

Conclusões e Sugestões Para Trabalhos Futuros

A pesquisa desenvolvida buscou aumentar a taxa de reconhecimento em aplicações onde nada se sabe sobre o tipo e a natureza do ruído. Nesta tese foram propostas duas estratégias para a combinação das respostas dos múltiplos classificadores em sub-bandas. A primeira proposta utiliza o cálculo das energias dos sinais passa-banda usados no treino para gerar pesos a serem aplicados às medidas de verossimilhanças. A segunda, emprega o cálculo do espaço nulo a partir de uma matriz de energias obtidas dos sinais passa-banda de treino para gerar um conjunto de pesos que são alternados na fase de teste. Essa última estratégia é a que apresentou os melhores resultados, mas requer mais espaço de memória para armazenar os vetores de pesos gerados pelo cálculo do espaço nulo. O desempenho da identificação de locutor, em termos da taxa de acerto, devido à utilização das estratégias propostas foi significativamente mais elevado que o das técnicas da literatura, para situações onde a voz de teste estava severamente contaminada por ruído branco e ruído não-branco. Isso evidencia a importância do uso de ponderações não-uniformes para a combinação das medidas de verossimilhanças, possibilitando favorecer às contribuições mais importantes para o reconhecimento. Nota-se também que a alternância na escolha de pesos dada pelo uso de vetores do espaço nulo, contribui para a melhoria da combinação das medidas de verossimilhança, tendendo a reduzir o efeito do ruído e a aumentar a taxa de acerto. Também foi examinada a utilização do treinamento em múltiplas condições, no esquema de reconhecimento que emprega o espaço nulo. Foi mostrado que esta estratégia pode melhorar o desempenho de reconhecimento, principalmente, quando a voz usada no teste está contaminada por ruído branco. Por outro lado, como uma sugestão de trabalho posterior, seria interessante a elaboração de um estudo no sentido de melhorar a modelagem do ruído a ser usado no treinamento tal que leve a melhorar o desempenho da identificação quando a voz de teste está contaminada por ruído não-branco. Além disso, foi observado que o emprego dos atributos dinâmicos (delta e delta-delta) no esquema de reconhecimento utilizando o espaço nulo melhora na grande maioria dos casos

o desempenho da identificação. Isso foi devido às informações dinâmicas e à robustez que esses atributos forneceram quando utilizados em conjunto com os atributos estáticos nos esquemas de reconhecimento que utilizam o espaço nulo. Adicionalmente, foi analisado o emprego dos atributos AMFCC na identificação de locutor realizada pelo esquema que usa o espaço nulo. Foi observado que a aplicação desse atributo produz, de modo geral, taxas de acerto comparáveis às das demais técnicas de múltiplos classificadores propostas nessa tese, sendo que são sempre as melhores quando a voz de teste está contaminada por ruído branco.

Esta tese mostrou que a técnica do espaço nulo em conjunto com a utilização de atributos dinâmicos pode melhorar significativamente o desempenho da identificação, em termos da taxa de acerto, quando a voz de teste está severamente contaminada por ruídos. Note-se que o aumento na taxa de acerto resultante da aplicação das estratégias propostas de múltiplos classificadores, foi alcançado sem a utilização de técnicas de extração de ruídos. Vários experimentos foram realizados a fim de avaliar o desempenho das novas propostas.

6.1

Trabalhos Futuros

Além das técnicas exploradas nessa tese, como continuidade das pesquisas, pode-se sugerir a experimentação de outros tipos de filtros para realizarem a decomposição em sub-bandas, de tal forma que proporcionem uma melhor separação inicial entre as contribuições da voz e as do ruído. Adicionalmente, pode-se avaliar o emprego de outros algoritmos de classificação, buscando gerar modelos mais precisos de cada locutor, principalmente, em situações envolvendo a presença de ruído, como as mostradas nesta tese. Além disso, também pode-se buscar uma melhor modelagem do ruído para o treinamento em múltiplas condições. Finalmente, é interessante testar os sistemas de reconhecimento propostos em aplicações de verificação de locutor e em reconhecimento independente do texto. Essas são algumas propostas para trabalhos futuros, visando melhorar os bons resultados já obtidos.