

MFCC Implementierung

Stefan Scherer

Abteilung Neuroinformatik
Universität Ulm

16. Juni 2006

Inhalt

- 1 Generelle Vorgehensweise
- 2 MFCC Merkmale

Generelle Vorgehensweise

- Aufteilen in Frames (32 ms)
- Hamming Fenster
- Features für einzelne Frames berechnen

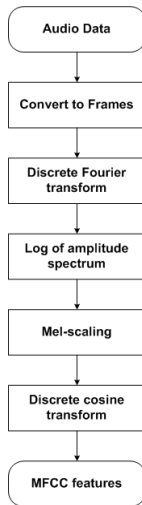
Gliederung

- 2 MFCC Merkmale
 - Grundlagen
 - Implementierung

Mel Frequency Cepstral Coefficients I

- angelehnt an das menschliche Gehörsorgan
- Verdopplung des Mel Wertes entspricht Verdopplung der Wahrgenommenen Tonhöhe
- 0 - 500 Hz gleichzusetzen mit Mel, weiter nicht mehr linear
- Beispiele: 100 Hz = 100 Mel, 200 Hz = 200 Mel →
Verdopplung des Mel Wertes = Verdopplung der Tonhöhe;
1500 Hz = 1100 Mel, 10000 Hz = 2200 Mel
- Comirva Bibliothek

Mel Frequency Cepstral Coefficients II



Gliederung

- 2 MFCC Merkmale
 - Grundlagen
 - Implementierung

Fast Fourier Transform

Benutzt wird die FFT implementiert in der CoMIRVA Bibliothek.
Vorgehensweise:

- FFT Objekt erstellen: `FFT powerFFT = new FFT(FFT.FFT_POWER, windowSize);`
- FFT durchführen (Transformation wird in-place durchgeführt): `powerFFT.transform(window,null);`

Filterdesign I

Filterbank mit M Dreiecksfiltern ($m = 1, 2, \dots, M$ mit $M = 24$),
Filter m ist definiert durch:

$$H_m[k] = \begin{cases} 0 & k < f[m-1] \\ \frac{2(k-f[m-1])}{(f[m+1]-f[m-1])(f[m]-f[m-1])} & f[m-1] \leq k \leq f[m] \\ \frac{2(f[m+1]-k)}{(f[m+1]-f[m-1])(f[m+1]-f[m])} & f[m] \leq k \leq f[m+1] \\ 0 & k > f[m+1] \end{cases}$$

Filterdesign II

Die Grenzpunkte $f[m]$ sind gleich verteilt in der Mel-Skala nach:

$$f[m] = \frac{N}{F_s} B^{-1} \left(B(f_l) + m \frac{B(f_h) - B(f_l)}{M + 1} \right),$$

mit $f_l = 20$ die tiefste Frequenz in der Filterbank, $f_h = 16000$ die höchste Frequenz, F_s Abtastrate in Hz, N die Länge der FFT.

Filterdesign III

Mel-Skala:

$$B(f) = 1125 \ln(1 + f/700)$$

Inverse der Mel-Skala:

$$B^{-1}(b) = 700(\exp(b/1125) - 1)$$

Log-Energy für den Output von jedem Filter

$$S[m] = \ln \left[\sum_{k=0}^{N-1} |X_a[k]|^2 H_m[k] \right], \quad 0 < m \leq M$$

Diskrete Cosinus Transformation

$$c[n] = \sum_{m=0}^{M-1} S[m] \cos(\pi n(m - 0.5)/M), \quad 0 < n \leq M$$