



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° 5

21 Março 2011 a 27 Março 2011

### Tarefas Realizadas:

- Concluído o estudo dos valores likelihood devolvidos pela função `gmdistribution.fit`, utilizando conjuntos de dados criados aleatoriamente com distribuições gaussianas. Foi feita a variação da dimensão destes dados, simulando situações de treino com conjuntos de dados maiores e dados de teste menores. A conclusão retirada deste estudo foi que, nestas situações ideais, a diminuição do tamanho dos dados de teste não influencia significativamente os valores de likelihood.
- Na sequência da conclusão indicada anteriormente foi iniciado o estudo de técnicas de normalização de coeficientes MFCC. Foram feitos alguns testes com a normalização Z e com a cepstral mean subtraction, ou cepstral mean normalization.
- Estudo prático da influência da variação do número de coeficientes utilizados. Estudo da dependência entre esses coeficientes e selecção adequada destes.

### Dificuldades Encontradas:

- Resultados até à data obtidos não se revelam satisfatórios, apesar de na literatura haver a indicação de que técnicas como a cepstral mean normalization aumentarem significativamente o desempenho dos sistemas de identificação.

## **Próximas Tarefas:**

- Continuação do estudo de técnicas de normalização.
- Estudo de métodos mais avançados de classificação como Support Vector Machines.

## **Bibliografia:**

Kleynhans, N., Barnard, E., "A channel normalization technique for speech recognition in mismatched conditions.", 2008

Lima, C., Tavares, A., Silva, C., Oliveira, J., "Spectral Normalisation MFCC Derived Features for Robust Speech Recognition", 2004

Garcia, A., Mammone, R., "Channel-robust speaker identification using modified-mean cepstral mean normalization with frequency warping", 1999