

Reconocimiento de instrumentos musicales

9 de febrero de 2014

Rodrigo Arias Mallo

Andrés Baamonde Lozano

Lois Alberte Gómez Sánchez

Jean Paul Pereira Príncipe

Carlos Pérez Ramil

Índice

1. Introducción	3
2. Descripción del problema	3

1. Introducción

En este documento vamos a recoger los resultados de la realización de un trabajo perteneciente a la asignatura de Aprendizaje Automático basado en el reconocimiento auditivo de instrumentos musicales. Se realizará distintas aproximaciones para la resolución, comenzando con casos simples (basados en el reconocimiento de dos instrumentos) hasta los más complejos, empleando técnicas de Redes Neuronales, Máquinas de Soporte Vectorial y la técnica kNN.

La motivación principal para la realización de este trabajo es la de facilitar las tareas de aprendizaje musical que puedan tener lugar en centros en los que se imparta algún tipo de educación musical así como para aprendizaje personal. Mediante las técnicas empleadas en este trabajo se podría desarrollar software de reconocimiento auditivo en el que la evaluación del alumnado se pudiese realizar de modo automático. Otro campo de aplicabilidad sería el de la accesibilidad, permitiendo la elaboración de software que realice algún tipo de traducción sensorial para permitir el que personas con discapacidad pudiesen disfrutar de espectáculos sonoros. La clasificación de instrumentos también podría permitir una futura clasificación por géneros u otro tipo de clasificaciones musicales.

Los objetivos iniciales del trabajo son los de alcanzar un reconocimiento de instrumentos fiable en fragmentos musicales simples. Según se desarrolle el trabajo se realizarán cada vez aproximaciones más complejas que permitan diferenciar un mayor número de instrumentos, así como lograr su clasificación en condiciones menos óptimas.

2. Descripción del problema

Para el desarrollo del trabajo se emplearán archivos de audio que permitan distinguir el instrumento que aparece en él. En los archivos de audio empleados inicialmente sólo aparecerá un único instrumento que nos permitirá definir las características más relevantes de cada uno. Los ficheros de audio empleados deberán ser convertidos a .wav y sólo emplearemos un canal de sonido, bien porque ambos canales contengan la misma información o en el caso de contener distinta información se realizará una conversión a mono, mezclando ambos canales.

La base de datos será realizada por el grupo tomando como partida muestras pertenecientes al sitio web www.philharmonia.co.uk. Cada una de las muestras tomadas se dividirá en fragmentos de dimensión potencia de dos para poder ser empleada por la función fft de Matlab. La base de datos se realiza mediante un script de Matlab que clasifica las muestras de los distintos instrumentos y los divide como hemos explicado. Por tanto para cada uno de los instrumentos tendremos una colección de fragmentos de los que almacenaremos tanto su valor original como el de su transformada de Fourier. Estos datos se organizan en una estructura de estructuras, en la que la estructura principal agrupa las estructuras para los instrumentos y para cada instrumento se almacena cada archivo de sonido troceado dentro de otra estructura que almacena el nombre de la canción, la señal de audio original, la señal troceada y la transformada de Fourier de cada uno de los fragmentos.

A cada muestra inicialmente representada en función del tiempo se le aplica la transformada de Fourier para obtener su representación en frecuencia que resulta de utilidad para la extracción de características. La frecuencia de muestreo de los archivos de audio empleados es de 44100Hz, dato que hemos de tener en cuenta a la hora de realizar las distintas aproximaciones.