

Escuela de Informática – Universidad Nacional de Trujillo
Tópicos Especiales en Ciencia de la Computación
Reconocimiento Automático del Habla
Jorge Luis Guevara Diaz

LABORATORIO NUMERO 3

1 Introducción

El laboratorio consta de las siguientes partes

1. Implementación del MFCC
2. Implementación del DTW
3. Experimentos
4. Propuesta de Mejora

1 Implementación del MFCC

Para implementar MFCC implementarán los siguientes algoritmos:

- Implementación filtro preémfasis
- Ventaneamiento: utilizar ventana hamming

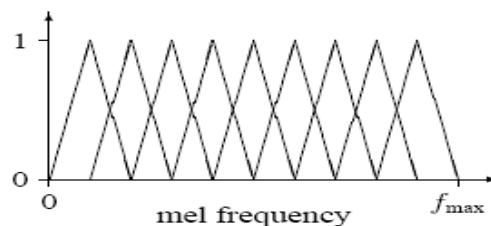
$$S_i = \left(0.54 - 0.46 \cos \frac{2\pi i}{N-1} \right) s_i$$

$\{s_i, i = 0, \dots, N-1\}$ son las muestras de la señal

- Mel binning
 Calcular FFT de cada porcion de señal ventaneada, y solo utilizaremos la mitad de estos vectores, la frecuencia puede ser hallada con $i*F_s/N$
 Se debe tener en cuenta que los bins están igualmente espaciados en la escala mel de acuerdo a

$$\text{Mel}(f) = 1127 \ln\left(1 + \frac{f}{700}\right)$$

Los bins deberan ser perfectamente triangulares en el espacio mel



siendo la parte derecha de cada bin la parte central del siguiente bin, para calcular el aporte de la magnitud de los valores del FFT a cada bin, se debera mapear estos a la

frecuencia Mel y encontrar su respectivo peso, luego se tomará el logaritmo de cada valor

$$S_i = \sum_f |X(f)| H_i(\text{Mel}(f))$$

- Transformada discreta del coseno, Aplicar la transformada discreta del coseno a cada frame de la siguiente manera

•

Valores de entrada: $\{s_i, i = 0, \dots, N - 1\}$

Valores de salida $\{S_j, j = 0, \dots, M - 1\}$

M es el número de coeficientes cepstrales

$$S_j = \sqrt{\frac{2}{N}} \sum_{i=0}^{N-1} s_i \cos \left(\frac{\pi(j+1)}{N} (i+0.5) \right)$$

- Al finalizar el algoritmo MFCC se debe obtener una matriz de $13 \times n$ (o en todo caso $n \times 13$) elementos donde cada vector de 13 elementos son los coeficientes MFCC representativos de cada frame
- Agregar características dinámicas a los coeficientes : deltas y dobles deltas, en este punto se tendrá una matriz de $39 \times n$ (o en todo caso $n \times 39$) elementos

2 Implementación del DTW

Construir un clasificador basado en DTW, utilizar $p=1$ y distancia euclidiana

3 Experimento

Realice los siguientes experimentos tome solo los coeficientes MFCC como base

- Como varia la tasa de reconocimiento usando MFCC + delta
- Como varia la tasa de reconocimiento usando MFCC + delta +delta delta
- Influye el factor de preemfasis en la tasa de reconocimientos? De ser así con cual factor obtiene mas reconocimientos? Experimentar solo con parte b
- Como varia la tasa de reconocimiento usando ventana Hamming y ventana rectangular, experimentar solo con parte b

4 Mejora

- Haga una propuesta de mejorar (solo parte MFCC), (investigación: papers, o algo totalmente nuevo inventado por usted)
- Implemente su mejora

Nota

Las voces de entrenamiento y de prueba serán suministradas por el profesor del curso