

Clasificador de canciones según géneros musicales

Gianfranco Fagioli, Victor Matzkin, Gaspar Oberti

Procesamiento Digital de Señales
Facultad de Ingeniería y Ciencias Hídricas
Universidad Nacional del Litoral

- 1 Introducción
 - Géneros musicales
- 2 Extracción de Características
 - Características del Timbre
 - Características del Ritmo
- 3 Clasificación
- 4 Resultados
- 5 Referencias

Género

Forma de categorizar a las canciones a partir de una determinada percepción del ser humano.

Clasificador de géneros

Distinguir entre géneros utilizando características propias de la señal, que se pueden obtener a partir de mediciones en el dominio temporal y frecuencial. Haciendo uso de un clasificador estadístico, luego se procede a estimar el género de un conjunto de canciones.

Extracción de Características

Caracterizar un segmento de audio mediante una representación numérica compacta. Se extraerán dos tipos de características: de timbre y de ritmo.

Timbre

El timbre es una de las cuatro cualidades que caracteriza al sonido, y es el que nos permite diferenciar dos sonidos de igual frecuencia fundamental e intensidad (por ejemplo, la misma nota tocada con dos instrumentos diferentes).

- Análisis de Fourier
- Tamaño de la ventana
- Textura del timbre
- Media y varianza de características

- Centro de masa del espectro
- *Brillo* del sonido

$$C_t = \frac{\sum_{n=1}^N M_t[n] * n}{\sum_{n=1}^N M_t[n]}$$

- Concentración de energía

$$C_t = \sum_{n=1}^{R_t} M_t[n] = 0,85 * \sum_{n=1}^N M_t[n]$$

- Cuánto cambia la potencia del espectro de una ventana a otra.
- Norma-2.

$$F_t = \sum_{n=1}^N (N_t[n] - N_{t-1}[n])^2$$

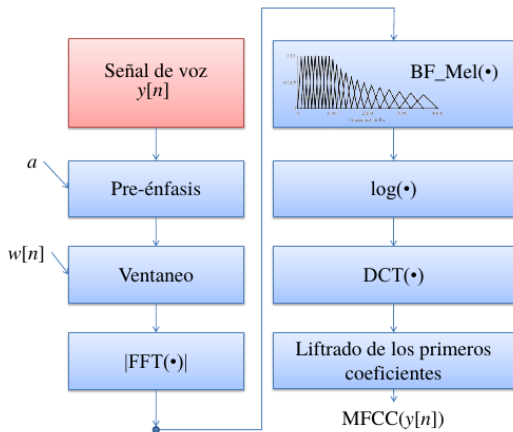
Cruces por cero en el dominio del tiempo

- Medida del ruido de la señal.

$$Z_t = \frac{1}{2} \sum_{n=1}^N |\text{signo}(x[n]) - \text{signo}(x[n-1])|$$

Coeficientes cepstrales en la escala de Mel (MFCC)

- Cepstrum.
- Banco de filtros (5 primeros coeficientes).



Ritmo

Fenómenos temporales de pequeña y mediana escala. Se define por el *Tempo* (una medida de la rapidez con la que fluye el ritmo). Esto se mide en golpes por minuto (BPM).

- Descomposición en bandas de frecuencia (DDWT).
- Cálculo de las envolventes
- Autocorrelación

- 1 Rectificación de onda completa: $y[n] = |x[n]|$
- 2 Filtrado pasa bajos: $y[n] = (1 - \alpha)x[n] + \alpha y[n - 1]$
- 3 Submuestreo: $y[n] = x[kn]$
- 4 Remover la media: $y[n] = x[n] - E[x[n]]$

- Descomposición en bandas de frecuencia (DDWT).
- Cálculo de las envolventes
- Autocorrelación: $y[k] = \sum_n x[n]x[n - k]$

Conjunto final de características

- Media y varianza de:
 - Flujo espectral
 - Roloff espectral
 - Centroide espectral
 - Cantidad de cruces por cero
 - 5 primeros coeficientes cepstrales en la escala de mel
- Características del histograma de ritmo:
 - Amplitud del primer pico
 - Amplitud del segundo pico
 - Relación entre las amplitudes mayores
 - BPM del primer pico
 - BPM del segundo pico
 - Suma

- Clasificador estadístico
 - Análisis discriminante
- Conjunto de Datos

- Géneros del conjunto de datos.
- Porcentaje de aciertos: 72 %

Cuadro: Matriz de confusión de los géneros.

	cla	dis	hip	reg	roc
cla	9	0	0	0	1
dis	0	6	4	0	0
hip	0	1	8	1	0
reg	1	1	3	4	1
roc	1	0	0	0	9



G. Tzanetakis, "*Musical Genre Classification of Audio Signals*", in *IEEE Trans Speech Audio Processing*, vol. 10, no. 5, pp. 293-302. Jul. 2002.



E. Scheirer, "*Tempo and beat analysis of acoustic musical signals*", in *J. Acoust. Soc. Amer.*, vol. 103, no. 1, pp. 588-601. Jan. 1998.

¿Preguntas?