

Przetwarzanie i Rozpoznawanie Dźwięku

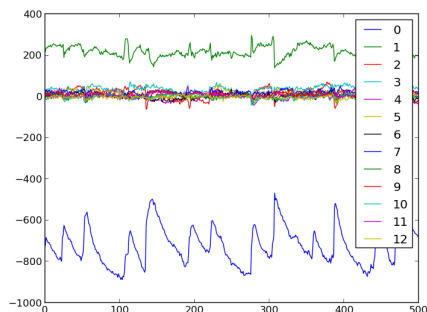
Klasyfikacja utworów muzycznych ze zbioru gtzan.

30 stycznia 2014

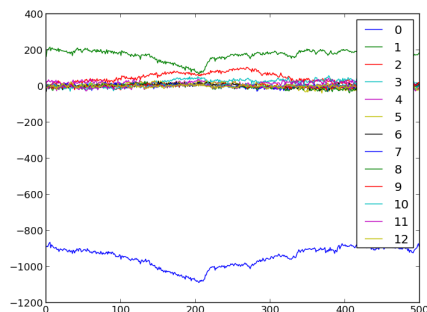
Autor:	Paweł Rychły	inf94362	ISWD	pawelrychly@gmail.com
	Dawid Wiśniewski	inf94387	ISWD	wisniewski.dawid@gmail.com

1 Analiza zmian wartości współczynników cepstralnych w czasie.

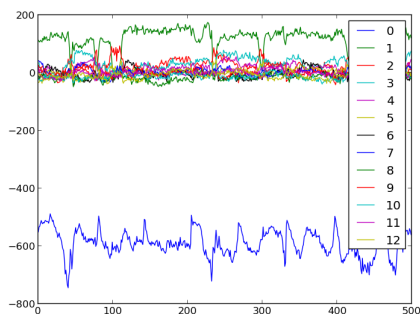
W celu doborzenia najbardziej odpowiednich metod agregacji danych opisujących poszczególne próbki dźwięku, wygenerowano zbiór wykresów prezentujących zmiany wartości współczynników cepstralnych w kolejnych próbkach.



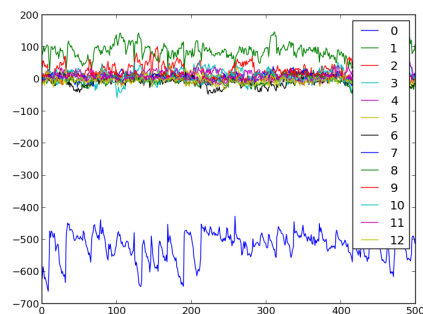
(a) blues



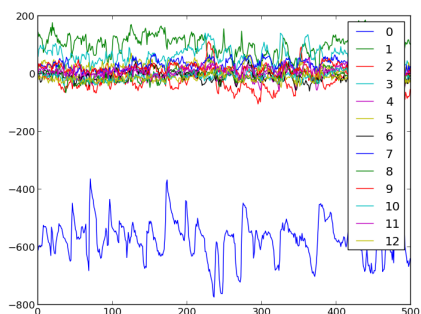
(b) classical



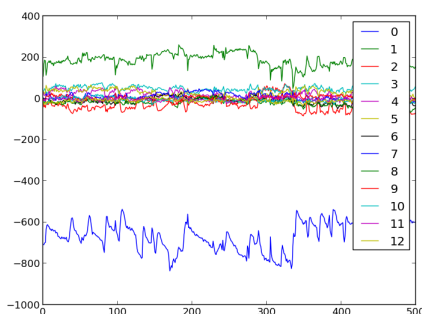
(c) country



(d) disco

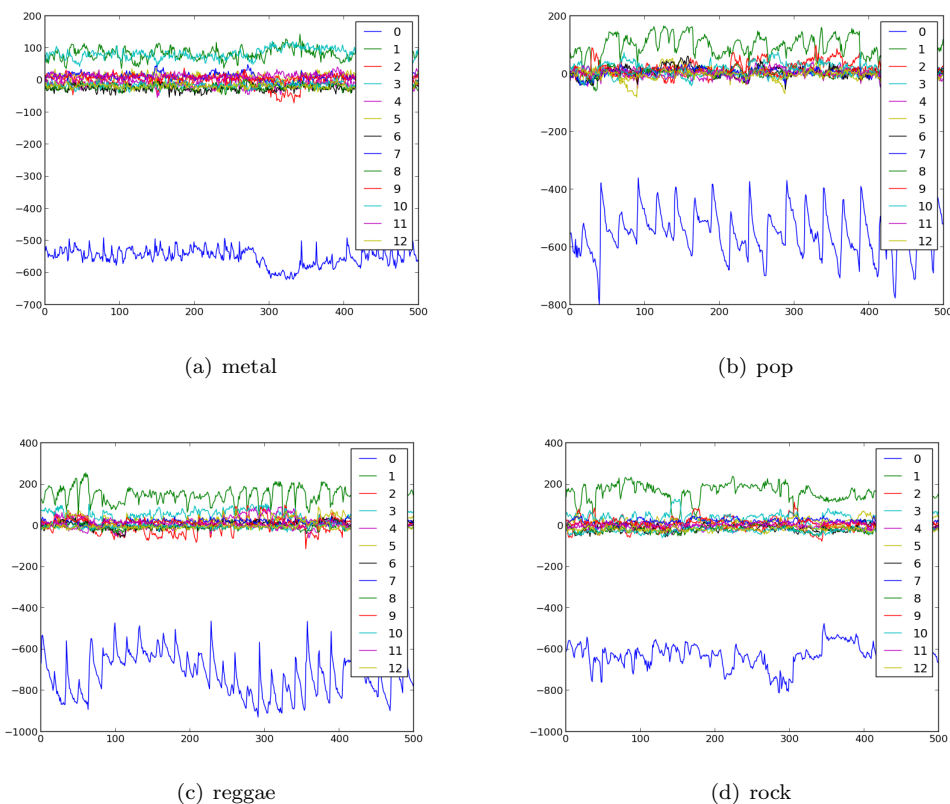


(e) hip-hop



(f) jazz

Rysunek 1: Porównanie zmian wartości współczynników cepstralnych dla przykładowych plików reprezentujących różne gatunki muzyczne.



Rysunek 2: Porównanie zmian wartości współczynników cepstralnych dla przykładowych plików reprezentujących różne gatunki muzyczne.

Jak można zauważyć, kształty wykresów ilustrujących zmiany współczynników cepstralnych, dla różnych gatunków muzycznych różnią się pomiędzy sobą. Wydaje się, że miara agregująca te dane, powinna odzwierciedlać w maksymalnym stopniu, nie tyle średnie wartości funkcji co ich kształt. Bardzo charakterystyczny jest wykres będący wynikiem analizy utworu muzyki klasycznej. Wartości współczynników cepstralnych zmianały się tutaj w sposób zdecydowanie bardziej łagodny niż w przypadku innych wykresów.

2 Metody agregacji danych dotyczących współczynników cepstralnych.

W celu agregacji danych pochodzących z różnych okien zastosowano kilka miar. Były to takie podstawowe miary jak: wartość minimalna, wartość maksymalna, wariancja oraz średnia arytmetyczna. Ich znaczenie jest oczywiste, dlatego w dalszej części zawarto opis trzech mniej popularnych ocen sygnału. Dwie pierwsze metody wykorzystują Dyskretną transformę Fouriera sygnału, oznaczaną literą F .

2.1 Spectral Flatness Measure

Miara ta opisuje jak bardzo analizowany sygnał zbliżony jest do szumu białego. Jeżeli wartość współczynnika jest duża, oznacza to, że wartość analizowanego sygnału (w tym przypadku wartość współczynnika cepstralnego) nie zmienia się lub, że zmienia się w sposób nieregularny. Mała wartość miary mówi o tym, że zmiany wartości współczynników cepstralnych mogłyby zostać opisane za pomocą kilku sinusoid.

$$sfm = \frac{e^{\overline{\ln(F)}}}{\overline{F}} \quad (1)$$

2.2 The most significant frequency

Miara ta zawiera informacje o częstotliwości sinusoidy, która ma największy wpływ na kształt funkcji wykreślonej przez zmianę współczynnika cepstralnego w czasie.

$$f_{max} = \operatorname{argmax}(F) \quad (2)$$

2.3 Maksymalna korelacja sygnału z własnym przesunięciem.

Ostatnia zastosowana miara, określa maksymalną korelację sygnału (Kolejnych wartości współczynnika cepstralnego), z własnym przesunięciem.