

Reconhecimento de Emoções em Sinais de Fala utilizando Redes Neurais e Análise de Componentes Independentes

Elmo A. T. Borges Junior, Eduardo F. Simas Filho, Jês J. F. Cerqueira, Antônio C. L. Fernandes Jr. e Arthur L. C. Weyll

Programa de Pós-Graduação em Engenharia Elétrica; Universidade Federal da Bahia; Rua Prof. Aristides Novis, 2, Federação, Salvador-BA

Email: elmoalberto.teixeiraborjes@gmail.com, (eduardo.simas, antonio.lopes, jes)@ufba.br

1-Histórico, Motivação e Objetivo

Com o uso cada vez maior de máquinas é necessário o desenvolvimento e aprimoramento de técnicas para reconhecimento de padrões de características físicas, objetivando uma maior interação entre usuários finais e as máquinas requerendo, dessas últimas, a percepção das ações e reações humanas e respondendo a elas de maneira apropriada.

O objetivo desse trabalho foi desenvolver um sistema de classificação de emoções a partir de sinais de voz, comparar diversas combinações de descritores de sinais de música e voz, realizar processamento estatístico por *ICA* dos vetores de parâmetros extraídos dos sinais de voz e classificação por meio de um classificador neural supervisionado com arquitetura *MLP* (*Multilayer Perceptron*).

2-Contribuição e Métodos

Uma contribuição importante do trabalho foi a utilização de estatística de alta ordem (*ICA*) em uma etapa de processamento estatístico com objetivo de redução da redundância entre os parâmetros utilizados. A técnica de Análise de Componentes Independentes foi aplicada ao vetor de característica antes da etapa de classificação, contribuindo para a redução da dependência estatística e a compactação dos dados, já que, o algoritmo da *ICA* calcula a análise de componentes principais (*PCA*) em uma etapa preliminar.

3-Resultados, discussão e conclusões

Os resultados encontrados indicam que os coeficientes *MFCC* “estáticos” e “dinâmicos”, são características que possuem maior capacidade de discriminação das emoções contidas no sinal de fala do que quando combinada às características propostas. Contudo, observa-se que alguns componentes, dentre os 15 extraídos para cada descritor, totalizando 45 coeficientes, possuem maior capacidade de diferenciação entre as emoções do que outros.

O algoritmo da *ICA* utilizado reduziu a dimensionalidade do vetor de características, pois verificou-se que 99,9% da energia (variância) estava armazenada nos 20 primeiros componentes da *PCA*, obtendo-se um aumento de eficiência do classificador neural com diminuição do vetor de característica. Houve aumento da eficiência (acurácia) do classificador após aplicação da *ICA* em 2,8 pontos percentuais. Para 50% das emoções houve aumento da eficiência e para 33,3% das emoções o resultado permaneceu constante.

Em uma comparação com outros trabalhos foi possível observar que, em alguns casos, obtiveram maior acerto do sistema de classificação, mas reduzindo-se o número de emoções, o que significa um problema de separação entre classes muito mais simples, ou o número de parâmetros que compõem o vetor apresentado a entrada do classificador.

Em futuros trabalhos pretende-se abordar diferentes tipos e arquiteturas de classificadores neurais, como redes profundas, testar novos descritores, coeficientes *MFCC*, *MFCC Delta* e *Delta Delta* específicos e utilizar outros bancos de dados de falas.