

Análisis y Clasificación de Sonidos de Motores de Carros, para Diagnosticar Fallas, con Ayuda de Redes Neuronales

Luis Andrés Méndez Ortega¹ y Orion Fausto Reyes Galaviz²

^{1,2}Universidad Autónoma de Tlaxcala

Facultad de Ciencias Básicas, Ingeniería y Tecnología

Apizaco, Tlaxcala. México

siul222@hotmail.com, <http://ingenieria.uatx.mx/>

²Instituto Tecnológico de Apizaco

orionfrg@itapizaco.edu.mx, <http://www.itapizaco.edu.mx/>

Resumen La detección de posibles fallas en un motor de automóvil, pueden ser capturadas a través de la identificación de los cambios en la vibración, los cuales son producidos mientras el motor está trabajando. Una persona con experiencia mecánica, es capaz de reconocer un desperfecto en la máquina, solamente al escuchar el ruido y percibiendo las vibraciones producidas. Para un propietario de un vehículo, es fácil detectar cualquier cambio, o si hay un ruido anormal en el motor y, consecuentemente, detectar alguna falla. Aún así, es difícil identificar qué tipo de falla produce ciertos ruidos y vibraciones en el motor, dificultando el proceso de trabajar directamente en las fallas del mismo. En este trabajo se propone un método para clasificar y detectar, directamente, algunas fallas comunes en los motores, usando solo el sonido grabado de la máquina. Para lograr este objetivo, se usaron Coeficientes Cepstrales de Predicción Lineal (LPCC) como el extractor principal de características acústicas, así como Análisis de Componentes Principales y una Red Neuronal con Retraso en el Tiempo para el reconocimiento de patrones. Se obtuvieron resultados de reconocimiento del 95.58 % cuando se diagnostican 3 fallas comunes en el motor.

1. Introducción

Si se ignora cualquier falla en alguna máquina, se pueden producir vibraciones suaves o violentas que pueden dañarla. Estas vibraciones pueden causar fisuras, desgaste y/o sobrecalentamiento de partes importantes y reducir el rendimiento del motor. Además, las vibraciones son un buen indicador del funcionamiento mecánico y son también sensibles a la evolución de la anomalía. En general estos tipos de desperfectos son precedidos por cambios en el motor, tales como: sonidos, vibraciones, pérdida de potencia, etc. Estos indicadores son una señal de algún tipo de falla futura en el desempeño de la máquina. En general, un oído acostumbrado detecta rápidamente la diferencia existente entre los ruidos producidos que podríamos llamar normales y una serie de ruidos anormales,

tales como silbidos, golpeteos en la máquina, el castaño de los engranes que vibran, etc. Si este tipo de información relevante existe oculta dentro del ruido de la máquina, la extracción, reconocimiento y clasificación de las características acústicas podría ser posible, usando mecanismos automáticos, y de esta manera obtener un diagnóstico confiable. El análisis presentado en este trabajo unido con los sistemas de monitoreo que existen actualmente proporcionarían más herramientas para ayudar a los especialistas de motores a obtener un diagnóstico robusto cuando se reparen motores de vehículos. Tomando esto en cuenta, se ha propuesto capturar y analizar los sonidos producidos por este tipo de maquinaria y el uso de técnicas de extracción de características acústicas y reconocimiento de patrones para obtener una clasificación de los diferentes tipos de ruidos que son producidos por motores en buen y mal estado.

En la siguiente sección se presentará una revisión de estudios anteriores comparativos en este campo. En la Sección tres se detallan las bases fundamentales en el proceso de reconocimiento de ruido y se describe el sistema propuesto. La sección cuatro trata del procesamiento acústico y el método de extracción de características implementado, el cual fue Coeficientes Cepstrales de Predicción Lineal. Una teoría fundamental en la clasificación de patrones, Análisis de Componentes Principales y Redes Neuronales con Retraso en el Tiempo, son revisadas en la sección seis. Nuestros resultados experimentales que alcanzaron un 98.58 % de reconocimiento, son mostrados en la sección siete, y las conclusiones se presentan en la sección ocho.

2. Estado del Arte

Se han propuestos varios sistemas en el campo de reconocimiento de patrones acústicos desde hace pocos años. Dependiendo del problema que se ha intentado resolver, un punto en común de estos enfoques se centra en un análisis espectral detallado de la señal acústica. En [1], Mario E. Munich usa los Coeficientes Cepstrales de Frecuencia Mel para el análisis acústico, modelos de Mezcla Gaussiana (GMM, por sus siglas en inglés) y modelos ocultos de Markov (HMM), combinados con Métodos de Subespacio Bayesiano aplicado al reconocimiento automático de características acústicas de vehículos: esto es usado en operaciones militares para propósitos de vigilancia, registrando una precisión del 83 %. Huadong Wu [2] y sus colegas, también trabajaron en el campo de reconocimiento de vehículos. Propusieron que cada tipo de modelo de vehículo emite el mismo tipo de ruidos, vibraciones, fricción con el viento, etc. Ellos usaron un método llamado eigenfaces, usado comúnmente en el campo de reconocimiento de rostros, esta metodología se usa para caracterizar los patrones de ruido y usarlos para reconocer el vehículo; este método es también conocido como una extensión de Korhunen-Loeve o como Análisis de Componentes Principales. Los resultados de este trabajo fueron prometedores. En [3], Edgar A. Estupiñán y Pedro N. usaron análisis de vibraciones mecánicas, como parte de su mantenimiento predictivo para establecer un estado mecánico de sus máquinas, previniendo con esto, problemas futuros. Propusieron técnicas basadas en la Transformada Rápida de

Fourier para analizar el ruido producido por la vibración de máquinas rotatorias de baja velocidad. Pedro N. Saavedra [3], [4], presenta algunas técnicas aplicadas al análisis de vibraciones, particularmente se analizó la detección de grietas en ejes y vigas. Para el estudio teórico, fue usada la metodología de Elementos Finitos y las grietas fueron modeladas usando la Teoría de la Fractomecánica Lineal. Es difícil comparar estos sistemas; porque las bases de datos, condiciones experimentales, tamaño de segmentos y tipo de reconocimiento son diferentes. A pesar de esto, el análisis de diferentes técnicas sugieren interesantes posibilidades para combinarlas, para que de esta manera, se puedan obtener mejores resultados en el reconocimiento.

3. Proceso de Reconocimiento del Ruido de Motor de Automóvil

El proceso de reconocimiento de ruido de motor del vehículo, es básicamente un problema de reconocimiento de patrones y esto es similar en el reconocimiento del habla. El objetivo es tomar unas señales sonoras del motor como entrada, y al final reconocer la falla en el motor. Generalmente este proceso se hace en dos pasos; el primero es el procesamiento acústico o extracción de características, mientras que el segundo es conocido como procesamiento o clasificación de patrones. El sistema propuesto puede verse en la Figura 1. Para este caso, en el análisis acústico, la señal del motor es procesada para extraer las características relevantes en función del tiempo. El conjunto de características obtenidas de cada muestra de ruido es representada por un vector y cada vector es tomado como un patrón. Con respecto al reconocimiento de patrones, cuatro principales enfoques han sido tradicionalmente usados: comparación de patrones, modelos estadísticos, sistemas basados en conocimiento y modelos conexionistas. Nos enfocamos al uso de estos últimos.

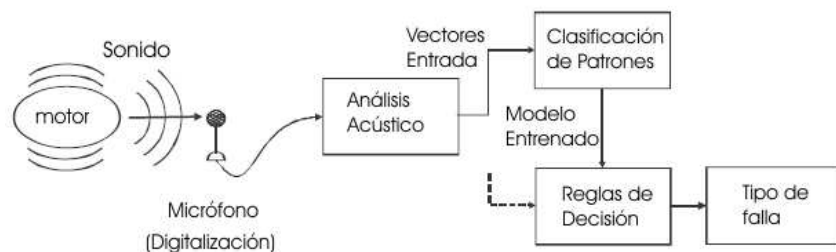


Figura 1. Proceso de reconocimiento del Ruido de Motor.

4. Procesamiento Acústico

Una señal acústica de motor es muy compleja y se codifica mas información que la necesaria para ser analizada y procesada en aplicaciones en tiempo real.

Por esta razón, en el proceso de reconocimiento de ruido de motor, se usó una función de extracción de características como procesador de primer plano. Su entrada es la señal acústica de motor y su salida es un vector de características que representa elementos clave de la onda sonora. En este trabajo se usaron los Coeficientes Cepstrales de Predicción Lineal (LPCC) [5], como método de extracción de características.

4.1. Coeficientes Cepstrales de Predicción Lineal

El modo de predicción lineal (LP), es históricamente uno de los métodos más importantes usados para el análisis de la voz [5]. Su base fundamental es la de establecer un modelo filtrado para la fuente de sonido. Con el suficiente número de parámetros el modelo de LP puede establecer una buena aproximación a la estructura espectral de cualquier tipo de sonido. Es por eso que usamos esta metodología para analizar nuestras muestras acústicas. El método de LP recibe este nombre por que pretende extrapolar el valor de la siguiente muestra de sonido $x(n)$ como la suma ponderada de las muestras previas $x(n-1), x(n-2), \dots, x(n-k)$:

$$x(n) = \sum_{i=1}^K a_i x(n-1) \quad (1)$$

Para ello se debe realizar el cálculo de los coeficientes, minimizando con alguna función de error E , concretamente de mínimos cuadrados sobre una ventana de tamaño N .

$$\begin{aligned} E &= \sum_{n=0}^{N-1} e^2(n) \\ &= \sum_{n=0}^{N-1} \left(x(n) - \sum_{i=1}^K a_i x(n-1) \right)^2; \quad 0 \leq n \leq N-1 \end{aligned} \quad (2)$$

Partiendo del Análisis de Predicción Lineal, es posible obtener la expresión de los coeficientes cepstrales asociados (LPCC):

$$c(0) = \log(1) = 0 \quad (3)$$

$$c(i) = -a(i) - \sum_{j=1}^{i-1} \left(1 - \frac{j}{i} \right) a(j) c(i-j); \quad 1 \leq i \leq N_c \quad (4)$$

Una transformación usual sobre este tipo de coeficientes es lo que se conoce como coeficientes cepstrales delta o coeficientes delta cepstrum. La expresión que permite obtener estos últimos es:

$$\Delta c_j(i) = \frac{1}{2T+1} \sum_{k=-T}^T k \cdot c_{j+k}(i) \quad (5)$$

5. Clasificación de las Muestras de Ruido de Motor

Después de extraer las características acústicas de cada muestra de ruido, los vectores de características son obtenidos; cada uno de estos vectores representa un patron. Estos vectores son usados después para el proceso de clasificación. Este trabajo, se concentró en los modelos conexionistas, mejor conocidos como Redes Neuronales Artificiales (RNA), para clasificar estos vectores(patrones). La RNA fue combinada con Análisis de Componentes Principales para reducir la dimensionalidad del vector y de esta manera mejorar el tiempo de procesamiento de entrenamiento/prueba, obteniendo con esto, un sistema de clasificación más eficiente.

5.1. Análisis de Componentes Principales

El objetivo del ACP es reducir la dimensionalidad de un conjunto de variables p a un conjunto m con menos variables para mejorar la legibilidad de los datos. El conjunto de variables o características originales, usualmente son redundantes, por esta razón es útil reducir la dimensionalidad de los datos originales sin la perdida de información relevante. El ACP intenta encontrar los componentes que describen exitosamente la mayoría de la varianza total; los datos que tienen una alta varianza son conservados, eliminando los componentes que contribuyen a la perdida de varianza del conjunto de datos. Un conjunto reducido es más fácil de analizar e interpretar. Esta reducción tiene importantes beneficios: primeramente, el costo computacional de las etapas de procesamiento subsecuentes es reducido, y segundo, en una proyección dentro de un subespacio de pocas dimensiones es más fácil de visualizar los datos [6].

5.2. Redes Neuronales Artificiales

Las Redes Neuronales Artificiales (RNA) son ampliamente usadas en tareas de clasificación de patrones, mostrando buen desempeño y resultados de alta precisión. En general, una RNA esta representada por un conjunto de nodos y conexiones(pesos). Los nodos son una simple representación de una Red Neuronal Natural, mientras las conexiones representan el flujo de los datos entre neuronas. Estos pesos son dinámicamente actualizados durante el entrenamiento de la RNA. En este trabajo, usamos el modelo de la Red Neuronal Feed Forward con Retrazo en el Tiempo(TDNN), ya que este modelo ha mostrado buenos resultados en tareas de reconocimiento de voz[7] y no-voz [6][8].

5.3. Red Neuronal Feed Forward con Retraso en el Tiempo

El retraso de entrada se refiere a un retraso en el tiempo, en otras palabras, si retrasamos la señal de entrada una unidad de tiempo y permitimos a la red neuronal recibir tanto la señal original y la retrasada, obtendremos una simple red neuronal con retraso en el tiempo. Esta red neuronal fue desarrollada para clasificar fonemas en 1987 por Weibel y Hanazawa [9]. La Red Neuronal Feed-Forward con retraso en el tiempo no fluctúa con los cambios; las características dentro de la señal de sonido pueden ser afectadas sin importar en que posición se encuentre. Los retrasos en el tiempo permiten a la red neuronal encontrar una relación temporal directamente en la señal de entrada y una representación más abstracta en las capas ocultas. Esto lo hace usando los mismos pesos para cada paso en el tiempo [7].

5.4. Entrenamiento Retropropagación por Gradiente Conjugado Escalado

El entrenamiento de la red neuronal se hace a través de una técnica conocida como Retropropagación (Backpropagation). El método de Gradiente Conjugado Escalado está basado en una estrategia de optimización general. Se ha utilizado Retropropagación por Gradiente Conjugado Escalado (SCGBP), para entrenar la Red Neuronal; ya que este algoritmo muestra una convergencia para decidir hasta donde ir en una dirección específica y evita el consumo de tiempo en búsquedas en línea para cada iteración, haciéndolo un algoritmo de segundo orden rápido [6].

6. Implementación del Sistema

En la mayoría de las poblaciones y ciudades en el territorio Mexicano, el uso de transporte público es muy usual. Un modelo de vehículo que es frecuentemente usado para este propósito es la *Combi Van*, de la Volkswagen, modelos '82, '83, '86, '88 y '90. Estas unidades trabajan la mayor parte del día, de las 6:00 am a las 9:00 pm, y hay algunas horas en las que llega a transportar aproximadamente hasta 16 personas, viajando varios kilómetros con este peso. Algunos talleres mecánicos, en ciudades donde estas unidades son típicamente usadas para el transporte público, se han llegado a especializar en el mantenimiento y reparación de este tipo de máquinas, dando a estos especialistas la habilidad para diagnosticar rápidamente la mayoría de las fallas comunes del motor, además de que trabajan directamente con los problemas de estos tipos de motores. Por estas razones se optó por este tipo de motores. Las condiciones elegidas para grabar un motor fueron:

- Modelos viejos ó unidades que han sido previamente reparadas.
- Características similares de motores.
- Modelos con el mismo motor.

Cuadro 1. Muestras grabadas, limpias y unidas por clase

Problema identificado	casos	Duración Unidas
Clase 1 (buen estado)	7	18.036 seg.
Clase 2 (cigüeñal)	6	18.502 seg.
Clase 3 (pistones)	6	17.664 seg.
Clase 4 (válvulas)	7	18.006 seg.
TOTALES	26	72.208 seg.

Se grabaron los sonidos de motor en dos ocasiones, cuando la unidad llegaba al taller mecánico y después de que fue diagnosticado y reparado. Este método ayudó a etiquetar las grabaciones y comparar el ruido producido por el motor antes y después de la reparación. Todo los ruidos de la máquina fueron grabados en dos fases: acelerando la máquina y en velocidad mínima (estado *Ralenti*), también algunos ruidos fueron grabados con ambas fases.

6.1. Muestras de Ruido

Para grabar el ruido de motor, algunos talleres fueron recomendados por conductores del transporte público, tres de ellos se especializan en este tipo de vehículos. Cada muestra grabada fue obtenida con ayuda de una grabadora digital ZicPlay y los archivos de salida fueron guardados en formato WAV, con las siguientes características: 8 khz, 16 bits y canal monoaural. Cada uno de estos archivos tiene un periodo de tiempo de 3 a 5 segundos, y fueron grabados en total 26 diferentes máquinas. Estas muestras fueron clasificadas en 4 categorías: motores en buenas condiciones (clase 1), problemas relacionados con el cigüeñal (clase 2), con pistones (clase 3) y con problemas en válvulas (clase 4), ver Cuadro 1. En el Cuadro 2, podemos ver el número de muestras obtenidas de motores en estado *Ralenti*.

Cuadro 2. Muestras grabadas, limpias y unidas por clase sin Aceleración

Problema identificado	casos	Duración Unidas
Clase 1 (buen estado)	7	16.396 seg.
Clase 2 (cigüeñal)	6	15.430 seg.
Clase 3 (pistones)	6	15.744 seg.
Clase 4 (válvulas)	7	16.590 seg.
TOTALES	26	64.208 seg.

Hay que tomar en cuenta que estas muestras fueron difíciles de obtener; por que en algunos casos, unas grabaciones fueron desechadas, debido a la pobre calidad de sonido, falla mínima y otras cuestiones que hicieron la muestra inutilizable para los objetivos del sistema. Se usaron sólo 3 fallas (clases), más otra

clase que representaba a los motores en buenas condiciones, ya que, de acuerdo a los expertos humanos (Mécánicos), estas fueron las fallas más comunes.

Los sonidos grabados de motores, obtenidos tanto en la aceleración y en estado *ralenti*, fueron segmentados en muestras de 1 segundo y etiquetados de la clase 1 a la 4; para cada muestra segmentada, se extrajeron diferentes números de LPCC (como se muestra en la tabla 3 y 4). En un experimento, los vectores fueron usados sin ninguna reducción, y en otro experimento, los vectores resultaron ser reducidos en un 52 % de su tamaño original usando ACP. Las configuraciones LPCC y reducción con ACP fueron obtenidas heurísticamente y los que se presentan en este trabajo, fueron los que dieron mejores resultados. Los vectores reducidos fueron usados para entrenar y probar la Red Neuronal con Retraso en el Tiempo; se usó un 80 % de las muestras segmentadas, para entrenar y un 20 % para probarla. La arquitectura de la Red Neuronal consiste de: una capa de entrada con el número de nodos correspondiente al tamaño del vector de entrada, una primera capa oculta con un tamaño correspondiente al 60 % de los nodos usados en la capa de entrada y una segunda capa oculta con el 40 % correspondiente al mismo tamaño del vector entrada, finalmente se encuentra la capa de salida, que le corresponde el número de clases a predecir en sus nodos, en este caso cuatro. La RNA fue implementada con la ayuda del *AuditoryToolbox* [10], para los vectores característica y junto con *NeuronalNetwork Toolbox*, integrado en Matlab v7, para la RN.

7. Resultados Experimentales

En el Cuadro 3, se pueden ver los resultados obtenidos usando muestras de máquinas en estado *ralenti*, el Cuadro 4, muestra los resultados obtenidos de muestras grabadas de máquinas en aceleración. Cada resultado, es el obtenido de la media aritmética de 10 experimentos previos cada uno. En ambos cuadros, la primera columna muestra el número de características extraídas por cada segmento de tiempo (ms).

Cuadro 3. Resultados obtenidos con muestras de motores en estado *ralenti*.

Configuración	Precisión(sinACP)	Precisión(ACP)
LPCC 36c 100ms	95.22 %	94.37 %
LPCC 26c 100ms	95.41 %	91.62 %
LPCC 20c 100ms	89.74 %	78.55 %

Los resultados en ambas tablas muestran que la mejor precisión fue dada cuando usamos 36 y 26 LPCC por cada 100 ms, cuando usamos muestras de

motores con aceleración; podemos decir que con 36 coeficientes por segmento de tiempo, el lapso de entrenamiento fue mayor y no había mucha diferencia en los resultados de reconocimiento, mostrando con esto, que 26 LPCC dan mejores resultados globales.

Cuadro 4. Resultados obtenidos con muestras de motores en aceleración.

Configuración	Precisión(sinACP)	Precisión(ACP)
LPCC 36c 100ms	95.55 %	94.27 %
LPCC 26c 100ms	93.01 %	95.58 %
LPCC 20c 100ms	86.47 %	82.58 %

Tomando esto en cuenta, se seleccionaron al azar doce muestras de sonido en formato WAV, provenientes de ruido no usado en la etapa de entrenamiento, y separado para este fin; tomando una RNA previamente entrenada, se alimentaron estas muestras para demostrar la precisión que logra alcanzar dicha red. Los resultados obtenidos se muestran en el Cuadro 5.

Cuadro 5. Comparación entre clases diagnosticadas y clases reales.

Test	Clase1	Clase2	Clase3	clase4	ClaseR
1	94.5945	0.0	5.4054	0.0	Si
2	100.0	0.0	0.0	0.0	Si
3	89.1891	0.0	0.0	10.8108	Si
4	0.0	75.6756	21.6216	2.7027	Si
5	5.4054	94.5945	0.0	0.0	Si
6	2.70270	67.5675	8.1081	21.6216	Si
7	0.0	0.0	97.2972	2.7027	Si
8	0.0	5.4054	86.4864	8.1081	Si
9	0.0	0.0	2.7027	97.2972	No
10	0.0	0.0	0.0	100.0	Si
11	0.0	0.0	8.1081	91.8918	Si
12	0.0	0.0	8.1081	91.891892	Si

8. Conclusiones y Trabajos Futuros

Los resultados muestran que en un motor en malas condiciones existen diferencias acústicas, comparándolo con un motor afinado o en buenas condiciones. Al menos estas diferencias entre sonidos pueden ser reconocidas por una Red Neuronal con Retraso en el Tiempo, claro esta, dependiendo del tipo de falla que

esté afectando la eficiencia del motor. Esta información puede ayudar al especialista del motor a dar diagnósticos rápidos, precios estimados por la reparación y permitirle trabajar directamente con el problema del motor. También podemos concluir que es viable trabajar en este problema e implementar un sistema barato que pueda ser usado por fabricantes de motores o dueños de talleres. Estamos trabajando en otras metodológicas de extracción de características como MFCC y LPCC, y métodos de reducción de ruido, usando filtros Wavelets, y de esta manera comparar los resultados mostrados con estos métodos. Otro de los objetivos es diseñar una interfaz de usuario, independiente de Matlab, que pueda ser usada en talleres mecánicos y predecir fallas mecánicas de motor con tan solo introducir directamente el ruido por micrófono o pre-grabando las muestras acústicas del motor.

Referencias

1. Mario E. Munich., *Bayesian Subspace Methods for Acoustic Signature Recognition of Vehicles.*, Published at the 12th European Signal Processing Conference (EUSIPCO-2004), Vienna, Austria. September 6-10, (2004).
2. Huadong Wu, Mel Siegel, y Pradeep khosla, *Vehicle Sound Signature Recognition by Frecuency Principle Component Analysis.*, The proccedings of IMTC'1998, selected in the IEEE Transaction on Instrumentation and Measurement Vol. 48, No. 5, page 1005-1009, (ISSN 0018-9456), October 1999.
3. Pedro Nelson Saavedra y Edgar A. Estupiñan, *Análisis de Vibraciones Aplicado A Las Máquinas Rotatorias de Baja Velocidad.* <http://internal.dstm.com.ar/sites/mmnew/bib/notas/Estupi1.pdf>. Chile 2000.
4. Pedro Nelson Saavedra. *La Medición y Análisis de las Vibraciones como Técnica de Inspección de Equipos y Componentes, Aplicaciones, Normativas y Certificación.*, Facultad de Ingeniería - Universidad de Concepción, casilla 160 - Concepcion - Chile. <http://www.aacende.org.ar/sitio/biblioteca/material/CONFCHILE.pdf>, 2001.
5. Agustín Álvarez Marquina. *Fundamentos del Reconocimiento Automático de la Voz, Algoritmos de Extracción de Características.* Facultad de Informática. Universidad Politécnica de Madrid (UPM). Noviembre 22, 2004.
6. Orozco-Garca, Jos. *Extracción y Análisis de Características Acústicas del Llanto de Bebés para su Reconocimiento Automático Basado en Redes Neuronales.* Tesis. INAOE, México. Febrero 2004.
7. Hilera J., Martínez V. *Redes Neuronales Artificiales Fundamentos, modelos y aplicaciones.* Alfaomega, Madrid, España. (2000).
8. Reyes-Galaviz O. F., Reyes-Garca, C.A., *Infant Cry Classification to Identify Hypoacoustics and Asphyxia with Neural Networks*, in LNCS 2972: Advances in Artificial Intelligence. Springer, Berlin, 2004, pp 69-78, ISBN 3-540-21459-3, ISSN 0302-9743.
9. A. Weibel, T. Hanazawa, G. Hinton, K. Shikano, and K.J. Lang, *Phoneme Recognition Using Time Delay Neural Networks*, IEEE Trans. Acoustics, Speech, Signal Proc., ASSP-37: 32'339, (1989).
10. Malcolm Slaney, *Auditory Toolbox.*, Interval Research Corporation <http://rvl4.ecn.purdue.edu/malcolm/interval/1998-010/AuditoryToolbox.zip>