



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° 14/15

23 Maio 2011 a 5 Junho 2011

Tarefas Realizadas:

- Continuação dos testes indicados no último relatório, nomeadamente o de variação de número de GMMs e o uso de coeficientes Δ MFCCs e $\Delta\Delta$ MFCCs em adição aos 13 coeficientes MFCCs já usados. Um dos objectivos deste teste era retirar conclusões sobre a performance dos Δ MFCCs e $\Delta\Delta$ MFCCs e sobre o número de gaussianas que optimiza essa performance, de forma a apurar os parâmetros mais adequados a utilizar no estudo da base de dados PJ.
- Visto que os resultados obtidos com as características Δ MFCCs e $\Delta\Delta$ MFCCs foram inferiores aos resultados obtidos apenas com MFCCs, foi equacionada a hipótese de as características delta e delta-delta serem mais adequadas a um cenário de reconhecimento de orador dependente de texto – em que os oradores dizem uma palavra chave, por exemplo. Como as bases de dados disponíveis não compreendem a mesma frase dita duas vezes pelo mesmo orador, foram gravados no laboratório alguns segmentos com 3 oradores. Num cenário tão simples como este a performance foi de 100% independentemente das características usadas.
- Segmentação das vogais da base de dados TIMIT através das anotações que são providenciadas com esta. Teste do sistema de reconhecimento utilizando apenas as vogais. Comparação da performance, sob várias condições de treino e teste, com a performance do sistema usando a voz completa.

Dificuldades Encontradas:

- Já indicadas na descrição das tarefas.

Próximas Tarefas:

- Continuação dos testes indicados no último relatório, relativamente aos “tipos” de voz – vozeada, apenas vogais, não vozeada, voz completa.
- Início dos testes com NRDs.

Bibliografia:

[1] Sousa, R. e Ferreira, A., *Importance of the relative delay of glottal source harmonics*, 2010

[2] Sousa, R. e Ferreira, A., *Singing Voice Analysis Using Relative Harmonic Delay*, 2010