



Dissertação: **Reconhecimento de Orador em Dois Segundos**

Mestrado Integrado em Engenharia Electrotécnica e de Computadores, ramo Telecomunicações

Orientador: Prof. Aníbal Ferreira

Relatório realizador por: Diana Rocha Mendes

Relatório N° 3

7 Março 2011 a 13 Março 2011

Tarefas Realizadas:

- Continuação do estudo teóricos de GMM e algoritmo EM.
- Continuação do estudo dos algoritmos de cálculo de MFCCs disponíveis da VoiceBox e na Auditory Toolbox.
- Estudo dos coeficientes MFCCs, em especial da sua distribuição estatística e da correlação entre os coeficientes. Averiguação prática através do desenho dos histogramas de coeficientes extraídos de gravações de voz, usando o Matlab.
- Início do estudo prático da função `gmdistribution.fit` da Statistics Toolbox do Matlab, através da criação de dados aleatórios com distribuição gaussiana, criação do modelo GMM desses dados e visualização do fitting deste aos dados. Estudo da relação entre o fitting e o log-likelihood calculado.

Dificuldades Encontradas:

- Dificuldade na identificação das diferenças entre os algoritmos de cálculo de MFCCs que provoquem diferenças tão acentuadas nos resultados obtidos com cada um deles.

Próximas Tarefas:

- Definição de critérios de análise dos diferentes algoritmos de forma a seleccionar os mais indicados para estudo mais aprofundado e implementação preliminar.
- Selecção de técnicas e algoritmos após estudo dos algoritmos encontrados, segundos os critérios referidos acima, e implementação.

- Estudo teórico mais aprofundado das técnicas mais importantes actualmente: MFCCs e GMMs
- Continuação da escrita do estado da arte.

Bibliografia:

<http://www.cedar.buffalo.edu/~srihari/CSE574/>

Rose, P., "Forensic Speaker Identification", New York: Taylor & Francis, 2002

Sigurdsson, S., Petersen, K. B., Lehn-Schiøler, T., "Mel Frequency Cepstral Coefficients: An Evaluation of Robustness of MP3 Encoded Music", 2006