DESENVOLVIMENTO DE ALGORITMOS DE SUPORTE AO PROJETO EMOSONICO

Larissa Aparecida Pereira Schueda [1], Rafael Kingeski,[2] Aleksander Sade Paterno[3]

[1] Acadêmico(a) do Curso de Engenharia Elétrica CCT - bolsista PIBIC/CNPq
[2] Mestrando do Curso de Engenharia Elétrica CCT
[3] Orientador, Departamento de Engenharia Elétrica CCT – Aleksander.paterno@udesc.br.

Palavras-chave: Coleta, Processamento de dados, Sinais.

O objetivo deste trabalho é a fundamentação teórica ao projeto *Emosonico*, que visa criar uma base de dados em formato de áudios separados em masculino e feminino e, para cada gênero uma classificação em 6 emoções. Tais emoções são: felicidade, medo, nojo, raiva, surpresa, tristeza e o estado neutro, tendo sido definidas por Paul Ekman [1]. Com a finalidade de validar programas analisadores de emoções na interface de comunicação, linguagem, psicologia e em pesquisa com dispositivos eletrônicos, esta pesquisa foi dividida em 3 partes: coleta de dados de voz, catalogação e classificação.

A coleta de áudios foi feita em português brasileiro de vídeos do *Youtube,* filmes e gravações em estúdios com autorizações dos envolvidos. As gravações em estúdio são formadas por uma média de 2 horas com 2 a 3 voluntários por sessão num total de 12 voluntários, estes sendo instruídos a falarem sobre vídeos, previamente selecionados, e para emoções como medo e raiva foram usados jogos que poderiam fazer o voluntario alcançar estas emoções com maior facilidade. Após a extração dos áudios, a catalogação dos mesmos começou tendo sido utilizada um modelo em árvore para emoções visto na Fig.1

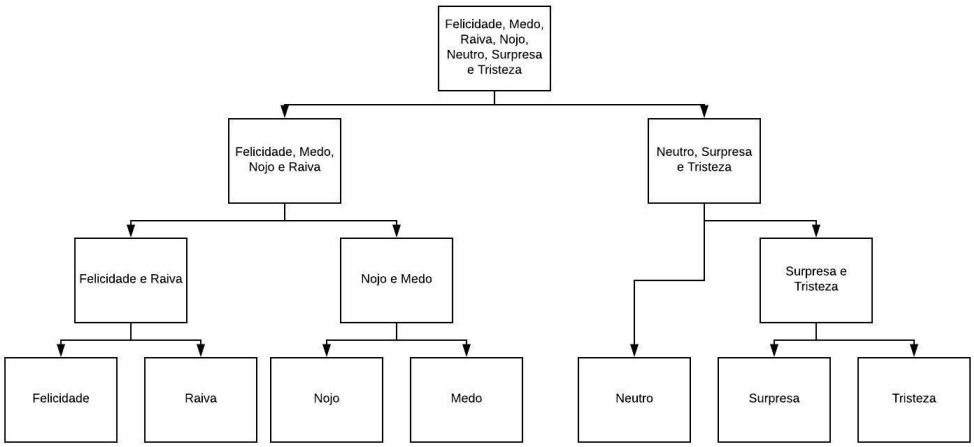

Fig.1 *Esquema de classificação proposto por Iriya (2014) [5]*

Os áudios originais tinham por volta de 2horas de duração e foram reduzidos para áudios de 5 segundos em média, estes seguindo as teorias de Plutchik[2], Ekman[1], Thalmann[3] e Goddard[4] onde, por exemplo, a surpresa é gerada por um evento inesperado ou pela interrupção súbita de um estímulo, provocando uma pausa e permitindo que o indivíduo tenha tempo para se orientar. Já o medo é despertado frente a um evento causado por outra pessoa ou pelo ambiente,

avaliado como ameaçador, resultando numa resposta de fuga, sendo que para a tristeza a resposta mais típica é o choro, o afastamento e o silêncio.

A extração de parâmetros, sendo os algoritmos implementados no software *MATLAB*, é baseada na revisão bibliográfica [6] totalizando 29 parâmetros sendo que para a extração da frequência fundamental foi utilizado uma sobreposição de 50 milissegundos com o janelamento de 60 milissegundos, extraindo os coeficientes LPC (*Linear prediction coding)* dos quadros de voz. A partir disso foram obtidos os valores do primeiro, segundo e terceiro formante. Obteve-se 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.

Ao se apresentarem os resultados, analisamos 35 áudios para cada emoção, obtendo-se assim um total de 245 áudios femininos. A matriz de confusão da avaliação dos resultados permite calcular uma exatidão de 82% segundo classificação com máquinas de vetor de suporte (SVM), apresentado em Fig.2 com áudios disponíveis em: https://drive.google.com/open?id=1yQEAQmYkWCTTD-i4JBwdTeC85KZUKQJ2

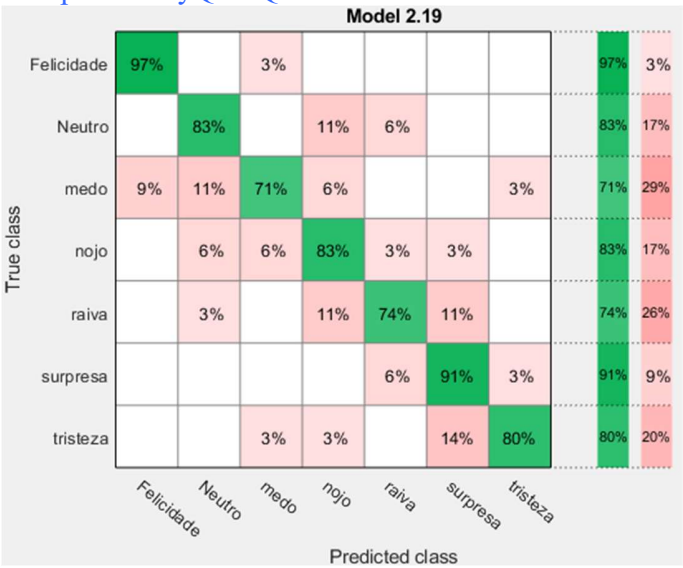

Fig.2 *Matriz de confusão da classificação de emoções em áudio usando SVM*

Referências bibliográficas:

[1] EKMAN, P.; FRIESEN, W. V. Unmasking the face: A guide to recognizing emotions from facial clues. Oxford, Prentice-Hall, 1975

[2] PLUTCHIK, R. (2002). Emotions and life: Perspectives from psychology, biology and evolution. Washington, DC: American Psychological Association.

[3] THALMANN, Y. A. (2013). Le décodeur des émotions. Paris: First-Gründ.

[4]GODDARD, Cliff. Contrastive semantics and cultural psychology:'Surprise'in Malay and English. Culture & Psychology, v. 3, n. 2, p. 153-181, 1997.

[5] IRIYA, R. Análise de sinais de voz para reconhecimento de emoções. Dissertação (Mestrado) — Curso de Engenharia e Sistemas Eletrônicos, Universidade de São Paulo, 2014.

[6] 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