

1

Introdução

Desde o começo da humanidade, a fala foi o meio de comunicação por excelência. Devido a isto, é que tem sido utilizada como interface entre o homem e a máquina, constituindo-se deste modo, em um objeto de pesquisa de grande interesse.

Os homens sempre procuraram meios de comunicação que promovessem a interação homem - máquina. Devido aos rápidos avanços tecnológicos na área de processamento digital de sinais, foi possível desenvolver sistemas de reconhecimento de locutor (habilidade de reconhecer uma pessoa unicamente por sua voz) de alto desempenho, baixo custo e velozes.

O reconhecimento automático de locutor pode-se dividir em identificação e verificação. Na identificação se identifica uma pessoa através de sua voz em um conjunto conhecido de N locutores e na verificação avalia-se um sinal de voz pertence ou não a uma determinada pessoa .

Dentre as aplicações de reconhecimento de locutor, pode-se destacar: Controle de acesso a dispositivos de dados, segurança e redes de trabalho, Auxílio a deficientes auditivos, Aplicações forenses, Uso militar, Comunicação com computador, Autenticação de transações comerciais para evitar e prevenir fraudes: operações bancárias e compras com cartão de crédito por telefone, etc.

Apesar de apresentarem alto desempenho, quando os sistemas de reconhecimento de locutor são utilizados em presença de ruído ou em canais de comunicações, suas contribuições degradam-se muito rápido. Esta ausência de robustez aparece quando os sistemas de reconhecimento, que são projetados para funcionar em um laboratório, são utilizados em situações reais. Torna-se necessário, portanto, desenvolver e avaliar métodos que sejam robustos a essas perturbações. Esta dissertação oferece uma contribuição nessa direção.

1.1

Motivação e Objetivos da Dissertação

Há varias motivações para o desenvolvimento deste trabalho, das quais podemos destacar o interesse por avaliar os sistemas com a base TIMIT que apresenta locuções de curta duração assim como os atributos MFCC, SSCH, TECC, que apresentaram um bom desempenho na tarefa de reconhecimento de voz, mas não são conhecidos em reconhecimento de locutor, além de uma proposta que usa o atributo PAC-MFCC com classificadores em sub-bandas.

Os objetivos principais desta dissertação são o estudo, análise e comparação dos atributos MFCC(Coeficientes Cepstrais de Frequência Mel), SSCH(Histogramas de Centróides de Sub-Bandas Espectrais), PAC-MFCC(Fase Autocorrelação-MFCC), PAC-MFCC com Classificador em Sub-Bandas e TECC (Coeficientes Cepstrais da Energia Teager) para reconhecimento de locutor independente do texto em ambientes ruidosos. Serão discutidas as principais vantagens e desvantagens desses sistemas.

1.2

Organização da Dissertação

O presente trabalho está organizado da seguinte maneira: No Capítulo 2 são apresentados conceitos básicos de reconhecimento de locutor e descritos atributos promissores para o reconhecimento de locutor independente de texto em presença de ruído. No Capítulo 3 é proposto o atributo PAC-MFCC com Classificador de Sub-Bandas. No Capítulo 4 são apresentados os resultados obtidos de uma comparação geral dos diferentes sistemas e, finalmente, as conclusões e sugestões para trabalhos futuros são fornecidos no Capítulo 5