

## DESENVOLVIMENTO DE ALGORITMOS DE SUPORTE AO SUBPROJETO EMOSONICO

Larissa Aparecida Pereira Schueda<sup>1</sup>, Rafael Kingeski,<sup>2</sup> Aleksander Sade Paterno<sup>3</sup>

<sup>1</sup> Acadêmico(a) do Curso de Engenharia Elétrica CCT - bolsista PIBIC/CNPq

<sup>2</sup> Mestrando do Curso de Engenharia Elétrica CCT

<sup>3</sup> Orientador, Departamento de Engenharia Elétrica CCT – Aleksander.paterno@udesc.br.

Palavras-chave: Emoções, Sinais, Processamento.

O objetivo deste trabalho é a fundamentação teórica ao subprojeto Emosonico, que visa criar uma base de dados em formato de áudios separados em masculino e feminino e em cada gênero uma classificação em 6 emoções sendo elas: felicidade, medo, nojo, raiva, surpresa, tristeza e o estado neutro. Com a finalidade de validar programas analisadores de emoções na interface de comunicação, linguagem, psicologia e pesquisa com dispositivos eletrônico.

As emoções são algo do cotidiano do ser humano, que são reações naturais a estímulos e com a evolução da tecnologia é comum a utilização de dispositivos eletrônicos, cada dia se procurando mais características humanas em máquinas, crescendo assim a importância no reconhecimento conciso e o mais preciso possível de emoções na interface homem-máquina. Para analisar um sinal de voz e identificar as emoções existem diversas técnicas de processamento de sinal de voz. Esta pesquisa foi dividia em três partes: coleta de dados de voz, parâmetros de análise e classificação[1]. Neste contexto foram estudadas as emoções Felicidade, Medo, Raiva, Nojo, Tristeza, Surpresa e o estado Neutro sendo estas as emoções bases definidas por Paul Ekman [2].

A coleta de áudios foi feita em português brasileiro de vídeos do *Youtube*, filmes e gravações em estúdios com autorizações dos envolvidos. A extração de parâmetros, sendo os algoritmos implementados no software Matlab, é baseada na revisão bibliográfica [3] totalizando 29 parâmetros sendo que para a extração da frequência fundamental foi utilizado uma sobreposição de 50 milissegundos com o janelamento de 60 milissegundos, extraíndo os coeficientes LPC (*Linear prediction coding*) dos quadros de voz, e então obtidos os valores do primeiro, segundo e terceiro formante. Obtendo também a largura de banda de cada formante, para a energia de curto termo foi utilizado um janelamento de 20 milissegundos. Para frequência fundamental, formantes e largura de banda de formantes foram utilizados: valor médio, valor mínimo, valor máximo e desvio padrão, já o valor médio quadrático foi utilizado para a energia de curto termo.

A classificação dos áudios segundo os 29 parâmetros foi feita no método máquinas de vetores de suporte (SVM, do inglês *Support vector machine*), com função de *Kernel* cúbica, sendo este o que apresentou melhor resultado (O método SVM classifica os dados encontrando o melhor hiperplano, ou seja, aquele com a maior margem entre as duas classes).

Para exemplificar os resultados analisamos 24 áudios para cada emoção sendo retiradas de 5 locutores masculinos teremos assim totalizando 168 áudios tendo uma exatidão de 88,7% segundo método SVM sendo visto pela Fig.1 sendo que estes áudios estão disponíveis no link: [https://drive.google.com/drive/folders/1yVgjFKj\\_mYwJVZviIvKGmnthC3sZ05z?usp=sharing](https://drive.google.com/drive/folders/1yVgjFKj_mYwJVZviIvKGmnthC3sZ05z?usp=sharing).



Valor Verdadeiro (%)	Valor Previsto (%)						
	Felicidade	Medo	Neutro	Nojo	Raiva	Surpresa	Tristeza
Felicidade	75			13	13		
Medo		100					
Neutro			96				4
Nojo	8			83		8	
Raiva	8			21	71		
Surpresa				4		96	
Tristeza							100

**Fig.1 Confusion Matrix (Matrix de Erro)**

Referências bibliográficas:

- [1] KAMINSKA, D. PELIKANT, A. Recognition of human emotion from a speech signal based on plutchik's model. In Kaminska, D. Pelikant, A., editor, Journal of Electronics and Telecommunications Vol.58 No. 2, pages 165,170. 2012
- [2] EKMAN, P.; FRIESEN, W. V. Unmasking the face: A guide to recognizing emotions from facial clues. Oxford, Prentice-Hall, 1975
- [3] SHARMA, R. NEUMANN, U. AND KIM, C. Emotion Recognition In Spontaneous Emotional Utterances From Movie Sequences, in Proceedings of the WSEAS International Conference on Electronics, Control & Signal Processing, 2002
- [4] Christianini, N. e J. Shawe-Taylor. Uma introdução ao suporte a máquinas de vetores e outros métodos de aprendizado baseados em kernel. Cambridge, UK: Cambridge University Press, 2000.