de una crepitancia, en (b) se muestra una de las crepitancias.

En una primera etapa de simulación, los parámetros temporales de las crepitancias IDW y 2CD se mantuvieron constantes, posteriormente, se generaron crepitancias con diferencias en los parámetros temporales. La variación de los parámetros temporales permitió considerar posibles cambios en los estertores crepitantes por cambios en los diámetros de las vías aéreas. Los valores para IDW y 2CD utilizados se muestran en la siguiente tabla:

Tabla 5.2.
Parámetros utilizados en la creación de diferentes tipos de crepitancias.

	IDW (ms)	2CD (ms)
1	0.07	0.01
2	0.4	3.5
3	0.7	6.8

Los diferentes tipos de crepitancias finas generadas con diferentes IDW, 2CD y amplitud, se muestran en la figura 5.7.

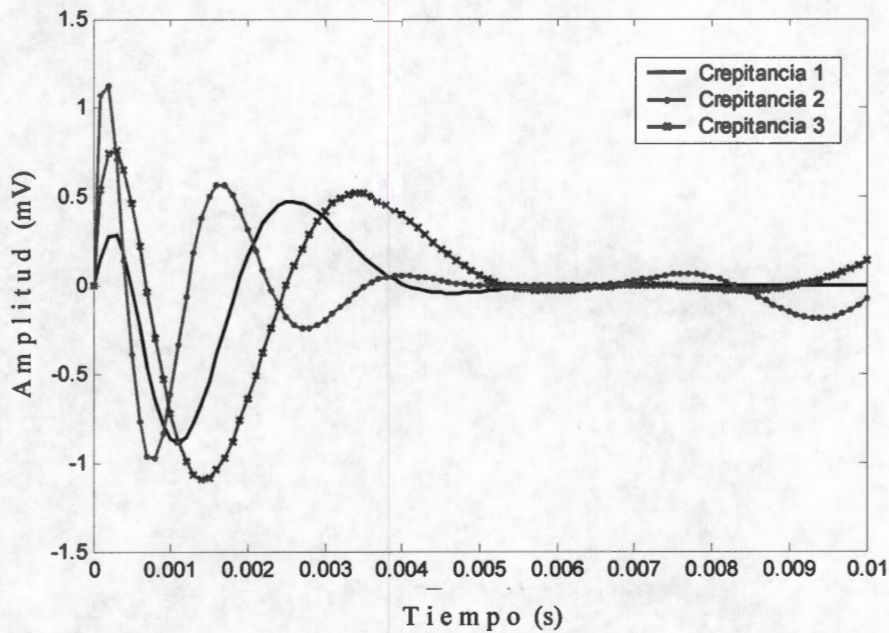


Figura 5.7. Gráfica de 3 crepitancias finas con diferentes IDW y 2CD.

5.4 GENERACIÓN DE LA BASE DE DATOS

La generación de la base de datos se realizó con el objetivo de probar los diferentes esquemas de procesamiento propuestos en esta tesis, dado que se tiene control de parámetros tales como: la relación SNR entre la crepitancia y el sonido de base, el número de crepitantes, el tipo de crepitantes y la zona pulmonar afectada.

Se adquirió la información acústica de 10 sujetos sanos, a partir de estos registros se generaron 3 registros adicionales agregando ruido blanco, de esta forma se obtuvo una base de datos con 40 registros de *sujetos sanos*. La desviación estándar del ruido blanco fue la desviación estándar de cada canal adquirido multiplicada por 0.04 y la media fue 0. Las señales respiratorias de *sujetos enfermos* se simularon insertando en forma aleatoria un número conocido de crepitancias en la segunda mitad de la fase inspiratoria. Se generó una base de datos con crepitancias finas y otra con crepitancias gruesas. Posteriormente se generó otra base de datos, insertando 3 crepitancias con diferentes IDW y 2CD y 2 crepitancias reales. En esta base de datos se insertaron 2 crepitancias de cada tipo, en total 10 crepitancias en cada canal. Esta base de datos sólo involucró crepitancias finas.

5.5 TÉCNICAS DE PROCESAMIENTO PARA GENERAR ÍNDICES DE CREPITANCIAS

5.5.1 DIMENSIÓN FRACTAL

Se han reportado resultados satisfactorios en la detección de crepitancias, utilizando dimensión fractal [4], por esa razón se exploró la técnica, bajo los criterios utilizados en esta investigación, es decir, con crepitancias simuladas y con un rango de relaciones señal a ruido. Con el uso de la dimensión fractal no se intenta identificar las características fractales de la señal pero si se intenta identificar las variaciones en la complejidad de la señal. La dimensión fractal fue calculada utilizando el algoritmo propuesto por Katz [5]:

$$DF = \frac{\log(n)}{\log(n) + \log\left(\frac{d}{L}\right)} \quad (3)$$

Se seleccionó una fase inspiratoria en cada canal de adquisición y fue dividida en ventanas de tamaño $0.006 \times$ frecuencia de muestreo [6], la ventana se traslapó un 99%, es decir, la dimensión fractal se calculó muestra a muestra.

Los valores de dimensión fractal obtenidos en las ventanas de la inspiración se promediaron y fueron interpolados para generar una imagen.

Posteriormente se obtuvo la dimensión fractal en una inspiración de cada uno de los 25 canales de adquisición y se interpoló para generar una imagen a la que se le denominó Mapa de Complejidad.

En la figura 5.8 se muestran tres señales básicas cuya dimensión fractal es conocida, con las cuales se comprobó el cálculo realizado de la dimensión fractal.

5.5.2 MODELO AR Y REDES NEURONALES ARTIFICIALES

La información acústica de la fase inspiratoria de la señal multicanal de sonido respiratorio fue segmentada en 30 ventanas temporales, para asegurar estacionaridad, y traslapadas un 25%. La parametrización de la información acústica en cada ventana se

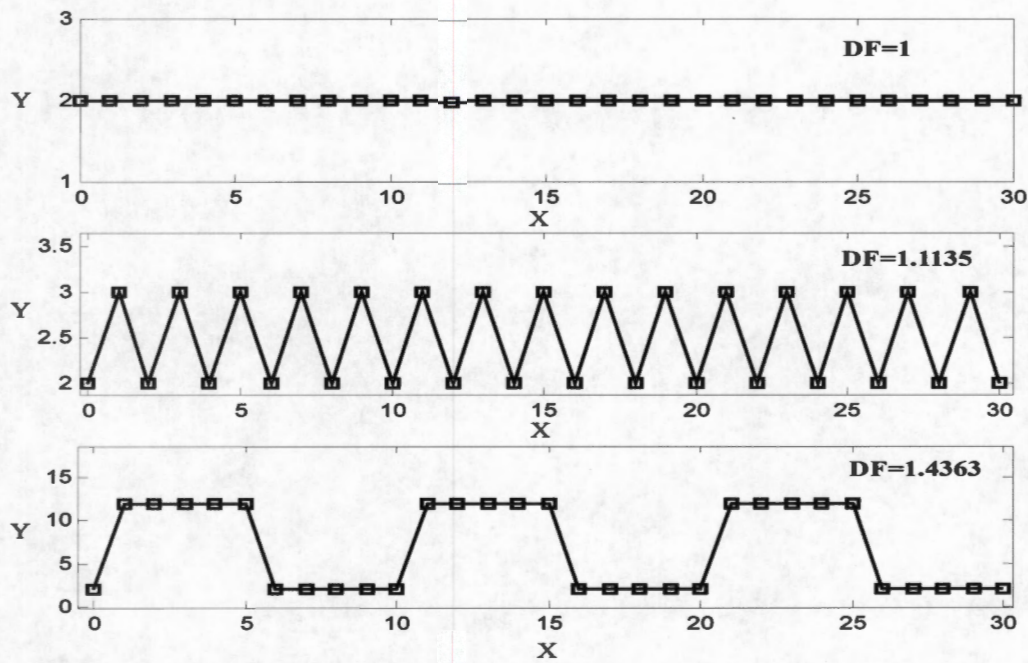


Figura 5.8. Series de tiempo con su respectiva dimensión fractal (DF).

realizó a través de un Modelo Autorregresivo (AR) y el orden del modelo AR se determinó a través del criterio de Akaike. Además, el orden del modelo se comprobó calculando el error residual del modelo, graficando la función de autocorrelación, y comprobando el comportamiento del error como ruido blanco. Debido a los diferentes escenarios de simulación se generaron modelos AR de orden 4 y 6. Los parámetros del modelo AR se calcularon en cada ventana temporal a través de la función ARX de MATLAB®.

$$A(q)x(t) = e(t)$$

donde $A(q) = 1 + a_1q^{-1} + \dots + a_pq^{-p}$

$$x(n) = -a_1x(n-1) - a_2x(n-2) - \dots - a_px(n-p) + e(n)$$

Se seleccionó una red neuronal artificial (RNA), con la finalidad de clasificar entre ventanas temporales enfermas y sanas utilizando un vector de características que contiene sólo los coeficientes del modelo AR. La arquitectura de la red neuronal fue tomada de un esfuerzo previo, en el cual se realizó la clasificación de la información acústica de sujetos sanos y enfermos con neumopatías intersticiales difusas [6]. La arquitectura seleccionada fue 20 neuronas en la capa de entrada, 10 neuronas en la capa oculta y una neurona en la

capa de salida, utilizando una función de transferencia tangencial sigmoidea. El entrenamiento de la RNA fue de tipo supervisado, el cual consiste en presentar un conjunto de entrenamiento a la RNA, encontrar la salida de la red y ajustar los pesos sinápticos para que la salida actual se aproxime a las etiquetas deseadas y definidas *a priori*. Los vectores de entrenamiento se presentaron de forma aleatoria y grupal a la RNA. La asignación de etiquetas a los vectores de características consistió en determinar cada ventana temporal como enferma o sana, dado que se conoce donde se ha insertado un crepitante. Se consideró como ventana temporal enferma, la ventana en donde se encontrara una o más crepitancias y para considerar a una crepitancia dentro de una ventana temporal, por lo menos el 50% de la extensión temporal de la crepitancia debería estar dentro de la ventana temporal.

El entrenamiento de la RNA se realizó utilizando la interfaz gráfica de MATLAB® NNTOOL, que permite al usuario desempeñar funciones como:

- ◇ Importar los datos de entrada, vectores de características, valores deseados o etiquetas y RNAs del espacio de trabajo o de un archivo.
- ◇ Exportar datos como salidas y RNAs entrenadas al espacio de trabajo o un archivo.
- ◇ Crear una RNA.
- ◇ Inicializar datos de la RNA.
- ◇ Configurar parámetros de entrenamiento del clasificador.
- ◇ Entrenar y simular las RNAs creadas.

El entrenamiento se realizó utilizando el algoritmo de aprendizaje de retropropagación (*FeedForward Backpropagation*), los valores de los pesos y el sesgo se actualizaron de acuerdo a la optimización de Levenberg-Marquardt, se utilizó una función de aprendizaje de gradiente descendente con momento. El entrenamiento se llevó a cabo durante 100 o menos épocas de acuerdo al error alcanzado y el desempeño de la RNA fue evaluado con el error cuadrático medio, con base a la curva de aprendizaje. La figura 5.9 muestra la parte inicial de la interfaz NNTOOL, donde se aprecian las funciones como importar y exportar datos, crear, inicializar, simular, entrenar y adaptar redes neuronales.

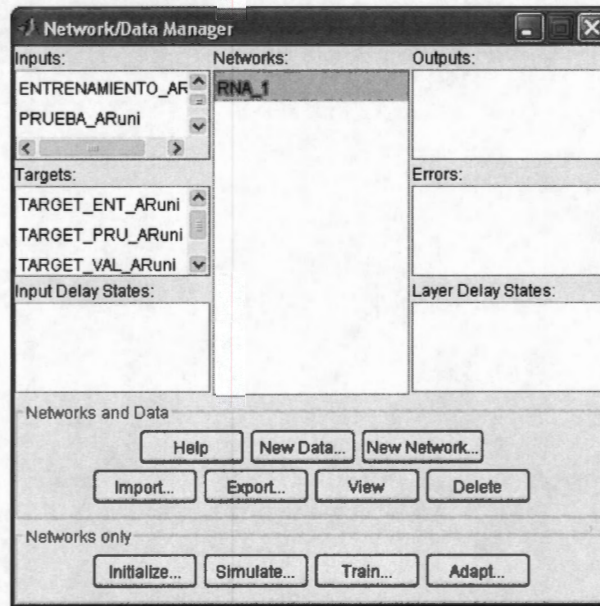


Figura 5.9. Interfaz gráfica NNTOOL de MATLAB® para realizar el entrenamiento de RNAs.

La parte importante de esta sección es la forma en que se suministraron los datos o vectores de características a la RNA para el entrenamiento. Se realizó un entrenamiento por regiones y por canales. En el entrenamiento por regiones, el arreglo de 5 x 5 canales o micrófonos fue dividido en 7 regiones, cada región incluye la parametrización de 4 canales y sólo se consideran 2 regiones traslapadas, como se observa en la figura 5.10. Por ejemplo, la región 1 incluye la información de los canales 1, 2, 3 y 4 y la región 2 contiene la información de los canales 2, 3, 4 y 5.

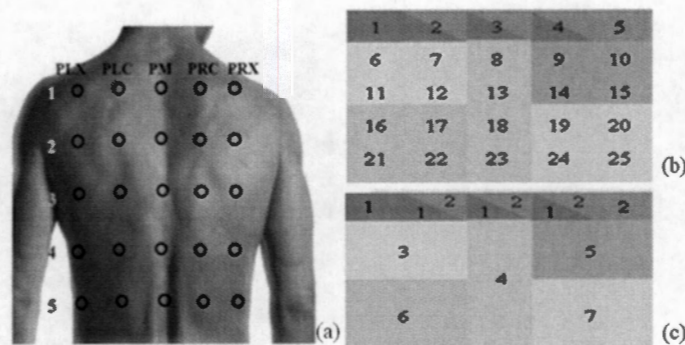


Figura 5.10. (a) Arreglo de micrófonos, (b) Número de sensores y definición de regiones de acuerdo al color, (c) Regiones enumeradas. Por ejemplo, la región 5 incluye los canales 9, 10, 14 y 15.

El entrenamiento por regiones se realizó con el propósito de distinguir entre regiones enfermas y sanas dentro del arreglo de 5 x 5 canales de adquisición. En la figura 5.11 se muestra la nomenclatura utilizada en el *toolbox* NNTOOL. P_1 es el vector de entrada con un número determinado de parámetros o características. En el caso del entrenamiento por regiones, el vector de entrada es de tamaño 16 (4 coeficientes AR por cuatro canales por región); en el caso del entrenamiento por canales, el vector de entrada es de tamaño 4. La nomenclatura IW y LW representan las matrices de pesos que ponderan a un valor de entrada, a_1 y a_2 son las salidas de la capa oculta y capa de salida, respectivamente.

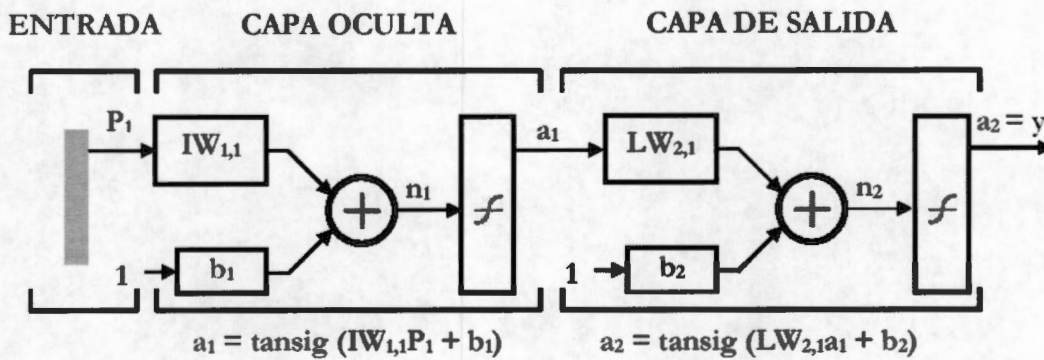


Figura 5.11. Esquema de la arquitectura de una RNA y la nomenclatura utilizada en la interfaz NNTOOL.

En el caso del entrenamiento por regiones se tienen vectores de características de 16×1050 ; el valor de 16 = [4 parámetros AR \times 4 canales que componen a una región] y 1050 = [30 ventanas temporales \times 5 inspiraciones \times 7 regiones]. En la figura 5.12 se muestra la forma de concatenar los vectores de características en el entrenamiento por regiones, donde cada vector de entrada es una de las columnas de coeficientes por ejemplo, $a_{1,1,1}$ hasta $a_{R,1,p}$ corresponde a la primera ventana temporal.

En el caso del entrenamiento por canales, los vectores de características se concatenaron como se muestra en la figura 5.13. Por ejemplo, un vector de características de tamaño 6×3000 , corresponde a 6 parámetros autorregresivos y $3000 = [30 \text{ ventanas temporales} \times 4 \text{ inspiraciones} \times 25 \text{ canales}]$.

Los valores deseados a la salida de la RNA, se consideraron de +1 para las ventanas temporales sanas o sin crepitancias y -1 para las ventanas temporales con presencia de crepitancias. La asignación a una de las dos clases de los valores de la salida de la RNA se realizó a través de un umbral. Los vectores de características que generaron un valor de salida fuera de este umbral fueron descartados y por lo tanto, no se asignaron a ninguna clase. En la tabla 5.3 se presenta el umbral utilizado.

Tabla 5.3.
Umbral para determinar a los vectores de salida
como sanos o enfermos.

Vector	Valor deseado	Umbral de clasificación
Sano	1	$\geq + 0.8$
Enfermo	-1	$\leq - 0.8$

Para detener el entrenamiento y el ajuste de pesos sinápticos de la RNA no existe un criterio bien definido. En la presente tesis se utilizó el método donde se dividen los datos disponibles en tres subconjuntos para brindar una mejor generalización. Los tres subconjuntos son entrenamiento, validación y prueba. El subconjunto de entrenamiento es usado para calcular el gradiente y actualizar los pesos y términos de sesgo de la RNA. En el subconjunto de validación se monitorea el error alcanzado durante el proceso de entrenamiento, que normalmente decrece durante la fase inicial de entrenamiento; el clasificador se entrena hasta que el error del subconjunto de validación alcanza un mínimo. Sin embargo, cuando la red comienza a sobre ajustar los datos, el error con el conjunto de validación comenzará a incrementarse. Si esto sucede para un número específico de iteraciones, se sugiere detener el entrenamiento, esto se ejemplifica en la figura 5.14. El conjunto de prueba no es usado durante el entrenamiento, pero es utilizado para comparar diferentes modelos o arquitecturas de red [7].

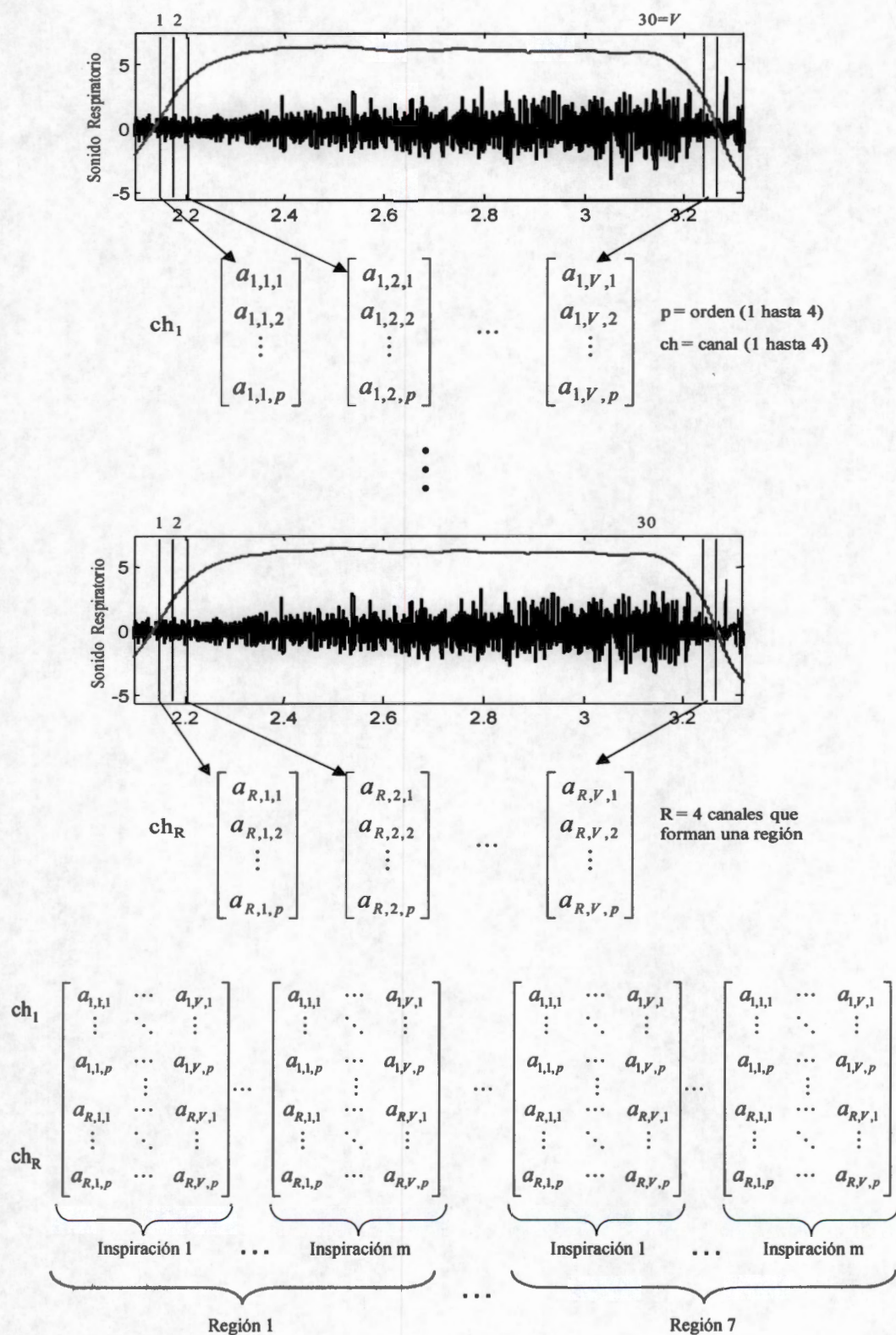


Figura 5.12. Concatenación de los vectores de características para entrenamiento por regiones.

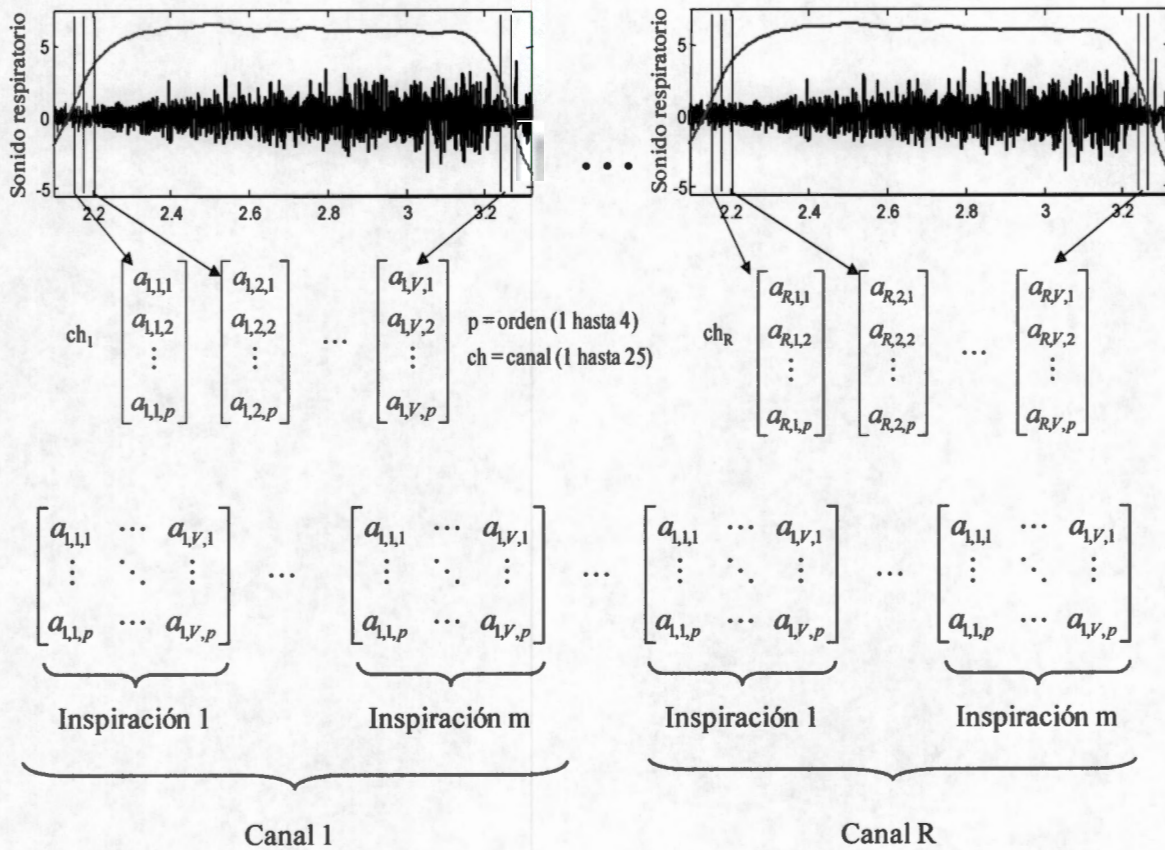


Figura 5.13. Concatenación de los vectores de características para entrenamiento por canales.

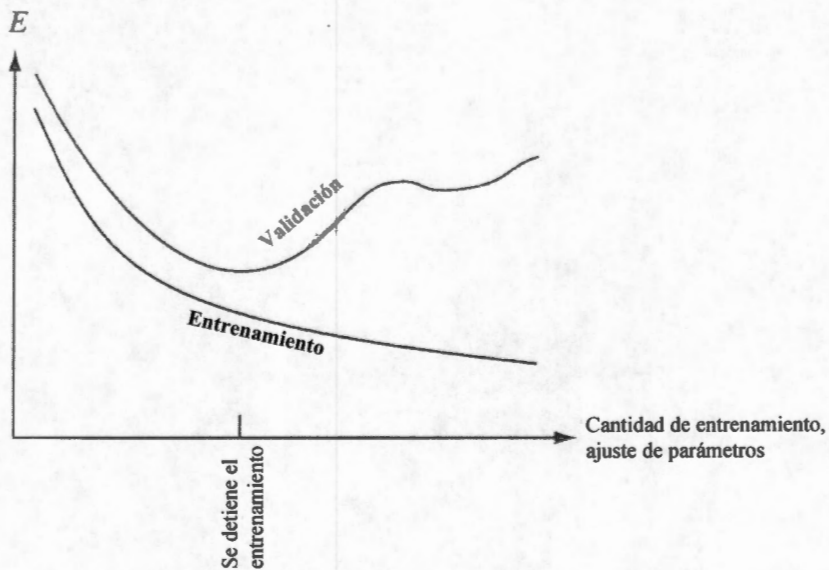


Figura 5.14. Curva de aprendizaje donde se muestra un criterio para detener el entrenamiento.

El conjunto total de vectores de características se dividió aleatoriamente en tres subconjuntos: entrenamiento (80 % de los datos), prueba (10 %) y validación (10 %). El desempeño de la red se evaluó en base a las siguientes medidas:

$$SENSIBILIDAD = \frac{\text{normal_normal}}{\text{normal_normal} + \text{enfermo_normal}} \times 100 \quad (4)$$

$$ESPECIFICIDAD = \frac{\text{enfermo_enfermo}}{\text{enfermo_enfermo} + \text{normal_enfermo}} \times 100 \quad (5)$$

$$DESEMPEÑO = \frac{\text{Sensitividad} + \text{Especificidad}}{2} \quad (6)$$

Después del entrenamiento de la RNA, se propuso un índice para generar una imagen, usando una interpolación Hermite, para evidenciar la presencia de crepitancias. El índice utilizado fue el siguiente:

$$\text{Indice} = \frac{\% \text{ventanas temporales clasificadas como enfermas}}{\% \text{ventanas temporales clasificadas como sanas}} \quad (7)$$

Para analizar el comportamiento del esquema de procesamiento generado al combinar el modelo AR con un clasificador mediante RNA, se generaron diferentes distribuciones espaciales de regiones enfermas con crepitancias simuladas que la red desconocía. Los resultados fueron evaluados con base a la similitud entre las regiones enfermas y las regiones detectadas como enfermas a través de la RNA. En las imágenes generadas un color rojo intenso indica una gran cantidad de crepitancias, mientras que un color azul indica ausencia de crepitancias.

El comportamiento del esquema AR-RNA se analizó con datos de sujetos desconocidos por el clasificador. Además, se evaluó el esquema AR-RNA con datos de sujetos enfermos, para comprobar si era posible distinguir ventanas temporales con crepitancias reales.

5.5.3 MODELO AR VARIANTE EN EL TIEMPO

Otra técnica utilizada en la parametrización de la información acústica fue un modelo AR variante en el tiempo, donde los coeficientes del modelo se estimaron usando un esquema de filtrado adaptable a través del algoritmo de mínimos cuadrados recursivos (RLS). El orden del modelo AR fue 4 y fue determinado por el criterio de Akaike. Los coeficientes autorregresivos se obtuvieron a partir de la información acústica de cada fase inspiratoria. El factor de olvido del modelo AR variante en el tiempo fue determinado con base a prueba y error, y se determinó que el mejor valor fue $\lambda = 0.97$.

El procedimiento para estimar el número de crepitancias en cada canal fue el siguiente:

- ◇ Se estimaron los coeficientes autorregresivos a través del algoritmo RLS.
- ◇ Se obtuvo la derivada de cada serie de tiempo del coeficiente estimado.
- ◇ La derivada fue segmentada en ventanas de tamaño 40 muestras, el tamaño del segmento fue determinado por la duración de la crepitancias.
- ◇ En cada segmento, se calculó la desviación estándar.
- ◇ Si la desviación estándar de cada segmento rebasaba un umbral establecido, se consideró que la desviación de la señal representa una crepitancia.
- ◇ Se determinó la presencia de una crepitancia, siempre que en la derivada de las cuatro series de tiempo de los coeficientes AR, la desviación estándar del segmento sobrepasara el umbral.

Después de estimar un número de crepitancias en cada canal, se generó una imagen acústica a partir de la matriz de información de 5 x 5 utilizando un interpolador tipo Hermite. En la figura 5.15 se esquematizan los pasos previamente explicados para la detección de crepitancias.

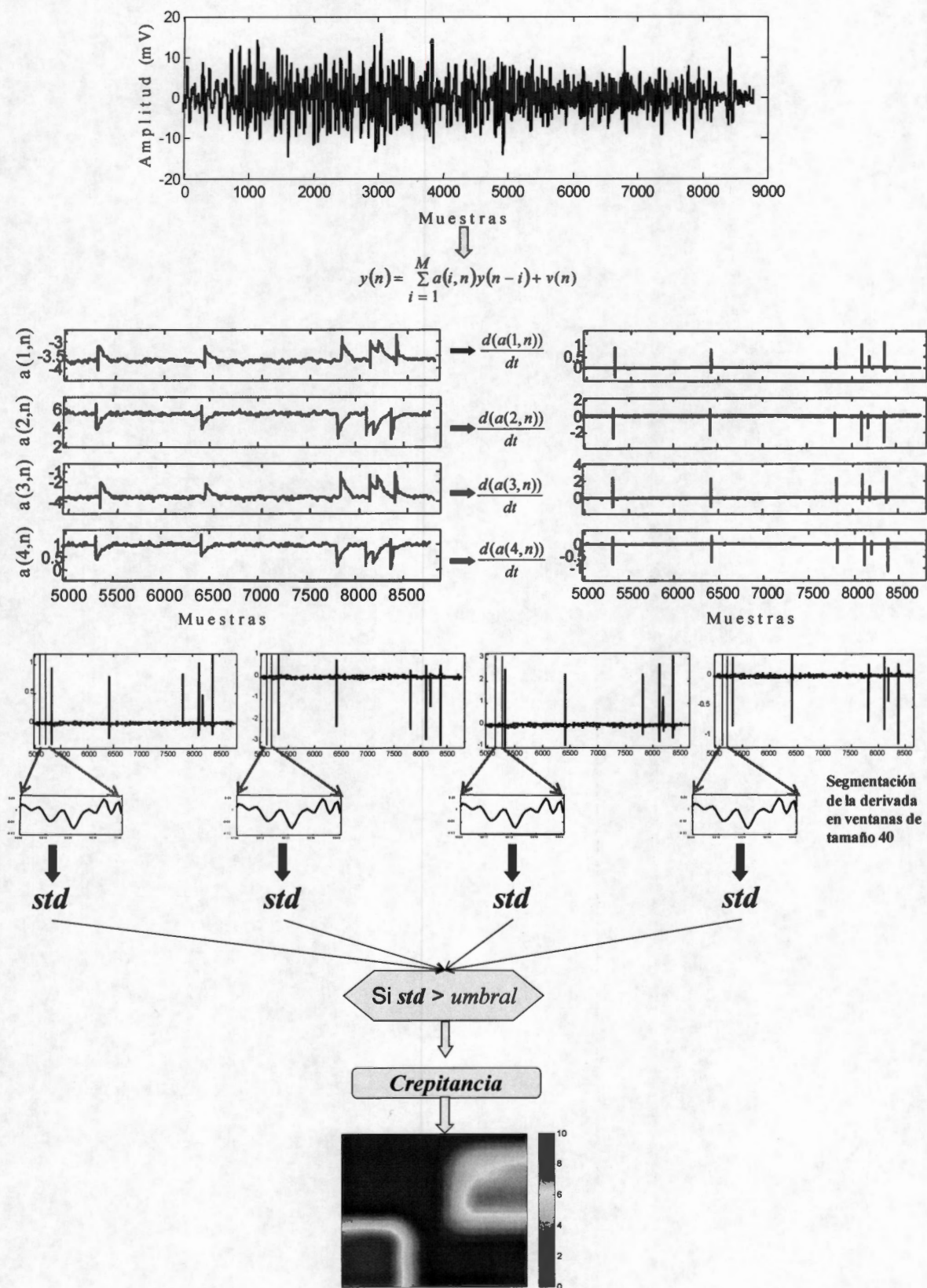


Figura 5.15. Esquema de detección de crepitancias, utilizando un modelo AR variante en el tiempo.

5.6 EVALUACIÓN DEL DESEMPEÑO DE LOS ALGORITMOS PROPUESTOS

El desempeño de los algoritmos se evaluó probando con crepitancias simuladas y crepitancias reales, es decir, con sujetos enfermos simulados y reales. Con el propósito de evaluar los esquemas de procesamiento utilizados en la presente tesis, se simularon 3 sujetos enfermos. En la tabla 5.4, se reportan las regiones enfermas y el número de crepitancias para cada sujeto. La representación gráfica de la tabla 5.4 se muestra en la figura 5.16.

Tabla 5.4

Regiones con crepitancias simuladas en sujetos de prueba

Sujeto 1	
Región	No. Crepitancias
5	6
6	10
Sujeto 2	
Región	No. Crepitancias
5	6
6	4
7	6
Sujeto 3	
Región	No. Crepitancias
5	8
7	10

Los resultados obtenidos a través de los algoritmos propuestos en este trabajo, se evaluaron con base a la aproximación de las tres imágenes patrón generadas a partir de la Tabla 5.4. Es importante mencionar que las imágenes patrón se obtuvieron con base al número de crepitancias, por lo tanto, son fácilmente comparables con las imágenes obtenidas por medio del modelo AR variante en el tiempo, ya que este método brinda un número estimado de crepitancias; mientras que el procedimiento seguido con redes neuronales genera una imagen a través de un índice con base al porcentaje de ventanas temporales clasificadas como sanas y enfermas. Sin embargo, las imágenes obtenidas deben

ser similares a las imágenes patrón en términos de la distribución de crepitancias y tal vez con diferencias en cuanto a la intensidad de color en la imagen.

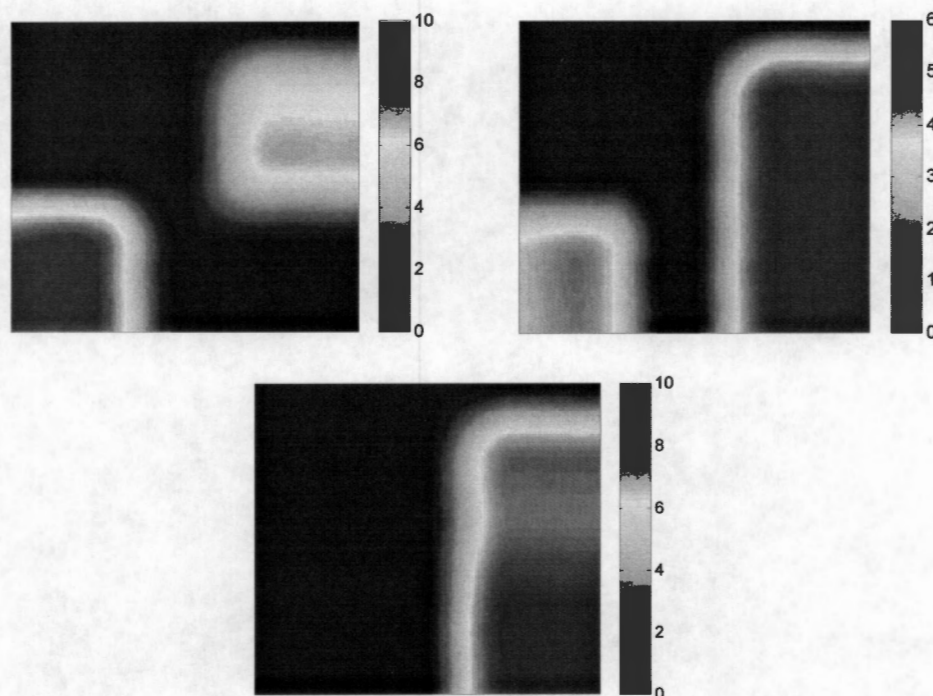


Figura 5.16. Imágenes patrón generadas para la evaluación de los algoritmos propuestos.

Además de la evaluación a través de sujetos enfermos simulados de prueba, se realizó una evaluación mediante un experto médico. La evaluación consta de dos esquemas; el primero de ellos consistió en brindarle al experto una señal de sonido respiratorio con crepitancias simuladas, donde el factor de escala, el número y la posición temporal de las crepitancias fueron desconocidas por el experto. El registro de sonido respiratorio en el que el experto realizó el conteo contiene dos canales PRC2 y PLX4. En el canal PRC2 se insertaron 6 crepitancias simuladas y en PLX4 se insertaron 10 crepitancias simuladas, con un factor de escala de 3.5. El segundo esquema consistió en comparar el conteo realizado por el experto y por el algoritmo del modelo AR variante en el tiempo, en dos registros de sonidos respiratorios, con crepitancias reales de pacientes diagnosticados con alveolitis extrínseca alérgica. En ambos esquemas la detección del experto se basó en el algoritmo de Murphy, TEWA, aunado a su experiencia.

5.7 IMAGENOLÓGÍA DE SONIDOS CREPITANTES

La formación de imágenes es reciente en el estudio de los sonidos respiratorios [8,9]. En este trabajo se introdujo el concepto de la imagenología de sonidos crepitantes, como ayuda al experto, dado que la imagen puede indicar la localización, distribución espacial o el número aproximado de crepitancias. A partir de los resultados obtenidos con los algoritmos propuestos, se obtuvo una matriz de 5×5 , ya sea con un número estimado de crepitancias a través del modelo AR variante en el tiempo o con un índice del porcentaje de ventanas temporales enfermas o sanas, utilizando redes neuronales. Esta matriz de 5×5 fue interpolada utilizando un interpolador tipo Hermite. Además se asignó una barra de color a la imagen generada, donde un color rojo intenso indica un número o un porcentaje mayor de crepitancias, y un color azul indica ausencia de crepitancias.

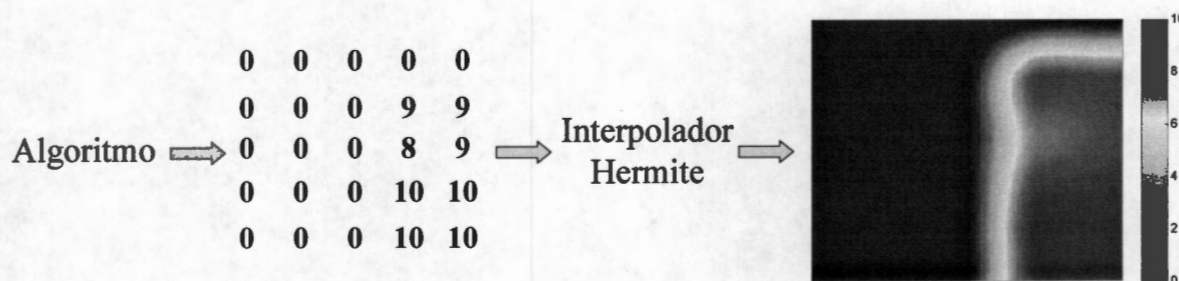


Figura 5.17. Esquema de la generación de una imagen a partir del índice de crepitancias.

REFERENCIAS

-
- [1] González Flores Tomás, "Sistema de Adquisición de Sonidos Respiratorios en un ambiente LabView," *Seminario de proyectos I y II*, UAM-I, 2004.
- [2] H. Kiyokawa, M. Greenberg, K. Shiota, and H. Pasterkamp, "Auditory detection of simulated crackles in breath sounds," *Chest*, vol. 119, no. 6, pp. 1886-1892, 2001.
- [3] A. R. A. Sovijärvi, L. P. Malmberg, G. Charbonneau, J. Vanderschoot, F. Dalmaso, C. Sacco, M. Rossi, J. E. Earis, "Characteristics of breath sounds and adventitious respiratory sounds." *European Respiratory Review*, 10(77), pp. 591-596, 2000.
- [4] L. J. Hadjileontiadis, I. T. Rekanos, "Detection of explosive lung and bowel sounds by means of fractal dimension," *IEEE Signal Processing Letters.*, vol. 10, no. 10, pp. 311-314, 2003.
- [5] Katz J. Michael, "Fractals and the analysis of waveforms," *Computers in Biology and Medicine*, vol. 18, no. 3, pp. 145-156, 1988.
- [6] Martínez Hernández Graciela E., "Evaluación acústica de las neumopatías intersticiales difusas utilizando análisis multicanal," Tesis de Maestría en Ciencias, UAM-I, 2006.
- [7] Duda O. Richard, Hart E. Meter, Store G. David, "Pattern Classification," Edit. Wiley-Interscience Publication, pp. 483-485, 2ª ed., 2001.
- [8] S. Charleston-Villalobos, S. Cortés-Rubiano, R. González-Camarena, G. Chi-Lem, T. Aljama-Corrales, "Respiratory acoustic thoracic imaging (RATHI): assessing deterministic interpolation techniques," *Medical & Biological Engineering & Computing*, vol. 42, pp.618-626, Sept 2004.
- [9] R.P. Dellinger, J.E. Parrillo, A. Kushnir, M. Rossi, I. Kushnir, "Dynamic visualization of lung sounds with a vibration response device: A case series," *Respiration*, vol. 75, pp. 60-72, 2008.

CAPITULO VI

RESULTADOS

En los párrafos subsecuentes se presentan los resultados obtenidos al aplicar los esquemas de procesamiento digital descritos en el capítulo referente a la metodología utilizada en la presente tesis. Para la presentación de los resultados es conveniente recordar lo siguiente: a) con el propósito de evaluar los diferentes esquemas propuestos, se simularon sujetos enfermos con diferentes características. En los escenarios que se simularon se controló el número de crepitancias, su distribución espacial, el tipo de crepitancia, y la relación de amplitud entre la crepitancia y el sonido pulmonar de fondo, b) considerando que los crepitantes se insertaron de forma aleatoria, en las fases inspiratorias y espiratorias, existió traslape temporal de dos o más crepitantes, c) se definió un índice para generar una imagen de sonidos adventicios. En el caso de la metodología que involucra redes neuronales el índice se definió con base al porcentaje entre ventanas temporales clasificadas como sanas y enfermas, d) En el caso del esquema con base a ARVT se comparó su desempeño con respecto al desempeño de un experto en fononeumografía al contar el número de crepitancias simuladas. El orden de la presentación de resultados es la siguiente:

1. Influencia de la relación de amplitud entre la señal (crepitancia) y el ruido (sonido pulmonar de fondo) en la morfología del crepitante.
2. Detección de crepitancias mediante dimensión fractal.
3. Detección de crepitancias mediante parametrización con un modelo AR invariante en el tiempo y clasificación por redes neuronales artificiales.
4. Detección de crepitancias con un modelo AR variante en el tiempo.
 - a. Detección de crepitancias finas y gruesas vía el modelo AR variante en el tiempo.
 - b. Desempeño del modelo AR variante en el tiempo vs. experto médico en la detección de crepitancias simuladas y casos reales.
 - c. Imagenología de sonidos adventicios discontinuos en casos reales.

6.1 INFLUENCIA DE LA RELACIÓN DE AMPLITUD ENTRE LA SEÑAL (CREPITANCIA) Y EL RUIDO (SONIDO PULMONAR DE FONDO) EN LA MORFOLOGÍA DEL CREPITANTE.

Uno de los aspectos importantes en la detección de crepitantes es su magnitud. Es importante recordar que en el caso de la técnica de TEWA es requisito que los crepitantes presenten una amplitud del doble de la señal pulmonar de fondo para posteriormente, evaluar la morfología y extraer parámetros como 2CD e IDW. Sin embargo, la amplitud de los crepitantes con relación al sonido pulmonar dependerá de varios factores, entre ellos el flujo respiratorio, el calibre de las vías aéreas anormalmente cerradas y las propiedades elásticas de las vías aéreas. Por lo tanto, es conveniente analizar el efecto de diferentes SNRs. En esta sección únicamente se realiza un análisis visual de las crepitancias insertadas y posteriormente, las señales simuladas con diferentes SNR se someten a la metodología establecida para su detección.

Para modificar la relación señal a ruido (SNR) de las crepitancias, se utilizaron tres factores: 1.5, 2.5 y 3.5; es decir, la amplitud de la crepitancia se determinó multiplicando la desviación estándar del sonido pulmonar de base, proveniente de un sujeto sano y del segmento temporal donde se insertó la crepitancia, por alguno de los tres factores de escala mencionados. En la figura 6.1 se observan señales con crepitancias, insertadas de forma aleatoria, con los tres factores de escala, la señal de sonido respiratorio es del mismo sujeto. Se pueden apreciar diferencias evidentes entre los factores de escala 1.5, donde algunas crepitancias son imperceptibles, y no serían detectadas aplicando el criterio TEWA, y el factor de escala 3.5, donde algunas crepitancias se pueden identificar visualmente con facilidad; por ejemplo, la flecha en color negro indica una de las crepitancias insertadas y detectadas visualmente. Así también, es necesario enfatizar el cambio de morfología que sufre el sonido crepitante al ser insertado, utilizando un modelo aditivo, al sonido pulmonar de base. Como puede observarse, la morfología de los crepitantes finos y gruesos no se conserva, si la comparamos con la morfología de la figura 5.4 del capítulo de metodología.

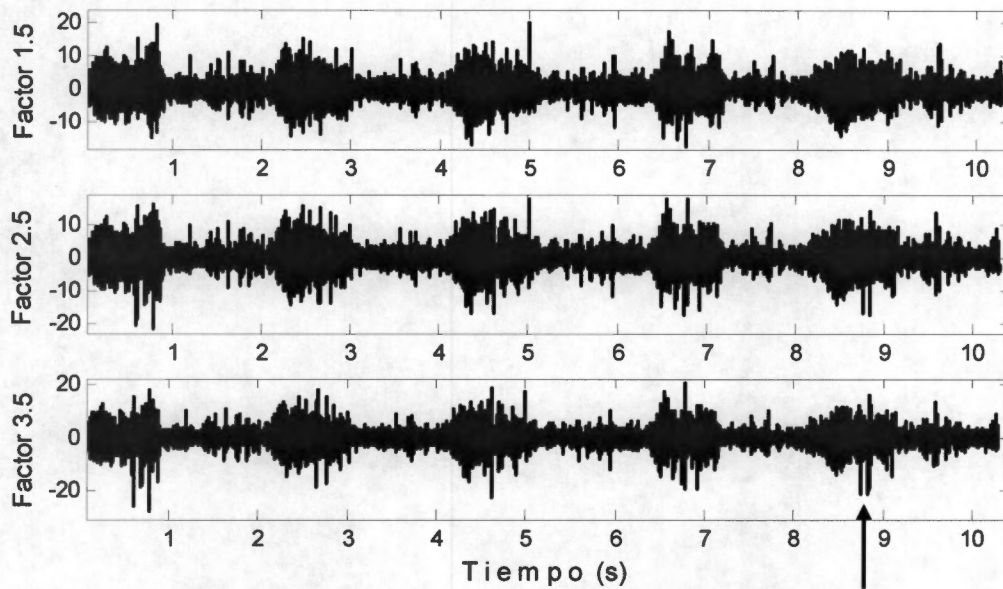


Figura 6.1. Señales de sonido respiratorio con crepitancias simuladas insertadas con diferentes factores de escala. (a) factor 1.5, (b) factor 2.5, (c) factor 3.5.

En la figura 6.2 se observa la expansión en el tiempo de la información acústica de la quinta fase inspiratoria de las señales en la figura 6.1, nuevamente las crepitancias con un factor de escala de 1.5 son difíciles de ubicar visualmente, mientras que algunas crepitancias con factor de escala de 2.5 ó 3.5 se pueden apreciar fácilmente. Las pequeñas líneas de color rojo indican las crepitancias insertadas en este ejemplo.

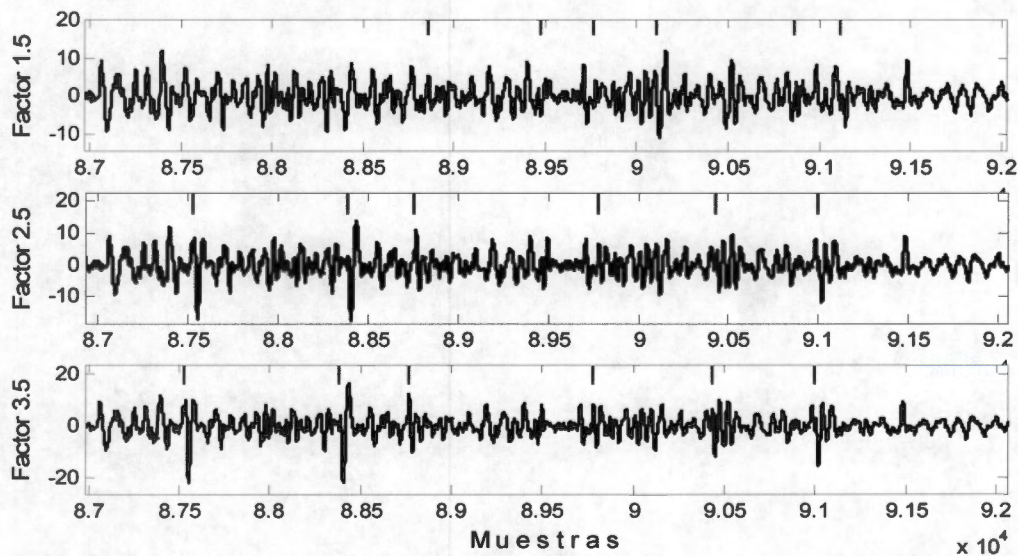


Figura 6.2. Quinta inspiración de la figura 1, con los tres factores de escala, las líneas rojas indican el inicio de un estertor crepitante.

6.2 DETECCIÓN DE CREPITANCIAS POR MEDIO DE DIMENSIÓN FRACTAL

Dado que se ha presentado un interés reciente en técnicas de análisis no lineal de sistemas y señales biológicas, en esta tesis se exploró el uso de la dimensión fractal para la detección de crepitancias, con base a los esfuerzos y experiencia de otros autores en este campo.

Para obtener la dimensión fractal (DF), se aplicó el algoritmo de Katz, que procesa las señales directamente en el dominio del tiempo, y se tomó la información acústica correspondiente a una fase inspiratoria en cada canal de adquisición. La serie de tiempo se dividió en ventanas con un traslape del 99% y se calculó la dimensión fractal de cada ventana, generando una serie de tiempo. La figura 6.3 muestra una inspiración con crepitancias simuladas con un factor de escala de 1.5 y la DF asociada. Los resultados muestran que ciertos picos de la DF indican la ubicación de los estertores crepitantes insertados, como se observa en la figura 6.3. Sin embargo, la DF de la señal procesada también refleja cambios en la primera mitad de la fase inspiratoria a pesar de que no se

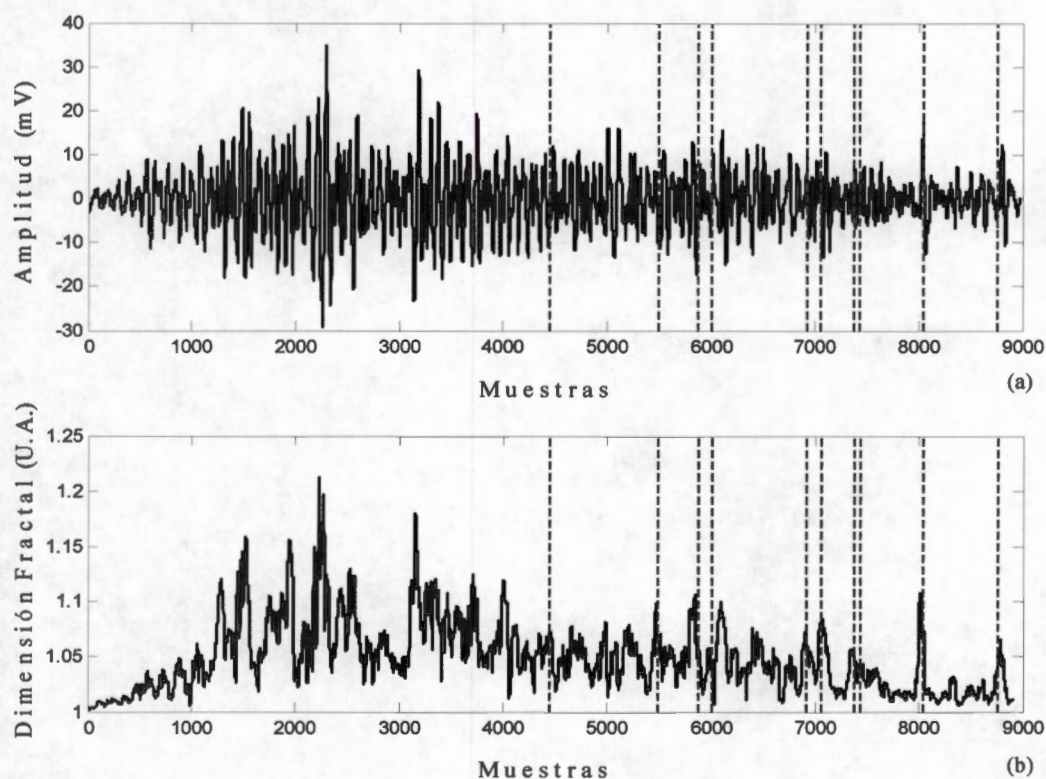


Figura 6.3. (a) Información acústica de la fase inspiratoria del canal PLX4, con crepitancias insertadas con un factor de escala de 1.5, (b) Serie de tiempo generada por el cálculo de dimensión fractal. Las líneas punteadas indican la ubicación temporal de las crepitancias.

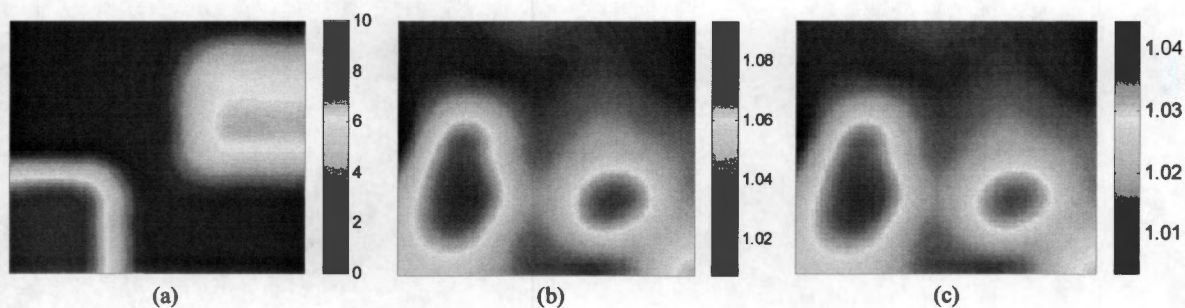


Figura 6.4. Imágenes de complejidad de casos simulados. (a) Imagen patrón que representa la distribución espacial de crepitancias con factor de escala 1.5. (b) Imagen obtenida promediando los valores de DF correspondientes a las ventanas temporales de una fase inspiratoria, (c) Imagen obtenida calculando la DF de la fase inspiratoria completa.

insertaron crepitancias en esas posiciones temporales. Los cambios de la DF en la primera mitad podrían confundirse con crepitancias si se intentara identificarlas por medio de un umbral.

En las figuras 6.4 y 6.5 se muestran dos ejemplos de *Mapas de Complejidad* de sonidos adventicios al *mapear* los valores de la DF de las 25 series de tiempo. Las figuras 6.4 y 6.5 muestran las imágenes correspondientes a dos patrones de distribución espacial de crepitancias finas, con un factor de escala de 1.5. En las figuras 6.4(a) y 6.5(a) se muestran los patrones generados con su respectiva barra de color que indica el número de estertores crepitantes insertados. La imagen corresponde a la parte posterior del tórax del sujeto. En el caso de la figura 6.4(a) el número máximo de estertores crepitantes (10) se localiza en la región pulmonar basal izquierda, mientras que en la figura 6.5(a) el número máximo de crepitancias (6) se localiza en el hemitórax derecho del sujeto simulado. Ninguno de los procedimientos utilizados genera una imagen cercana a la imagen patrón. Sin embargo, es importante recordar que la imagen patrón se relaciona directamente con el número de

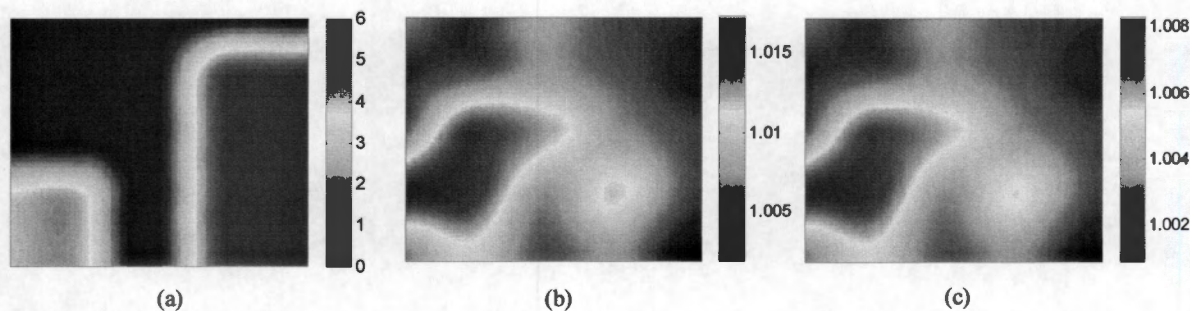


Figura 6.5. Imágenes de complejidad de casos simulados. (a) Imagen patrón que representa la distribución espacial de crepitancias con factor de escala 1.5. (b) Imagen obtenida promediando los valores de DF correspondientes a las ventanas temporales de una inspiración, (c) Imagen obtenida calculando la DF de la inspiración completa.

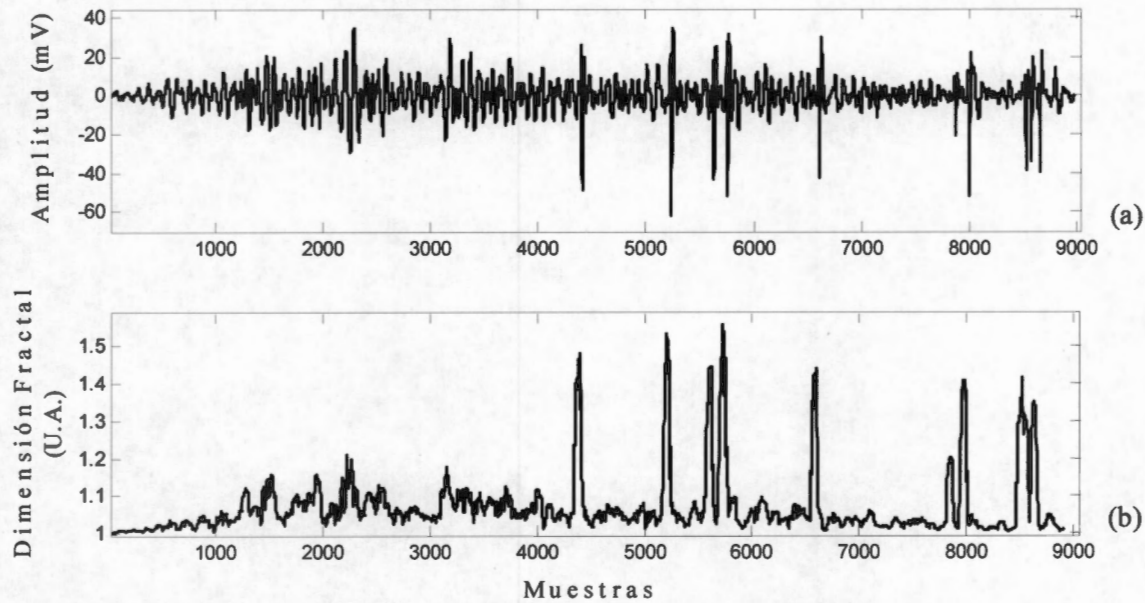


Figura 6.6. (a) Fase inspiratoria del canal PLX4 con crepitancias insertadas con un factor de escala de 10. (b) Dimensión Fractal correspondiente.

crepitantes insertados mientras que las imágenes de los incisos b) y c) se asocian con el grado de complejidad de la señal.

Con una mayor SNR entre las crepitancias y el sonido pulmonar, factor de escala de 10, la DF define mejor en sus cambios morfológicos la presencia de estertores crepitantes, en comparación con el intervalo de tiempo en donde no existen crepitancias y sólo se encuentra el sonido respiratorio de base. Si la SNR de la crepitancia se aumenta demasiado, la información que aparece en la primera mitad de la inspiración se reduce notablemente como se observa en la figura 6.6 y las crepitancias son fácilmente reconocibles. Sin embargo, con el factor de escala de 3.5, la información adicional en la primera mitad de la fase inspiratoria sigue siendo de una amplitud relevante con relación a los cambios producidos por los crepitantes.

En la figura 6.7 se observa el resultado de aplicar la DF a sonidos crepitantes con un factor de escala de 10, posteriormente la señal de DF se sometió a un proceso similar al seguido en la detección de crepitancias a través del modelo ARVT. La figura 6.7 (a) muestra las dos zonas con crepitancias con una sobreestimación de estas, sin embargo, es importante señalar que es posible generar una imagen que refleje la localización de las crepitancias, siempre que se tenga una SNR grande.

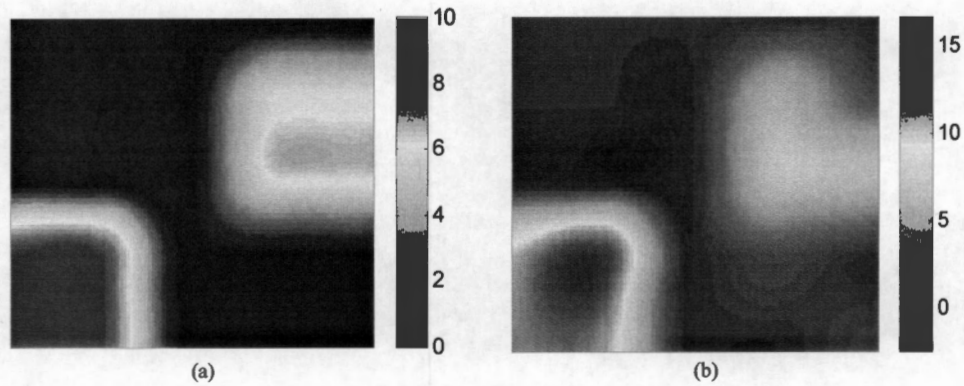


Figura 6.7. (a) Patrón # 1, (b) Detección de crepitancias con un factor de escala de 10, utilizando DF.

6.3 DETECCIÓN DE CREPITANCIAS MEDIANTE PARAMETRIZACIÓN CON UN MODELO AR INVARIANTE EN EL TIEMPO Y CLASIFICACIÓN CON REDES NEURONALES ARTIFICIALES

Las señales de sonido respiratorio multicanal con crepitancias simuladas se parametrizaron a través de un modelo AR. Se probaron dos ordenes del modelo AR, 4 y 6, y se obtuvieron mejores resultados con un orden 6, por lo tanto, se muestran los resultados utilizando un orden 6 y sólo en algunos casos se proporcionan los resultados con orden 4 para fines comparativos. El entrenamiento de las RNAs se realizó utilizando la interfaz gráfica de MATLAB® NNTOOL, con las características indicadas en la sección de metodología.

Los parámetros de entrenamiento de las RNAs utilizadas se reportan en la tabla 6.1.

Tabla 6.1

Parámetros de entrenamiento de la RNA

PARÁMETRO	VALOR
Épocas	100
Goal	0
Max_fail	5
Mem_reduc	1
Min_grad	1.00E-10
mu	0.001
mu_dec	0.1
mu_inc	10
mu_max	10000000000
Show	1
Time	inf

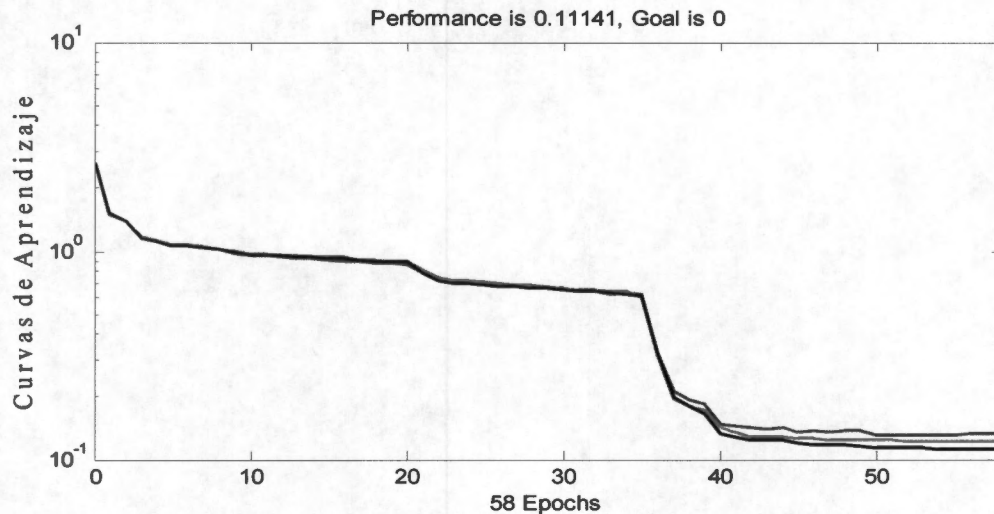


Figura 6.8. Curvas de aprendizaje de los conjuntos de entrenamiento (línea azul), prueba (línea roja) y validación (línea verde), de una RNA entrenada por regiones, con crepitancias finas y con un factor de escala de 3.5.

Las RNAs entrenadas por regiones utilizaron información de 28 sujetos y en el caso del entrenamiento por canales se utilizaron 20 sujetos, ya que el requerimiento de memoria en la computadora fue más alto en el entrenamiento por canales. El entrenamiento se realizó durante 100 épocas en promedio. En la figura 6.8 se muestran las curvas de aprendizaje para el conjunto de entrenamiento, prueba y validación de la RNA entrenada por regiones con crepitancias finas y con un factor de escala de 3.5. En éste ejemplo el error alcanzado fue de 0.11141 después de 58 épocas de entrenamiento. Obsérvese como todas las curvas van disminuyendo lo que indica una buena generalización de la red.

En la tabla 6.2 se muestra una relación de los errores alcanzados en los diferentes entrenamientos realizados.

En ésta tabla se observa que el error al final de la etapa de entrenamiento va disminuyendo conforme la SNR se mejora; es decir, conforme el factor de escalamiento aumenta. Así también, el error final es menor en el caso de las crepitancias finas que en el caso de las gruesas. Para el entrenamiento por regiones, el número de épocas necesarias también es menor en el caso de crepitancias finas que para el de crepitancias gruesas.

Tabla 6.2
Error alcanzado durante el entrenamiento de cada RNA

REGIONES					
Crepitancias Finas			Crepitancias Gruesas		
Factor de escala	Error alcanzado	No. épocas	Factor de escala	Error alcanzado	No. épocas
1.5	0.243662	92	1.5	0.40757	64
2.5	0.126865	54	2.5	0.296645	100
3.5	0.11141	58	3.5	0.195035	100

CANALES					
Crepitancias Finas			Crepitancias Gruesas		
Factor de escala	Error alcanzado	No. épocas	Factor de escala	Error alcanzado	No. épocas
1.5	0.367531	100	1.5	0.451375	100
2.5	0.220307	100	2.5	0.382642	100
3.5	0.130585	100	3.5	0.297235	100

6.3.1 RESULTADOS DEL ENTRENAMIENTO POR REGIONES

Se entrenaron RNAs con los tres factores de escala mencionados para crepitancias finas y gruesas, utilizando tres patrones de regiones enfermas. A continuación se muestran la imagen patrón y las imágenes estimadas a través de la RNA entrenada por regiones.

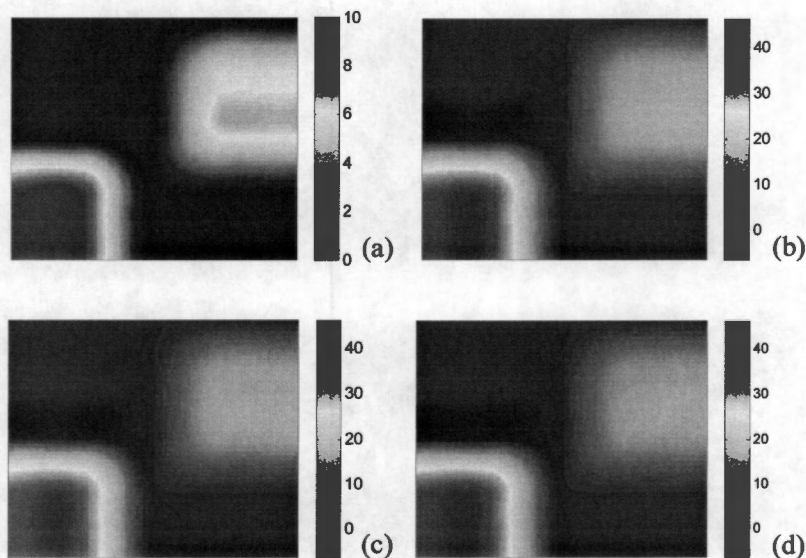


Figura 6.9. Patrón # 1. RNA entrenada por regiones, considerando crepitancias finas. (a) Imagen patrón, (b) Imagen estimada, factor de escala de 1.5, (c) Imagen estimada, factor de escala 2.5, (d) Imagen estimada, factor de escala 3.5.

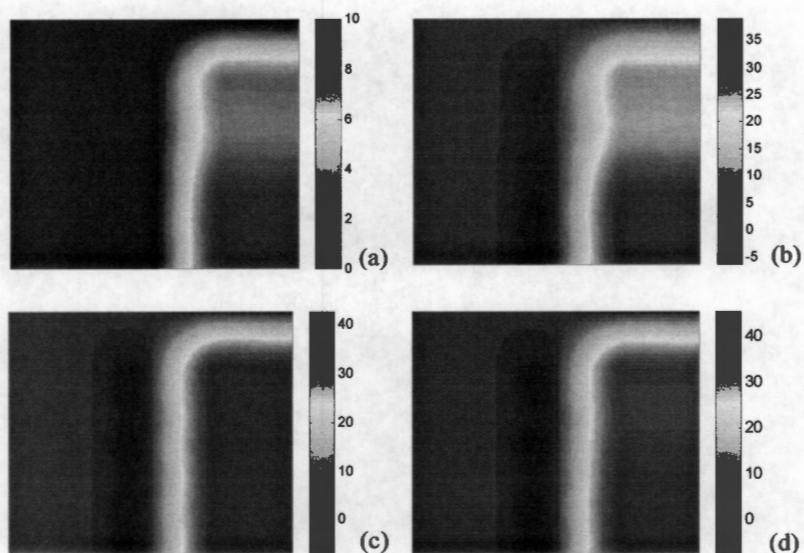


Figura 6.10. Patrón # 2. RNA entrenada por regiones considerando crepitancias finas. (a) Imagen patrón, (b) Imagen estimada, factor de escala 1.5, (c) Imagen estimada, factor de escala 2.5, (d) Imagen estimada, factor de escala 3.5.

Los resultados mostrados en las figuras 6.9 a 6.11 indican que la estimación de las zonas con crepitancias se aproximaron bastante al patrón establecido. En el caso de la figura 6.10 se obtiene una mejor aproximación a las zonas enfermas con un factor de escala 1.5 y 3.5, reflejando las zonas con mayor presencia de crepitancias. Es importante mencionar que las imágenes patrón fueron generadas con base al número de crepitancias insertadas y las imágenes estimadas por medio de las RNAs muestran un porcentaje del número de ventanas temporales clasificadas como enfermas o sanas. Es decir, para la imagen patrón la barra de color indica el número de crepitantes insertados y para las imágenes estimadas el porcentaje de ventanas temporales enfermas detectadas. Por lo tanto, se espera que en las imágenes estimadas se puedan distinguir zonas con diferente cantidad de crepitancias, con la salvedad que la escala numérica de la barra de colores es diferente a la escala numérica de la imagen patrón. Las imágenes estimadas con crepitancias gruesas muestran de igual forma una buena aproximación a las zonas enfermas. Estos resultados se muestran en las figuras 6.12 a 6.14 con las mismas distribuciones espaciales utilizadas para crepitancias finas.

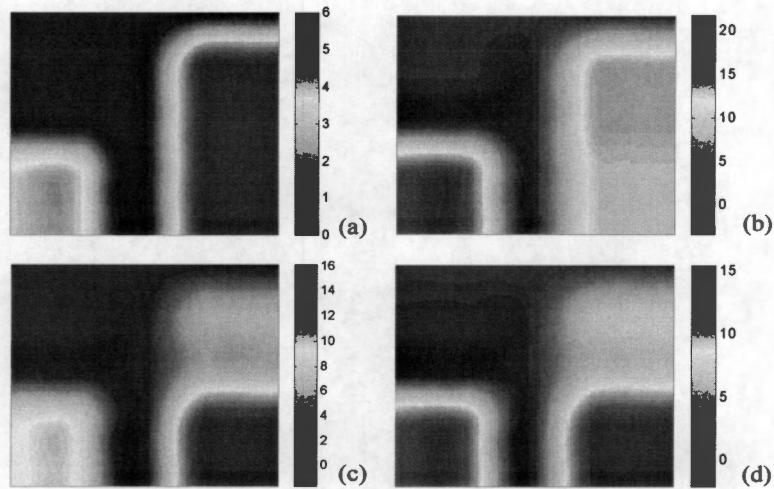


Figura 6.11. Patrón # 3. RNA entrenada por regiones considerando crepitanacias finas. (a) Imagen patrón, (b) Imagen estimada, factor de escala 1.5, (c) Imagen estimada, factor de escala 2.5, (d) Imagen estimada, factor de escala 3.5.

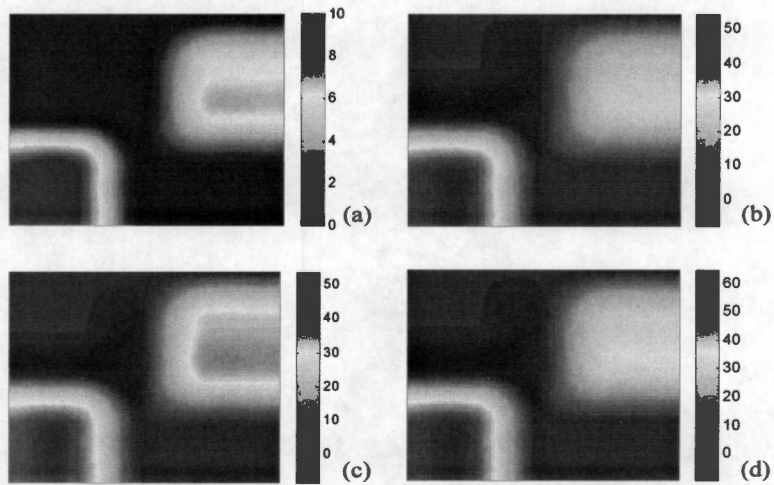


Figura 6.12. Patrón # 1. RNA entrenada por regiones, considerando crepitanacias gruesas. (a) Imagen patrón, (b) Imagen estimada, factor de escala de 1.5, (c) Imagen estimada, factor de escala de 2.5, (d) Imagen estimada, factor de escala de 3.5.

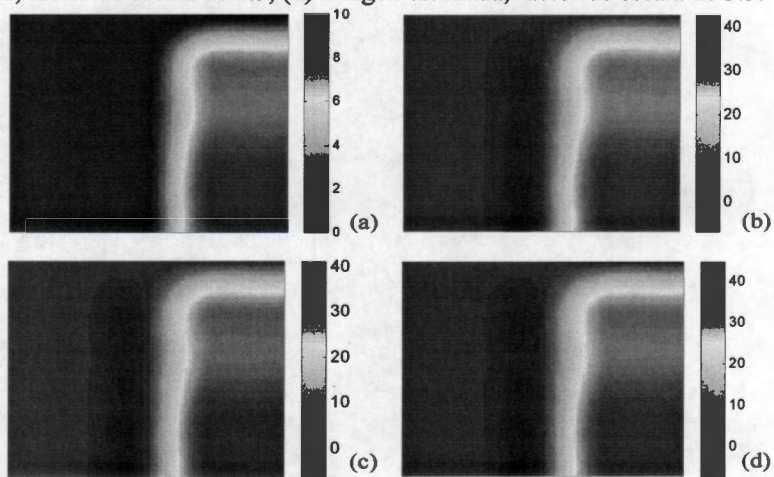


Figura 6.13. Patrón # 2, RNA entrenada por regiones, considerando crepitanacias gruesas. (a) Imagen patrón, (b) Imagen estimada, factor de escala de 1.5, (c) Imagen estimada, factor de escala de 2.5, (d) Imagen estimada, factor de escala de 3.5.

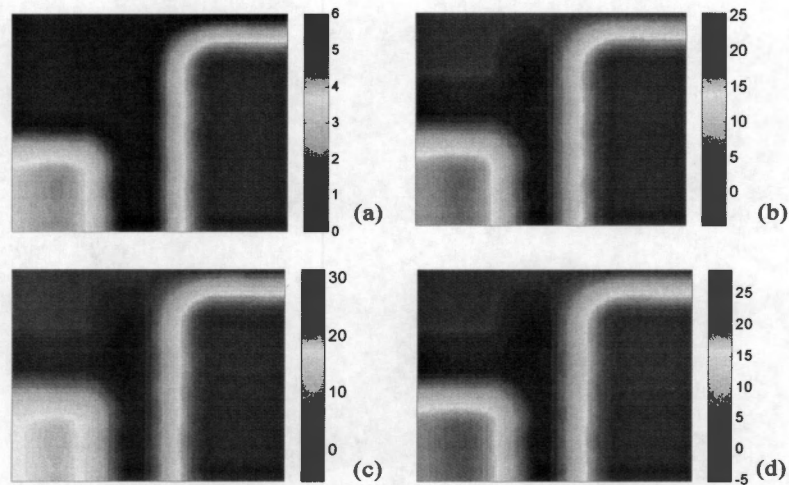


Figura 6.14. Patrón # 3. RNA entrenada por regiones considerando crepitancias gruesas. (a) Imagen patrón, (b) Imagen estimada, factor de escala de 1.5, (c) Imagen estimada, factor de escala de 2.5, (d) Imagen estimada, factor de escala de 3.5.

6.3.2 RESULTADOS DEL ENTRENAMIENTO POR CANALES

El entrenamiento de las RNAs también se realizó por canales. Las imágenes estimadas a través del entrenamiento por canales se muestran en las figuras 6.15 a 6.17 para crepitancias finas y de la 6.18 a 6.20 para las crepitancias gruesas. Para este tipo de entrenamiento, se ve en mayor medida el efecto de la SNR, a mayor SNR la imagen estimada es más próxima a la imagen patrón.

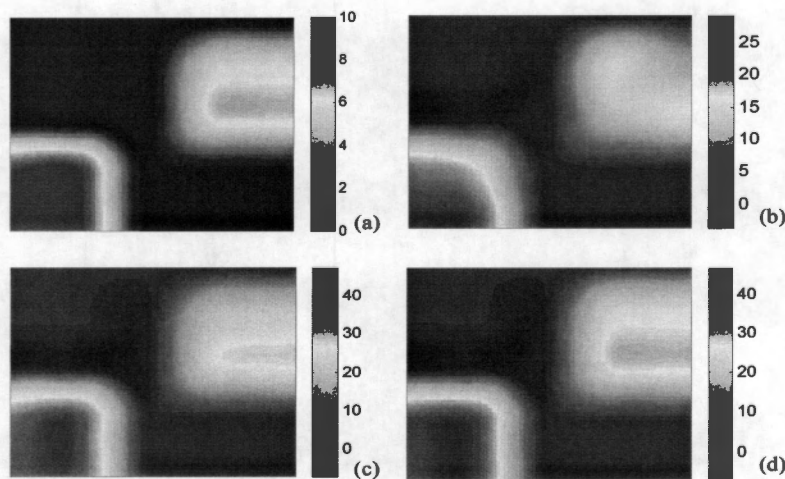


Figura 6.15. Patrón # 1. RNA entrenada por canales, considerando crepitancias finas. (a) Imagen patrón, (b) Imagen estimada, factor de escala de 1.5, (c) Imagen estimada, factor de escala de 2.5, (d) Imagen estimada, factor de escala de 3.5.

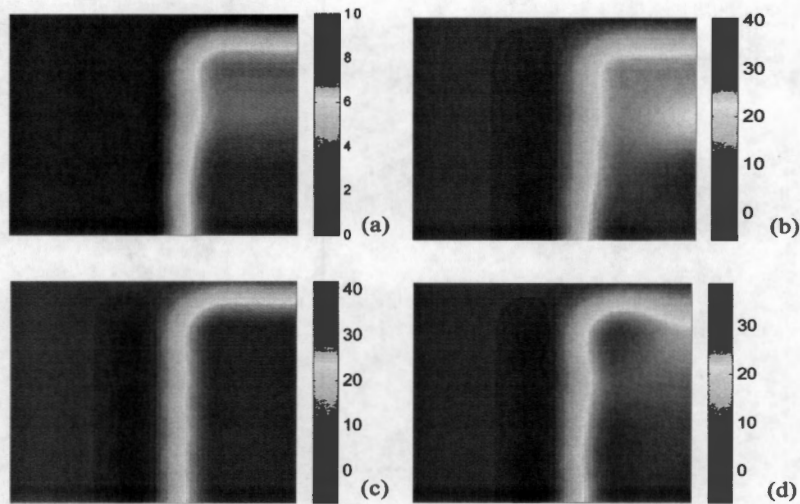


Figura 6.16. Patrón # 2. RNA entrenada por canales considerando crepitancias finas. (a) Imagen patrón, (b) Imagen estimada, factor de escala de 1.5, (c) Imagen estimada, factor de escala de 2.5, (d) Imagen estimada, factor de escala de 3.5.

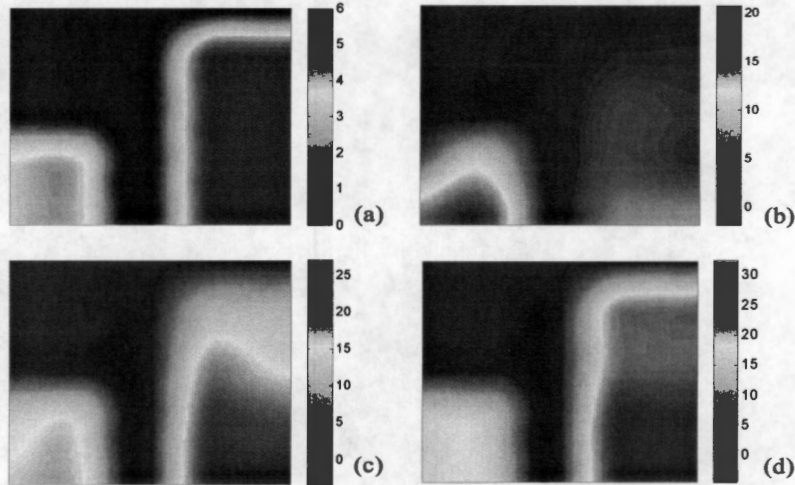


Figura 6.17. Patrón # 3. RNA entrenada por canales considerando crepitancias finas. (a) Imagen patrón, (b) Imagen estimada, factor de escala de 1.5, (c) Imagen estimada, factor de escala de 2.5, (d) Imagen estimada, factor de escala de 3.5.

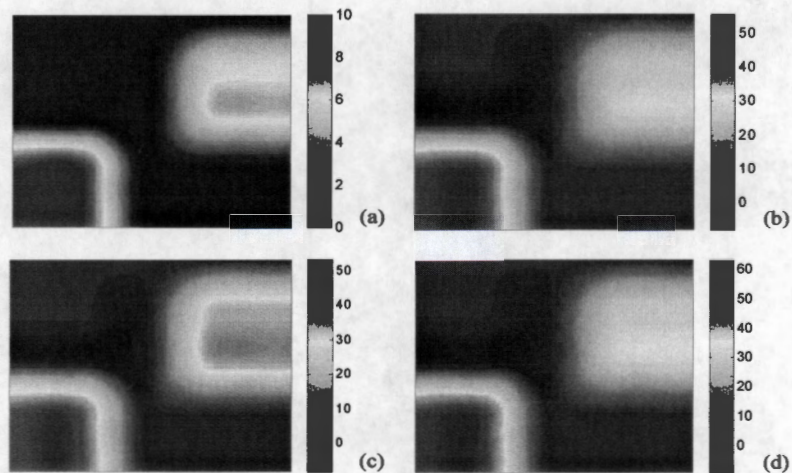


Figura 6.18. Patrón # 1. RNA entrenada por canales, considerando crepitancias gruesas. (a) Imagen patrón, (b) Imagen estimada, factor de escala de 1.5, (c) Imagen estimada, factor de escala de 2.5, (d) Imagen estimada, factor de escala de 3.5.

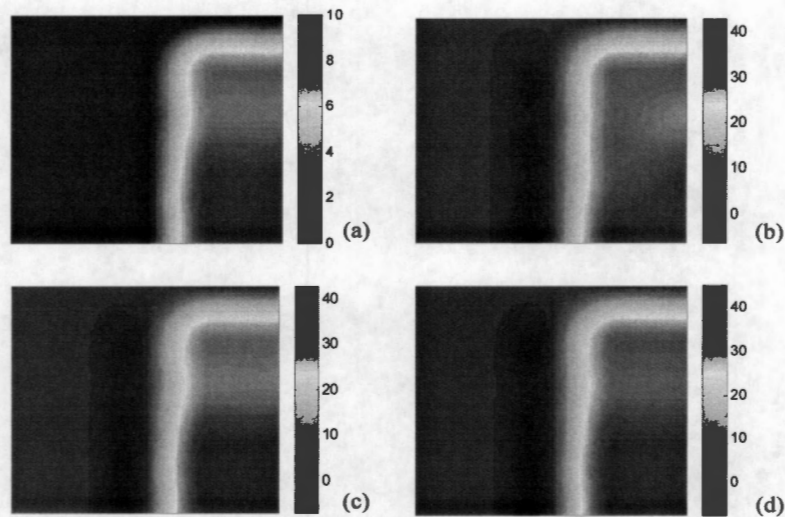


Figura 6.19. Patrón # 2. RNA entrenada por canales, considerando crepitancias gruesas. (a) Imagen patrón, (b) Imagen estimada, factor de escala de 1.5, (c) Imagen estimada, factor de escala de 2.5, (d) Imagen estimada, factor de escala de 3.5.

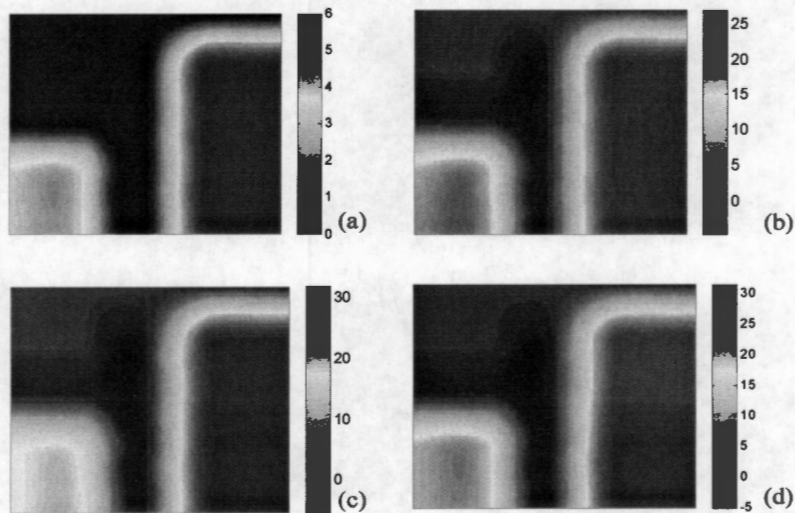


Figura 6.20. Patrón # 3. RNA entrenada por canales, considerando crepitancias gruesas. (a) Imagen patrón, (b) Imagen estimada, factor de escala de 1.5, (c) Imagen estimada, factor de escala de 2.5, (d) Imagen estimada, factor de escala de 3.5.

En la figura 6.21 se muestran los resultados correspondientes a parametrizaciones con modelos AR de orden 4 y 6 para la imagen patrón # 1. Se puede observar que las imágenes estimadas correspondientes a la parametrización con orden 4 se alejan más de la imagen patrón. Sin embargo, las imágenes con factor de escala 1.5 son similares en ambos casos.

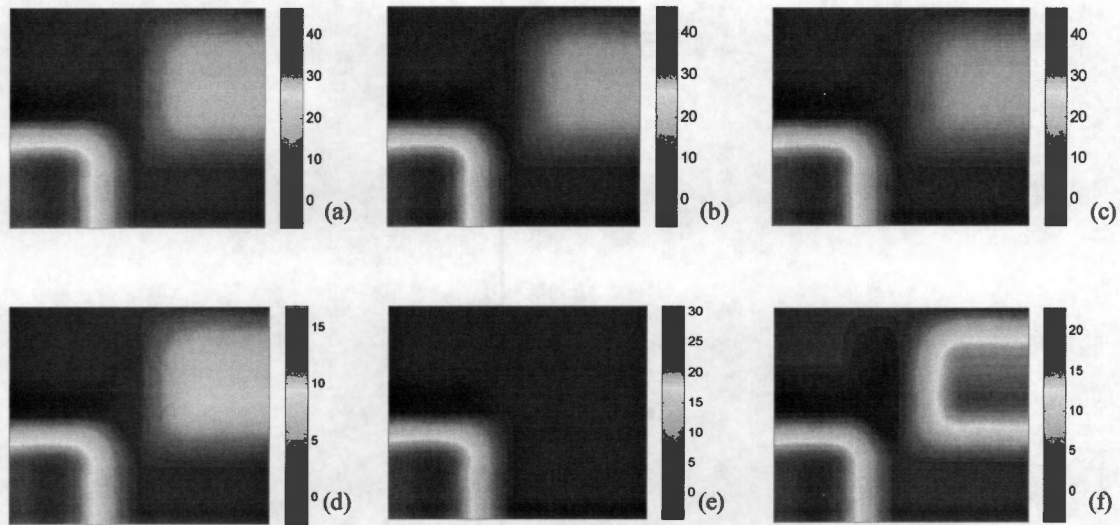


Figura 6.21. Imágenes generadas con diferente orden del modelo AR. (a)-(c) Imágenes estimadas con orden de parametrización 6 y factor de escala 1.5, 2.5 y 3.5, de izquierda a derecha. (d)-(f) Imágenes con un orden 4.

Se obtuvieron las gráficas de los valores de las neuronas de salida de las RNAs para los conjuntos de entrenamiento, prueba y validación con el objetivo de visualizar como fueron clasificados los vectores de características. En las figuras 6.22 y 6.23 se muestran los valores del conjunto de prueba de la RNA entrenada por regiones con un factor de escala de 1.5 y 3.5, respectivamente. Se observa que para el caso del factor de 1.5 un buen número de vectores de características no generan a la salida de la RNA un valor mayor al umbral establecido de ± 0.8 y por lo tanto, no logran clasificarse adecuadamente. Los cuadros rojos con la etiqueta *Sano-sc*, en la gráfica de la figura 6.22, indican sujetos sanos que no

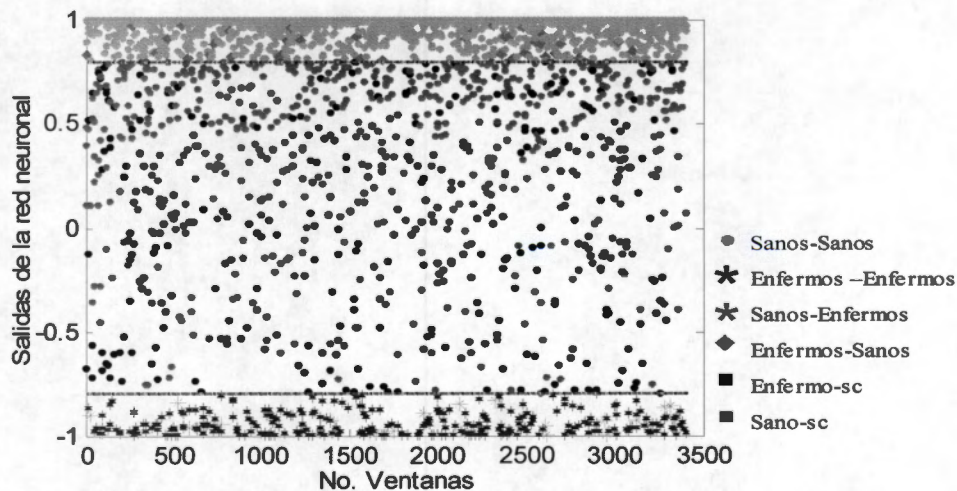


Figura 6.22. Valores de la neurona de salida para el conjunto de prueba de la RNA entrenada por regiones, con crepitancias finas y con factor de escala de 1.5.

alcanzaron el umbral de 0.8 y en consecuencia no se asociaron a ninguna clase. Es evidente, que si se establece el umbral en un valor menor, un mayor número de ventanas temporales serían clasificadas correctamente. En contraste, para el factor de 3.5 el número de vectores de características no clasificados disminuye sensiblemente entre el valor de umbral de 0.8 a -0.8.

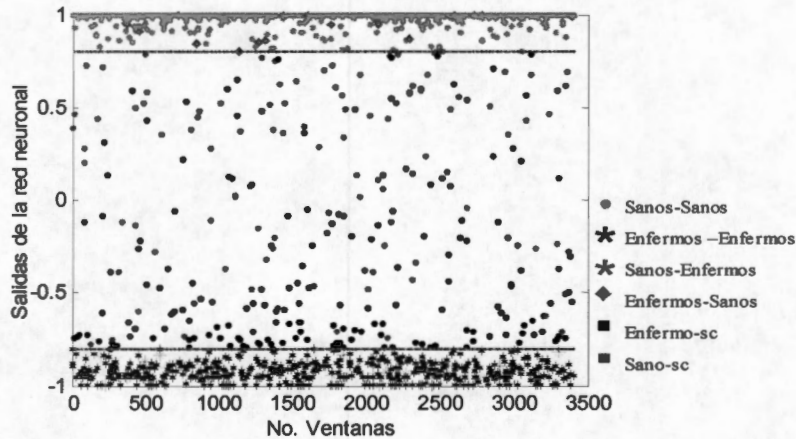


Figura 6.23. Valores de la neurona de salida para el conjunto de prueba de la RNA entrenada por regiones, con crepitancias finas y con factor de escala de 3.5.

Otra forma de visualizar la clasificación de los vectores de características fue graficando éstos en 3D; es decir, los ejes x , y , z de la representación lo definen los primeros tres valores del vector de características formado por los coeficientes AR. De esta forma se visualizó la forma en que se agrupan los vectores de características. En las gráficas de la figura 6.24 se aprecian 3 conjuntos, los círculos amarillos corresponden a los vectores de características clasificados como sanos, los círculos rojos son los vectores de características clasificados como enfermos y el conjunto de círculos azules corresponden a los vectores de características restantes. La distribución de los vectores de características presentada en la figura 6.24(a) corresponde al conjunto de prueba, durante el entrenamiento de una RNA por regiones, con crepitancias finas de factor de escala de 1.5, mientras que la figura 6.24(b) corresponde al factor de escala 3.5. En la figura 6.25 se muestran las salidas de la RNA entrenada por regiones para un sujeto con crepitancias insertadas con base en el patrón # 1, con factor de escala 1.5, para crepitancias finas, figura 6.25(a), y crepitancias gruesas, figura 6.25 (b). La figura 6.26 presenta un ejemplo similar al de la figura 6.25, pero basado en una RNA entrenada por canales.

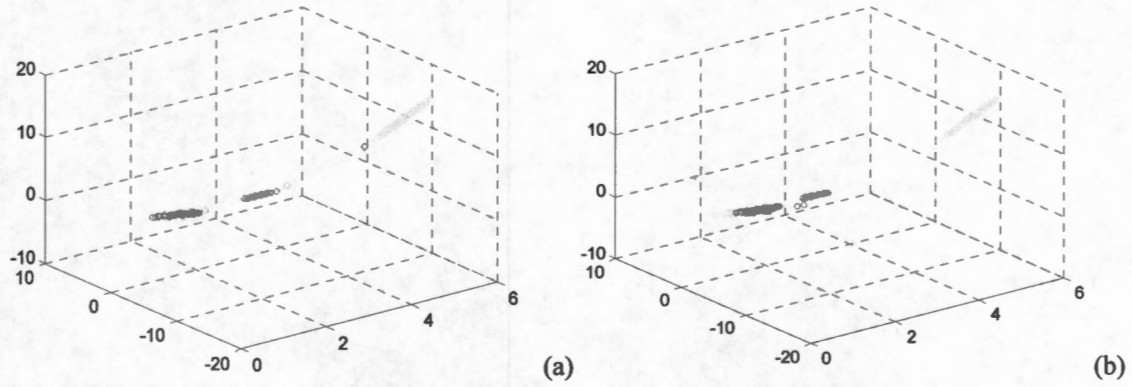


Figura 6.24. Salidas del conjunto de prueba de la RNA entrenada por regiones, con crepitancias finas, (a) con factor de escala de 1.5 y (b) con factor de escala 3.5.

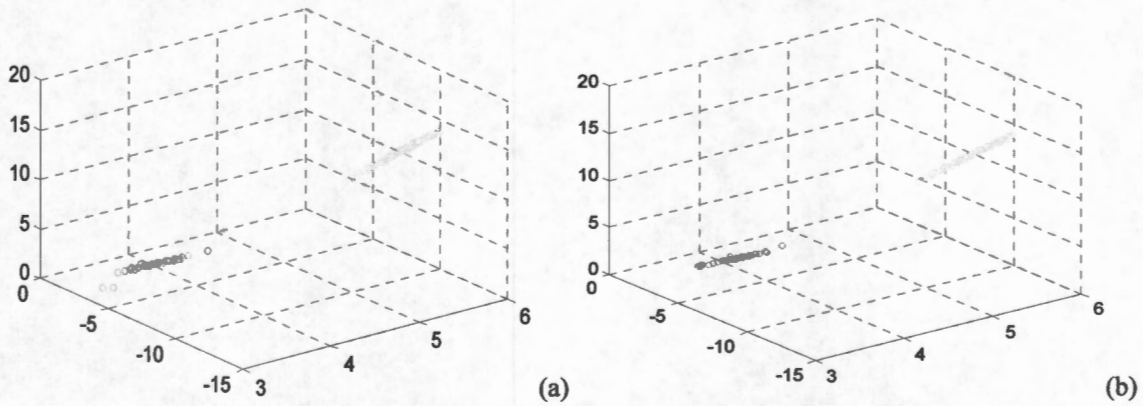


Figura 6.25. Salidas del sujeto enfermo con el patrón # 1, clasificado con la RNA entrenada por regiones, con factor de escala de 1.5, (a) con crepitancias finas y (b) con crepitancias gruesas.

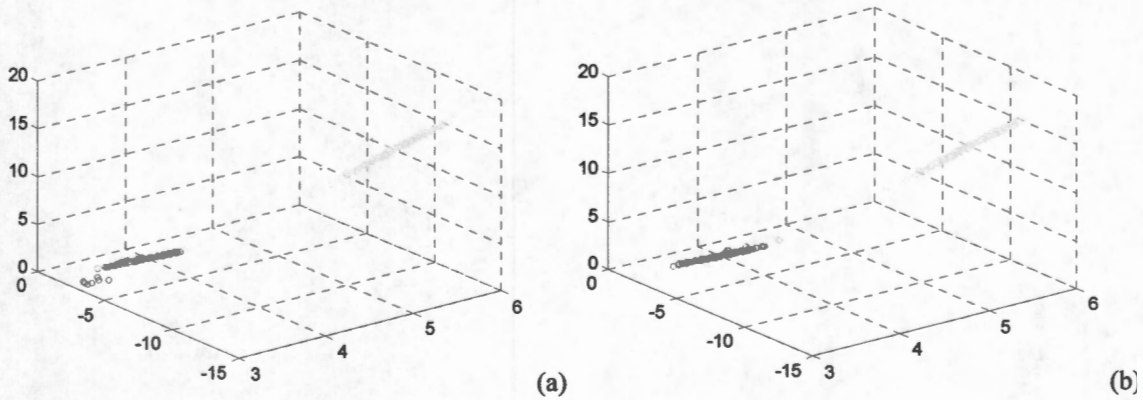


Figura 6.26. Salidas del sujeto enfermo con el patrón # 1, clasificado con la RNA entrenada por canales, con factor de escala de 1.5, (a) con crepitancias finas y (b) con crepitancias gruesas.

RESULTADOS

En las figuras 6.24, 6.25 y 6.26 se observa que los conjuntos de vectores de características sanos y enfermos se encuentran ampliamente separados, en algunos casos los vectores se mezclan, sin embargo, es posible separar en dos conjuntos a los vectores de características.

En las tablas 6.3, 6.4 y 6.5 se muestran los porcentajes de sensibilidad (S), especificidad (E) y desempeño (D), de las RNAs entrenadas por canales, por regiones así como el promedio del desempeño de las RNAs. En general el desempeño de las RNAs entrenadas es bueno. Las tablas de desempeño indican que el entrenamiento por canales es ligeramente mejor que el entrenamiento por regiones. Además, con base a las imágenes obtenidas se puede decir que el entrenamiento por regiones generó aproximaciones más cercanas a la distribución espacial de las crepitancias insertadas.

TABLA 6.3

DESEMPEÑO DE LA RED NEURONAL ENTRENADA POR REGIONES

REGIONES									
FINOS									
FACTOR DE ESCALA	ENTRENAMIENTO			PRUEBA			VALIDACIÓN		
	%S	%E	%D	%S	%E	%D	%S	%E	%D
1.5	98.29	95.05	96.67	98.35	93.92	96.13	97.91	91.89	94.90
2.5	99.06	95.35	97.21	99.24	94.48	96.86	98.72	95.32	97.02
3.5	99.35	92.84	96.10	99.19	90.46	94.83	99.12	92.93	96.03

GRUESOS									
FACTOR DE ESCALA	ENTRENAMIENTO			PRUEBA			VALIDACIÓN		
	%S	%E	%D	%S	%E	%D	%S	%E	%D
1.5	99.56	95.04	97.30	99.04	94.05	96.55	99.40	93.89	96.65
2.5	98.08	94.63	96.36	98.26	95.28	96.77	97.49	94.31	95.90
3.5	98.75	95.68	97.21	98.53	94.72	96.63	98.43	93.81	96.12

TABLA 6.4
DESEMPEÑO DE LA RED NEURONAL ENTRENADA POR CANALES

CANALES									
FINOS									
FACTOR DE ESCALA	ENTRENAMIENTO			PRUEBA			VALIDACIÓN		
	%S	%E	%D	%S	%E	%D	%S	%E	%D
1.5	98.87	96.52	97.69	99.02	96.17	97.59	98.92	96.71	97.82
2.5	98.70	96.15	97.43	98.89	97.01	97.95	98.56	96.78	97.67
3.5	99.19	96.29	97.74	99.21	95.49	97.35	99.11	95.29	97.20

GRUESOS									
FACTOR DE ESCALA	ENTRENAMIENTO			PRUEBA			VALIDACIÓN		
	%S	%E	%D	%S	%E	%D	%S	%E	%D
1.5	99.54	94.93	97.23	99.55	95.44	97.50	99.50	94.48	96.99
2.5	99.29	95.87	97.58	99.35	95.75	97.55	99.30	94.17	96.74
3.5	98.51	95.28	96.90	98.38	92.82	95.60	98.50	94.99	96.75

TABLA 6.5

PROMEDIO DEL DESEMPEÑO DE LAS REDES NEURONALES

REGIONES				
FINOS				
FACTOR DE ESCALA	ENTRENAMIENTO	PRUEBA	VALIDACIÓN	% PROMEDIO
	%DESEMPEÑO	%DESEMPEÑO	%DESEMPEÑO	
1.5	96.67	96.13	94.90	95.90
2.5	97.21	96.86	97.02	97.03
3.5	96.10	94.83	96.03	95.65

GRUESOS				
FACTOR DE ESCALA	ENTRENAMIENTO	PRUEBA	VALIDACIÓN	% PROMEDIO
	%DESEMPEÑO	%DESEMPEÑO	%DESEMPEÑO	
1.5	97.30	96.55	96.65	96.83
2.5	96.36	96.77	95.90	96.34
3.5	97.21	96.63	96.12	96.65

CANALES				
FINOS				
FACTOR DE ESCALA	ENTRENAMIENTO	PRUEBA	VALIDACIÓN	% PROMEDIO
	%DESEMPEÑO	%DESEMPEÑO	%DESEMPEÑO	
1.5	97.69	97.59	97.82	97.70
2.5	97.43	97.95	97.67	97.68
3.5	97.74	97.35	97.20	97.43

GRUESOS				
FACTOR DE ESCALA	ENTRENAMIENTO	PRUEBA	VALIDACIÓN	% PROMEDIO
	%DESEMPEÑO	%DESEMPEÑO	%DESEMPEÑO	
1.5	97.23	97.50	96.99	97.24
2.5	97.58	97.55	96.74	97.29
3.5	96.90	95.60	96.75	96.41

6.4 DETECCIÓN DE CREPITANCIAS CON UN MODELO AR VARIANTE EN EL TIEMPO

Para identificar crepitancias simuladas en el sonido respiratorio, se utilizó un modelo AR variante en el tiempo (ARVT), los coeficientes de este modelo se obtuvieron a través de un esquema de filtrado de predicción adaptable.

Para evaluar el desempeño del esquema de procesamiento se insertaron crepitancias simuladas en ciertos canales del registro de sonidos respiratorios, simulando zonas enfermas. Por otra parte, se generaron imágenes acústicas patrón que posteriormente fueron comparadas con las imágenes obtenidas a través del modelo AR variante en el tiempo. En la figura 6.27 se muestran las imágenes patrón y el número de crepitancias que fueron insertadas en cada canal.

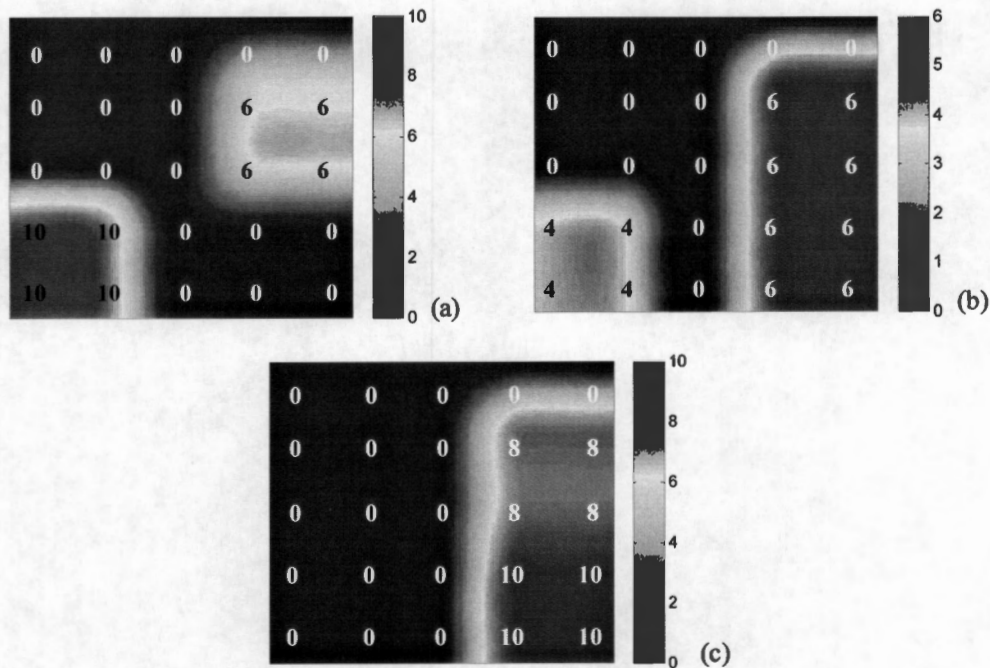


Figura 6.27. Imágenes patrón, (a) imagen patrón # 1 donde la región 5 tiene 6 crepitancias y la región 6 tiene 10 crepitancias. (b) Imagen patrón # 2 donde las regiones 5 y 7 tienen 6 crepitancias y la región 6 tiene 4 crepitancias. (c) Imagen patrón # 3 donde la región 5 tiene 8 crepitancias y la región 7 tiene 10 crepitancias.

El modelo AR variante en el tiempo se aplicó a cada una de las señales multicanales simuladas; el factor de olvido fue seleccionado con base a prueba y error, el valor utilizado fue $\lambda = 0.97$. El número estimado de crepitancias se asoció con cambios abruptos en la

derivada de los coeficientes ARVT. Se estableció un umbral en 0.035 veces la desviación estándar de los coeficientes AR en señales simuladas y reales. El procedimiento de umbralización se aplicó a las derivadas de las series de tiempo generadas por ARVT y sólo si el cambio abrupto era superior al umbral en las cuatro series de tiempo, la información fue considerada producida por una crepitancia. La habilidad en estimar el número de crepitancias se evaluó por la concordancia entre el número total de crepitancias insertadas y el número de crepitancias estimadas por el modelo ARVT.

Se obtuvieron resultados con los 3 patrones de crepitancias y con los tres factores de escala, para crepitancias finas y gruesas. Es importante mencionar que debido a que la inserción de las crepitancias fue aleatoria, el traslape temporal de crepitancias fue evidente en algunos casos simulados. Por ejemplo, con la imagen patrón # 1, considerando un factor de escala de 1.5 e incluyendo crepitancias finas, los resultados reflejan una imagen estimada semejante a la imagen patrón. Sin embargo, en la región 6, donde se insertaron 10 crepitancias, el algoritmo sólo detectó 9, ya que dos de éstas crepitancias se encuentran traslapadas temporalmente. La imagen estimada con esta imagen patrón se muestra en la figura 6.28.

En la figura 6.29 se observa el comportamiento de los coeficientes estimados vía el modelo ARVT en el canal PLC4. En la figura 6.29(a) se pueden contar claramente 9 cambios abruptos correspondientes a las 9 crepitancias insertadas y en la figura 6.29 (b) se muestra una expansión en el tiempo del tercer coeficiente AR, $a(3, n)$, con la finalidad de visualizar fácilmente el traslape temporal de crepitancias. Una de éstas crepitancias se

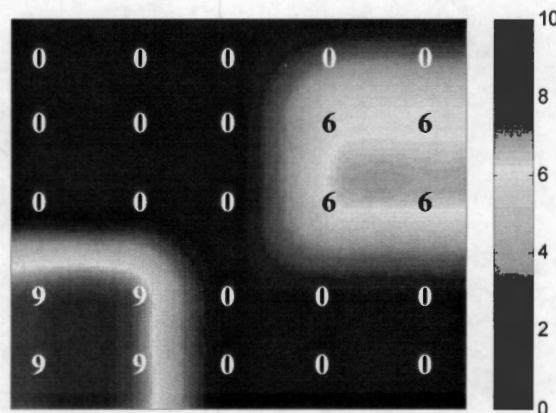


Figura 6.28. Imagen estimada con distribución espacial de crepitancias del patrón # 1 y con un factor de escala de 1.5.

insertó en la muestra 5414 y la otra en la muestra 5416. El algoritmo detectó sólo un cambio no estacionario, lo cual se traduce en la cuenta de una sola crepitancia. En la figura 6.29(c) se muestra las derivadas de las señales de los coeficientes AR de la figura 6.29(a), se observa que es posible detectar de mejor forma los cambios de los coeficientes.

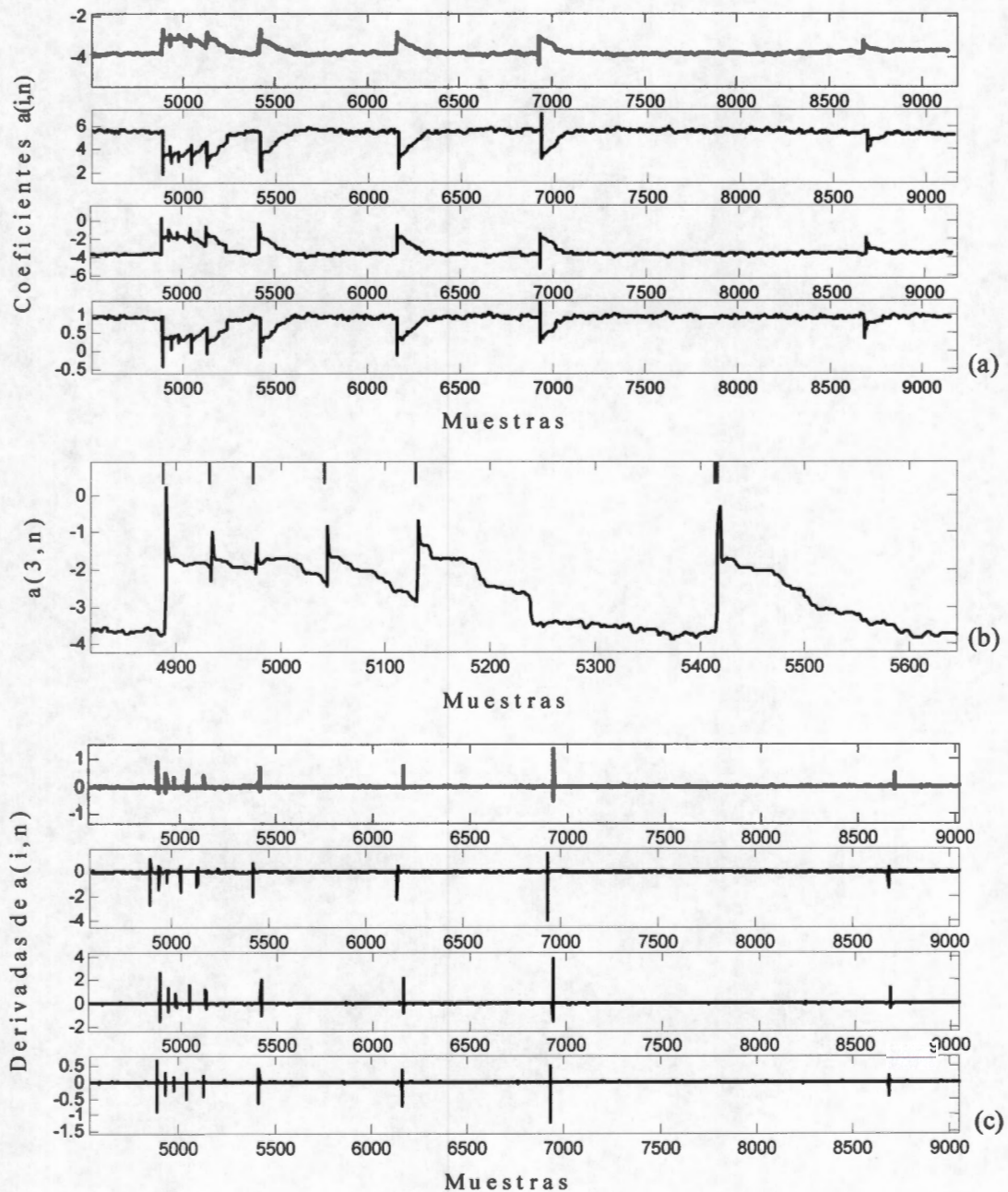


Figura 6.29. Resultados del patrón # 1 con un factor de escala de 1.5. (a) Coeficientes estimados a través del modelo ARVT. (b) Expansión en el tiempo del tercer coeficiente. (c) Derivadas de los coeficientes estimados.

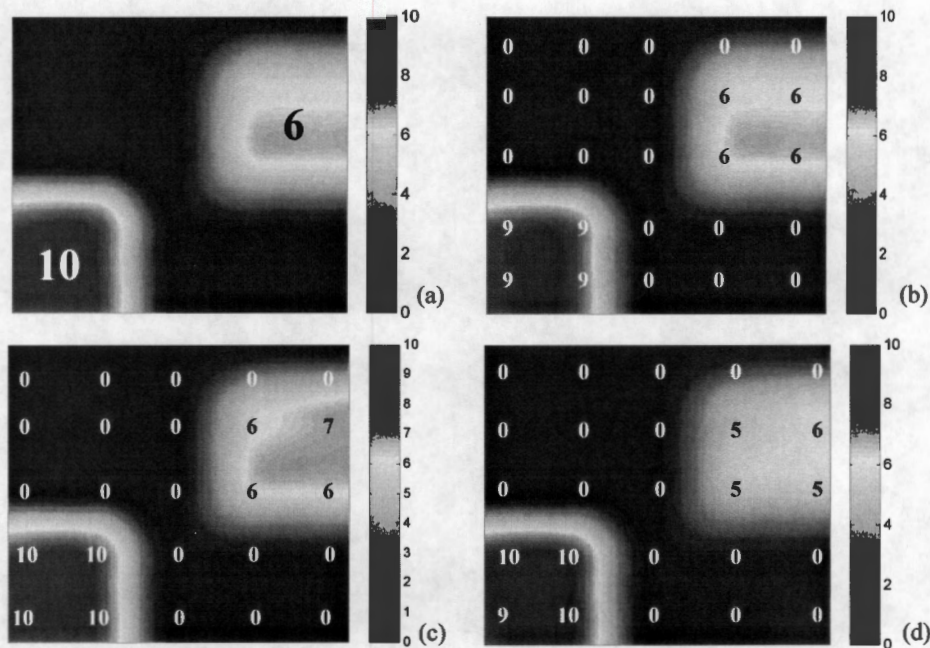


Figura 6.30. Imágenes obtenidas a partir del patrón # 1 con los tres diferentes factores de escala y con crepitancias finas. (a) Imagen patrón, (b) Imagen estimada con factor de escala de 1.5, (c) Imagen estimada con factor de escala de 2.5, (d) Imagen estimada con factor de escala de 3.5.

En la figura 6.30 se muestran las imágenes estimadas a través del modelo ARVT con crepitancias finas y con los tres diferentes factores de escala. En la figura 6.30 (b) la imagen estimada con un factor de escala de 1.5, es muy aproximada a la imagen patrón en la figura 6.30(a). Sin embargo, en la región 6 se estimaron 9 de 10 crepitancias debido probablemente a un traslape temporal de las mismas, los coeficientes estimados se muestran en la figura 6.29. Para el factor de escala 2.5 (figura 6.30(c)) el algoritmo sobreestimó una crepitancia en la región 5 y en la figura 6.30 (d) en la región 5 se estimaron sólo 5 crepitancias debido a un traslape temporal de las mismas. Este traslape temporal de las crepitancias insertadas se muestra en el comportamiento de las señales correspondientes a las series ARVT, figura 6.31. En las figuras 6.32 y 6.33 se muestran las imágenes estimadas para las dos imágenes patrón generadas.

Algunas estimaciones no corresponden exactamente al número insertado de crepitancias, esto nuevamente puede ser debido al traslape temporal de las mismas. Por ejemplo, en la figura 6.32 (b), en la región 7 se insertaron 10 crepitancias en los 4 canales y el algoritmo estimó 9; en la figura 6.34 se muestran los coeficientes ARVT

correspondientes al canal PRC4 de la figura 6.32 (b), donde se aprecia que la 4ª - 5ª y 8ª - 9ª crepitancias, se encuentran traslapadas. En la figura 6.35 se aprecia una expansión en tiempo del primer coeficiente ARVT de la figura 6.34, con la finalidad de observar las crepitancias traslapadas.

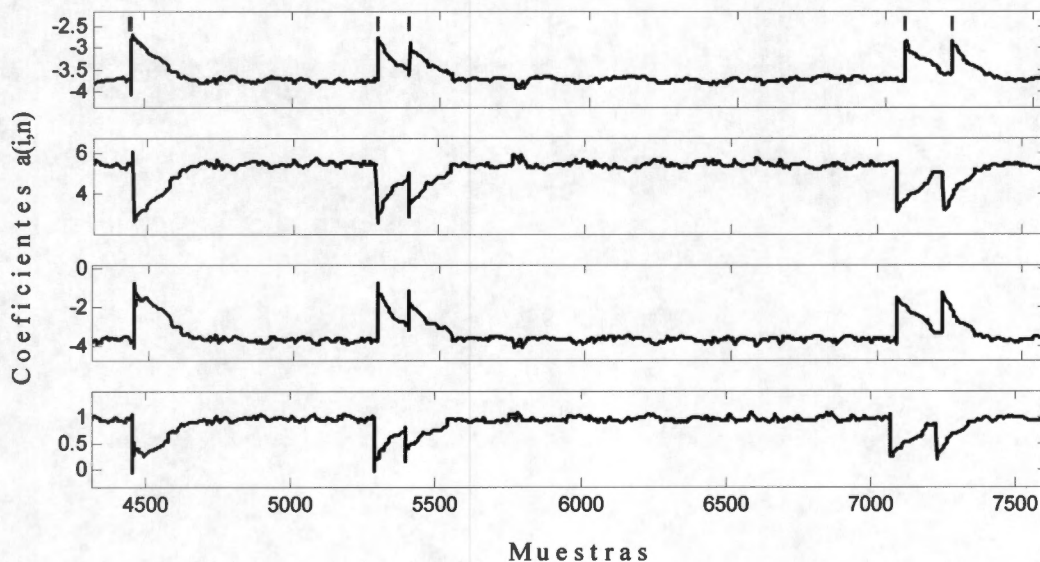


Figura 6.31. Coeficientes AR correspondientes al canal PRC2 del sujeto de prueba con el patrón no. 1 de inserción de crepitancias y con un factor de escala de 3.5.

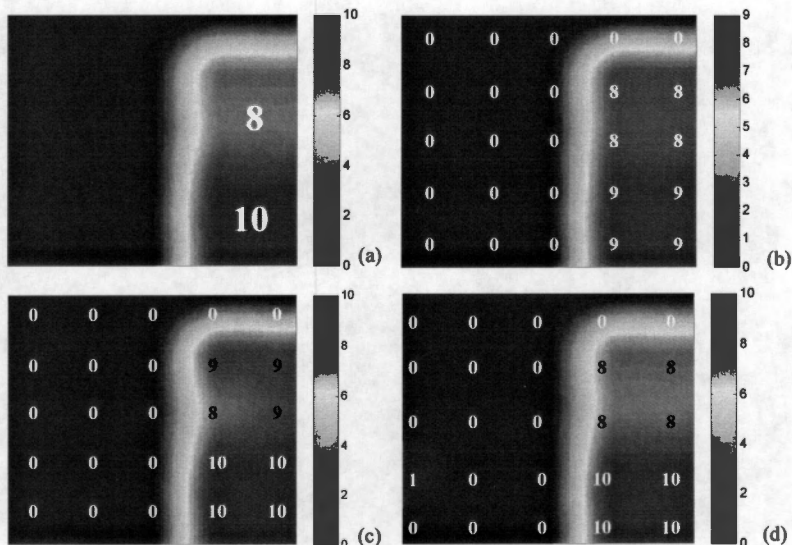


Figura 6.32. Imágenes obtenidas a partir del patrón # 2 con los tres diferentes factores de escala y con crepitancias finas (a) Imagen patrón, (b) Imagen estimada con factor de escala de 1.5, (c) Imagen estimada con factor de escala de 2.5, (d) Imagen estimada con factor de escala de 3.5.

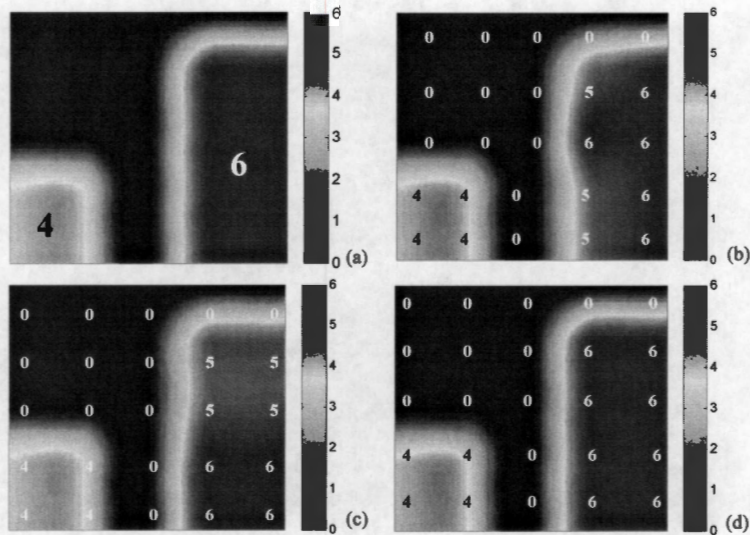


Figura 6.33. Imágenes obtenidas a partir del patrón # 3 con los tres diferentes factores de escala y con crepitancias finas. (a) Imagen patrón, (b) Imagen estimada con factor de escala de 1.5, (c) Imagen estimada con factor de escala de 2.5, (d) Imagen estimada con factor de escala de 3.5.

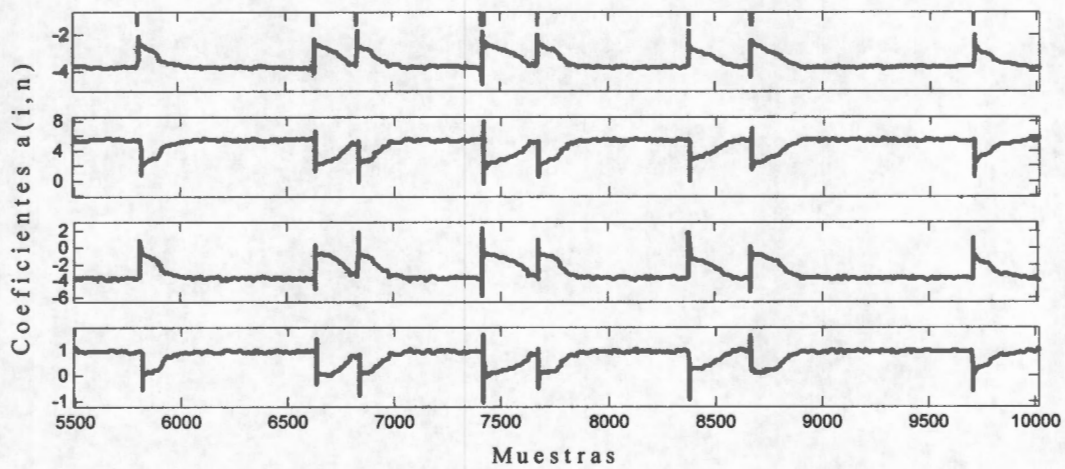


Figura 6.34. Coeficientes ARVT del patrón # 2 del canal PRC4, con un factor de escala de 1.5. La 4ª, 5ª y 8ª, 9ª crepitancias se encuentran traslapadas.

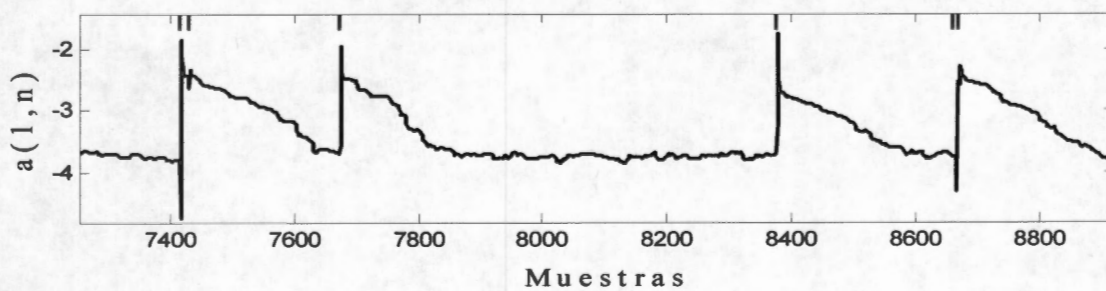


Figura 6.35. Zoom del primer coeficiente ARVT de la figura 6.34, donde se muestra el traslape de crepitancias.

Posteriormente se realizaron las mismas pruebas con los tres factores de escala, pero ahora con registros que contenían crepitancias gruesas; los resultados se muestran en las figuras 6.36-6.38. En la Tabla 6.6 se muestra el conteo de crepitancias por cada imagen patrón y factor de escala. Además, se incluye el número de crepitancias que no estimó el algoritmo y el número de crepitancias que sobre-estimó en el arreglo de 25 canales de adquisición. A partir de esta tabla se concluye que las crepitancias con factor de escala 1.5 fueron más difíciles de detectar en comparación con los factores de escala 2.5 y 3.5.

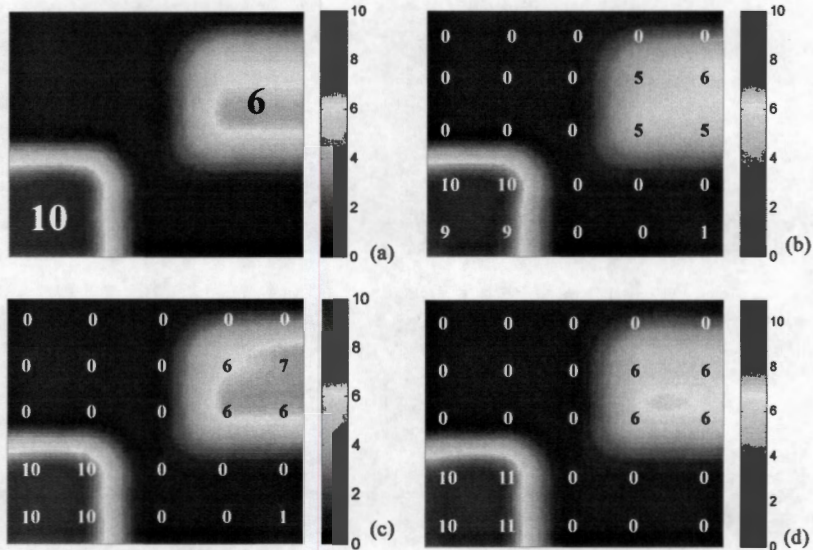


Figura 6.36. Imágenes obtenidas a partir del patrón # 1 considerando crepitancias gruesas. (a) Imagen patrón, (b) Imagen estimada con factor de escala de 1.5, (c) Imagen estimada con factor de escala de 2.5, (d) Imagen estimada con factor de escala de 3.5.

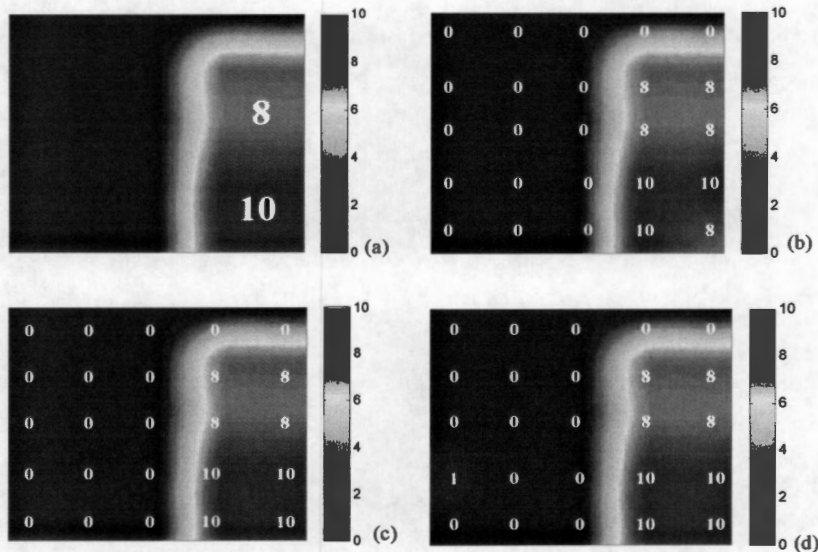


Figura 6.37. Imágenes obtenidas a partir del patrón # 2 considerando crepitancias gruesas. (a) Imagen patrón, (b) Imagen estimada con factor de escala de 1.5, (c) Imagen estimada con factor de escala de 2.5, (d) Imagen estimada con factor de escala de 3.5.

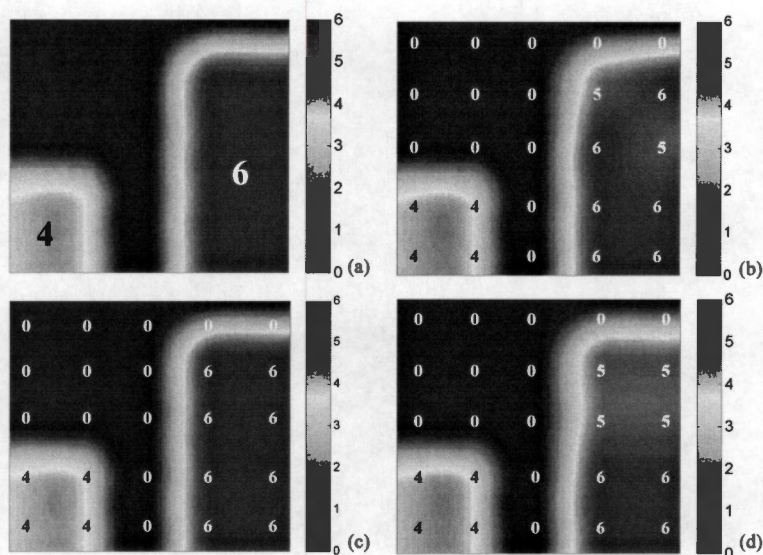


Figura 6.38. Imágenes obtenidas a partir del patrón # 3 considerando crepitancias gruesas. (a) Imagen patrón, (b) Imagen estimada con factor de escala de 1.5, (c) Imagen estimada con factor de escala de 2.5, (d) Imagen estimada con factor de escala de 3.5.

Tabla 6.6

Conteo de crepitancias por patrón de inserción y factor de escala.

Crepitancias Finas

	Crepitancias estimadas			Crepitancias insertadas
	Factor 1,5	Factor 2,5	Factor 3,5	
Patrón 1	60	65	60	64
Patrón 2	68	75	72	72
Patrón 3	61	60	64	64
No. crepitancias sin estimar	11	4	4	
No. crepitancias sobre estimadas	0	4	0	

Crepitancias Gruesas

	Crepitancias estimadas			Crepitancias insertadas
	Factor 1,5	Factor 2,5	Factor 3,5	
Patrón 1	60	66	66	64
Patrón 2	70	72	73	72
Patrón 3	62	64	60	64
No. crepitancias sin estimar	8	0	4	
No. crepitancias sobre estimadas	0	2	3	

6.5 DETECCIÓN DE CREPITANCIAS COMBINADAS FINAS Y GRUESAS VÍA EL MODELO AR VARIANTE EN EL TIEMPO

Se insertaron crepitancias finas y gruesas en una misma señal de sonido respiratorio, dos crepitancias finas y dos gruesas en las regiones 5 y 6. Se utilizaron los factores de escala 1.5 y 3.5. En la figura 6.39 se muestran la imagen patrón y la estimada con el factor de escala de 1.5. Obsérvese que las imágenes son muy semejantes.

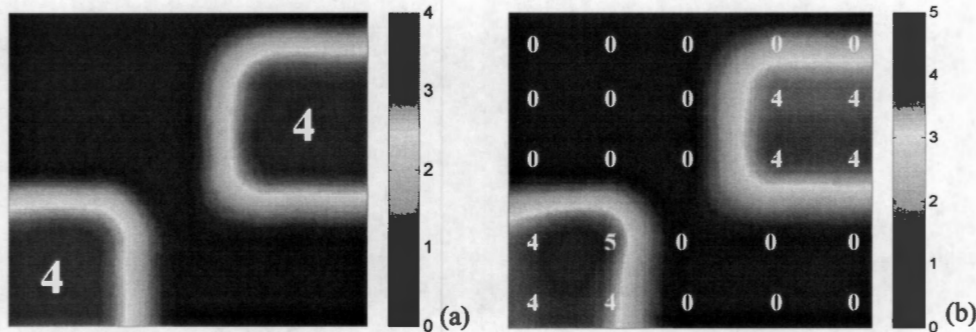


Figura 6.39. (a) Imagen patrón, (b) Imagen estimada con factor de escala de 1.5 con crepitancias finas y gruesas.

En la figura 6.40 se muestra la señal de sonido respiratorio correspondiente al canal PLX4 de la imagen estimada en la figura 6.39(b), las líneas en rojo indican la posición de las crepitancias insertadas de las cuales la 1ª y 3ª crepitancias son finas y la 2ª y 4ª son gruesas. A partir de la morfología de la señal en la figura 6.40 no es posible indicar la presencia de crepitancias, sin embargo, en este canal el algoritmo estimó 4 crepitancias. En la figura 6.41 se muestran las derivadas de la serie de los coeficientes ARVT del canal PLX4. El modelo ARVT fue capaz de detectar crepitancias finas y gruesas inmersas en una misma señal, a pesar de que las crepitancias gruesas suponen mayor dificultad en su detección.

6.6 MODELO AR VARIANTE EN EL TIEMPO VS. EXPERTO MÉDICO

Se utilizaron dos esquemas para comparar el desempeño del modelo ARVT contra el desempeño de un experto médico. El primer esquema incluyó crepitancias simuladas, cuyas características como el factor de escala, número y posición temporal de las crepitancias no fueron conocidas por el experto. Las crepitancias simuladas se insertaron en 2 canales,

PRC2 y PLX4 en fases inspiratorias; 6 crepitancias en PRC2 y 10 crepitancias en PLX4, con un factor de escala de 3.5. En el segundo esquema se utilizaron dos registros multicanales con crepitancias reales, pertenecientes a pacientes diagnosticados con alveolitis alérgica extrínseca. La estimación de crepitancias se realizó a través del modelo ARVT y la estimación del experto médico se basó en el criterio TEWA en combinación con su experiencia clínica.

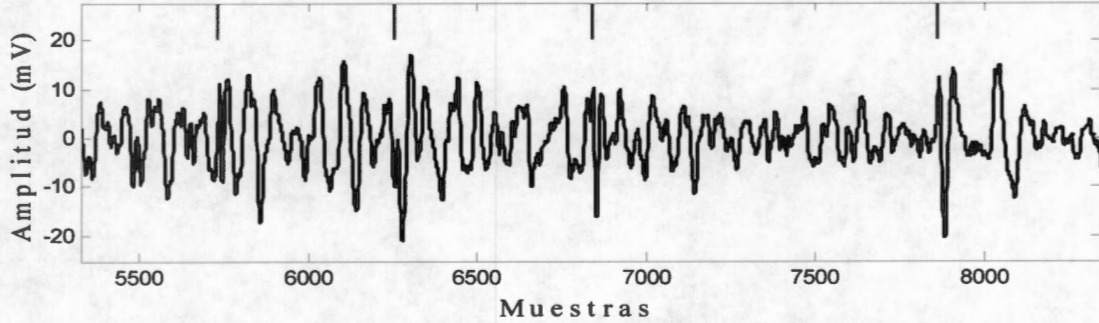


Figura 6.40. Señal de sonido respiratorio con crepitancias finas y gruesas, las líneas rojas indican el inicio de las crepitancias. La 1ª y 3ª crepitancias son finas y las restantes son crepitancias gruesas.

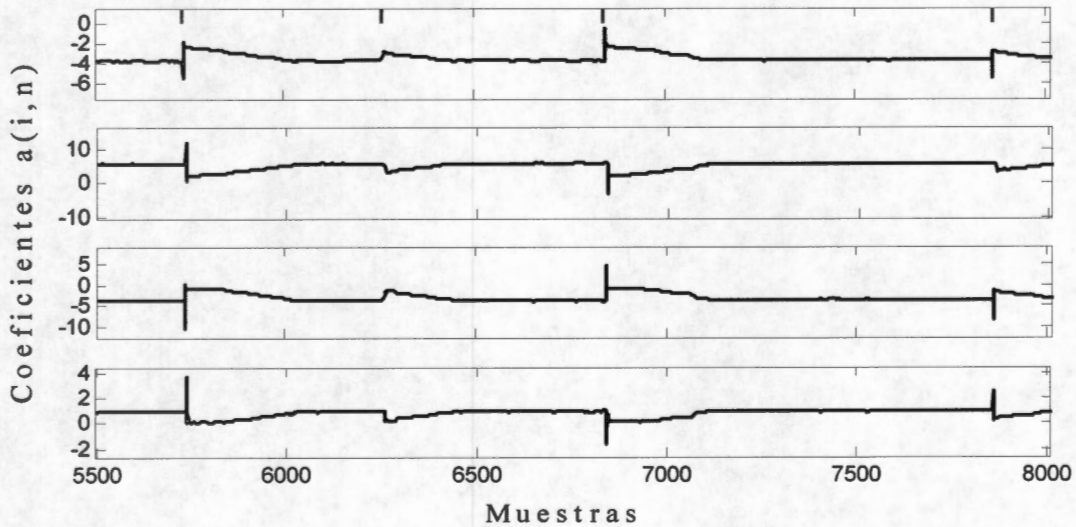


Figura 6.41. Derivadas de los coeficientes ARVT del canal PLX4.

En la figura 6.42 se muestra la señal del canal PRC2 con 6 crepitancias finas insertadas, durante una inspiración, las posiciones de las crepitancias son: 4451, 4456, 5287, 5393, 7065 y 7225 y se señalan mediante líneas rojas. En este caso, las dos primeras crepitancias están traslapadas, por lo tanto, el modelo ARVT sólo detectó 5 de 6 crepitancias. En la figura 6.43 se observa el canal PLX4 con 10 crepitancias finas insertadas

RESULTADOS

durante una inspiración, en este caso las crepitancias no se encuentran traslapadas por lo tanto el algoritmo basado en el modelo ARVT fue capaz de detectar las 10 crepitancias.

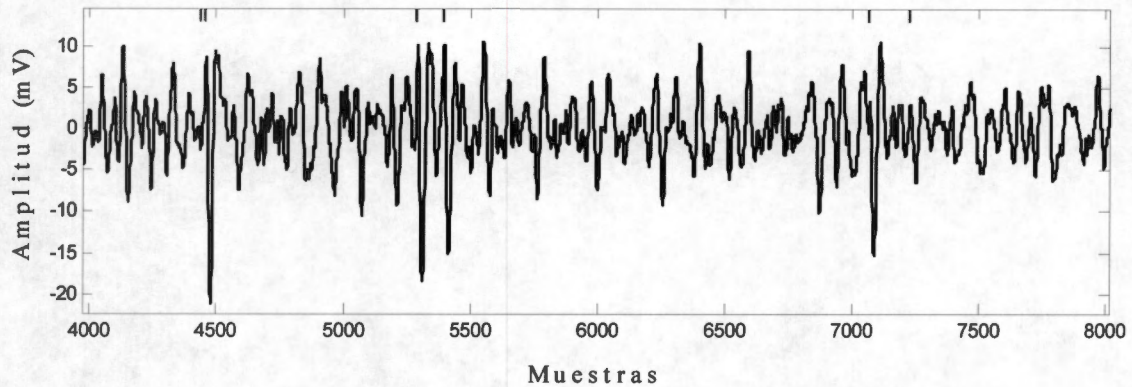


Figura 6.42. Señal del canal PRC2 con 6 crepitancias insertadas durante una inspiración

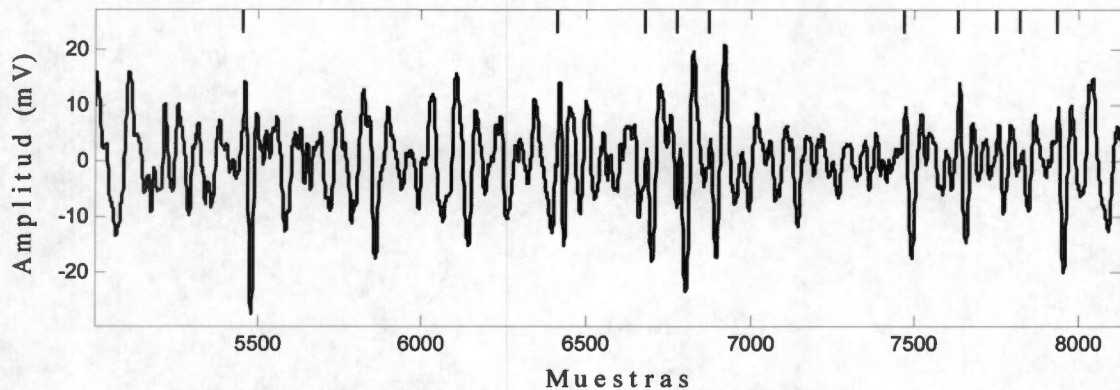


Figura 6.43. Señal del canal PLX4 con 10 crepitancias insertadas durante una inspiración

La tabla 6.7 resume los resultados derivados de los dos esquemas para comparar la capacidad de detección de crepitancias entre el modelo ARVT y el experto. En el primer esquema, el cual utilizó un factor de escala de 3.5, los datos muestran que el modelo ARVT no tiene diferencias significativas ($p > 0.05$) con el número de crepitancias insertadas, mientras que la detección del experto médico fue limitada con diferencias significativas ($p < 0.05$). En el segundo esquema, con crepitancias reales, el número de crepitancias detectadas fue similar ($p > 0.05$) entre el modelo ARVT y el experto. A pesar de que la SNR no fue medida en los casos reales, se estima que la SNR fue mucho más alta que la producida por el factor de escala 3.5.

A continuación se muestra la imagen obtenida al aplicar el modelo ARVT en un sujeto enfermo simulado considerando la información de sólo una fase inspiratoria, en contraste

con la imagen resultado del conteo realizado por el experto en los 25 canales, figura 6.44. Obsérvese la cercanía en la distribución espacial de los crepitantes detectados en ambas imágenes, no así en el número de ellos.

Tabla 6.7. Capacidad de detección de crepitancias entre el modelo ARVT y el experto médico

ESQUEMA A (Simulación)			
Posición del sensor	Numero de crepitancias insertadas	Crepitancias estimadas ARVT promedio±desv. est.	Crepitancias estimadas Experto promedio±desv. est.
PRC2 (n=5 Insp)	6	5.6 ± 0.6	2.0 ± 1.8*
PLX4 (n=5 Insp)	10	10.2 ± 1.1	2.7 ± 1.9*
ESQUEMA B (Real)			
		Crepitancias estimadas ARVT promedio±desv. est.	Crepitancias estimadas Experto promedio±desv. est.
CASO 1 (n=7 Insp)	-	5.7 ± 1.1	6.0 ± 1.0
CASO 2 (n=6 Insp)	-	3.2 ± 1.9	3.0 ± 1.7

* Diferencias significativas, $p < 0.05$, comparadas con el número de crepitancias insertadas. n se estableció para un número de fases inspiratorias analizadas.

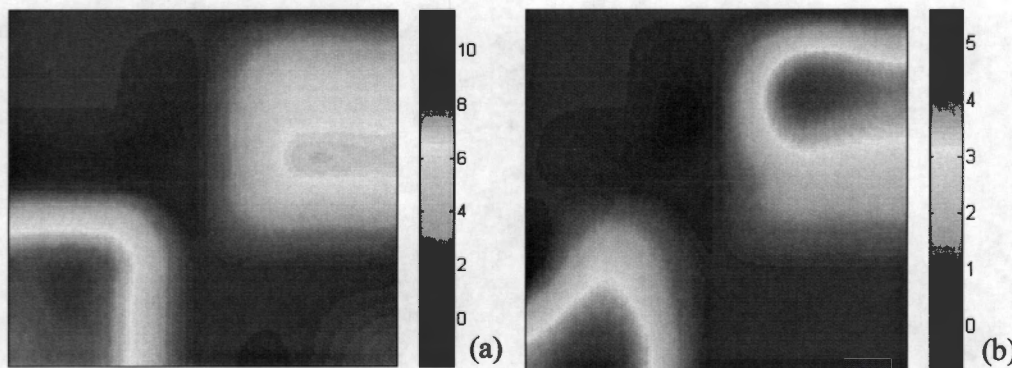


Figura 6.44. (a) Imagen obtenida a través del modelo ARVT, (b) conteo del experto en los 25 canales plasmado en una imagen.

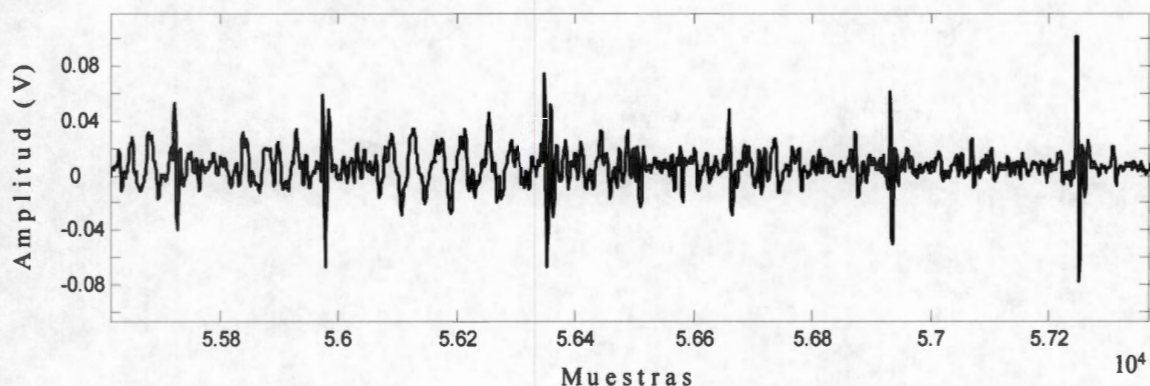


Figura 6.45. Señal de sonido respiratorio con crepitancias reales.

En la figura 6.45 se muestra una señal de sonido respiratorio de un sujeto enfermo con Neumonía Intersticial Difusa y es evidente la diferencia en amplitud de las crepitancias reales y las simuladas en la presente tesis.

6.7 IMAGEN DE SONIDOS ADVENTICIOS DISCONTINUOS EN CASOS REALES

El modelo ARVT se probó con sonidos respiratorios de un sujeto enfermo que previamente fue auscultado por dos expertos médicos. Los expertos generaron una representación de la información que escucharon. La representación por auscultación se muestra en la figura 6.46(a), donde la diagonal a 45° representa la fase inspiratoria y la diagonal a 135° representa la fase espiratoria, el número de diagonales representa la intensidad acústica escuchada y los círculos representan la cantidad de estertores crepitantes escuchados en el sentido de escasos, moderados y abundantes. Por ejemplo, en el canal PRC4, uno de los expertos escuchó la espiración más intensa que la inspiración y en la inspiración escuchó el mayor número de crepitancias. Posteriormente se le solicitó al experto en neumología que aplicando TEWA y su experiencia clínica contara el número de crepitancias en PRC4, el resultado coincidió con el número estimado por el modelo ARVT.

Los sonidos respiratorios con crepitancias reales se procesaron con el modelo ARVT y el resultado se muestra en la imagen 6.46 (b). La imagen estimada resalta algunas zonas que coinciden con el mapa que generaron los expertos. El algoritmo detectó más crepitancias que los expertos posiblemente debido a las limitaciones del oído humano. Es

RESULTADOS

importante remarcar que el objetivo principal es resaltar zonas con crepitancias, lo cual se logró a través de este método con crepitancias reales a pesar de que en sonidos respiratorios de un sujeto enfermo las crepitancias tienen diferentes SNR.

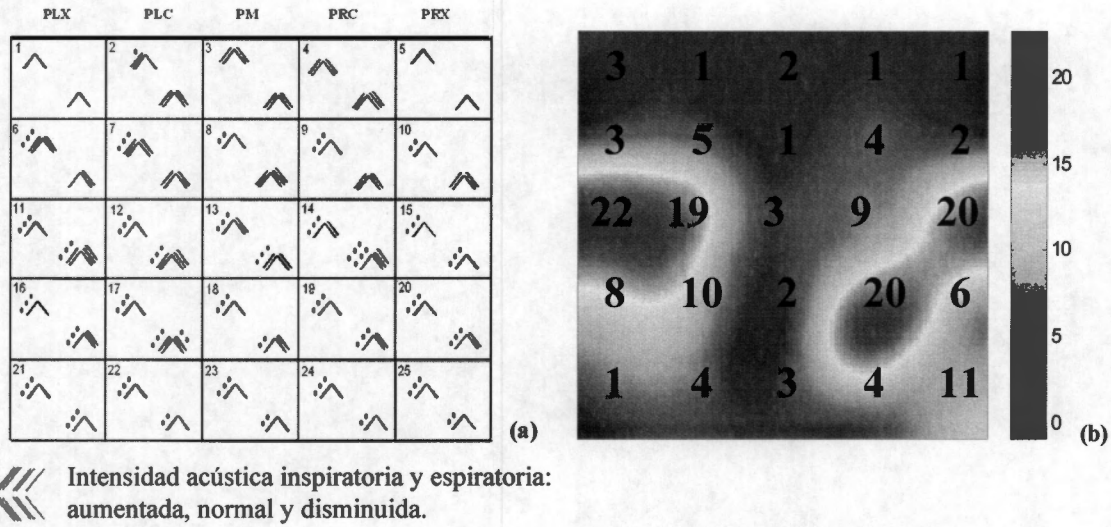


Figura 6.46. Imagen de sonidos adventicios discontinuos de un paciente con neumopatía intersticial difusa. (a) Resultados de la auscultación realizada por dos expertos y (b) imagen obtenida a través del modelo ARVT incluyendo el número de crepitancias estimadas en cada canal.

CAPITULO VII

DISCUSIÓN DE RESULTADOS

Los sonidos adventicios discontinuos están fuertemente relacionados con diversas patologías y su uso clínico para la interpretación de enfermedades respiratorias depende en gran parte de su eficiente detección de los sonidos vesiculares. En el caso particular de los sonidos crepitantes su análisis se ha abordado de diferentes formas. Sin embargo la experiencia indica que la detección de estertores crepitantes resulta una tarea compleja, desde el punto de vista del procesamiento digital de señales, ya que intervienen diversos factores como: la no-estacionaridad de los sonidos respiratorios y las crepitancias, la relación señal a ruido entre crepitancias y sonidos respiratorios (SNR), el posible traslape temporal de las crepitancias, y la modificación de la forma de onda de las crepitancias debido a la influencia del sonido respiratorio de base.

Como se mencionó en el capítulo de antecedentes, se han desarrollado diferentes esquemas de procesamiento para identificar o aislar las crepitancias [1-3]. Sin embargo, existen varias limitaciones a las propuestas realizadas, por ejemplo, para evaluar el desempeño de los esquemas se ha tomado como patrón los resultados obtenidos por médicos especialistas al aplicar la técnica TEWA a señales de sonido pulmonar. Sin embargo, la técnica TEWA requiere que la amplitud de las crepitancias sea del doble de la amplitud del sonido respiratorio de base y por lo tanto, esto conduce a que las crepitancias sean relativamente fácil de visualizar realizando la expansión en el tiempo de la señal adquirida. Además, en las investigaciones reportadas se han seleccionado posiciones en la espalda de sujetos enfermos donde las crepitancias son muy evidentes para el médico especialista. Además, ya se ha demostrado que es difícil para los médicos detectar en forma auditiva las crepitancias, incrementando el grado de dificultad para las crepitancias de tipo grueso [4].

En contraste, en este trabajo se decidió trabajar con crepitancias simuladas ya que de esta forma se tiene control de su número y su distribución espacial, entre otros parámetros.

En este capítulo se discuten los resultados obtenidos al agregar crepitancias simuladas al sonido respiratorio de sujetos sanos, previamente adquiridos en forma

multicanal, y de aplicar técnicas de procesamiento digital de señales con la finalidad de proporcionar una imagen que refleje la localización o distribución espacial así como su relevancia, de los sonidos adventicios discontinuos (sonidos estertores crepitantes). Además, se discuten los resultados de los esquemas propuestos en sonidos respiratorios con crepitancias reales; es decir, con señales provenientes de sujetos enfermos y el desempeño de un experto médico con uno de los esquemas propuestos para la detección de estertores crepitantes.

La discusión de los resultados obtenidos en esta tesis aborda los siguientes puntos:

1. Detección de crepitancias por medio de dimensión fractal.
2. Detección de crepitancias mediante parametrización con un modelo AR invariante en el tiempo y clasificación por redes neuronales artificiales.
3. Detección de crepitancias mediante un modelo AR variante en el tiempo.
4. Comparación entre una detección computarizada y una detección visual realizada por expertos.

7.1 ANÁLISIS NO LINEAL MEDIANTE DIMENSIÓN FRACTAL

A pesar de que en la literatura se menciona a la dimensión fractal (DF) como una técnica adecuada para detectar la presencia de crepitancias, en este trabajo se demuestra que a partir de la serie de tiempo que se genera con los valores de dimensión fractal resulta difícil detectarlas, ya sea por medios automáticos o visuales. Además, en la presente tesis se establece la dependencia de la técnica de dimensión fractal con la amplitud del estertor. Con una relación crepitancia a sonido pulmonar de base baja no es posible discernir sin ambigüedades las posiciones de las crepitancias; la curva de DF presenta información de amplitud relevante en tiempos donde no se insertaron crepitancias. Si la amplitud de las crepitancias aumenta es decir, se incrementa la relación crepitancia a sonido pulmonar se observa que en la curva de DF aparecen cambios que se pueden asociar fácilmente a las crepitancias, y por lo tanto susceptibles de detectarse automáticamente. En investigaciones donde han utilizado la dimensión fractal [3] se han obtenido buenos resultados en la detección de crepitancias. Sin embargo, es necesario considerar que la SNR es mayor a la utilizada en el presente

trabajo, y seguramente es la razón principal de los resultados reportados en la presente comunicación. En la práctica clínica la amplitud de las crepitancias no siempre es mayor a la amplitud del sonido pulmonar de base, como lo demuestra nuestra base de datos de sujetos enfermos. Sin embargo, el concepto de irregularidad o complejidad asociado con la teoría de la dimensión fractal resulta atractivo en la evaluación de la condición pulmonar de un sujeto y es posible que combinando la dimensión fractal con otras técnicas de procesamiento se obtengan resultados que sean de utilidad en el diagnóstico médico. Más aún, es posible que en enfermedades que no se caracterizan por la presencia de estertores finos o gruesos el mapa de complejidad revele aspectos patológicos pulmonares.

Un hecho que puede resultar interesante en futuras investigaciones es que los mapas de complejidad promediando la dimensión fractal de cada ventana temporal de la inspiración y calculando la dimensión fractal de toda la inspiración resultan similares; es posible que en otras condiciones pulmonares el comportamiento dinámico de la DF contenga información útil.

7.2 DETECCIÓN DE CREPITANCIAS MEDIANTE PARAMETRIZACIÓN CON UN MODELO AR INVARIANTE EN EL TIEMPO Y CLASIFICACIÓN POR REDES NEURONALES ARTIFICIALES

Es necesario recordar que esta propuesta se basa en un investigación previa que demostró que es posible realizar una clasificación entre sonido respiratorio de sujetos sanos y sonido respiratorio de sujetos enfermos con neumopatía intersticial difusa [5]. En esta investigación se requirió etiquetar la información acústica de ventanas temporales, dado que se utilizó una RNA del tipo supervisado. Para simplificar la tarea de asignación de etiquetas se consideró que la información acústica, en el caso de los sujetos enfermos, en la segunda mitad de la fase inspiratoria representaba la parte enferma. Lo ideal es tener una base de datos etiquetada por médicos especialistas que permita conocer en que posiciones se encuentran los sonidos adventicios, sin embargo el realizarla representa una cantidad de trabajo fenomenal. Desafortunadamente no existen bases de datos de sonidos respiratorios y en consecuencia, en la presente tesis se decidió generar bases de datos simuladas para asignar etiquetas de una forma fácil y correcta, dado que se conoce la posición de las crepitancias insertadas. A pesar de que la base de datos generada en este trabajo no fue tan extensa, las imágenes obtenidas

muestran buenos resultados.

El orden del modelo autorregresivo se estableció, utilizando el criterio de Akaike, en cuatro y seis. Otros investigadores como Sankur *et al.* [6] han reportado que un orden seis del modelo AR es adecuado en la parametrización de los sonidos respiratorios; nuestros resultados corroboran que efectivamente con un orden seis se obtienen mapas de sonidos adventicios que se correlacionan de mejor forma con la imagen patrón simulada.

La generación de la imagen de sonidos adventicios con el esquema de procesamiento AR-RNA requirió la definición de un índice de crepitancias, basado en la relación del porcentaje de ventanas de información temporal clasificadas como enfermas y sanas.

Se generaron imágenes mediante dos protocolos de entrenamiento, denominados por regiones y por canales y los resultados indican que las imágenes estimadas con entrenamiento por regiones presentan una distribución espacial más cercana a las imágenes patrón, mientras que las imágenes estimadas con entrenamiento por canales se alejan un poco más de las imágenes patrón. Probablemente debido a que en el entrenamiento por canales se utilizó un menor número de vectores de características, la generalización no fue tan buena como la del entrenamiento por regiones. Además, hay que considerar que se insertó el mismo número de crepitancias en los cuatro canales que conformaron una región, esto probablemente generó mejores resultados en el caso por regiones. La similitud de la distribución espacial entre las imágenes estimadas y patrón se evaluó visualmente sin embargo, en una futura investigación y con una base de datos etiquetada es recomendable cuantificar las imágenes obtenidas, por ejemplo mediante la utilización de técnicas de procesamiento de imágenes como la información mutua. Un comentario adicional sobre los procedimientos de entrenamiento se relaciona con la posible utilidad de entrenar por regiones en investigaciones que incluyan el fenómeno de transmisión de los sonidos crepitantes.

En cuanto a los errores alcanzados en las etapas de entrenamiento de las RNAs, la tabla 6.2 del capítulo de Resultados muestra que al considerar crepitancias finas se alcanzan errores menores en comparación a los errores generados con crepitancias gruesas. Lo anterior, sugiere que existe una mayor dificultad de la RNA para clasificar las crepitancias gruesas. Además, los errores de los entrenamientos por regiones fueron menores a los alcanzados por los entrenamientos por canales.

Las imágenes estimadas por el método AR-RNA utilizando el entrenamiento por regiones con crepitancias finas muestran una buena aproximación a la imagen patrón. Por ejemplo, en el caso de la figura 6.9 del capítulo de Resultados se insertaron un máximo de 10 crepitancias en la imagen patrón y si consideramos que existe una crepitancia en cada ventana temporal de la información acústica, el valor del índice sería $(\% \text{ ventanas enfermas}) / (\% \text{ ventanas sanas}) = 10/20 \times 100 = 50\%$. Obsérvese el valor del índice de las imágenes estimadas de la figura 6.9, cuyo rango de valores se muestran en la paleta de color, los valores se encuentran cercanos al 50%; en general, el valor del índice puede ser menor a 50% debido al posible traslape temporal de las crepitancias.

Por otro lado, los resultados con entrenamiento por regiones y crepitancias gruesas muestran valores del índice mayores, lo cual se puede interpretar como que la RNA clasificó más ventanas temporales de la información acústica en la clase enferma que en el caso de crepitancias finas. Es decir, las características del sonido respiratorio de base se encuentra más cercanas a las características de las crepitancias gruesas vía la parametrización de la información con el modelo AR.

Las imágenes estimadas a partir del entrenamiento de la RNA por canales con crepitancias finas muestran índices menores a los esperados para la SNR más baja. Nuestros resultados muestran que el esquema AR-RNA con entrenamiento por canales se afecta en mayor medida por la SNR. Si se comparan las imágenes estimadas de crepitancias finas y gruesas del entrenamiento por canales, se observa que con crepitancias gruesas los índices estimados fueron un poco mejores. En el futuro, las simulaciones se podrían realizar pensando en generar escenarios más cercanos a la realidad, por ejemplo insertar un número diferente de crepitancias en un grupo de canales mientras que en otros, fenómeno de transmisión, un número igual de crepitancias. Es importante señalar que los valores negativos del índice, que aparecen en la barra de color, en las imágenes estimadas son debidos a la interpolación realizada.

El efecto de la SNR en el desempeño de la RNA, cuando se utiliza el entrenamiento por regiones, quedó demostrada en los resultados de las figuras 6.22 y 6.23. El número de ventanas temporales de información acústica sin clasificar es menor para el factor de escala 3.5, es decir, la mejor SNR.

Los porcentajes de desempeño de las RNAs, sensibilidad y especificidad en la etapa de prueba, en general fueron buenos por arriba del 92%. Sin embargo, es posible que los resultados mejoren si la base de datos es de mayor tamaño en el sentido de incluir más

ejemplos considerando más tipos de crepitancias; es decir, en este trabajo se consideraron sólo dos tipos de crepitancias, finas y gruesas, con parámetros en tiempo fijos. No obstante, en datos adquiridos las crepitancias pueden tener un rango de valores para IDW, 2CD, etc, generando crepitancias más angostas o más expandidas en tiempo que las consideradas en este trabajo. De hecho, se probó agregar diferentes tipos de crepitancias, modificando sus parámetros característicos y agregando también crepitancias reales, sin embargo, los resultados con datos reales no mostraron las zonas esperadas en base a los expertos.

7.3 DETECCIÓN DE CREPITANCIAS MEDIANTE UN MODELO AR VARIANTE EN EL TIEMPO

Los resultados obtenidos al aplicar el esquema de procesamiento con base en ARVT en los sujetos de prueba, muestran que la estimación del número de crepitancias fue satisfactoria, ya que la estimación se aproximó al número original de crepitancias insertadas, en la mayoría de los casos. En la tabla 6.6 del capítulo de Resultados se muestra un conteo de crepitancias por cada imagen obtenida, así como una relación del número de crepitancias sin estimar y sobre-estimadas con respecto al patrón generado. La información de la tabla 6.6, permite inferir que las crepitancias más difíciles de detectar fueron las crepitancias con factor de escala 1.5. Es importante señalar que en las figuras 6.1 y 6.2 del capítulo de Resultados algunas crepitancias con factor de escala 1.5 son imperceptibles visualmente, y en consecuencia es muy complicado detectarlas, a través del criterio TEWA, para un experto en el campo de la neumología.

Con respecto a la dependencia con la SNR, el esquema ARVT fue capaz de detectar las crepitancias con los tres factores de escala utilizados, es decir ARVT es más robusto en la estimación del número de crepitancias, en comparación con el esquema que utiliza el cálculo de la dimensión fractal. Una ventaja adicional del esquema ARVT se relaciona con la capacidad de discernir crepitancias que se encuentran no completamente traslapadas pero temporalmente muy cercanas. A pesar de existir un cierto porcentaje de traslape entre crepitancias fue posible generar imágenes representativas de la información presente en la imagen patrón.

Dado que el esquema ARVT proporcionó los mejores resultados en los casos simulados, el esquema se evaluó en un escenario simulado con la combinación de

crepitancias finas y gruesas. A pesar de que las crepitancias gruesas son más difíciles de identificar, debido a que su morfología se mezcla en mayor medida con el sonido de base, los resultados demuestran que el esquema ARVT es capaz de identificarlas.

La ventaja del modelo ARVT sobre la metodología AR-RNA, es que no se requiere generar un banco de datos etiquetado así como efectuar una etapa de entrenamiento.

7.4 COMPARACIÓN ENTRE UNA DETECCIÓN COMPUTARIZADA Y UNA VISUAL REALIZADA POR EXPERTOS

En la comparación de la capacidad de detección de crepitancias entre el esquema ARVT y un experto médico, los resultados muestran la dificultad que tiene el experto para detectar crepitancias con las características que se utilizaron en este trabajo. La tabla 6.7 resume los resultados obtenidos, en el primer esquema con un factor de escala 3.5 el modelo ARVT no presentó diferencias significativas ($p > 0.05$) y la detección del experto mostró un desempeño pobre con diferencias significativas ($p < 0.05$). En el esquema con crepitancias reales, el número de crepitancias detectadas fue similar ($p > 0.05$) entre el modelo ARVT y el experto. La SNR en los casos reales no se midió pero se estima que la SNR es mayor que la producida por el factor de escala 3.5.

En la figura 6.45 del capítulo de Resultados se muestra un ejemplo de un caso real donde las crepitancias son evidentes. Por lo tanto, se asume que si el factor de escala fuera menor que el de 3.5, el experto tendría más limitaciones para detectar crepitancias aplicando el criterio TEWA. Este problema puede ser debido a la distorsión que genera el sonido de base a la crepitancia y al traslape temporal. De acuerdo a los resultados obtenidos el modelo ARVT parece tener un buen desempeño aún con traslape temporal y posibles cambios de la morfología de la señal.

La figura 6.46, en el capítulo de Resultados, muestra una imagen de sonidos adventicios generada a partir del registro multicanal de un enfermo con Neumopatía Intersticial Difusa, la imagen presenta zonas que indican la presencia importante de estertores. Al solicitarle al médico especialista que revisara los registros de sonido respiratorio se comprobó la presencia de estertores y se confirmó que el esquema ARVT estimó de forma adecuada su número. Este resultado es importante, ya que mediante el esquema ARVT se cumple el objetivo de proporcionar una idea al médico de las zonas afectadas con crepitancias.

REFERENCIAS

- 1 L.J. Hadjileontiadis, S.M. Panas, "Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter," *IEEE Trans. Biomed. Eng.*, vol. 44, no. 12, pp. 1269–1281, 1997.
- 2 Y.A. Tolia, L.J. Hadjileontiadis, and S.M. Panas, "Real-time separation of discontinuous adventitious sounds from vesicular sounds using a fuzzy rule-based filter," *IEEE Trans. Inf. Technol. Biomed.*, vol. 2, no. 3, pp. 204–215, 1998.
- 3 L. J. Hadjileontiadis, I. T. Rekanos, "Detection of explosive lung and bowel sounds by means of fractal dimension," *IEEE Signal Proc. L.*, vol. 10, no. 10, pp. 311-314, 2003.
- 4 H. Kiyokawa, M. Greenberg, K. Shirota, and H. Pasterkamp, "Auditory detection of simulated crackles in breath sounds," *Chest*, vol. 119, no. 6, pp. 1886-1892, 2001.
- 5 H.G. Martinez-Hernandez, T. Aljama-Corrales, R. Gonzalez-Camarena, S. Charleston-Villalobos, G. Chi-Lem, "Computerized classification of normal and abnormal lung sounds by multivariate linear autoregressive model," *Proc. 27th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1464-1467, Shanghai, China 2005.
- 6 B. Sankur, Y.P. Kahya, E.Ç. Güler, T. Engin, "Comparison of AR based algorithms for respiratory sounds classification," *Comput. Biol. Med.*, vol. 23, pp. 67-67, 1994.

CAPITULO VIII

CONCLUSIONES

Una de las aportaciones de la presente tesis es establecer el concepto de imagenología de sonidos adventicios discontinuos, que posiblemente represente una ayuda al diagnóstico y seguimiento de enfermedades pulmonares. La imagen generada podría en un futuro cercano integrarse al expediente del paciente y sería posible seguir los cambios en la distribución espacial de los crepitantes, así como la disminución o su aumento, a lo largo del tiempo bajo los efectos del tratamiento médico. Hay que hacer notar que este tipo de imagen se encuentra ligada a la función pulmonar. Otra de las aportaciones es explotar la importancia de generar escenarios simulados que permiten explorar y evaluar el desempeño de esquemas de procesamiento en el análisis de sonidos discontinuos.

Los resultados muestran que el esquema de procesamiento con base a la *Dimensión Fractal* de los sonidos respiratorios, no mostró resultados satisfactorios con las relaciones de magnitud de la crepitancia a la magnitud del sonido de base (SNR), utilizadas en la presente investigación. Es importante mencionar que otros autores han obtenido buenos resultados al utilizar la Dimensión Fractal, sin embargo la magnitud reportada de los sonidos adventicios es por lo menos de tres veces la magnitud máxima de los estertores incluidos en el presente trabajo. Es evidente que con estertores de gran magnitud es posible generar mapas de sonidos adventicios de una manera confiable. La Dimensión Fractal y el concepto relacionado de complejidad de una serie de tiempo podrían tener impacto en la generación de diversos mapas afines con la acústica pulmonar de un sujeto y es por esto que se sugiere realizar más investigación al respecto.

En el caso de los esquemas de procesamiento con base en AR-RNA y ARVT los resultados muestran que se pueden detectar crepitancias finas y gruesas y en consecuencia, generar una imagen de sonidos adventicios. El esquema AR-RNA construye la imagen relacionada con la presencia de crepitancias a través del desempeño de la clasificación, mientras que el modelo ARVT realiza el mapeo del número estimado de crepitancias. Un parámetro fundamental que se consideró en la presente tesis es la SNR de la crepitancia al sonido de base, comprobando su impacto y utilizándolo para validar la robustez de los métodos propuestos. Aún con el factor de escala más pequeño, valor de 1.5, la imagen acústica de sonidos adventicios muestra satisfactoriamente las zonas pulmonares donde se

insertaron las crepitancias. Lo anterior es importante considerando que en la práctica clínica el experto aplica el criterio TEWA en señales equivalentes a un factor de escala de 3.5 o mayor; es posible que los métodos expuestos en este trabajo sean más sensibles a detectar crepitancias con SNRs bajas si se comparan con los resultados obtenidos con el médico especialista.

Una ventaja importante del modelo ARVT es que no necesita una etapa de entrenamiento, debido a que el mapa se genera directamente a partir del número estimado de crepitancias, ventaja que es importante desde el punto de vista del tiempo de procesamiento en el análisis de los sonidos respiratorios.

El esquema ARVT es una alternativa adecuada para estimar el número de crepitancias finas y gruesas, incluso en presencia de crepitancias traslapadas y con una amplitud equivalente a 1.5 de la desviación estándar del sonido de base. El criterio TEWA en estas circunstancias revela limitaciones evidentes. Además, también se comprobó que las crepitancias gruesas con un factor de escala de 1.5 son las más difíciles de diferenciar del sonido de base, a pesar de esto, el modelo ARVT y el método AR-RNA fueron capaces de obtener buenos resultados.

El esquema ARVT, con la base de datos disponible y los factores de escala utilizados, fue el único capaz de detectar zonas enfermas con datos reales. Para mejorar los resultados con RNAs, probablemente es necesario una base de datos más extensa que incluya una mayor cantidad de crepitancias y con mayor variedad de su tipo. Es decir, tomar en cuenta más variaciones en los parámetros temporales de las crepitancias. En el futuro, la detección de crepitancias se podría complementar diferenciando entre crepitancias finas y gruesas, a través del modelo ARVT, ya que en los coeficientes se podría medir la duración de los cambios no estacionarios y correlacionarse con el tipo de crepitancias con base a sus parámetros en el tiempo. La inserción de crepitancias por regiones podría ser retomada tratando de simular la transmisión de las crepitancias; el número de crepitancias que conforman la región probablemente no tendría que ser el mismo.

Evidentemente, es necesario realizar más investigación que considere sujetos enfermos con otro tipo de padecimiento pulmonar y de cómo modificar o combinar esquemas de procesamiento digital para apoyar de mejor forma al diagnóstico médico vía imágenes de sonidos adventicios. Sin embargo, en la presente tesis se cumplió con el objetivo de generar una imagen de sonidos adventicios capaz de brindar asistencia al experto clínico en la detección y conteo de crepitancias.

ANEXO

PARTICIPACIONES EN CONGRESOS

“Crackles Detection Using a Time-Variant Autoregressive Model”, G. Dorantes-Méndez, S. Charleston-Villalobos, T. Aljama-Corrales, G. Chi-Lem, R. González-Camarena. IEEE Engineering in Medicine and Biology Society, "Personalized Healthcare through Technology", Vancouver Convention & Exhibition Centre Vancouver, British Columbia, Canada, Agosto 20-24, 2008.

“Imaging of Simulated Crackle Sounds Distribution on the Chest”, G. Dorantes-Méndez, S. Charleston-Villalobos, T. Aljama-Corrales, G. Chi-Lem, R. González-Camarena. IEEE Engineering in Medicine and Biology Society, "Personalized Healthcare through Technology", Vancouver Convention & Exhibition Centre Vancouver, British Columbia, Canada, Agosto 20-24, 2008.

“Discontinuous Adventitious Sounds Imaging”, G. Dorantes-Méndez, S. Charleston-Villalobos, R. González-Camarena, G. Chi-Lem, T. Aljama-Corrales. International Lung Sounds Association, Boston, MA, Octubre 10-11, 2008.

Crackles Detection Using a Time-Variant Autoregressive Model

G. Dorantes-Méndez, S. Charleston-Villalobos, R. González-Camarena, G. Chi-Lem, J. G. Carrillo, T. Aljama-Corrales.

Abstract— Several techniques have been explored to detect automatically fine and coarse crackles; however, the solution for automatic detection of crackles remains insufficient. The purpose of this work was to explore the capacity of the time-variant autoregressive (TVAR) model to detect and to provide an estimate number of fine and coarse crackles in lung sounds. Thus, simulated crackles inserted in normal lung sounds and real lung sounds containing adventitious sounds were processed with TVAR and by an expert that based crackle detection on time-expanded waveform-analysis. The coefficients of the TVAR were obtained by an adaptive filtering prediction scheme. The adaptive filter used the recursive least squares algorithm with a forgetting factor of 0.97 and the model order was four. TVAR model showed an efficiency to detect crackles over 90% even with crackles overlapping and amplitudes as low as 1.5 of the standard deviation of background lung sounds, where expert presented an efficiency around 30%. In conclusion, TVAR model is a proper alternative to detect and to provide an estimate number of fine and coarse crackles, even in presence of crackles overlapping and crackles with low amplitude, conditions where crackles detection based on time-expanded waveform-analysis reveals evident limitations.

I. INTRODUCTION

Lung sound (LS) analysis from electronic auscultation has been of great interest for several decades, as a diagnostic tool of pulmonary diseases. In particular, crackle sounds has been associated to both cardiac and pulmonary diseases [1]. Among the crackle features mentioned as useful to differentiate such pathologies, the waveform, the spectral content, the pulmonary distribution, the number, and timing during the breathing cycle have been the most explored [2].

Murphy *et al.* [3] and Hoevers and Loudon [4] have described crackles through their time-expanded waveform-

analysis (TEWA), whose visual analysis defined useful time parameters as the initial deflection width (IDW), the two cycles duration (2CD) and the largest deflection width (LDW). Based on such parameters, fine and coarse crackles were differentiated and associated to different pathologies [5]. Mori *et al.* [6] and Munakata *et al.* [7] also studied spectral characteristics of crackle sounds and reported their spectral content mainly in the range of 0.1 to 1000 Hz.

Several techniques have also been explored to detect automatically fine and coarse crackles. Among such techniques are nonlinear digital filters [8], the spectral stationarity of LS [9], high-order statistics AR modeling [10], wavelet transform [11], neurofuzzy modeling [12], fractal dimension [13], and empirical mode decomposition [14]. However, the solution for automatic detection of crackles remains insufficient.

Crackles detection by computer is a complex task due to factors such as: a) nonstationarity of LS and crackle sounds, b) low magnitude relation between crackles and LS (SNR), c) crackles overlapping, d) crackle waveform distortion by LS, and e) difficulty to establish the time domain parameters such as IDW, 2CD, and LDW.

Taking into account the former problems, we hypothesize that time-variant autoregressive (TVAR) model is a good alternative to detect and to provide an estimation of the number of crackles, fine and coarse, immersed in LS.

Therefore, the goal in the present work is to explore the capacity of the TVAR modeling to detect and to provide an estimate number of fine and coarse crackles in LS. TVAR model was tested on simulated and real crackles.

II. METHODOLOGY

A. Acquisition of Multichannel LS

Multichannel LS signals were taken from a database that includes healthy subjects whose acquisition protocol was carried out at the Digital Signal and Image Processing Laboratory of the Metropolitan Autonomous University in conjunction with the Acoustic Laboratory of the National Institute of Respiratory Diseases. The sensor array of 5 by 5 consisted in electret microphones (BT-1834 Knowles®) inserted in a plastic bell with a flat frequency response from 50 Hz to 3 kHz. To digitalize the multichannel LS and airflow signals an A/D card was used at 12 bits and sampling frequency of 10 kHz. The subjects were seated, breathing through a calibrated Fleish pneumotachograph and

G. Dorantes Méndez is a Master student of the Biomedical Engineering Program, Universidad Autónoma Metropolitana, Mexico City 09340, Mexico (e-mail: lupita.dorantes@gmail.com).

S. Charleston-Villalobos is with the Electrical Engineering Department, Universidad Autónoma Metropolitana, Mexico City 09340, Mexico (email: schv@xanum.uam.mx).

R. González-Camarena is with the Health Science Department, Universidad Autónoma Metropolitana, Mexico City 09340, Mexico (email: rgc@xanum.uam.mx).

G. Chi-Lem is with the National Institute of Respiratory Diseases, Mexico City, Mexico (email: gchilem@iner.gob.mx)

G. Carrillo is with the National Institute of Respiratory Diseases, Mexico City, Mexico (email: memo2carrillo@hotmail.com)

T. Aljama-Corrales is with the Electrical Engineering Department, Universidad Autónoma Metropolitana, Mexico City 09340, Mexico (corresponding author email: alja@xanum.uam.mx).

wearing a nose clip. The sensors array nomenclature is described elsewhere [14]. In brief, we selected from the posterior surface of the thorax, two LS sensors; one corresponding to the apical region, at the second row of the array and at the right mid-clavicular line projection (PRC2). The second one was located on the basal region at the fourth row and at the left posterior axilar line (PLX4).

Signals were processed by a band pass filter with cutoff frequencies of 75 and 1000 Hz. The lower cutoff frequency was selected in order to attenuate the heart sounds interference.

B. Simulation of Abnormal LS

Fine and coarse crackles were simulated using the mathematical function proposed by Kiyokawa et al. [15]:

$$y_O(t) = \sin(4\pi t^\alpha), \quad \text{with} \quad \alpha = \frac{\log(0.25)}{\log(t_0)}, \quad (1)$$

where $y_O(t)$ has two cycles with the first positive time interception at t_0 . The modulation function of $y_O(t)$ is given by:

$$m(t) = 0.5 \{ 1 + \cos[2\pi(t^{0.5} - 0.5)] \}. \quad (2)$$

The main power of $y(t)$ is concentrated in the initial part of the signal. This mathematical model keeps the time domain characteristics for fine crackles as IDW of 0.5 ms and 2CD of 5ms, while for coarse crackles IDW was 1.2 ms and 2CD of 9 ms.

To simulate abnormal LS, the simulated crackles were embedded within the real LS by an additive model, from half of the inspiratory phase in a random fashion, allowing temporal overlapping of crackles.

One standard deviation (SD) was computed from the segment of LS where crackles were added, and factors of 1.5, 2.5 and 3.5 times the SD were used to provide different crackle amplitudes. Fig. 1 shows an example of fine and coarse crackles inserted in a LS segment from PRC2, with two overlapped crackles, first the coarse and then the fine.

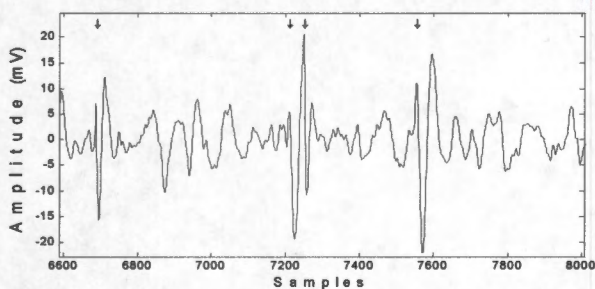


Fig. 1. Lung sound signal with immersed fine and coarse crackles in sensor position PRC2. The first and third arrows point out to fine crackles while the second and fourth arrows point out to coarse crackles.

C. Time Variant AR Model.

The coefficients of the time variant AR model were obtained by adaptive filtering prediction scheme and the representation of the abnormal LS was established by:

$$y(n) = \sum_{i=1}^M a(i,n)y(n-i) + v(n) \quad (3)$$

where M is the model order, $a(i,n)$ represents the TVAR coefficients and $v(n)$ is a white noise error term [16]. The adaptive filter used the recursive least squares (RLS) algorithm to compute the TVAR coefficients [17] with a forgetting factor of 0.97, while the model order was selected according to the Akaike criterion and was fixed at four. The estimated number of crackles was associated with abrupt changes in the derivative of the TVAR coefficient time series. A threshold, based on 0.035 times the standard deviation of the AR coefficient series from simulated and real signals, was established. The threshold procedure was applied to the four TVAR coefficient time series and only if the abrupt change in the four time series was above the threshold, the information was considered as produced by a crackle. Finally, the ability to detect and to estimate the number of crackles was assessed by the concordance between the total number of inserted crackles and the number of estimated crackles by the TVAR model.

D. Detection of Real Crackles by Expert.

Two schemes were used to compare TVAR model and the performance of one expert. The first one included simulated abnormal LS where scale factor, number and temporal position of crackles were unknown by the expert; however, the scale factor was 3.5, and the number of inserted crackles was 6 for PRC2 and 10 for PLX4, for inspiratory phases. On the second one, two LS recordings with real crackles, from patients diagnosed with extrinsic allergic alveolitis were used. In both schemes, crackles detection by the expert was based on TEWA criteria [3, 4] in conjunction with expert experience to identify crackles.

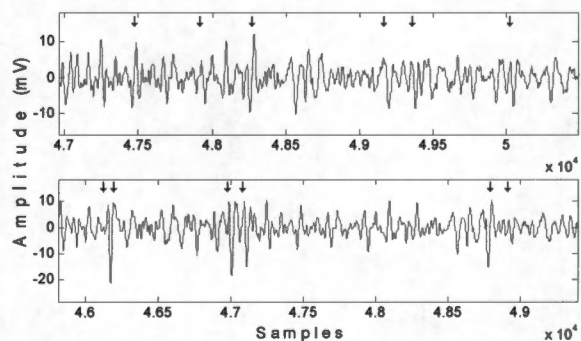


Fig. 2. Embedded fine crackles where arrows point out crackles position at scale factors of (a) 1.5 and (b) of 3.5.

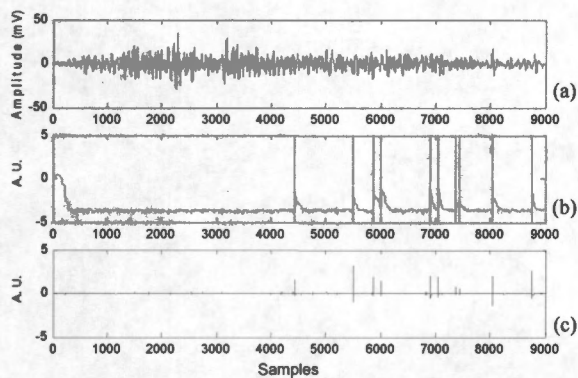


Fig. 3. Immersed fine crackles. (a) LS inspiratory phase signal, (b) TVAR coefficients behavior in Arbitrary Units (A.U.), (c) Derivative of the TVAR coefficients.

III. RESULTS AND DISCUSSION

A. Influence of Crackles to LS Ratio (SNR)

In fig. 2, embedded fine crackles with scale factors of 1.5 and 3.5 are shown. It can be observed that the visual detection of crackle position is relatively easy with a factor of 3.5, but with the factor of 1.5, some of the crackles are imperceptible and they even do not achieve the TEWA criteria.

B. Detection of Crackles by TVAR Model

In fig. 3(a) a signal with 10 fine crackles at scale factor of 1.5 is depicted. In fig. 3(b) the behavior of the third TVAR coefficient is displayed where sudden jumps are possible to observe as time evolving, remarked with dashed vertical lines. Fig. 3(c) shows the derivative of the third TVAR coefficient time series where 10 jumps are visually detected. Note that particularly jumps 7 and 8 correspond to overlapped crackles.

For the case of coarse crackles with a factor of 1.5, the algorithm also detected changes related with the 10 inserted crackles, as shown in fig. 4. Also, note that crackles 1-2, and

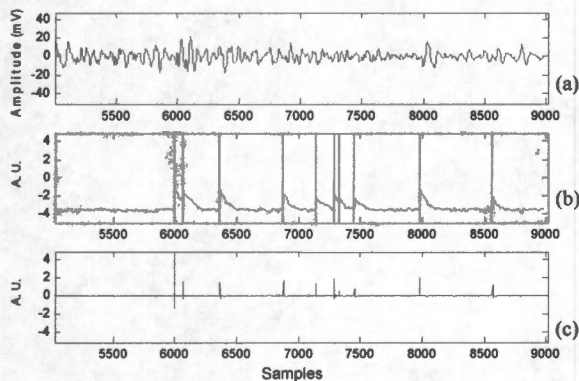


Fig. 4. Immersed coarse crackles. (a) LS inspiratory phase signal, (b) TVAR coefficients behavior, (c) Derivative of the TVAR coefficients.

6-7 are overlapped, and that crackle number 7 is of difficult visualization.

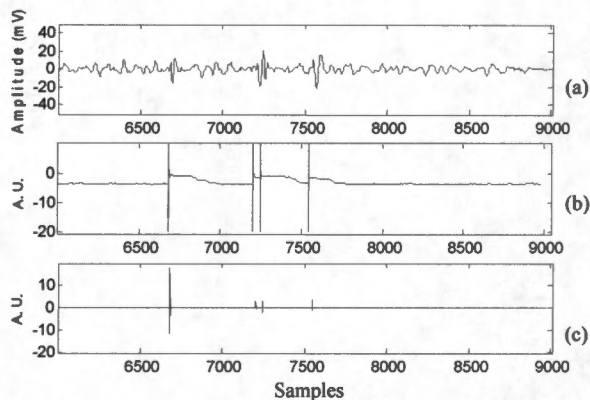


Fig. 5. Immersed fine and coarse crackles. (a) LS inspiratory phase signal, (b) TVAR coefficients behavior, (c) Derivative of the TVAR coefficients.

C. Detection of combined fine and coarse crackles

Fig. 5 shows an example of the behavior of the TVAR model when a combination of two fine and two coarse crackles are embedded in PLX4 with scale factor of 3.5. The model was tested for crackles inserted in PRC2 and PLX4 signals, and four crackles were detected for PRC2 and five for PLX4. These results remark the effect of the LS characteristics where crackles are inserted. Note again that in this example, crackles number 3 and 4 are overlapped.

TVAR Model vs. Expert Performance.

Table I resume the results derived from the two schemes for comparing the capacity for crackles detection between TVAR model and the expert. From the first scheme that used a scale factor of 3.5, data shows that TVAR model had non significant differences ($p > 0.05$) with the number of inserted crackles while expert detection had a poor performance, with significant differences ($p < 0.05$). On the other scheme, with real crackles, the number of detected crackles was similar ($p > 0.05$) between TVAR model and the expert. Although the SNR was not measured in real cases, we estimate that SNR was much higher than the one produced by the scale factor 3.5. Fig. 6 presents an example

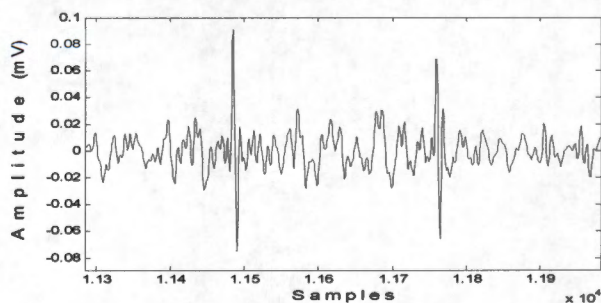


Fig. 6. Acquired LS signal with two evident crackles.

of one of the used real cases where the crackle sounds are evident. Therefore, we assume that if the scale factor is even lower than 3.5 then experts would have more limitations to detect crackles applying TEWA. The reason of this problem may be due to crackle waveform distortion by the background LS, and the temporal crackles overlapping. It seems that TVAR model has a good performance to deal with temporal overlapping and possible waveform changes.

TABLE I.- CAPACITY FOR CRACKLES DETECTION BETWEEN TVAR MODEL AND THE EXPERT

SCHEME A (Simulation)			
Sensor position	Number of inserted crackles	Estimated crackles TVAR	Estimated crackles Expert
PRC2 ($n=5$ Insp)	6	5.6 ± 0.6	2.0 ± 1.8*
PLX4 ($n=5$ Insp)	10	10.2 ± 1.1	2.7 ± 1.9*

SCHEME B (Real)			
		Estimated crackles TVAR mean±s.d.	Estimated crackles Expert mean±s.d.
CASE 1 ($n=7$ Insp)	-	5.7 ± 1.1	6.0 ± 1.0
CASE 2 ($n=6$ Insp)	-	3.2 ± 1.9	3.0 ± 1.7

* Significant difference, $p < 0.05$, compared with number of inserted crackles. n stands for number of inspiratory phases analyzed.

IV. CONCLUSION

We concluded that the time-variant autoregressive model is a proper alternative to detect and to provide an estimate number of fine and coarse crackles, even in presence of crackles overlapping and crackles with amplitude equivalent to 1.5 the standard deviation of the background lung sound. Under these conditions, crackles detection based on time-expanded waveform-analysis reveals evident limitations.

REFERENCES

- [1] P. Piirila, A.R. Sovijärvi, T. Kaisla, H.M. Rajala, T. Katila, "Crackles in patients with fibrosing alveolitis, bronchiectasis, COPD, and heart failure", *Chest*, vol. 99, pp. 1076-1083, 1991.
- [2] P. Piirila, A.R.A. Sovijärvi, "Crackles: recording, analysis and clinical significance", *Eur. Respir. J.*, vol. 8, pp. 2139-2148, 1995.
- [3] R.L. Murphy, S. K. Holford, W. C. Knowler, "Visual lung sound characterization by time-expanded waveform analysis", *N. England. J. Med.*, vol. 296, no. 17, pp. 968-971, 1977.
- [4] J. Hoevers, R. Loudon, "Measuring crackles", *Chest*, Vol. 98, pp. 1240-1243, 1990.
- [5] N. al Jarad, S.W. Davies, R. Logan-Sinclair, R.M. Rudd, "Lung crackle characteristics in patients with asbestosis, asbestos-related pleural disease and left ventricular failure using a time-expanded waveform analysis--a comparative study", *Respir. Med.*, vol. 88, pp. 37-46, 1994.
- [6] M. Mori, K. Kinoshita, H. Morinari, T. Shiraishi, S. Koike, S. Murao, "Waveform and spectral analysis of crackles", *Thorax*, Vol. 35, pp. 843-850, 1980.
- [7] M. Munakata, H. Ukita, I. Doi, Y. Ohtsuka, Y. Masaki, Y. Homma, Y. Kawakami, "Spectral and waveform characteristics of fine and coarse crackles", *Thorax*, Vol. 46, pp. 651-757, 1991.

- [8] M. Ono, K. Arakawa, M. Mori, T. Sugimoto, H. Harashima, "Separation of Fine Crackles from Vesicular Sounds by a Nonlinear Digital Filter", *IEEE T. Biomed. Eng.*, vol. 36, no. 2, pp. 286-291, 1989.
- [9] T.K. Kaisla, A. Sovijärvi, P. Piirila, H.M. Rajala, S. Haltsonen, and T. Rosqvist, "Validated method for automatic detection of lung sounds crackles", *Med. Biol. Eng. Comput.*, vol. 29, pp. 517-521, 1991.
- [10] L. J. Hadjileontiadis and S. M. Panas, "Nonlinear separation of crackles and squawks from vesicular sounds using third-order statistics", *Proc. IEEE-EMBS 18th Conf.*, vol. 5, pp. 2217-2219, Amsterdam, The Netherlands, 1996.
- [11] L. J. Hadjileontiadis and S. M. Panas, "Separation of discontinuous adventitious sounds from vesicular sounds using a wavelet-based filter", *IEEE T. Biomed. Eng.*, vol. 44, no. 12, pp. 1269-1281, 1997.
- [12] Y. A. Talias, L. J. Hadjileontiadis, and S. M. Panas, "Real-time separation of discontinuous adventitious sounds from vesicular sounds using a fuzzy rule-based filter", *IEEE T. Inform. Technol. Biomed.*, vol. 2, no. 3, pp. 204-215, 1998.
- [13] L. J. Hadjileontiadis, I. T. Rekanos, "Detection of explosive lung and bowel sounds by means of fractal dimension", *IEEE Signal Proc. L.*, vol. 10, no. 10, pp. 311-314, 2003.
- [14] S. Charleston, R. González, G. Chi-Lem, T. Aljama, "Crackle sounds analysis by empirical mode decomposition", *Proc. Int. Conf. IEEE-EMBS*, pp. 40-47, 2007.
- [15] H. Kiyokawa, M. Greenberg, K. Shirota, and H. Pasterkamp, "Auditory detection of simulated crackles in breath sounds", *Chest*, vol. 119, no. 6, pp. 1886-1892, 2001.
- [16] Proakis G. J., Manolakis G. D., *Digital Signal Processing, Principles, Algorithms and Applications*, pp. 853-860, Prentice Hall, 3a ed., 1996.
- [17] S. Haykin, *Adaptive filter theory*, Prentice Hall, Upper Saddle River, N J 07458, 1996.

Imaging of Simulated Crackle Sounds Distribution on the Chest

G. Dorantes-Mendez, S. Charleston-Villalobos, R. González-Camarena, G. Chi-Lem, T. Aljama-Corrales

Abstract—Crackles sounds have been associated with several pulmonary pathologies and diverse algorithms have been proposed for extracting and counting them from the acquired lung sound. These tasks depend among other factors, of the relation between the magnitude of the crackle and the background lung sound. In this work, we explore multivariate signal processing to deal with the tasks and propose a new concept, the discontinuous adventitious sounds imaging. The image formation is founded on the results of two proposed methodologies that use an autoregressive (AR) model. In the first case, the AR coefficients feed an artificial neural network (ANN) to classify temporal acoustic information as healthy or sick and; in the second case, a time-variant AR (TVAR) model, obtained by the RLS algorithm, permits to detect changes in the TVAR coefficients to be associated with the number of crackles. For AR-ANN, the ratio of the temporal windows classified as sick to the classified as healthy is used as an index to form the adventitious image, while for TVAR-RLS, an estimation of the number of crackles is obtained to form the corresponding image. The results indicated that fine and coarse crackles could be detected and counted even with very low crackle magnitude so that the formation of a crackle distribution image was consistent.

I. INTRODUCTION

It is recognized that crackles sounds are adventitious respiratory sounds relevant for the diagnosis of different pulmonary pathologies [1] and several attempts have been proposed to detect and to count them. For example, Murphy *et al.* proposed the time expanded waveform analysis (TEWA) to detect crackles [2], but the method requires certain criteria that are difficult to achieve in most of the acquired lung sounds (LS) [3]. Extracting and counting crackles, fine and coarse, represent a difficult task since it depends, among other factors, on the amplitude relation between crackles and the background LS, and of the temporal overlapping of two or more crackles.

The majority of the proposed methodologies for crackle

G. Dorantes Méndez is a Master student of the Biomedical Engineering Program, Universidad Autónoma Metropolitana, Mexico City 09340, Mexico (e-mail: lupita.dorantes@gmail.com).

S. Charleston-Villalobos is with the Electrical Engineering Department, Universidad Autónoma Metropolitana, Mexico City 09340, Mexico (email: schv@xanum.uam.mx).

R. González-Camarena is with the Health Science Department, Universidad Autónoma Metropolitana, Mexico City 09340, Mexico (email: rgc@xanum.uam.mx).

G. Chi-Lem is with the National Institute of Respiratory Diseases, Mexico City, Mexico (email: gchilem@iner.gob.mx)

T. Aljama-Corrales is with the Electrical Engineering Department, Universidad Autónoma Metropolitana, Mexico City 09340, Mexico (corresponding author email: alja@xanum.uam.mx).

detection test them on databases that include acquired LS where physicians have identified the crackles using TEWA [2]. However, it has been demonstrated that these kinds of sounds are difficult to detect for them [4]. Hence, we consider that a previous important stage is the use of simulated signals to evaluate the performance of the proposed methodology where the magnitude, number and position of crackles, temporally and spatially, need to be controlled [3].

Image formation is of great interest in different fields, anatomic or functional modalities, and for the study of LS, the application is recent [5, 6]. In this work a new concept is introduced, the imaging of crackle sounds for the assistance of the physicians since the image would indicate the location, distribution or the approximate number of crackles. The proposed methodology is tested using multichannel simulated adventitious sounds and multivariate signal processing.

II. METHODOLOGY

A. Lung Sounds Multichannel Acquisition

LS from 7 healthy subjects were acquired using a microphone array of 5 by 5 sensors attached to the back of the subjects, as is shown in fig. 1.(a). The sensors nomenclature was described elsewhere [3, 5] and the distance between sensors was 5 cm, vertically and horizontally. The sensors included subminiature electret microphones air coupled to the thoracic surface by plastic bells, and with frequency response up to 3.0 kHz. The LS and airflow were acquired by A/D of 12 bits with a sampling frequency of 10 kHz; afterwards, LS were filtered keeping the frequency band from 75 Hz to 1 kHz. During

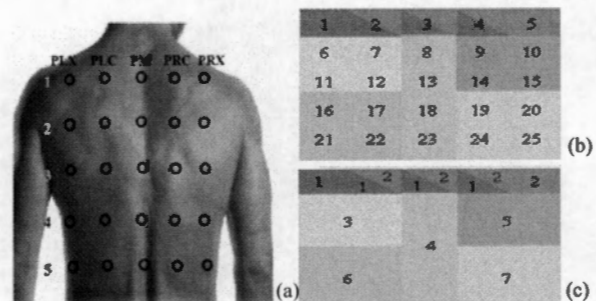


Fig. 1. (a) Microphone array, (b) Sensor number and regions definition (colors) according to the microphone array for ANN training phase by regions. For example, region 5 in (c) includes the information in channels 9, 10, 14 and 15.

acquisition, the subjects were seated, breathing through a Fleish pneumotachometer at 1.5 L/s, and wearing nose-clips, in a quiet room.

B. Simulated Abnormal Multichannel RS Signals

To simulate multichannel LS signals, synthesized fine and coarse crackles were inserted in multichannel LS signals acquired from healthy subjects, using an additive model, from the middle to the end of each inspiratory phase in a random fashion, allowing temporal crackles overlapping. To simulate fine and coarse crackles, the mathematical model proposed by Kiyokawa *et al.* [4] was implemented:

$$y(t) = \sin(4\pi^\alpha), \quad \text{con} \quad \alpha = \frac{\log(0.25)}{\log(t_0)} \quad (1)$$

where $y(t)$ is conformed by two cycles with the first positive interception at t_0 . Besides, a modulation function given by:

$$m(t) = 0.5 \{ 1 + \cos[2\pi(t^{0.5} - 0.5)] \} \quad (2)$$

is used to concentrate the main power of $y(t)$ at the beginning. The mathematical model reproduce the temporal characteristics of fine and coarse crackles described by Murphy *et al.*, i.e., for fine crackles IDW is 0.5 ms and 2CD is 5ms, while for coarse crackles IDW is 1.2 ms whereas 2CD is 9 ms [2], as shown in fig. 2.

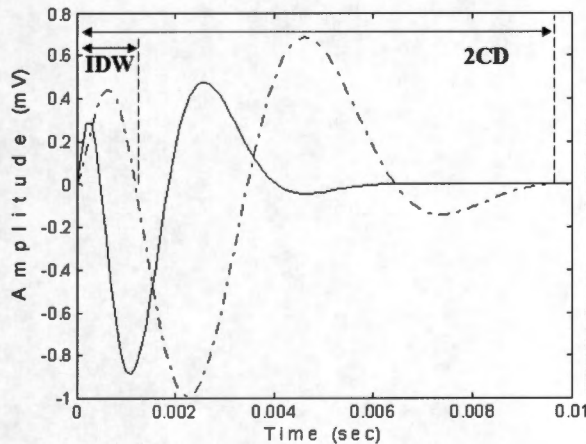


Fig. 2. Fine (continuous line) and coarse (broken line) crackles and their temporal characteristics.

The magnitudes of simulated crackles were 1.5, 2.5 and 3.5 times the variance of the LS surrounding the temporal crackle position. Crackles were inserted at a certain number of channels to provide simulated sick pulmonary zones called a pattern. In fig.3 an example of simulated fine crackles inserted into the base LS is shown, the scaling factor is the biggest used to synthesize the adventitious spatial patterns, that means 3.5 times the variance of the surrounding breathing sound. Even with this scaling factor,

the crackle sounds were difficult to define by TEWA.

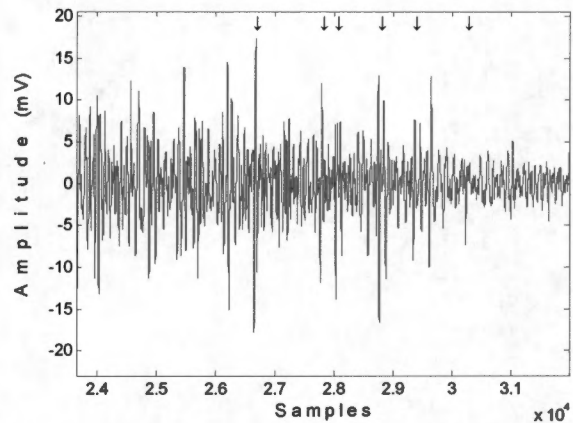


Fig 3. Simulated fine crackles inserted at the end of an inspiratory phase (scale factor 3.5), the crackles are pointed out by arrows.

C. Signal Processing by AR-ANN

To classify acoustical information, and considering our previous efforts [7] where an ANN provided good results for normal and abnormal LS classification, in the present work an ANN was used with a supervised training algorithm, including a tangential activation function and the gradient-descendent Levenberg-Marquardt method for adjusting its weights [8]. The feature vector, input to the ANN, was formed segmenting each inspiratory phase in 30 overlapped windows by 25 %. The AR model of order six, according to the Akaike criterion, for fine and coarse crackles was obtained from each temporal window. Each temporal window was labeled as healthy or ill according to the crackle's position and duration. In the training phase the temporal window was labeled as healthy if the ANN output was higher or equal to 0.8, otherwise, if ANN output is less than or equal to -0.8, the information was labeled as ill.

Two different training schemes were used in this work, by regions and by channels. By regions, the array of 25 sensors was divided in seven regions with two overlapped regions; sensors 1 to 4 shaped region 1 and sensors 2 to 5 shaped region 2, as shown in figure 1(b) and 1(c). In this scheme, an ANN with three layers was used; the size of the first layer was six AR coefficients by 4 channels, in each region. The second layer had 10 neurons selected and the last layer included one neuron. The ANN structure was selected by experimentation. The training phase included the acoustical information from 28 simulated patients generated from the acoustical information of 7 healthy subjects.

On the other hand, in the training scheme by channels, the ANN topology was modified only in the input layer since the feature vector is of dimension six. Due to computational facilities, the training phase was achieved with the simulated information provided by 20 simulated patients and 5 healthy subjects.

In both schemes, the database was divided in training data (80% of database), validation data (10% of database) and testing data (10% of database).

The proposed index to be mapped, using a Hermite interpolating function and showing the presence of crackle sounds, was the ratio:

$$\text{Index} = \frac{\% \text{ windows classified as ill}}{\% \text{ windows classified as healthy}} \quad (3)$$

D. Signal Processing by TVAR-RLS

The TVAR coefficients were estimated using a prediction adaptive filtering scheme by the recursive least squares (RLS) algorithm [9], with a forgetting factor of 0.97. The model order was four defined by the Akaike criterion. The estimated number of crackles was associated with abrupt changes in the derivative of the TVAR coefficient time series. A threshold, based on 0.035 times the standard deviation of the AR coefficient series, was established. The threshold procedure was applied to the four TVAR coefficient time series and only if the abrupt change in the four time series was above the threshold, the information was considered as produced by a crackle. The number of crackles was estimated for each of 25 sensors of information. With the estimated number, and using a Hermite interpolating function, an image was generated.

III. RESULTS

A. Imaging by AR-ANN

To test the method AR-ANN by regions, for instance crackle sounds were inserted at the right hemithorax, eight and ten fine crackles were inserted in regions five and seven, respectively. The synthesized adventitious spatial pattern is shown in fig. 4(a), where the letters L and R stand for the left and right side of subject if a viewer sees the subject by the back. The same visualization condition held for the rest of the figures. In figs. 4(b)-(d) the three images obtained by the method are shown, with 1.5, 2.5 and 3.5 scale factors,

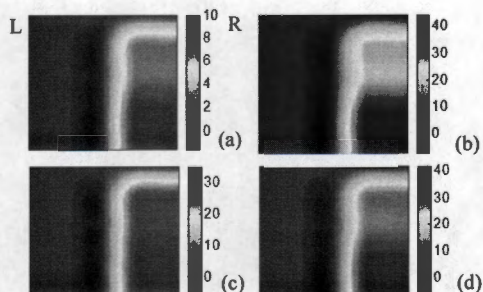


Fig. 4. Images by AR-ANN with training by regions. (a) Pattern image with crackles inserted in regions five and seven, (b)-(d) resulting images for scale factors 1.5, 2.5 and 3.5, respectively.

respectively. In figs. 4(b)-(d), the blue color indicates the absent of crackles whereas the red is associated with the

highest index defined at (3). It is worthy to note that in the obtained images there are two well defined zones, and more important, they are associated with the crackle number used to synthesize the pattern image. Obviously, the image in fig. 4(d) is the closest to the pattern image in fig. 4(a) due to the biggest scale factor; however, even with the lowest scale factor it is possible to observe that regions five and seven contain different number of crackles, as shown in fig. 4(b).

An alternative simulated adventitious spatial pattern was used, fig. 5(a), in this case six fine crackles were inserted in region five, whereas ten crackles were inserted in region six. As can be seen, figs. 5(b)-(d), two "illness regions" are revealed by imaging the AR-ANN method results.

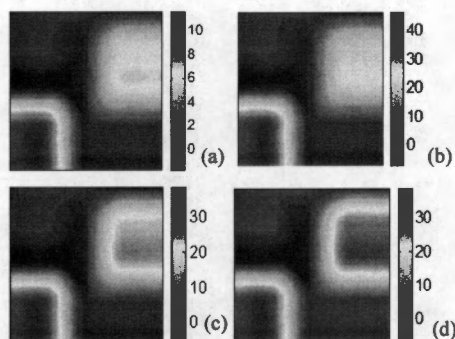


Fig. 5. Images AR-ANN by regions, (a) Pattern image with crackles inserted in regions five and six, (b)-(d) resulting images for scale factors 1.5, 2.5 and 3.5, respectively.

In order to tests the capability of the scheme to detect coarse crackles, the former spatial pattern was repeated but now inserting coarse crackles. Fig. 6 shows that this type of adventitious sounds can also be identified for the three scaling factors.

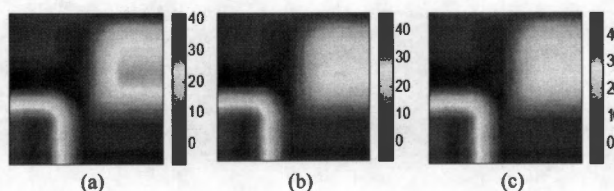


Fig. 6. Images AR-ANN by regions with coarse crackles, six in region five and ten in region six (pattern image in fig. 5(a)) for scale factors (a) 1.5, (b) 2.5 and (c) 3.5.

For the channels scheme, as an example the pattern used in fig. 4 was also used to test the AR-ANN method, figs. 7(a)-(c) show that two zones are still identified, the left hemithorax as the healthy, the right one as the sick one.

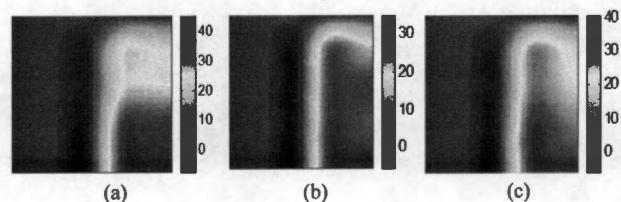


Fig. 7. Images AR-ANN by channels with fine crackles, eight and ten crackles in regions five and seven for scale factors (a) 1.5, (b) 2.5 and (c) 3.5, respectively.

B. Imaging by TVAR-RLS

The pattern of fig. 4, and repeated in fig. 8(a), was also used to probe the detection robustness of the TVAR-RLS method. In fig. 8(b), using a 1.5 scaling factor, the resulting image as well as the estimated crackle number can be observed. Even though the disagreement with the pattern, the proposed method turns out a good approximation to it.

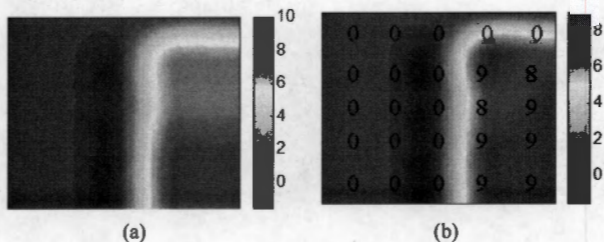


Fig. 8. (a) Spatial pattern with fine crackles for scale factor of 1.5 (b) Image TVAR-RLS and the estimated crackles number.

The TVAR-RLS, using the patterns of fig. 4(a) and fig. 5(a), was tested for the detection of coarse crackles, for the smallest scale factor. The results that are shown in fig. 9(a) and (b), reveal that for both cases there is a good estimation of the number of coarse crackles and consequently, the images obtained are in good agreement with the patterns.

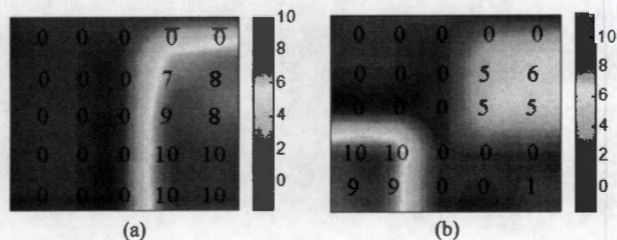


Fig. 9. Image TVAR-RLS for coarse crackles and the estimated crackles number for scale factor of 1.5.

IV. DISCUSSION AND CONCLUSIONS

The results show that fine and coarse crackles can be detected by both methodologies, AR-ANN and TVAR-RLS, for generating an image. The AR-ANN method intends to construct an adventitious image related to the presence of crackles using the classification performance of an ANN. On the other side, the TVAR-LS makes up the mapping from the estimated crackle number. As expected, the relation in magnitude of the crackle sounds to the lung sound impacts the goodness of the approximation to the spatial pattern, used to validate the robustness of the proposed methods. Even though the smallest scaling factor we used, it was possible to identify the pulmonary zones where the crackles were inserted. It is worthy to say that in clinical practice the physician, applying temporal criteria, identify crackles making use of signals as those generated with a scaling factor of 3.5 or higher; therefore, it seems that our methods are more sensible to detect crackles in low SNR. In the case of AR-ANN the results are better for regions than

for the channels scheme, nevertheless the results could be biased by the fact that the same number of crackles were inserted at the four channels to create a region. However, it is possible that the AR-ANN by regions could be useful at some researches where it is important to take into account the crackle transmission phenomenon. On the other hand, the TVAR-RLS results show that the proposed index allows to obtain the crackle distribution image in a good fashion, with the advantage that no training phase is needed and the mapping is generated directly from the estimated crackle number. Although, the coarse crackle morphology for the scale factor of 1.5 was more difficult to visualize from the base lung sound, both proposed methodologies were capable to produce good results.

It should be considered that one of the fundamental assumptions in this work was that adventitious sounds are added to base normal lung sounds but in patients, the base lung sounds could be altered. We emphasize that the intention of the new concept established in this work, the imaging of discontinuous adventitious sounds, is to provide some assistance to the physician for the crackle detection and counting task. In consequence, the statistical comparison regarding the performance of both schemes, AR-ANN and TVAR-RLS, is not reported in this work.

Finally, both methods AR-ANN and TVAR-RLS allowed the imaging of the crackles distribution on the chest in a robust and confident way. The concept of discontinuous adventitious sounds imaging established in this work is currently explored with real data in a clinical setting.

REFERENCES

- [1] H. Pasterkamp, S. S. Kraman, and G. R. Wodicka, "Respiratory sounds: advances beyond the stethoscope," *Am. J. Respir. Crit. Care Med.*, vol. 156, pp. 974-987, 1997.
- [2] Murphy, R. L., S. K. Holford and W. C. Knowler. "Visual lung sound characterization by time-expanded wave-form analysis", *N. Engl. J. Med.*, 296, pp. 968-971, 1977.
- [3] S. Charleston-Villalobos, R. González-Camarena, G. Chi-Lem, T. Aljama-Corrales, "Crackle sounds analysis by empirical mode decomposition", *IEEE Eng. Med. Biol. Mag.*, 26, pp. 40-47, 2007.
- [4] H. Kiyokawa, M. Greenberg, K. Shirota, and H. Pasterkamp, "Auditory detection of simulated crackles in breath sounds," *Chest*, vol. 119, pp. 1886-1892, 2001.
- [5] S. Charleston-Villalobos, S. Cortés-Rubiano, R. González-Camarena, G. Chi-Lem, T. Aljama-Corrales, 'Respiratory acoustic thoracic imaging (RATHI): assessing deterministic interpolation techniques', *Medical & Biological Engineering & Computing*, vol. 42, pp.618-626, Sept 2004.
- [6] R.P. Dellinger, J.E. Parrillo, A. Kushnir, M. Rossi, I. Kushnir, "Dynamic visualization of lung sounds with a vibration response device: A case series", *Respiration*, vol. 75, pp. 60-72, 2008.
- [7] H.G. Martínez-Hernández, T. Aljama-Corrales, R. González-Camarena, S. Charleston-Villalobos, G. Chi-Lem, "Computerized classification of normal and abnormal lung sounds by multivariate linear autoregressive model", *Proc. 27th Ann. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 1464-1467, Shanghai, China 2005.
- [8] O. R. Duda, E. P. Hart, and G. D. Store, *Pattern Classification*, Wiley, pp. 284-303, 2001.
- [9] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, pp. 562-587, 1998.