



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

Mestrado Integrado em Engenharia Electrotécnica e de Computadores, ramo Telecomunicações

Orientador: Prof. Dr. Aníbal Ferreira

Relatório realizador por: Diana Rocha Mendes

## Relatório N° 2

28 Fevereiro 2011 a 6 Março 2011

### Tarefas Realizadas:

- Pesquisa de métodos de comparação entre GMMs, medidas de distância aplicadas a GMMs;
- Estudo aprofundado teórico de GMMs e algoritmo EM;
- Estudo do conceito de *likelihood ratio*, pesquisa de implementação de cálculo deste em Matlab;
- Estudo aprofundado teórico de MFCCs;
- Comparação dos algoritmos de cálculo de MFCCs disponíveis na Auditory Toolbox e na VoiceBox, testes preliminares.

### Dificuldades Encontradas:

- Alguma dificuldade na compreensão do significado de *log-likelihood* calculado no algoritmo EM, aplicado a GMMs. Este não se encontra explícito na informação disponível sobre a função de cálculo de GMMs (`gmdistribution.fit`) da Statistics Toolbox. Já a VoiceBox não oferece nenhuma função que retorne este valor, o que dificulta a comparação entre os algoritmos.
- Informação disponível relativa a *likelihood ratios* pouco extensa e maioritariamente teórica – apenas um algoritmo foi encontrado [1].

### Próximas Tarefas:

- Continuação do estudo referente a GMMs, algoritmo EM e MFCCs.

- Concluir sobre as principais diferenças entre os algoritmos de cálculo de MFCCs da Auditory Toolbox e VoiceBox e qual o mais indicado para utilização numa primeira fase de implementação de uma solução;
- Implementação de solução relativa a cálculo de distância entre GMMs, testes preliminares.

### **Referências:**

[1] [http://geoff-morrison.net/Software/multivar\\_kernel\\_LR.m](http://geoff-morrison.net/Software/multivar_kernel_LR.m)

### **Bibliografia:**

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

Karpov, E., “Real-Time Speaker Identification”. Tese de mestrado, Universidade de Joensuu, 2003

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