FlexPersuade - Explorando uma abordagem flexível em softwares de persuasão: um estudo de caso com players de música

Leandro Yukio Mano Alves

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Leandro Yukio Mano Alves

FlexPersuade - Explorando uma abordagem flexível em softwares de persuasão: um estudo de caso com players de música

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientador: Prof. Dr. Jó Ueyama

USP – São Carlos
Julho de 2016

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

A634f

Alves, Leandro Yukio Mano
FlexPersuade - Explorando uma abordagem flexível em softwares de persuasão: um estudo de caso com players de música / Leandro Yukio Mano Alves; orientador Jó Ueyama. - São Carlos - SP, 2016.
81 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2016.

1. Comitê de Classificação. 2. Classificação da Emoção. 3. Interação Humano-Computador (IHC). 4. FaceTracker. I. Ueyama, Jó, orient. II. Título.

Leandro Yukio Mano Alves

FlexPersuade - Exploring a flexible approach to persuasion software: a case study with music players

Master dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master Program in Computer Science and Computational Mathematics. *FINAL VERSION*

Concentration Area: Computer Science and Computational Mathematics

Advisor: Prof. Dr. Jó Ueyama

USP – São Carlos
July 2016

Este trabalho é dedicado aos meus pais, Cleudo e Elza, pelo sacrifício ilimitado em todos os sentidos, que mais do que me proporcionar uma boa infância, formaram os fundamentos do meu caráter e me apontaram o caminho a seguir. Obrigado por serem a minha referência de tantas maneiras e estarem sempre presentes na minha vida de uma forma indispensável.

A minha irmã Camila e família, que apesar de tantas provações esteve sempre disposta a me estender a mão e me oferecer ajuda, carinho, companhia. Que eu possa servir de exemplo para Lívia e Mariana (sobrinhas), para que elas possam seguir um caminho de luta e trabalho para conquistas de seus sonhos.

Aos meus familiares, que me apoiaram de todas as formas possíveis, incentivando, aconselhando em toda minha vida.

Aos amigos de perto e de longe, pelo amor e preocupação demonstrados através de ligações, visitas e e-mails. Obrigado, vocês que aliviaram minhas horas difíceis, me alimentando de certezas, força e alegria.

Muito obrigado, nunca será suficiente para demonstrar a grandeza do que recebi de todos.

AGRADECIMENTOS

Ao Prof^{o}. Dr. Jó Ueyama, orientador, professor e amigo pelo desprendimento ao escolher me dar apoio, muito obrigado pela dedicação e ajuda durante essa nova etapa da minha vida.

Agradeço a Universidade de São Paulo - USP e ao Laboratório de Sistemas Distribuídos e Programação Concorrente - LaSDPC pelo incentivo, apoio e infraestrutura para a pós-graduação e o desenvolvimento deste projeto. Um agradecimento especial aos integrantes do *Wireless Networks Group - WinesGroup* pela disponibilidade, esforço e dedicação para com os membros do grupo, em especial a minha pessoa.

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio no início do curso de pós-graduação por disponibilizar uma bolsa de estudo destinado a alunos de mestrado.

Agradeço a Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) por apoiar este discente no restante e maior parte do curso de pós-graduação por disponibilizar uma bolsa de estudo destinado a alunos de mestrado - **processo n° 2014/10685-3, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)**.

"A educação é a arma mais poderosa
pela qual se pode mudar o mundo."
(NELSON MANDELA)

RESUMO

MANO, L. Y.. **FlexPersuade - Explorando uma abordagem flexível em softwares de persuasão: um estudo de caso com players de música**. 2016. 81 f. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

Estudos atuais na área de Interação Humano-Computador evidenciam a importância de se considerar aspectos emocionais na interação com sistemas computacionais. Acredita-se que ao permitir agentes artificiais identificar emoções de usuários, em uma interação humano-computador, torna-se possível induzir e despertar emoções a fim de estimulá-los em suas atividades. Um dos grandes desafios dos pesquisadores em Interação Humano-Computador é prover sistemas capazes de reconhecer, interpretar e reagir de modo inteligente e sensível às emoções do usuário, para atender aos requisitos do maior número possível de indivíduos; um dos caminhos que se apresenta é o desenvolvimento de sistemas flexíveis. O principal objetivo de se promover essa interação emotiva é contribuir para o aumento da coerência, consistência e credibilidade das reações e respostas computacionais providas durante a interação humana via interface humano-computador. Nesse contexto, surge a oportunidade de explorar sistemas computacionais capazes de identificar e inferir o estado emocional do usuário em tempo de execução. Este projeto tem como objetivo desenvolver e avaliar um modelo que possa: i.) identificar o estado emocional do usuário; ii.) prover um mecanismo de persuasão com vistas a mudar o estado emocional do usuário (com um estudo de caso em player de música) e; iii.) explorar a abordagem flexível na persuasão (de acordo com o estado emocional particular de cada usuário) através de mecanismos persuasivos que poderão variar entre um player de música, jogos e/ou vídeos. Assim, ao longo do estudo, o modelo baseado em Comitê de Classificação se mostrou eficiente na identificação das emoções básicas (alegria, aversão, medo, neutro, raiva, surpresa e tristeza) com média de acurácia superior a 80% e, ainda, observou-se a satisfação dos usuários mediante a aplicação do modelo com o player de música.

Palavras-chave: Comitê de Classificação, Classificação da Emoção, Interação Humano-Computador (IHC), FaceTracker.

ABSTRACT

MANO, L. Y.. **FlexPersuade - Explorando uma abordagem flexível em softwares de persuasão: um estudo de caso com players de música**. 2016. 81 f. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

Current studies in the field of Human-Computer Interaction highlight the relevance of emotional aspects while interacting with computers systems. It is believed that allowing intelligent agents to identify user's emotions, they can induce and awaken emotions in order to stimulate them while interacting with computers. A major challenge for researchers in human-computer interaction is to provide systems capable of recognizing, interpreting and reacting intelligently and sensitively to the emotions of the user, to meet the requirements of the largest possible number of individuals. One of the ways presented in this project is the development of flexible systems to meet a large number of emotions/behaviors. The main objective of promoting this emotional interaction is to contribute to increasing the coherence, consistency and credibility of reactions and computational responses provided during human interaction via human-computer interface. In this context, the opportunity arises to explore computational systems able to identify and infer the emotional state of the user at runtime. This project aims to develop and evaluate a model that can: i.) identify the emotional state of the user/developer; ii.) provide a mechanism of persuasion in order to change the emotional state of the user (with a case study in music player) and; iii.) explore the flexible approach in persuasion (according to the particular emotional state of each user) through persuasive mechanisms that may vary from a music player, games and/or videos. Thus, throughout the study, the Rating Committee based model is efficient for identification of basic emotions (happiness, disgust, fear, neutral, anger, surprise and sadness) with average accuracy higher than 80% and also noted himself to user satisfaction by applying the model to the music player.

Key-words: Ensemble of Classification, Emotion Classification, Human-Computer Interaction (HCI), FaceTracker.

LISTA DE ILUSTRAÇÕES

Figura 1 – Modelo *circumplex* proposto por Russell (1980). 28
Figura 2 – Emoções básicas propostas por Ekman (2006). 29
Figura 3 – Processo de combinação de decisão de um EC (SCHULLER *et al.*, 2005). . 32
Figura 4 – Módulo *FaceTracker* (SANTOS V.C.; COCA, 2013). 33
Figura 5 – Processo de mapeamento da face pelo *FaceTracker* (MANO *et al.*, 2015). . 43
Figura 6 – Módulo baseado em Cômite de Classificação para a classificação da emoção da face do usuário (MANO *et al.*, 2015). 43
Figura 7 – Cenário de interação de componentes do *FlexPersuade*. 45
Figura 8 – Estrutura para interação do usuário com o dispositivo computacional. 46
Figura 9 – Tela para configurar as preferências de artista e de emoções. 47
Figura 10 – Tela de captura e de classificação da emoção em tempo de execução. 47
Figura 11 – Tela de execução do Player (áudio e áudio/vídeo). 48
Figura 12 – Tela de *feedback* das emoções para o usuário. 48
Figura 13 – Boxplots das acurácias apresentadas pelos classificadores para determinar a emoção do usuário. Tais resultados foram obtidos com o uso da técnica *k-fold cross-validation* com $k = 10$. 51
Figura 14 – Gráfico ROC do modelo de classificação induzido pelo Comitê de Classificadores. 54
Figura 15 – Configuração: 12 indivíduos e 25 Interações. 55
Figura 16 – Configuração: 25 indivíduos e 50 Interações. 55
Figura 17 – Configuração: 50 indivíduos e 100 Interações. 56
Figura 18 – Acurácia alcançada pelo Comitê com o uso de pesos das diferentes abordagens avaliadas. 57
Figura 19 – Elementos configuráveis (níveis) da arquitetura SAHHc (MANO *et al.*, 2016). 58
Figura 20 – Avaliação de desempenho: *OnBoard* versus Nó Decisor. 60
Figura 21 – Avaliação de desempenho do processamento na nuvem. 62
Figura 22 – Nível de satisfação de interação do usuário. 65

LISTA DE TABELAS

Tabela 1 – Síntese das pesquisas identificadas na literatura. 38
Tabela 2 – Parâmetros utilizados nos classificadores selecionados do Weka. 50
Tabela 3 – Média (%) das acurácias e os valores-*p* dos conjuntos de resultados. 52
Tabela 4 – Valores-*p* da comparação de pares realizada com o teste *Wilcoxon Rank Sum*. Valores inferiores a 0,05 indicam diferença estatisticamente significativa entre os grupos de resultados. 52
Tabela 5 – Matriz de confusão referente à classificação do Comitê. 53
Tabela 6 – Detalhamento do custo de classificação apresentado pelo Comitê para cada emoção. 53
Tabela 7 – O método de *Shapiro-Wilk* aprova a hipótese de adequação dos conjuntos de amostras à distribuição normal com intervalo de confiança de 95%. Valores-*p* superiores a 0,05 indicam compatibilidade com a distribuição normal. . . . 57
Tabela 8 – Divisão dos grupos para os experimentos. 63
Tabela 9 – Matriz de confusão referente à classificação do Comitê das imagens dos usuários submetidos ao teste. 64

SUMÁRIO

1 INTRODUÇÃO 21
1.1 Contexto e Motivação 21
1.2 Objetivo 23
1.3 Organização da Dissertação 23

2 FUNDAMENTAÇÃO TEÓRICA 25
2.1 Teoria Componencial das Emoções 25
2.2 Discussão do Espaço Emocional Semântico 27
2.3 Mecanismos de Classificação da Emoção 29
2.4 Captura de Expressões Faciais 32

3 TRABALHOS RELACIONADOS 35
3.1 Discussão dos Principais Trabalhos 35
3.2 Síntese dos Trabalhos Relacionados 38

4 ABORDAGEM PARA PERSUASÃO FLEXÍVEL 41
4.1 Visão Geral do Projeto 41
4.2 Extração das Características da Face 42
4.3 Uso do Comitê de Classificação para Identificar Emoções 43
4.4 *FlexPersuade* 44
4.4.1 Modelo e Objetivo 44
4.4.2 Estudo de Caso: Player de Música 46

5 AVALIAÇÃO 49
5.1 Metodologia 49
5.2 Avaliação da Utilização do Comitê de Classificação 51
5.3 Avaliação dos Pesos Utilizados no Comitê 54
5.4 Avaliação do Desempenho Computacional do *FlexPersuade* 57
5.4.1 Desempenho em Rede de Sensores 59
5.4.2 Análise dos efeitos de transferência de imagens a serem processadas na nuvem 61
5.5 Avaliação da Interação com o Usuário (Player de música 62
5.5.1 Planejamento 63
5.5.2 Resultados da Observação 64

6 CONCLUSÕES 67
6.1 Conlcusões e Trabalhos Futuros 67
6.2 Atividades Acadêmicas e Complementares 68
6.2.1 Disciplinas Cursadas 68
6.2.2 Produção Científica 68
6.2.2.1 Periódico 68
6.2.2.2 Conferência 69

REFERÊNCIAS 71

ANEXO A TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO 77

ANEXO B AUTORIZAÇÃO DE CAPTAÇÃO E EXIBIÇÃO DE IMAGEM 79

ANEXO C FORMULÁRIO DE OBSERVAÇÃO DO PARTICIPANTE 81

CAPÍTULO

1

INTRODUÇÃO

1.1 Contexto e Motivação

A emoção é um aspecto fundamental do ser humano, que o permite desfrutar de experiências, além de influenciar suas decisões diárias. Esse aspecto é importante na interação entre as pessoas (ZHOU *et al.*, 2011) e pode ser utilizado para intervir na realização de uma tarefa, ou até mesmo para tomar decisões em prol do usuário; por exemplo, sugerindo um gênero de filme para aliviar o estresse com base na sua atual emoção (GONÇALVES *et al.*, 2013a). Os aspectos emocionais vêm sendo estudados há muito tempo na área da Psicologia (LICHTENSTEIN *et al.*, 2008). No entanto, somente nos últimos anos, houve um aumento significativo de pesquisas sobre o tema na área de Computação Afetiva (KLEIN; MOON; PICARD, 2002; BAILENSON *et al.*, 2008; GONÇALVES *et al.*, 2011; ZHOU *et al.*, 2011; PETER; URBAN, 2012; LIKAMWA *et al.*, 2013). Uma das divisões básicas da Computação Afetiva é a Interação Humano-Computador (IHC) (PICARD, 1997; PICARD, 2004), a área que estuda como os dispositivos computacionais podem reconhecer, modelar e responder as emoções humanas (dentre outros aspectos) e expressá-las mediante uma interface/interação computacional (GONÇALVES *et al.*, 2011).

Diante desse panorama, estudos atuais na área de IHC evidenciam a importância de se considerar aspectos emocionais na interação com sistemas computacionais (ESLAMBOLCHILAR *et al.*, 2011; OINAS-KUKKONEN, 2013). Analisada por Pesquisadores da *University of Southern California*, na matéria publicada pelo Tech (2013), a face desenvolve características diferentes de acordo com estados emocionais distintos. Tais características podem ser analisadas por sensores específicos, de maneira a atribuir uma emoção de acordo com a face do usuário. Esses sistemas computacionais podem ser introduzidos na descrição comportamental e emocional do usuário, utilizando-se desses dados para a melhor satisfação pessoal do indivíduo na interação com dispositivos computacionais.

O principal objetivo de se promover essa interação afetiva é contribuir para o aumento da

coerência, da consistência e da credibilidade das reações e das respostas computacionais providas durante a interação humana, via interface humano-computador. Por exemplo, se o computador "percebe" que o usuário está cansado, estressado ou triste, esse tenta "persuadí-lo" à mudança de estado a partir de um player de música, jogos, chat com amigos, filmes e/ou redes sociais.

Assim, sistemas capazes de reconhecer as emoções do usuário, interpretando-as e reagindo de modo inteligente e sensível, poderiam atender mais adequadamente aos requisitos de interação dos indivíduos. Ao permitir que agentes artificiais compreendam emoções em uma interação humano-computador, é possível "persuadir" (induzir e/ou despertar) emoções em usuários de sistemas computacionais (XAVIER; GARCIA; NERIS, 2012). O conhecimento sobre a persuasão permite identificar emoções provocadas no uso de sistemas interativos e pode ser aplicado na concepção e no desenvolvimento de tais sistemas, fomentando, assim, mudanças de atitude e de comportamento nos indivíduos (SAIDIN; MACAULAY; HINE, 2011). O fato de que as pessoas respondem socialmente a sistemas computacionais tem implicações significativas, que criam oportunidades para aplicar uma série de dinâmicas de persuasão (FOGG, 2002).

Estudos conduzidos por pesquisadores de diversas universidades americanas e européias indicam que as companhias perdem bilhões de dólares com os males da fadiga de seus funcionários (EXAME, 2013) e, diante disso, algumas delas preferem não deixar seus funcionários trabalharem quando estão cansados, na tentativa de diminuir a produção de softwares com erros, ou ainda, evitar desastres caso o usuário "manuseie" sistemas críticos (piloto de aeronaves, usina de reator nuclear, entre outros).

As emoções são complexas, difíceis de identificar e de mensurar; estão ligadas à reação de vários componentes que podem ser sintetizados em: avaliações cognitivas, sentimentos subjetivos, tendências comportamentais, reações fisiológicas e expressões motoras (SCHERER, 2005; MAHLKE; MINGE, 2008) e trazem à tona a discussão sobre o que as emoções são e o que precisa ser estudado quando se quer determinar o significado das mesmas (SCHERER, 2005). De maneira mais específica, o uso de imagens da face dos usuários para identificar as emoções permite uma interação pervasiva [1]. Isso porque tal informação pode ser obtida sem uma ação explícita do usuário e sem o uso de equipamentos específicos ou intrusivos. Além disso, no âmbito tecnológico, é comum que equipamentos computacionais (por exemplo, tablet, desktop, notebook e smartphone) possuam câmeras frontais. Tipicamente, esses dispositivos permanecem direcionados à face do usuário, podendo ser utilizados para a aquisição de imagens em tempo real.

Devido as características de identificação de emoções, técnicas de Aprendizado de Máquina, do inglês *Machine Learning* (ML), têm sido amplamente investigadas nesse contexto. Nessa abordagem, as emoções são consideradas classes e as expressões suas respectivas características. Assim, a identificação de emoções é tratada como um problema de classificação,

[1] Computação Pervasiva ou Computação Ubíqua tem como objetivo tornar a interação homem computador invisível, ou seja, integrar tecnologia com as ações e comportamentos naturais das pessoas.

comumente supervisionado. Estudos encontrados na literatura descrevem o uso de classificadores isolados para resolver tal problema (ØHRN; ROWLAND, 2000; CHANEL *et al.*, 2009; ZHOU *et al.*, 2011; GONÇALVES *et al.*, 2013a; RAMAKRISHNAN; EMARY, 2013). Contudo, poucos trabalhos abordam o uso de Comitê de Classificadores, do inglês *Ensemble of Classification* (EC), para classificar emoções (SCHULLER *et al.*, 2005), porém existem trabalhos que relatam bons resultados dessa técnica aplicada a outros problemas de classificação (CANUTO, 2001; DUDA; HART; STORK, 2012).

Assim, o desenvolvimento de sistemas que permitem alterações em seu comportamento durante a interação em tempo de execução e que oferecem a cada usuário a possibilidade de ter o sistema ajustado de acordo com suas preferências, necessidades e intenções de uso (NERIS; BARANAUSKAS, 2010), considerando requisitos para a persuasão e as emoções provocadas pelos sistemas nos diferentes perfis de usuários, configura-se um aspecto importante na concepção de novas soluções. Entretanto, ainda é um grande desafio medir o quão persuasiva uma solução é e o quanto ela pode influenciar na tomada de decisão ou na mudança de comportamento dos indivíduos. Um dos grandes desafios dos pesquisadores em IHC é encontrar formas de prover sistemas capazes de reconhecer, de interpretar e de reagir de modo inteligente e sensível as emoções do usuário, a fim de atender os requisitos do maior número possível de indivíduos (GONÇALVES; UEYAMA, 2012). Um dos caminhos que se apresenta é o desenvolvimento de sistemas flexíveis (GONÇALVES *et al.*, 2013a). Nesse contexto, a flexibilidade se refere a mudanças relativas ao comportamento do sistema, tais como adição ou remoção de componentes.

1.2 Objetivo

Este trabalho de mestrado teve como objetivo desenvolver e avaliar um modelo que possa: **i.)** identificar o estado emocional do usuário; **ii.)** prover um mecanismo de persuasão que objetiva mudar o estado emocional do usuário (com um estudo de caso em player de música) e **iii.)** explorar a abordagem flexível na persuasão (de acordo com o estado emocional particular de cada usuário), através de mecanismos persuasivos que poderão variar entre um player de música, jogos e/ou vídeos.

1.3 Organização da Dissertação

Esta dissertação é composta de 6 capítulos. No capítulo 2, apresenta-se a fundamentação teórica utilizada para o desenvolvimento deste trabalho, além de conceitos psicológicos e mecanismos computacionais para a identificação e a classificação das emoções. O capítulo 3 apresenta uma síntese do levantamento bibliográfico realizado, que aborda trabalhos e problemas encontrados por pesquisadores para a identificação, a classificação e a interação de dispositivos computacionais e de usuários. O capítulo 4 propõe a abordagem *FlexPersuade* para a interação entre dispositivos computacionais e usuários por meio da emoção. O capítulo 5 apresenta as

avaliações referentes à acurácia do Comitê de Classificação proposto, aos pesos utilizados nos algoritmos, o desempenho do *FlexPersuade* e uma avaliação comparativa de satisfação na utilização do mesmo. Por fim, o capítulo 6 apresenta as contribuições e conclusões desse trabalho de mestrado, além dos trabalhos futuros e das publicações alcançadas no decorrer do curso de mestrado.

CAPÍTULO

2

FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta alguns conceitos que serão utilizados neste trabalho. Inicialmente, são definidas a Teoria Componencial das Emoções e a Abordagem do Espaço Emocional Semântico, exploradas nesta pesquisa. Subsequentemente, discute-se técnicas da área de ML que auxiliarão no desenvolvimento do trabalho. E, finalmente, faz-se uma discussão sobre tecnologias e sua utilização no contexto proposto.

2.1 Teoria Componencial das Emoções

A emoção é uma reação complexa que envolve todo o organismo do indivíduo e tem uma relação direta com as suas necessidades, metas, valores e bem-estar em geral. Nesse sentido, a fim de determinar o significado das emoções, diferentes componentes são estudados em pesquisas sobre o tema (FONTAINE *et al.*, 2002; SCHERER, 2005). Alguns pesquisadores se concentram nos antecedentes das emoções e na avaliação de situações (FRIJDA, 1986; PARKINSON, 1995; SCHERER, 2001), em termos do bem-estar de um indivíduo (ELLSWORTH; SCHERER, 2003) e de padrões de reação psicofisiológica [1] (EKMAN; FRIESEN, 1981; STEMMLER, 2003). O comportamento manifesto em experiências emocionais também recebe atenção da literatura (FRIJDA; KUIPERS; SCHURE, 1989). Uma das pesquisas mais conhecidas da emoção centra-se nas suas experiências subjetivas (YIK; RUSSELL; BARRETT, 1999), bem como na inibição e no controle sobre as mesmas (FONTAINE *et al.*, 2002).

Diante do exposto, a Teoria Componencial das Emoções (STEMMLER, 2003; SCHERER, 2005; NIEDENTHAL; KRAUTH-GRUBER; RIC, 2006) entende as emoções como níveis variáveis de mudanças inter-relacionadas entre um conjunto de cinco componentes: avaliações cognitivas, experiências subjetivas, tendências comportamentais, reações fisiológicas e expressões motoras. Mahlke e Minge (2008) destacam que uma abordagem baseada em componentes

[1] Psicofisiológica também chamada de fisiopsicológica é o estudo das relações entre fenômenos psíquicos e fisiológicos.

oferece uma maneira mais abrangente para entender os diferentes aspectos das emoções em todos os tipos de situações. Assim, diferentes aspectos podem ser utilizados para a análise emocional sem, no entanto, desconsiderar a relação entre os diferentes componentes (SCHERER, 2001). A seguir, são detalhados os cinco componentes da Teoria Componencial das Emoções:

1. *Avaliação Cognitiva:* está ligada à interpretação de uma situação e contribui para o desenvolvimento das emoções. O indivíduo avalia constantemente o mundo ao seu redor e procura perceber as qualidades afetivas de eventos, objetos e recursos (FONTAINE *et al.*, 2002); um processo complexo, dinâmico e com diversos aspectos (SCHERER, 2001; GONÇALVES *et al.*, 2011; GONÇALVES *et al.*, 2013b). Essa constante avaliação implica em utilizar recursos sensoriais, perceptivos e cognitivos, incluindo o pensamento racional, ou seja, uma mesma situação pode induzir diferentes emoções, dependendo da interpretação do usuário (SCHERER, 2005; MAHLKE; MINGE, 2008).

2. *Experiência Subjetiva ou Sentimento Subjetivo:* é responsável pela avaliação de um episódio emocional, ou seja, da experiência subjetiva consciente que se relaciona com a capacidade do indivíduo de regular as emoções (DESMET, 2003). Tais experiências são consideradas percepções: sentir medo é perceber-se com medo. Mais tecnicamente, Yik, Russell e Barrett (1999) caracterizaram esse sentimento como uma meta-experiência, pois os dados brutos em que eles se baseiam incluem outras experiências, como, por exemplo, *feedback* somatossensorial, avaliação do evento provocado, atribuição, crenças, desejos, planos e comportamentos. O objetivo desse componente é monitorar o estado interno do corpo e a interação entre o organismo e o ambiente (SCHERER, 2005). De acordo com Scherer (2005), não há nenhum método objetivo capaz de avaliar os sentimentos subjetivos, pois a única maneira de se obter tais informações é perguntando ao próprio indivíduo.

3. *Tendências Comportamentais:* preparam e orientam o organismo para (ou durante) uma determinada ação (FONTAINE *et al.*, 2002). Esse componente aponta a direção e a energia de repostas emocionais específicas para o desempenho de tarefas que preparam o indivíduo para uma determinada situação (EKMAN, 2006). Exemplos desse componente são: o tempo necessário para completar uma tarefa, o número de erros, a precisão para se alcançar um objetivo, o número de pedidos de ajuda e o número de ideias criativas durante a interação (FONTAINE *et al.*, 2002; MAHLKE; MINGE, 2008).

4. *Reações Fisiológicas, também conhecidas como Mudanças Psicofisiológicas (FONTAINE et al., 2007), Neurofisiológicas (HAYASHI; BARANAUSKAS, 2011) ou Ativação Fisiológica (LAURANS; DESMET; HEKKERT, 2009):* são aspectos emocionais que permitem ao usuário revelar espontaneamente e inconscientemente suas emoções. De acordo com alguns autores (STEMMLER, 2003; SCHERER, 2005; MAHLKE; MINGE, 2008) esse componente é responsável por regular: (i) o Sistema Nervoso Central, que recebe informações dos cinco sentidos e direciona as ordens ao sistema motor; (ii) o Sistema Neuroendócrino,

que é responsável pela regularização das mudanças internas do corpo e (iii) o Sistema Nervoso Autônomo, que controla as funções do corpo, tais como: respiração, digestão, circulação, entre outros.

5. *Expressões Motoras, também conhecidas como Reações Expressivas:* são responsáveis por comunicar as tendências comportamentais (SCHERER, 2005; MAHLKE; MINGE, 2008; XAVIER; GARCIA; NERIS, 2012). Esse componente implica em alterações nas expressões faciais, vocais e gestuais que acompanham a experiência emocional do usuário. Assim, o rosto e a voz sofrem mudanças com o grau de excitação (LANG *et al.*, 1993). Em termos de respostas emocionais, têm-se as expressões motoras relacionadas as características do discurso da fala, como a velocidade, a intensidade, a melodia e o som (LEVENTHAL, 1984) ou relacionadas as características do rosto, como um olhar de ódio, o franzir da testa, a compressão dos lábios ou até mesmo um sorriso.

Considerando os aspectos multidimensionais do espaço semântico (SCHERER, 2005), que será abordado na seção seguinte, o uso dessa abordagem faz sentido, pois considera a abrangência que a proposta reflete sobre os aspectos emocionais do indivíduo, a saber: (i) avaliações cognitivas; (ii) sentimentos subjetivos; (iii) tendências comportamentais; (iv) reações fisiológicas; e (v) expressões motoras (SCHERER, 2005; MAHLKE; MINGE, 2008; XAVIER; GARCIA; NERIS, 2012). Este trabalho considera aspectos emocionais relacionados as **Expressões Motoras**, mais especificamente as características da face do usuário. Este componente foi selecionado devido a facilidade da captura das imagens dos usuários por dispositivos computacionais (por exemplo, tablet, desktop, notebook, smartphone e sistemas embarcados), visto que normalmente possuem uma câmera e a interação é realizada em frente ao dispositivo.

2.2 Discussão do Espaço Emocional Semântico

Como dito anteriormente, as emoções são complexas, difíceis de identificar e de mensurar. Nesse sentido, é importante identificar uma maneira de classificá-las, ou seja, relacionar as diferentes emoções existentes. Sabe-se que as emoções são distintas e alguns modelos foram propostos (RUSSELL, 1980; SCHERER, 2005) originalmente para auxiliar os psicólogos a representar a estrutura da experiência afetiva. De acordo com Russell (1980), o reconhecimento da emoção é baseado na ciência psicológica e cognitiva, que afirma haver duas visões primárias sobre a representação das emoções: (i) contínuas e (ii) categóricas.

A visão contínua aborda as emoções como pontos em um espaço multidimensional e utiliza escalas contínuas ou bases dimensionais. Assim, um estado emocional é caracterizado em termos de um pequeno número de dimensões latentes. Nesse ponto de vista, estados afetivos são sistematicamente relacionados entre si e não discretos e independentes uns dos outros. A abordagem dimensional mais comum é o modelo *circumplex* de Russell (1980) (Figura 1), o

qual argumenta que todas as emoções se encontram em um espaço bidimensional contínuo e suas dimensões são valência (positiva ou negativa, eixo X) e excitação (o nível de energia ou de excitação associado à emoção, eixo Y). Resumidamente, o modelo é composto por uma estrutura em formato de círculo, por meio de um escalonamento em dois domínios principais: prazer e desprazer.

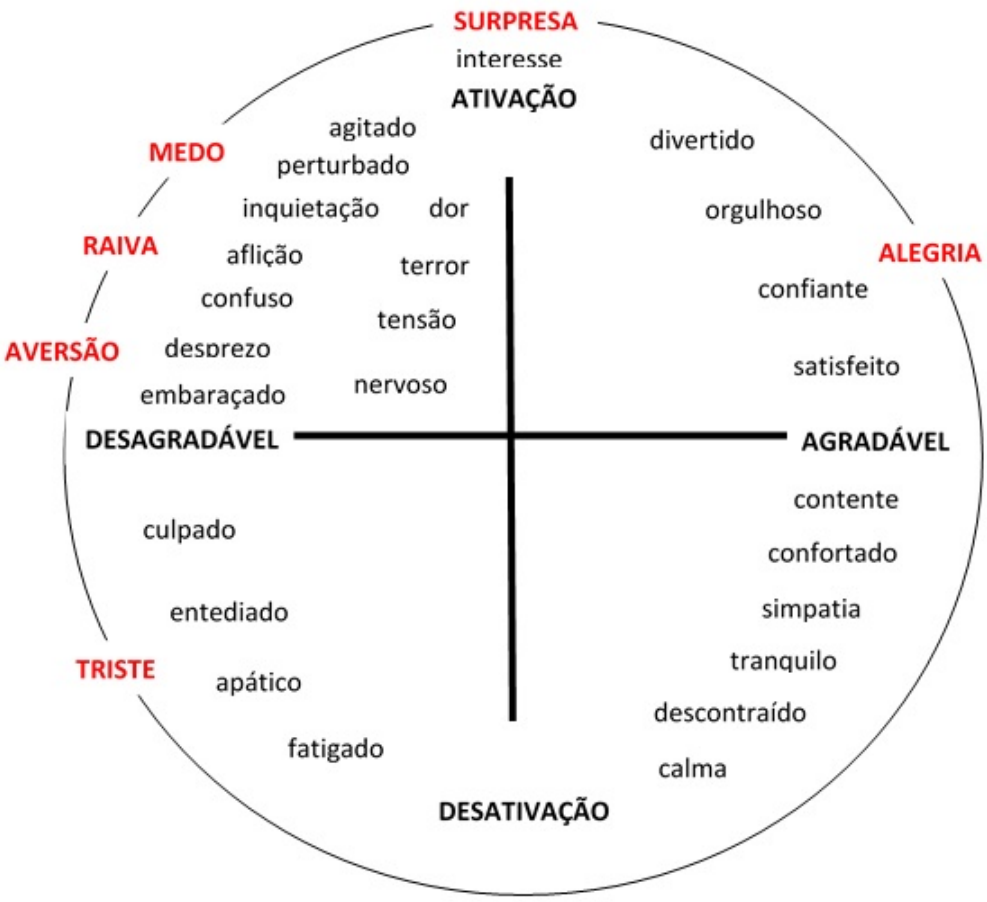

Figura 1 – Modelo *circumplex* proposto por Russell (1980).

Na representação categórica, diferentes emoções são mapeadas em categorias distintas que utilizam os componentes da Teoria Componencial das Emoções, descrita na seção anterior. Neste trabalho foram utilizadas as *Expressões Motoras*, o componente responsável por comunicar as tendências comportamentais do indivíduo (SCHERER, 2005; MAHLKE; MINGE, 2008; XAVIER; GARCIA; NERIS, 2012). Esse componente das emoções implica em alterações nas expressões faciais que acompanham a experiência emocional do usuário. O exemplo mais comum dessa descrição são as seis emoções básicas (alegria, aversão, medo, raiva, surpresa e tristeza) e o conjunto de expressões faciais relacionadas a elas, sendo inatas e universais entre as culturas. O termo *universal* indica que os mesmos movimentos musculares são produzidos ao realizar as expressões faciais. Todas as outras categorias emocionais são, então, construídas a partir de combinações dessas emoções básicas. Essa abordagem é apoiada pelos estudos transculturais realizados por Ekman (1973), os quais indicam que os seres humanos percebem certas emoções

básicas transmitidas por expressões faciais da mesma forma, independentemente da cultura, da etnia e dos costumes.

A principal vantagem de uma representação por meio de um esquema categórico é que as pessoas usam tal esquema para descrever demonstrações emocionais observadas no cotidiano de forma semelhante. O sistema de rotulagem baseado em categoria é muito intuitivo e, no geral, corresponde à experiência das pessoas.

De acordo com Martinez e Du (2012), as emoções são identificadas por seres humanos de maneiras distintas e os autores classificam as emoções em três grupos: (i) alegria e surpresa, fáceis de identificar; (ii) raiva e tristeza, um pouco mais difíceis e (iii) medo e aversão, ainda mais difíceis. A seguir, na Figura 2, encontram-se as emoções analisadas. O estado neutro também é considerado, o qual é utilizado como referência para a detecção de estados emocionais.

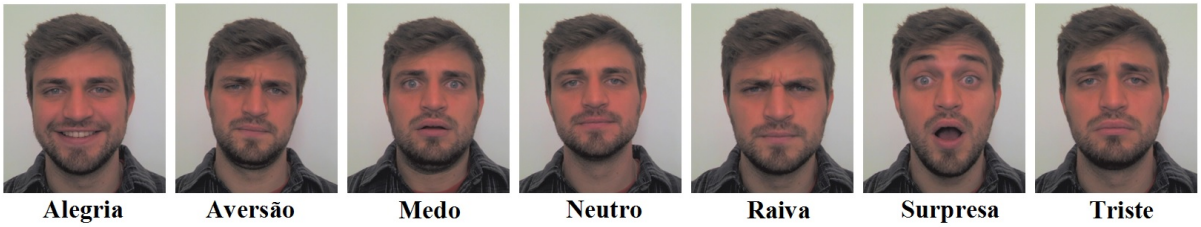

Figura 2 – Emoções básicas propostas por Ekman (2006).

É importante notar que ambas as representações de emoções têm sido utilizadas eficientemente em várias aplicações computacionais (SCHERER, 2005; GONÇALVES *et al.*, 2013a; LIBRALON; ROMERO, 2014). O largo espectro de aplicações e o constante aumento da capacidade de processamento computacional vêm motivando pesquisadores a tentar identificar emoções de usuários em sistemas computacionais e utilizar tal informação como base para a tomada de decisões e para a execução de tarefas cada vez mais complexas. De fato, procedimentos de classificação têm ajudado na análise de repostas emocionais, auxiliando médicos no diagnóstico da depressão, usuários que buscam uma mudança de comportamento em sua rotina, entre outros.

2.3 Mecanismos de Classificação da Emoção

Segundo Desmet (2003), cada emoção está associada a um determinado padrão de expressão motora. O autor apresenta um exemplo em que a "fúria vem com um olhar fixo, sobrancelhas contraídas, lábios comprimidos, movimentos vigorosos e rápidos e, geralmente, uma voz com tom elevado, quase gritando". Assim, as expressões motoras estão relacionadas a expressões faciais, a gestos corporais e a características do discurso da fala (DESMET, 2003). Complementando a ideia Mahlke e Minge (2008) sugerem que existe uma relação entre a dilatação da pupila e o prazer de um estímulo, sendo que estímulos agradáveis são suscetíveis a pupilas mais dilatadas.

Pesquisas na área de ML têm, nas últimas décadas, gerado uma ampla quantidade de métodos que "aprendem" a partir de exemplos e que podem ser utilizados para extrair, a partir de padrões, dados empíricos para a classificação (DUDA; RICHARD, 1973; CLARK; NIBLETT, 1989; CRISTIANINI; SHAWE-TAYLOR, 2000; UEYAMA *et al.*, 2014). Tais técnicas de classificação estão cada vez mais sendo aplicadas na análise de respostas emocionais (ØHRN; ROWLAND, 2000; CHANEL *et al.*, 2009; ZHOU *et al.*, 2011; GONÇALVES *et al.*, 2013a; RAMAKRISHNAN; EMARY, 2013; LIBRALON; ROMERO, 2014). A seguir, encontra-se uma súmula conceitual dos algoritmos de classificação utilizados na literatura (Seção 3 no reconhecimento de emoções:

- *k-Nearest Neighbor classification (kNN)* - Classificação dos k-vizinhos mais próximos. O kNN, utilizado em alguns trabalhos (BAILENSON *et al.*, 2008; YUAN *et al.*, 2010), é uma abordagem de classificação não-paramétrica, isto é, não assume nenhuma distribuição de probabilidade para os dados *a priori*. O algoritmo kNN encontra um grupo de k objetos no conjunto de treino que estão mais próximos do objeto de teste e o classifica com a classe predominante dos seus k-vizinhos (FACELI, 2011).

- *Support Vector Machine (SVM)* - O algoritmo SVM, utilizados por Ramakrishnan e Emary (2013), Peter e Urban (2012) e Littlewort *et al.* (2011) para o reconhecimento de emoções, é um dos métodos mais robustos e acurados entre os algoritmos conhecidos. Segundo Faceli (2011), o objetivo do SVM é encontrar a melhor função para distinguir duas ou mais classes dos dados de treinamento. A garantia da melhor função é atingida, pois o algoritmo maximiza a margem entre as classes e oferece melhor habilidade de generalização (CLARK; NIBLETT, 1989; LITTLEWORT *et al.*, 2011; RAMAKRISHNAN; EMARY, 2013).

- *Fuzzy Logic* - A Lógica Fuzzy, utilizada nos trabalhos de Katsis *et al.* (2008) e Peter e Urban (2012), é capaz de processar dados incompletos e de prover soluções aproximadas para problemas que outros métodos encontram dificuldades em solucionar (CLARK; NIBLETT, 1989; KATSIS *et al.*, 2008).

- *Decision or Regression Tree* – A árvore de decisão ou regressão, aplicada nos trabalhos de Bailenson *et al.* (2008) e Peter e Urban (2012), utiliza-se de uma árvore para realizar a classificação ou a estimação de um valor de um dado teste. Um dos algoritmos mais famosos para gerar essas árvores é o C4.5, cuja árvore cresce a partir de uma abordagem "dividir e conquistar" e é ramificada utilizando o atributo que obtiver o melhor ganho de informação (DUDA; RICHARD, 1973).

- *Bayesian Networks* – A Rede Bayesiana é um modelo de grafo probabilístico, aplicado aos trabalhos de Peter e Urban (2012), Ramakrishnan e Emary (2013) e Chanel *et al.* (2009), e que representa um conjunto de variáveis aleatórias e suas dependências condicionais

como um grafo acíclico [2] direto (FACELI, 2011). Por exemplo, uma rede bayesiana pode representar a relação entre os sintomas e as doenças. Dados os sintomas, pode-se inferir a probabilidade de várias doenças. No caso desse trabalho em particular, dadas as características da face pode-se inferir a probabilidade de determinada emoção.

No entanto, o trabalho apresentado por Santos (2008) evidencia o uso de Comitê de Classificadores para melhorar a classificação, o que é mostrado nos parágrafos a seguir.

Apesar de uma ampla utilização, os classificadores gerados a partir de métodos de ML dificilmente alcançam 100% de acurácia (CANUTO, 2001; MONARD; BARANAUSKAS, 2003; SANTOS, 2008). Isso se deve ao fato do desempenho de cada método depender do ajuste de vários parâmetros, bem como do grau de dificuldade associado a cada problema em particular. As várias configurações que podem ser dadas a esses parâmetros resultam em diferentes decisões de classificação. A escolha da melhor configuração normalmente ocorre a partir da construção de diversos classificadores com diferentes configurações. Em seguida, tais classificadores são testados e o que apresentar melhor desempenho é escolhido, enquanto os demais são descartados (CANUTO, 2001; SANTOS, 2008).

A seleção de um único classificador descarta uma significante quantidade de informações potencialmente relevantes. Por esse motivo, o conceito de EC tem sido identificado como uma solução promissora para o desenvolvimento de sistemas de alto desempenho na área de reconhecimento de padrões (SCHULLER *et al.*, 2005; SANTOS, 2008). Na literatura, um Comitê de Classificadores também possui as seguintes denominações: *Ensemble de Classificadores, Combinação de Especialistas, ou Classificadores Modulares* (DUDA; HART; STORK, 2012). Esse fato se deve à premissa de que a combinação de classificadores pode conduzir a uma melhoria no desempenho em termos de melhor generalização e/ou de aumento da acurácia (SHARKEY, 1999).

Um dos aspectos que garante um bom desempenho de um EC é a diversidade entre seus componentes (DUDA; HART; STORK, 2012). Para que os classificadores base sejam considerados diversos entre si, eles devem apresentar erros não correlacionados, ou seja, não devem cometer erros nos mesmos padrões. Desse modo, para que um Comitê obtenha um desempenho aceitável, ele deverá ser formado por classificadores que possuam acurácia razoável e que não cometam erros coincidentes, para que os erros de um classificador possam ser corrigidos pela opinião de todos os outros componentes (SHARKEY, 1999; SCHULLER *et al.*, 2005; SANTOS, 2008; DUDA; HART; STORK, 2012). A Figura 3 apresenta a estrutura do processo de combinação de um EC. A primeira camada consiste em classificadores individuais que podem ter diferentes arquiteturas, mas com saídas em uma forma comum, tal como um ordenamento das classes candidatas. A segunda camada consiste em um processo que opera nas decisões da camada anterior para gerar a decisão geral do EC.

[2] Acíclico - aquilo que não é cíclico, ou seja, sua ocorrência é aleatória, mas pode se prever a sua frequência, pelo uso da estatística.

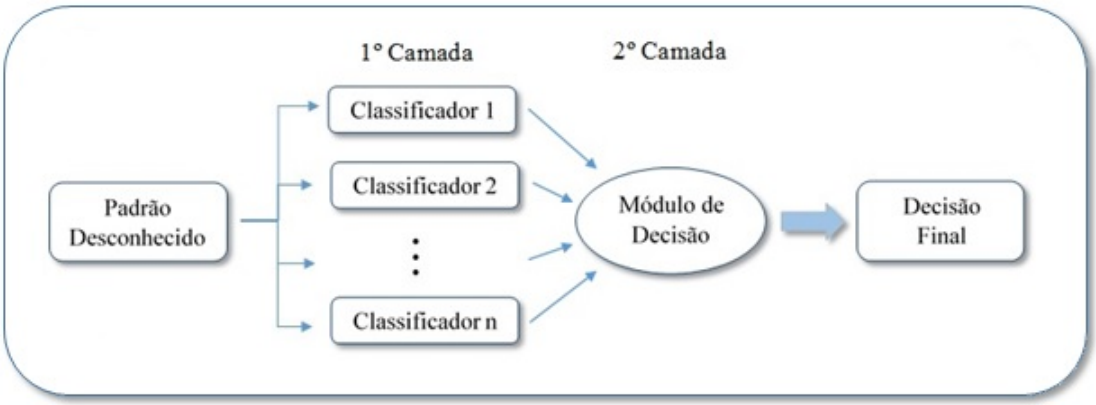

Figura 3 – Processo de combinação de decisão de um EC (SCHULLER *et al.*, 2005).

A área de classificação de dados visa construir sistemas que aprendam com os dados. No contexto de reconhecimento de emoções, os algoritmos de classificação utilizam o tipo de aprendizado supervisionado. Os algoritmos de aprendizado supervisionado são treinados com exemplos rotulados, isto é, a saída da entrada é conhecida, por exemplo, uma imagem que contém uma face com expressão já conhecida (CLARK; NIBLETT, 1989; CRISTIANINI; SHAWE-TAYLOR, 2000).

2.4 Captura de Expressões Faciais

Esta seção apresenta os métodos utilizados para o mapeamento da face do usuário. Inicialmente, apresenta-se o método para o reconhecimento da face e, posteriormente, o método para a extração de suas características.

Considerando os aspectos multidimensionais do espaço semântico proposto por Scherer (2005) e a utilização da abordagem baseada na Teoria Componencial da Emoção, atribuem-se aspectos emocionais relacionados às expressões motoras, mais especificamente as características da face do usuário, como dito anteriormente. Como demonstrado por Tian, Kanade e F. (2005), a análise da expressão facial pode ser dividida em três etapas principais: (i.) a detecção do rosto, (ii.) a extração de características faciais e (iii.) a classificação. A detecção do rosto é a fase de processamento que encontra automaticamente a região do rosto. Essa etapa não é trivial devido alguns aspectos importantes, tais como, a movimentação da cabeça, a iluminação das imagens, a presença de cabelo e de óculos. Após a detecção da face, pode-se extrair e representar as mudanças faciais produzidas por cada expressão, pois vários tipos de estímulos são perceptivos aos estados emocionais exibidos pela face.

Diante disso, o sistema de Reconhecimento Facial utilizado neste estudo e desenvolvido por SANTOS V.C.; COCA (2013) foi concebido em dois módulos distintos: a Extração Automática de Características Faciais (EACF) e o Reconhecimento de Expressões Faciais (REF)

(SANTOS V.C.; COCA, 2013). As imagens recebidas a partir da câmera são analisadas pelo módulo EACF, que extrai as características faciais. Tal módulo processa a imagem, procura por um rosto humano e o analisa de modo a produzir uma série de características. Além disso, ele envia esses dados para o módulo de REF, que, por sua vez, determina as características expressivas. Para isso, foi utilizado o sistema idealizado por Saragih, Lucey e Cohn (2011), que desenvolve uma estratégia de otimização de modelagem parametrizada de uma imagem, de modo que seus quadros correspondam aos locais de interesse. Assim, o EACF e o REF formam o módulo *FaceTracker*, ilustrado pela Figura 4.

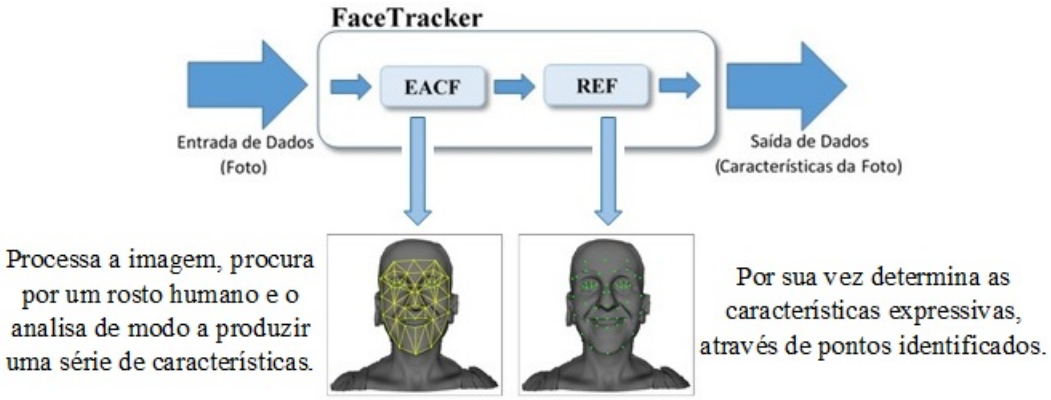

Figura 4 – Módulo *FaceTracker* (SANTOS V.C.; COCA, 2013).

O módulo *FaceTracker* é responsável por receber o conjunto de características produzido e efetuar um mapeamento adequado entre as características obtidas a partir da face humana.

CAPÍTULO

3

TRABALHOS RELACIONADOS

Considerando que este projeto de pesquisa propõe uma metodologia para desenvolver e avaliar um modelo capaz de identificar e inferir o estado emocional do usuário em tempo de execução. Neste capítulo, serão apresentados alguns trabalhos publicados em conferências, periódicos e em livros que estão relacionados a esta proposta.

3.1 Discussão dos Principais Trabalhos

Diante da pesquisa realizada em repositórios acadêmicos na área de IHC, Classificação de Emoção, Computação Persuasiva, Abordagem Flexível, foram selecionados os seguintes trabalhos relacionados por serem mais similares ao trabalho desenvolvido:

- *Real-time classification of evoked emotions using facial feature tracking and physiological responses* (BAILENSON *et al.*, 2008) (Qualis A1).
  A pesquisa proposta por Bailenson *et al.* (2008) apresenta um modelo baseado em algoritmos de classificação – Regressão Linear e Rede Neural Artificial – para a análise de emoções em tempo real. A análise foi feita a partir de vídeos de 41 indivíduos assistindo filmes que causavam emoções diferentes. Um dos sensores faz o reconhecimento das emoções faciais e o outro faz a análise da variabilidade da frequência cardíaca e da atividade eletrodérmica do usuário. Cinco especialistas fizeram a análise dos vídeos com o objetivo de definir as rótulos (saídas) para os algoritmos de classificação, por exemplo, analisando se o usuário estava triste em determinado momento. Bailenson *et al.* (2008) apresenta algoritmos baseados em pontos extraídos do rosto e em respostas fisiológicas do usuário. Essa pesquisa trabalha apenas com o eixo da Valência (eixo X do modelo proposto por Russell (1980), mais especificamente, com as emoções: diversão, neutro e tristeza.

  Neste trabalho de mestrado, obordou-se todos os quadrantes dos domínios do espaço semântico de Russell (1980), tendo como objetivo identificar emoções de usuários inte-

ragindo com sistemas computacionais em tempo real, considerando expressões motoras (face) como componentes da emoção.

- *Affect prediction from physiological measures via visual stimuli* (ZHOU *et al.*, 2011) (Qualis A1).
  Zhou *et al.* (2011) desenvolveram um estudo com o objetivo de prever diferentes estados emocionais do usuário por meio de reações fisiológicas, comparando três tipos de modelos computacionais. Um experimento foi desenvolvido para provocar respostas emocionais com fotos de emoções padronizadas ao passo que algumas reações fisiológicas foram medidas. Três algoritmos de classificação – Regras de Decisão, kNN e Árvore de Decomposição – foram aplicados à construção de modelos de previsão das características fisiológicas extraídas. A pesquisa desenvolvida por Zhou *et al.* (2011) utilizou apenas sensores que analisavam as reações fisiológicas e utilizam a fala do usuário para a definição das classes (saídas), por exemplo, quando o indivíduo dizia: "Ao olhar essa imagem, eu fico triste".

  Este trabalho, no entanto, se diferencia da pesquisa realizada já que tenta trazer uma visão mais abrangente e assertiva para a análise das emoções e considera somente um dos componentes apontados por Scherer (2005), nesse caso as expressões motoras.

- *This computer responds to user frustation: theory, design and results* (KLEIN; MOON; PICARD, 2002) (Qualis B1).
  Klein, Moon e Picard (2002) sugeriram uma solução computacional para auxiliar os usuários a gerenciar e a mudar o estado emocional da frustração. Durante a interação com um sistema computacional, foi feita uma análise da interação do usuário apenas com textos e botões. Um estudo de viabilidade do sistema considerou os momentos em que as emoções dos usuários foram ignoradas e se os usuários foram capazes ou não de informar indícios de seus estados emocionais para o computador. O sistema solicitava informações sobre o estado emocional do indivíduo durante toda a interação e fornecia *feedback*, especialmente sobre o conteúdo emocional. O objetivo era incentivar o usuário a sentir-se satisfeito e informar que ele não é o único a ter esse tipo de emoção.

  Klein, Moon e Picard (2002) propuseram uma abordagem que percebe o estado emocional do usuário por meio de mensagens fornecidas durante a interação. Esse caminho traz indícios do estado emocional do usuário e pode não ser o mais eficiente, tendo em vista que esse nem sempre consegue verbalizar o que está sentindo. Os autores propõem tratar apenas as questões relacionadas à frustração, presente no quadrante 1 da Estrutura do Espaço Semântico de Russell (1980), ao passo que, nesse trabalho, propõe-se abordar o maior número possível de emoções.

- *Emotion in human-computer interaction* (PETER; URBAN, 2012) (Livro).
  Peter e Urban (2012) apresentaram uma abordagem geral de reconhecimento de emo-

ções que utiliza sensores fisiológicos e desenvolveram o Sistema de Reconhecimento de Emoções (EREC). Esse sistema coleta dados de alguns tipos de sensores: atividade eletrodérmica, temperatura do ar atmosférico e variabilidade da frequência cardíaca. Os autores propuseram duas aplicações que utilizam o sistema EREC: uma ferramenta de avaliação de usabilidade emocional e um ambiente de emphe-learning [1] emocional. A primeira aplicação busca integrar a emoção reconhecida no processo de avaliação de usabilidade ao intuito de identificar sobrecarga cognitiva, cansaço ou desorientação, identificando, assim, problemas com a usabilidade da interface. A segunda aplicação utiliza o reconhecimento da emoção para evitar regiões emotivas que são negativas para o aprendizado, provendo outras formas de interação com o conteúdo a ser aprendido. No cenário real, sensores fisiológicos podem não ser usados corretamente, o que afeta a confiabilidade dos mesmos, ou ainda, não serem capazes de permanentemente prover informação afetiva do usuário na vida cotidiana.

Portanto, esta trabalho de mestrado visa utilizar os resultados de sensores visuais, com o objetivo de aumentar a acurácia do reconhecimento da emoção e reduzir situações em que os sensores podem apresentar dados errôneos.

- *MoodScope: Building a mood sensor from smartphone usage patterns* (LIKAMWA *et al.*, 2013) (Qualis A1).
  Esse trabalho propõe uma abordagem que infere o humor do usuário a partir do padrão de uso do smartphone, chamado *MoodScope*. LiKamWa *et al.* (2013) realizaram um estudo com 32 usuários durante dois meses, atingindo uma acurácia de 66% no início, com um modelo estatístico genérico e 93% após dois meses de treinamento personalizado. Segundo LiKamWa *et al.* (2013), a detecção automática de humor permite a criação de novos aplicativos e a melhoria de outros existentes. Um aplicativo que registra o humor do usuário e permite que seja consultado pelo próprio usuário em outras situações, por exemplo: local e tempo, pode melhorar a consciência do usuário sobre suas mudanças de humor. O *MoodScope* se diferencia de outras aplicações de detecção de humor (TECH, 2013; WESTERGREN, 2013) por não utilizar dados de sensores físicos dos smartphones, nem de áudio e nem de vídeo. LiKamWa *et al.* (2013) utilizam um modelo de reconhecimento de humor baseado em seis tipos de informações: SMS, e-mail, registro de ligações, uso de aplicações, navegação na Web e localização. O design do sistema apresenta duas aplicações: uma que coleta os dados em plano de fundo; e outra utilizada para o usuário reportar o seu humor. A pesquisa desenvolvida por LiKamWa *et al.* (2013) tem uma abordagem distinta deste trabalho de mestrado, pois não utiliza dados de sensores físicos (imagem). Outro fator interessante é que não há uma captura de dados em tempo real, tendo em vista que o usuário tem que reportar o seu humor quatro vezes ao dia à aplicação.

[1] O e-learning (do inglês *electronic learning*, "aprendizagem eletrônica") ou ensino eletrónico corresponde a um modelo de ensino não presencial apoiado em tecnologia.

3.2 Síntese dos Trabalhos Relacionados

A Tabela 1 apresenta uma síntese das pesquisas detalhadas anteriormente. Estão descritas as pesquisas, os tipos de sensores usados para a captura dos dados, o método utilizado para a extração das características, os tipos de algoritmos de classificação utilizados para a análise e para a classificação da experiência emocional do usuário e os sujeitos responsáveis pela definição dos rótulos (saídas) usadas para treinar esses algoritmos de classificação.

Tabela 1 – Síntese das pesquisas identificadas na literatura.

Referência	**Captura**	**Extração de Dados**	**Análise e Classificação**	**Como foi identificada o rótulo**
Bailenson et al., 2008	Atividade eletrodérmica, reconhecimento facial, frequência cardíaca	-	Regressão Linear e Rede Neural Artificial	Análise do especialista
Zhou et al., 2011	Atividade eletrodérmica, eletroencefalograma, eletromiografia facial, frequência respiratória	-	Regras de decisão, kNN e árvore de decomposição	Fala do usuário
Klein at al., 2002	Análise textual	-	-	-
Peter, C.; Urban, B., 2012	Atividade eletrodérmica, frequência e batimento cardíaco, temperatura da pele e do ar	Extração de Características (PICARD, 2004)	Árvores de Decisão, Lógica Fuzzy; SVM e Redes Bayesianas	-
LiKamWa et al., 2013	SMS, e-mail, ligações, uso de aplicações, navegação na web, localização	10 Frequências normalizadas do histograma	Regressão Multilinear e modelo personalizado	25 diferentes tipos de emoções
Proposta	Face do Usuário	Seleção de pontos da face (SARAGIH; LUCEY; COHN, 2011)	Comitê de Classificação	Imagens classificadas pelo usuário

Os trabalhos apresentados trazem capturas relacionadas as Reações Fisiológicas predominantemente, fato que pode comprometer a riqueza de detalhes de uma análise emocional. A partir desse levantamento bibliográfico, é possível perceber que não há consenso referente ao responsável pela definição dos rótulos (saídas) a serem usados no treinamento desses algoritmos de classificação. O trabalho desenvolvido por Bailenson *et al.* (2008) aponta que podem ser os especialistas (por exemplo, psicólogos); já outras pesquisas (ZHOU *et al.*, 2011; LIKAMWA *et al.*, 2013) dizem que o próprio usuário pode realizar essa definição. Grande parte dos trabalhos

encontrados na literatura têm focado nos domínios da Valência e da Excitação para a classificação das emoções (ZHOU *et al.*, 2011; BAILENSON *et al.*, 2008; LIKAMWA *et al.*, 2013). Avaliar a experiência emocional em termos de Valência e de Excitação em um sistema computacional é um processo importante nas decisões de design, uma vez que, é possível identificar, respectivamente, o grau de satisfação e de motivação do usuário durante a interação. Todavia, outros domínios da experiência emocional como Sentimento de Controle e Facilidade de Conclusão do Objetivo também devem ser considerados durante a avaliação das emoções. Isso se deve à possibilidade de que a associação das emoções aos quatro domínios da experiência emocional (Valência, Excitação, Sentimento de Controle e Facilidade de Conclusão do Objetivo) (SCHERER, 2005) permita que aplicações computacionais se flexibilizem a partir dessa abordagem. No entanto, o modelo proposto por Scherer (2005) é uma extesão do modelo proposto por Russell (1980) se tornando bastante complexo a pesquisa sem o apoio de um especialista (psicólogo) para auxiliar no desenvolvimento do trabalho.

Nesse sentido, o trabalho aborda a classificação da emoção baseada no Espaço Emocional Semântico por meio das emoções proposta por Ekman (2006) (visão categórica) e sua classificação baseada na visão contínua proposto por Russell (1980) (visão contínua) na interação dos usuários com dispositivos computacionais.

CAPÍTULO

4

ABORDAGEM PARA PERSUASÃO FLEXÍVEL

Este capítulo apresenta a visão geral do projeto desenvolvido ao longo do mestrado, os modelo utilizados e seus objetivos.

4.1 Visão Geral do Projeto

Aliado ao crescimento dos usuários de sistemas computacionais e ao acesso facilitado aos meios de telecomunicação, têm-se a possibilidade de capturar e de analisar dados de indivíduos em diversos contextos, assim como o processamento remoto (*offload* [1]). Diante disso, a detecção da emoção pode ser utilizada para adaptar o comportamento de sistemas computacionais as necessidades de cada usuário ou persuadí-lo à transição para uma emoção positiva (por exemplo, de neutro para feliz). Sendo uma proposta que visa a maior interação pessoal com equipamentos tecnológicos, fica evidente a inovação dos métodos de interação do usuário, dentre os vários softwares de uso diário.

O trabalho desenvolvido visa contribuir com a área de IHC frente a uma proposta que enfatiza a interação do usuário com sistemas computacionais, considerando seus aspectos emocionais. Destaca-se o interesse em investigar metodologias aplicadas à análise e à classificação de respostas emocionais durante a interação, visando permitir que aplicações computacionais possam se flexibilizar a partir de tal classificação e deixar as aplicações computacionais cientes do estado emocional do usuário. O trabalho se insere no contexto de *Design Emocional*, que considera a captura de respostas emocionais por meio de sensores e as analisa com o apoio de um *framework* que julga a Estrutura do Espaço Emocional Semântico para classificar o estado emocional do usuário Scherer (2005).

A detecção da emoção pode ser utilizada para o desenvolvimento de novas maneiras

[1] Na ciência da computação, *offload* refere-se à transferência de certas tarefas de computação para uma plataforma externa, tal como um *cluster* ou uma nuvem. *Offloading* pode ser necessário devido a limitações de um sistema de lidar com uma determinada tarefa em seu próprio *hardware*.

de interação do usuário e, posteriormente, caminhar para os dispositivos cognitivos (CAMPBELL; CHOUDHURY, 2012). Tecnologias da informação que possam intervir, caso o usuário esteja estressado ou mau humorado, indicando músicas que possam acalmá-lo, por exemplo e, assim, deixá-lo um pouco mais feliz e/ou reduzir o estresse é um dos caminhos para atender as preferências e as necessidades dos usuários.

4.2 Extração das Características da Face

Neste trabalho, a abordagem para a identificação do rosto e para a análise da expressão facial é baseada em características geométricas. Assim, é possível modelar a forma e a localização de componentes faciais (incluindo a boca, os olhos, as sobrancelhas e o nariz) pelo uso de pontos característicos e de elementos geométricos, como ângulos, distâncias ou áreas utilizadas para a representação geométrica do rosto. Por fim, para a realização da classificação da expressão facial, modelamos uma EC baseada nos algoritmos de ML.

Os métodos baseados em características geométricas são utilizados na modelagem facial (expressões motoras) com o objetivo de desenvolver uma abordagem que se assemelhe à maneira como os seres humanos interpretam os elementos da face. Assim, diferentes representações faciais (alegria, aversão, medo, neutro, raiva, surpresa e tristeza) são propostas para a identificação e a classificação das emoções, codificando a configuração facial de um indivíduo. O *FaceTracker* (SARAGIH; LUCEY; COHN, 2011) é caracterizado como um sistema de visão computacional utilizado para obter as informações faciais. Ele faz uso de uma estratégia de otimização que emprega aproximação linear e ajusta pontos de referência em localizações consistentes para registrar um modelo parametrizado. Com base em um modelo de referência facial composto por 66 pontos característicos, o algoritmo busca alinhar os elementos da face em análise aos pontos característicos do modelo de referência.

Nesse sentido, Libralon e Romero (2014) e Mano *et al.* (2015) modificam o algoritmo do *FaceTracker* de modo a mapear somente um subconjunto de 33 dos 66 pontos faciais obtidos inicialmente, o que diminui o custo computacional com a eliminação de possíveis redundâncias. A Figura 5 apresenta um exemplo do mapeamento da face realizada pelo algoritmo modificado do *FaceTracker*.

Os elementos considerados na representação facial têm como objetivo modelar as partes do rosto que são essenciais na relação dos movimentos e a emoção expressada de acordo com psicólogos (EKMAN, 2006). Portanto, os 33 pontos mapeados são distribuídos da seguinte forma: oito para a boca, seis para cada um dos olhos, três para cada sobrancelha e o queixo, dois para as narinas e delimitando duas extremidades laterais da face perto dos olhos. Além disso, para modelar a forma dos olhos, da boca e das regiões da face relacionadas aos movimentos musculares, oito áreas são mapeadas. Destaca-se que todas as combinações possíveis de pontos são obtidas por meio dos ângulos e das distâncias entre dois pontos distintos, o que gera uma

Figura 5 – Processo de mapeamento da face pelo *FaceTracker* (MANO *et al.*, 2015).

representação com dimensionalidade D_1 = 2 x 33 + 8 + 2 x 528 = 1130.

4.3 Uso do Comitê de Classificação para Identificar Emoções

Utilizando a hipótese citada na Seção 2.3, a Figura 6 apresenta a estrutura do módulo para a classificação da face no processo de combinação do EC. Nesse sentido, combinamos técnicas de classificação que são cada vez mais aplicadas à análise de respostas emocionais, sendo elas: *kNN, Árvore de Decisão, Lógica Fuzzy, Redes Bayesianas e SVM* (CHANEL *et al.*, 2009; GONÇALVES *et al.*, 2013a; ØHRN; ROWLAND, 2000; RAMAKRISHNAN; EMARY, 2013; ZHOU *et al.*, 2011) para o desenvolvimento do Comitê de Classificação proposto.

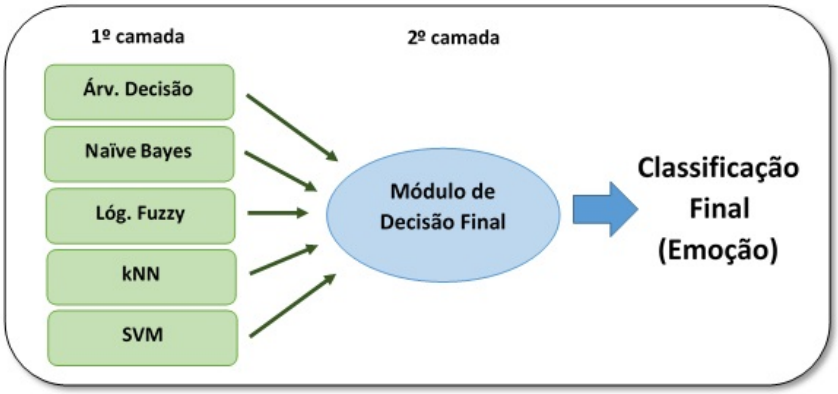

Figura 6 – Módulo baseado em Cômite de Classificação para a classificação da emoção da face do usuário (MANO *et al.*, 2015).

Tendo como entrada o resultado do mapeamento facial realizado pelo *FaceTracker*, a primeira camada de processamento consiste em classificadores individuais que podem ter diferentes arquiteturas, mas que apresentam saídas de forma comum, isto é, como a indicação de uma classe (emoção) a partir do viés de classificação do respectivo algoritmo. A segunda camada

de processamento consiste em uma tomada de decisão que opera nos resultados da camada anterior para gerar a decisão geral do EC (vide Figura 6). Vale salientar que, para a decisão final do Comitê, foi utilizada a Equação 4.1 no processo de ponderação e foram atribuídos pesos proporcionais a cada algoritmo de acordo com a sua taxa de acurácia. Na equação, $\overline{W}$ é a média ponderada da acurácia de cada algoritmo e AR corresponde à taxa média da acurácia do algoritmo. Assim, para cada algoritmo utilizado, um peso proporcional terá sido atribuído à sua taxa de acurácia.

$$\overline{Wx_i} = \frac{ARx_i}{\sum_j^n ARx_j} \tag{4.1}$$

Para avaliar o avanço e a acurácia obtidos com a utilização do módulo de classificação baseado na face, implementamos os algoritmos citados, bem como o módulo de decisão do EC. Com isso, coletamos dados sobre o comportamento e a avaliação do módulo de classificação como um todo.

4.4 *FlexPersuade*

4.4.1 Modelo e Objetivo

Esse trabalho discute o uso de uma abordagem genérica para a construção de aplicações adaptativas em dispositivos computacionais. Ueyama *et al.* (2009) afirmam que a reconfiguração em tempo de execução é uma característica fundamental para lidar com a heterogeneidade de hardware inerente a dispositivos com recursos escassos (dispositivos móveis - smartphone e nós sensores). Sendo assim, uma camada de software denominada *FlexPersuade* foi desenvolvida baseada no modelo de componentes do *OpenCom* (UEYAMA *et al.*, 2009), a qual apresenta as seguintes características: **i)** camada de software genérico; **ii)** arquitetura flexível e extensível; **iii)** independência de linguagem e; **iv)** baseada em *microkernel*, onde as funcionalidades são incrementadas quando necessário. Portanto, o *FlexPersuade* é um *framework* que apoia o desenvolvimento de sistemas que se adaptam ao estado emocional do usuário durante a interação (GONÇALVES; UEYAMA, 2012; GONÇALVES *et al.*, 2011; UEYAMA *et al.*, 2009; UEYAMA *et al.*, 2004; UEYAMA, 2005). A Figura 7 ilustra a arquitetura do *FlexPersuade*, na qual o *microkernel* instância e realiza a conexão dos componentes a fim de "persuadir" o usuário. No caso deste trabalho o *FlexPersuade* instância e conecta um player de música (em vermelho), mas podendo ser desenvolvidos outros softwares para interação com o usuário, por exempo jogos (em pontilhado).

A Figura 8 ilustra a estrutura para a interação do *FlexPersuade* com o usuário, tal como a interação do usuário com o dispositivo computacional, nesse exemplo o smartphone, e o reconhecimento da emoção realizado pelo *FlexPersuade*. Esse *framework* consiste no módulo *FaceTracker* (SANTOS V.C.; COCA, 2013), que extrai as características da face do usuário, e

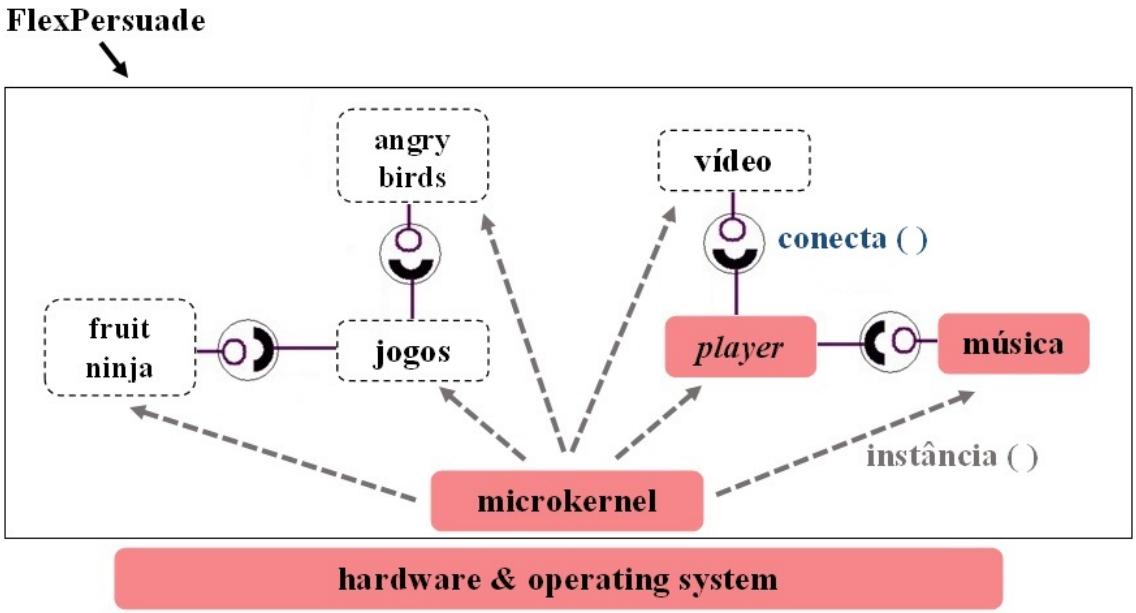

Figura 7 – Cenário de interação de componentes do *FlexPersuade*.

no Cômite de Classificação, que classifica uma emoção a partir de tais características. Assim, temos como objetivo possibilitar a interação de softwares, de maneira a "persuadir" o usuário, por exemplo, acesso a redes sociais, chats de bate-papo, jogos ou player de música. No caso do player de música, estudo de caso desta pesquisa, o modelo *FlexPersuade* carrega a playlist (conjunto de músicas) de gênero musical que melhor se adaptar ao estado emocional do usuário naquele momento, de acordo com os dados transmitidos pelo *framework* de classificação. Por exemplo, se o *framework* classifica a experiência emocional do usuário como "triste", o player irá carregar uma playlist para tentar persuadí-lo a baixar o nível de tristeza e até mesmo para uma mudança de estado emocional. Vale salientar que devido a abordagem estudada não se ater a um dispositivo computacional específico, mas com o intuito de abranger o maior número possível de dispositivos (por exemplo, tablet, desktop, notebook e smartphone) o módulo de reconhecimento e classificação da emoção foi implementado em um servidor remoto aplicando conceitos de *Cloud Computing*. Assim, independente da arquitetura e poder de processamento do dispositivo utilizado o usuário pode fazer uso do *FlexPersuade*.

Diante desse contexto, este trabalho teve como principal objetivo desenvolver e avaliar um modelo capaz de identificar e de persuadir o estado emocional do usuário em tempo de execução com base na classificação da emoção. O modelo pode interagir com o usuário para proporcionar a ele um sentimento de satisfação dentro do contexto em que se encontra. Esta proposta foi validada com um estudo de caso, onde um player de música propõe uma playlist com o objetivo de estimular o usuário de maneira positiva. O modelo desenvolvido considera como entrada um estado emocional do usuário, reconhecido por meio de um *framework* para a classificação das emoções, para a classificação automatizada de emoções através do reconhecimento das

Figura 8 – Estrutura para interação do usuário com o dispositivo computacional.

características da face do usuário.

4.4.2 Estudo de Caso: Player de Música

Desenvolveu-se uma página Web para os testes de reconhecimento da emoção e para a interação com o player de música. *Ruby on Rails* foi a linguagem de programação escolhida para a implementação devido à experiência e à familiaridade desse discente com tal linguagem. Seguindo a estrutura proposta para este trabalho (vide Figura 8), o desenvolvimento da página Web para a interação com o *FlexPersuade* foi dividido em módulos distintos, descritos a seguir:

- *Preferências do Usuário:* ao se conectar à página, é necessário que o usuário informe suas preferências musicais para que o software possa disponibilizar as músicas que mais se adaptam ao seu estado emocional, visto que cada usuário possui suas próprias preferências. Esse indica, por meio de um formulário (Figura 9), os artistas que costuma ouvir de acordo com o seu estado emocional. As emoções indicadas no formulário são as tratadas neste trabalho (alegria, aversão, medo, neutro, raiva, surpresa e tristeza).

- *Aquisição da Imagem do Usuário:* ao se conectar à página, é solicitado ao usuário o acesso à câmera de seu dispositivo computacional (móvel ou fixo) e suas imagens são adquiridas em intervalos de 5 segundos. As imagens são enviadas ao servidor, onde são processadas e classificadas, e sua emoção é identificada e armazenada em um banco de dados (para os testes, foi utilizado o banco de dados relacional MySQL). O tempo de 5 segundos foi utilizado de forma empírica, levando em conta somente o tempo para a transmissão e a

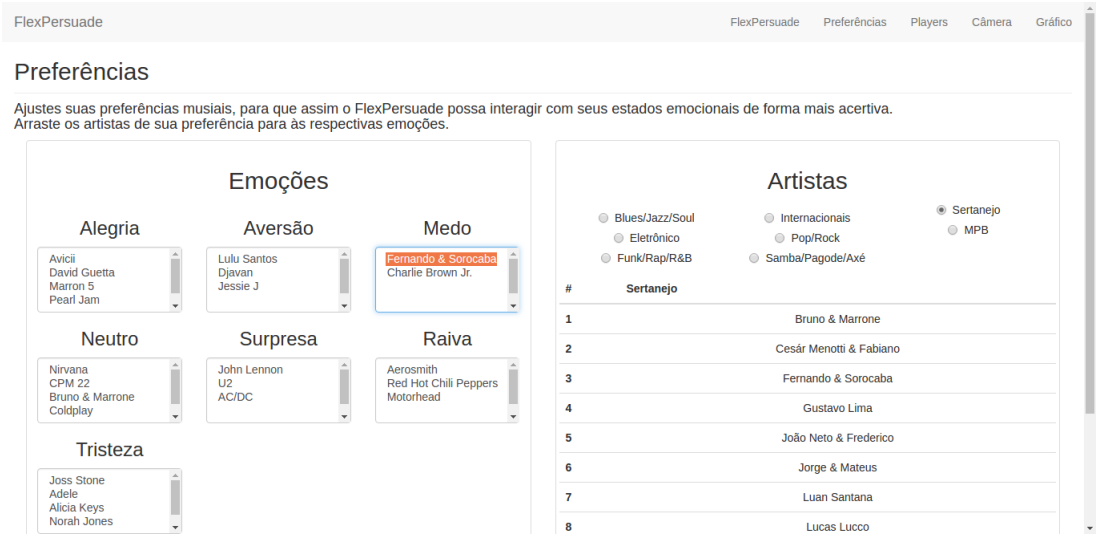

Figura 9 – Tela para configurar as preferências de artista e de emoções.

classificação da imagem. Além disso, o sistema disponibiliza ao usuário a possibilidade de verificar a classificação da sua face em tempo de execução (Figura 10).

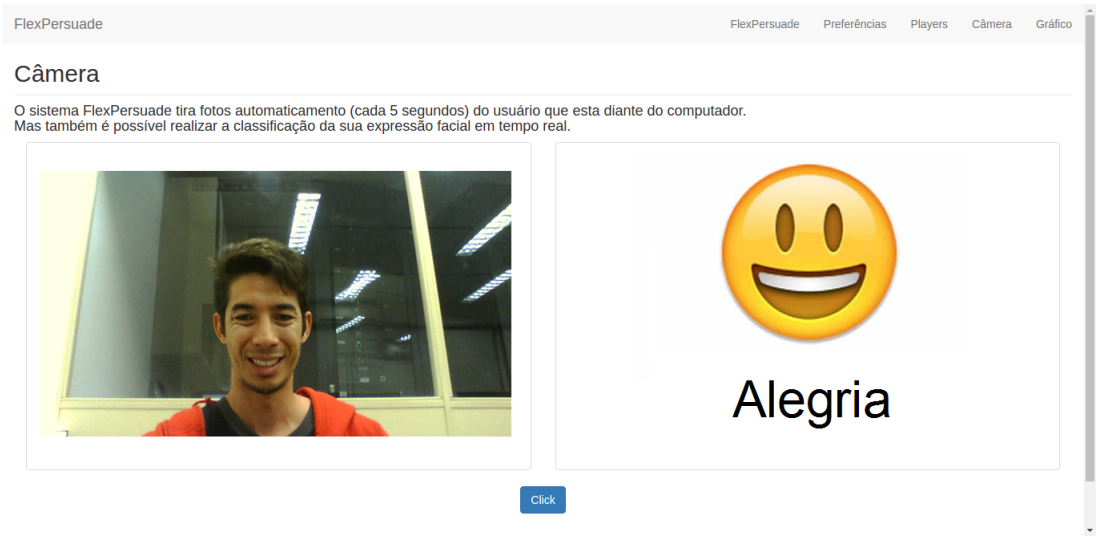

Figura 10 – Tela de captura e de classificação da emoção em tempo de execução.

- *Player de Música:* já de posse das preferências musicais e da imagem da face do usuário, o player executa as músicas que mais se adaptam ao seu estado emocional. Visto que as imagens são classificadas a cada 5 segundos, as emoções são armazenadas em um arquivo temporário e, ao término de cada música, o sistema verifica qual emoção foi mais detectada nesse intervalo de tempo e a próxima música a ser tocada correspondente as preferências informadas. O player disponibiliza as músicas de duas formas: i) áudio, por meio de um módulo do Spotify® e ii) áudio e vídeo, por meio do Youtube® (Figura 11).

Figura 11 – Tela de execução do Player (áudio e áudio/vídeo).

- *Visualização dos Dados:* o sistema ainda disponibiliza um gráfico (Figura 12), em tempo de execução e fornece ao usuário um *feedback* das emoções classificadas ao longo do tempo.

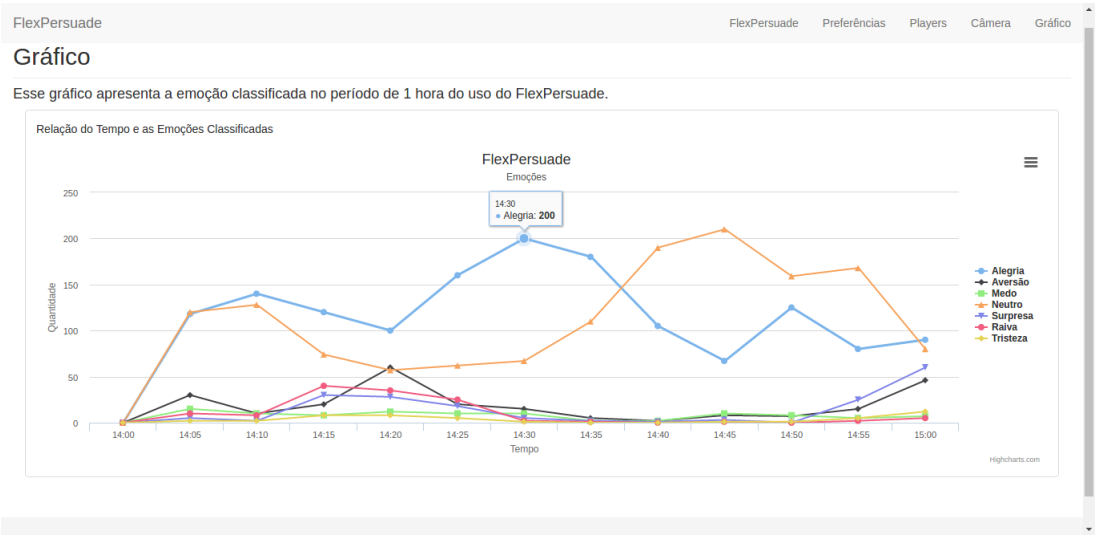

Figura 12 – Tela de *feedback* das emoções para o usuário.

CAPÍTULO

5

AVALIAÇÃO

5.1 Metodologia

Para validar o Comitê proposto nesse trabalho, utilizaram-se imagens de expressões faciais emocionais provenientes das bases de dados *Radboud Faces* (RaFD) (LANGNER *et al.*, 2010), *Extended Cohn-Kanade* (CK+) (LUCEY; COHN *et al.*, 2010), IMPA-FACE3D (MENA-CHALCO; CESAR-JR; VELHO, 2008) e FACES (EBNER; RIEDIGER; LINDENBERGER, 2010), todas de acesso livre.

A base de dados RaFD apresenta exemplos de expressões faciais de 67 atores, adultos e crianças, do sexo masculino e feminino, de nacionalidades americana e marroquina. Já, a base de dados CK+ apresenta expressões faciais de atores adultos, dentre os quais 69% são mulheres e 31% são homens. Ainda, 81% são europeus ou americanos, 13% afro-americanos e 6% pertencentes a outros grupos étnicos. Essa base possui 593 expressões faciais, de 123 atores. A IMPA-FACE3D, por sua vez, inclui aquisições de 38 indivíduos, sendo amostras correspondentes a seis expressões faciais universais e a expressão neutra. Ao todo, o conjunto de dados é composto por 22 homens e 16 mulheres e os indivíduos têm entre 20 e 50 anos de idade. Por fim, a base de dados FACES é composta por imagens de cada uma das seis expressões faciais de 171 indivíduos, divididos em 58 jovens, 56 pessoas de meia-idade e 56 idosos, entre homens e mulheres.

Para os experimentos apresentados, apenas um subconjunto das bases citadas anteriormente é utilizado. Foram selecionadas todas as imagens nas quais os atores estão com o olhar voltado para a frente e expressando as seguintes emoções: alegria, aversão, medo, neutro, raiva, surpresa e tristeza. 67 exemplos de cada emoção são obtidos da base de dados RaFD, 38 da IMPA-FACE3D e 171 da FACES, incluindo crianças, adultos e idosos. O subconjunto selecionado para a base CK+ inclui todas as expressões existentes, de diferentes atores, para cada uma das expressões faciais avaliadas. Então, para o nosso experimento foram utilizadas um total de 3.115

Tabela 2 – Parâmetros utilizados nos classificadores selecionados do Weka.

Classificador	*Modelo*	*Parâmetros*
Naive Bayes	NaiveBayes	displayModelInOldFormat = false, debug = false, useKernelEstimator = false, useSupervisedDiscretization = false
Decision Tree	J48	collapseTree = true, confidenceFactor = 0.25, debug = false, reducedErrorPruning = false, numFolds = 3, saveInstanceData = false subtreeRaising = true, unpruned = false, seed = 1, binarySplits = false, minNumObj = 2 useLaplace = false, useMDLcorrection = true
Fuzzy	FuzzyNN	KNN = 10, debug = false, fuzzifier = 3.0, similarity = SimilarityFNN -R first-last -T weka.fuzzy.tnorm.TNormLukasiewicz -C 0.0
kNN	IBk	KNN = 1, crossValidate = false, debug = false, distanceWeighting = No distance weighting, meanSquared = false, windowSize = 0, nearestNeighbourSearchAlgorithm = LinearNNSearch -A "weka.core.EuclideanDistance -R first-last"
SVM	SMO	buildLogisticModels = false, c = 1.0, checksTurnedOff = false, debug = false, epsilon = 1.0E-12, filterType = Normalize training data, kernel = PolyKernek -C 250007 -E 1.0, numFolds = -1, randomSeed = 1, toleranceParameter = 0.001

imagens, divididas em 445 imagens para cada emoção, de modo a alcançar um balanceamento ideal para os testes realizados. Vale salientar que a união das imagens selecionadas fornece um conjunto de testes heterogêneo, pois consiste em imagens de pessoas de diferentes regiões, o que possibilita que o modelo possa ser aplicado em qualquer dispositivo computacional independente da etnia, da cultura e dos costume do indivíduo que o utiliza.

Após a modelagem da base de dados, utilizou-se o *Framework* Weka (BOUCKAERT *et al.*, 2013), amplamente utilizado por pesquisadores da área de ML. O Weka é uma coleção de algoritmos que podem ser aplicados diretamente a um conjunto de dados ou chamado a partir de seu próprio código Java e que contém ferramentas para pré-processamento de dados, classificação, regressão, *clustering*, regras de associação e de visualização. É também adequado para o desenvolvimento de novos sistemas de ML (BOUCKAERT *et al.*, 2013). A Tabela 2 apresenta os parâmetros utilizados tanto para os testes individualizados dos algoritmos quanto para o desenvolvimento do Comitê de Classificação.

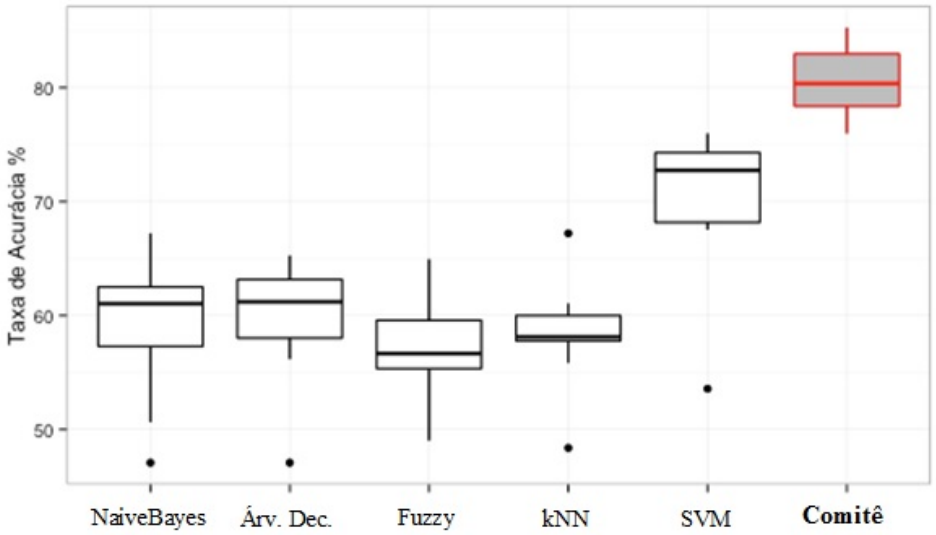

Figura 13 – Boxplots das acurácias apresentadas pelos classificadores para determinar a emoção do usuário. Tais resultados foram obtidos com o uso da técnica *k-fold cross-validation* com $k = 10$.

5.2 Avaliação da Utilização do Comitê de Classificação

Inicialmente, analisou-se o desempenho dos algoritmos separadamente por meio de uma técnica de Planejamento e de Avaliação de Experimentos. Nesse caso, foi o *k-fold cross-validation* [1] com $k = 10$, sendo $k - 1$ para treino e o restante para teste. Assim, é possível gerar uma estimativa de erro mais precisa, pois a média das estimativas tende a uma taxa de erro verdadeiro conforme o aumento de n, e, geralmente, é utilizado para pequenos conjuntos de exemplos. Desse modo, foi possível obter a taxa média da acurácia de cada algoritmo e a Equação 4.1 foi utilizada para a ponderação dos pesos para a decisão final do EC, como visto na seção anterior.

Os resultados demonstram que o Comitê proposto possibilita uma classificação mais precisa comparado ao uso dos classificadores de maneira isolada. Isso pode ser visto na Figura 13, a qual apresenta os Boxplots [2] referentes as classificações realizadas. É importante ressaltar que o Boxplot com cor destacada refere-se aos resultados do EC e exibe a mediana da acurácia superior alcançada pelos demais classificadores (os valores das medianas podem ser encontrados na Tabela 3 e a maior delas encontra-se destacada). É possível notar, pela amplitude interquartil do Boxplot, uma menor dispersão dos resultados obtidos pelo Comitê, o que demonstra uma maior estabilidade em suas decisões.

Com o intuito de validar tais resultados, diversas análises estatísticas foram realizadas.

[1] A validação cruzada é uma técnica para avaliar a capacidade de generalização de um modelo, a partir de um conjunto de dados. Esta técnica é amplamente empregada em problemas onde o objetivo da modelagem é a predição. Busca-se então estimar o quão preciso é este modelo na prática, ou seja, o seu desempenho para um novo conjunto de dados.

[2] O Boxplot, ou diagrama de caixa, é um gráfico que capta importantes aspectos de um conjunto de dados através do seu resumo dos cinco números, formado pelos seguintes valores: valor mínimo, primeiro quartil, segundo quartil, terceiro quartil e valor máximo.

Tabela 3 – Média (%) das acurácias e os valores-*p* dos conjuntos de resultados.

Classificadores	*Acurácia* Média (%)	*Shapiro Wilk* Valor-p
Naive Bayes	59,16	0,115
Árv. Dec.	59,77	0,076
Fuzzy	57,21	0,984
kNN	58,41	0,181
SVM	70,29	**0,007**
Comitê	**80,53**	0,594

Tabela 4 – Valores-*p* da comparação de pares realizada com o teste *Wilcoxon Rank Sum*. Valores inferiores a 0,05 indicam diferença estatisticamente significativa entre os grupos de resultados.

	N.B.	Árv.	Fuzzy	kNN	SVM
Árv.	1,000	-	-	-	-
Fuzzy	1,000	0,783	-	-	-
kNN	1,000	1,000	1,000	-	-
SVM	0,017	0,017	0,017	0,017	-
Comitê	**0,003**	**0,003**	**0,003**	**0,003**	**0,003**

Inicialmente, utilizou-se o método de *Shapiro Wilk* [3] para verificar a sua adequação à normalidade e, consequentemente, para conduzir testes paramétricos ou não-paramétricos. Já que os valores-*p* obtidos não foram todos maior que 0,05 (vide Tabela 3), consideramos recusada a hipótese de normalidade com confiabilidade de 95%. Portanto, o teste não-paramétrico é o mais indicado para as próximas análises.

As comparações de pares realizadas com o teste *Wilcoxon Rank Sum* [4] são exibidas na Tabela 4. Os valores-*p* obtidos com a técnica de *Wilcoxon* indicam que os classificadores Naive Bayes (N.B.), Árvore de Decisões (Árv.), Fuzzy e kNN não apresentam diferença estatisticamente significativa entre si na classificação de emoções. No entanto, a classificação realizada pelo SVM, além de maior acurácia, como apresentado na Figura 13, também apresenta diferença estatisticamente significativa em relação aos demais classificadores utilizados de maneira isolada. Por fim, é possível constatar que os resultados obtidos com o Comitê apresentam diferença estatisticamente significativa em relação aos demais classificadores.

[3] O teste de *Shapiro-Wilk* é um teste geral projetado para detectar todas os desvios a normalidade. O teste rejeita a hipótese de normalidade quando o valor *p* é menor ou igual a 0,05. Falhando o teste de normalidade permite afirmar com 95% de confiança que os dados não se encaixam a distribuição normal. Passar no teste de normalidade só lhe permite declarar que nenhum desvio significativo à normalidade foi encontrado.

[4] O teste de *Wilcoxon*, que se refere ao teste de *Rank Sum* ou o teste de classificação assinado, é um teste não paramétrico que compara dois grupos emparelhados. O teste essencialmente calcula a diferença entre cada conjunto de pares e analisa estas diferenças. O teste de *Wilcoxon Rank Sum* pode ser utilizado para testar a hipótese nula de que duas populações têm a mesma distribuição contínua.

Tabela 5 – Matriz de confusão referente à classificação do Comitê.

	A	B	C	D	E	F	G
Alegria = A	**426**	6	4	1	3	1	2
Aversão = B	17	**366**	5	10	29	2	17
Medo = C	5	8	**340**	34	10	13	35
Neutro = D	1	8	35	**326**	13	4	59
Raiva = E	8	54	7	25	**312**	2	37
Surpresa = F	1	1	13	2	3	**423**	2
Triste = G	7	16	24	47	31	4	**316**

Tabela 6 – Detalhamento do custo de classificação apresentado pelo Comitê para cada emoção.

Emoções	*TFP*	*TVP*
Alegria	0,015	0,962
Aversão	0,035	0,821
Medo	0,033	0,764
Neutro	0,044	0,731
Raiva	0,033	0,701
Surpresa	0,009	0,951
Triste	0,057	0,710

Tais resultados demonstram que entre os classificadores utilizados de maneira isolada, o SVM é o que apresenta maior acurácia na classificação de emoções. Contudo, tal desempenho é superado com o uso do Comitê de Classificadores, visto que essa abordagem proporciona acurácia superior comparada aos demais classificadores utilizados de maneira isolada na classificação de emoções. A Tabela 5, apresenta a matriz de erro, também conhecida como Matriz de Confusão do Comitê referente à classificação das emoções dos usuários. Na matriz de erro, as colunas representam os dados de referência e as linhas, a classificação gerada a partir da emoção. A Tabela 6, gerada a partir da Matriz de Confusão, mostra a relação entre custo (Taxa de Falso Positivo (TFP)) e benefício (Taxa de Verdadeiro Positivo (TVP)) alcançada pelo Comitê na classificação de cada emoção.

A partir dos dados descritos na Tabela 6, gerou-se o gráfico *Receiver Operating Characteristics* (ROC), exibido na Figura 14. Nesse gráfico, é apresentado o desempenho alcançado pelo modelo de classificação induzido pelo Comitê de Classificadores para cada emoção (cada ponto no espaço se refere à classificação de uma determinada emoção pelo mesmo modelo). Nota-se que o modelo baseado no Comitê possibilita uma melhor classificação das emoções “Alegria” e “Surpresa”. Contudo, apesar de inferior, se comparadas à das outras emoções, o modelo de classificação do Comitê classifica as demais emoções com menor custo de classificação e melhor taxa de acurácia comparado aos modelos induzidos pelas técnicas utilizadas de maneira isolada.

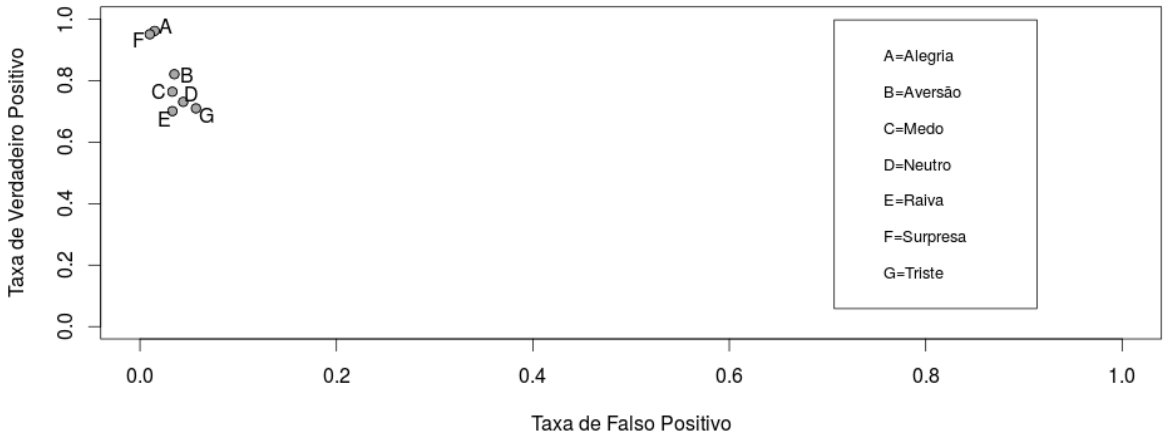

Figura 14 – Gráfico ROC do modelo de classificação induzido pelo Comitê de Classificadores.

5.3 Avaliação dos Pesos Utilizados no Comitê

Os problemas de otimização apresentam um nível de complexidade *NP*-difícil, o que os torna extremamente custosos para resolver utilizando técnicas exatas. Assim, Técnicas de Computação Evolutiva buscam modelar o comportamento evolutivo proposto por Darwin e Wallace (1858) tanto para resolver problemas complexos, quanto para aproximá-las de soluções ótimas em problemas computacionalmente intratáveis. De acordo com essa teoria, os indivíduos de uma população com maior capacidade de se adaptar a um dado ambiente possuem maiores chances de reproduzir suas características em gerações futuras (FACELI, 2011). Em outras palavras, as soluções que apresentam melhor resolução de um problema, possuem mais chances de terem suas características utilizadas em novas soluções.

Os Algoritmos Genéticos (AG) apresentam diversas vantagens sobre técnicas de otimização tradicionais, como seguem, de acordo com Goldberg, Barker e Perez-Grau (1989):

- Não é necessário o cálculo de derivados;
- Otimizar uma população de soluções, em vez de uma única solução, permite com que a pesquisa seja realizada em paralelo e procura, simultaneamente, diferentes partes do espaço de busca;
- Sendo estocástica, eles podem escapar de mínimos locais, e;
- A codificação do problema por cromossomos os torna mais fáceis de serem utilizados em qualquer problema de otimização.

Dadas tais vantagens, a técnica AG foi selecionada para realizar a otimização dos pesos utilizados no Comitê a fim de melhorar a sua acurácia. Apesar de ter obtido uma pequena

melhora, os resultados evidenciam que o uso da ponderação otimizada não proporcionou um aumento suficientemente significativo dos resultados obtidos com os pesos matemáticos. Tal cenário reforça a premissa que o Comitê que utiliza pesos matemáticos é robusto e pode ser escalável a outras emoções sem ter a sua acurácia prejudicada.

As Figuras 15, 16 e 17 exibem os melhores *Fitness* encontrados durante as evoluções do AG no grupo de calibração do *cross-validation*. As configurações do AG referentes à população e a gerações são escritas como "I**X**G**Y**", onde **X** é a quantidade de indivíduos e **Y**, a quantidade de gerações que o AG evolui a população. A Figura 15 exibe a evolução proporcionada pela configuração I12G25, a Figura 16 para I25G50 e Figura 17 para I50G100.

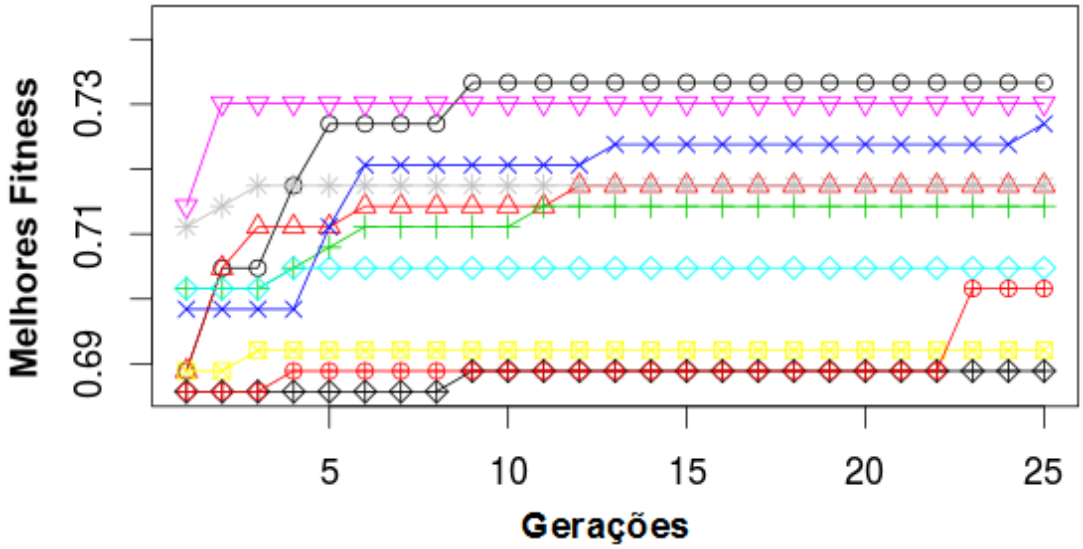

Figura 15 – Configuração: 12 indivíduos e 25 Interações.

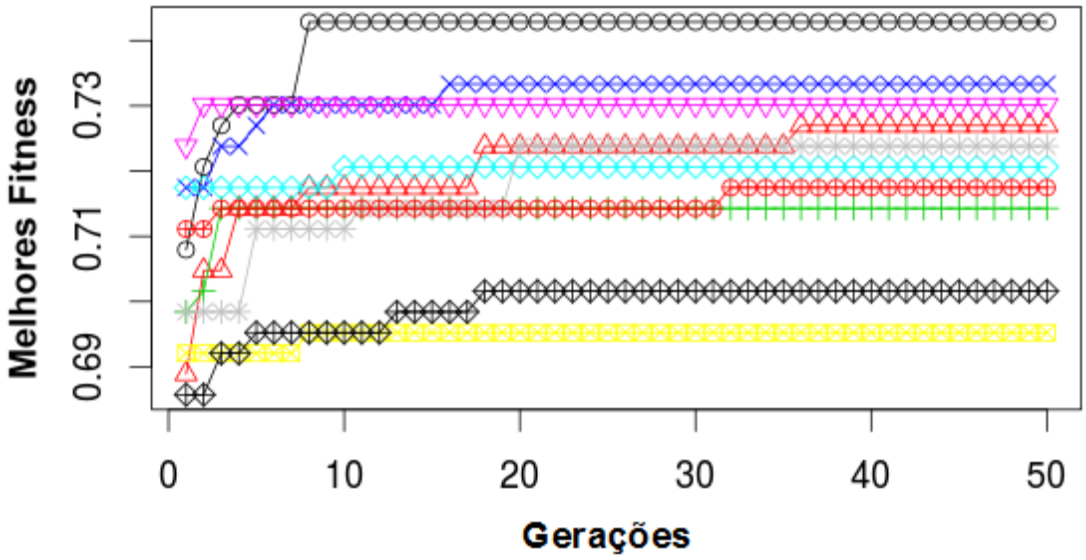

Figura 16 – Configuração: 25 indivíduos e 50 Interações.

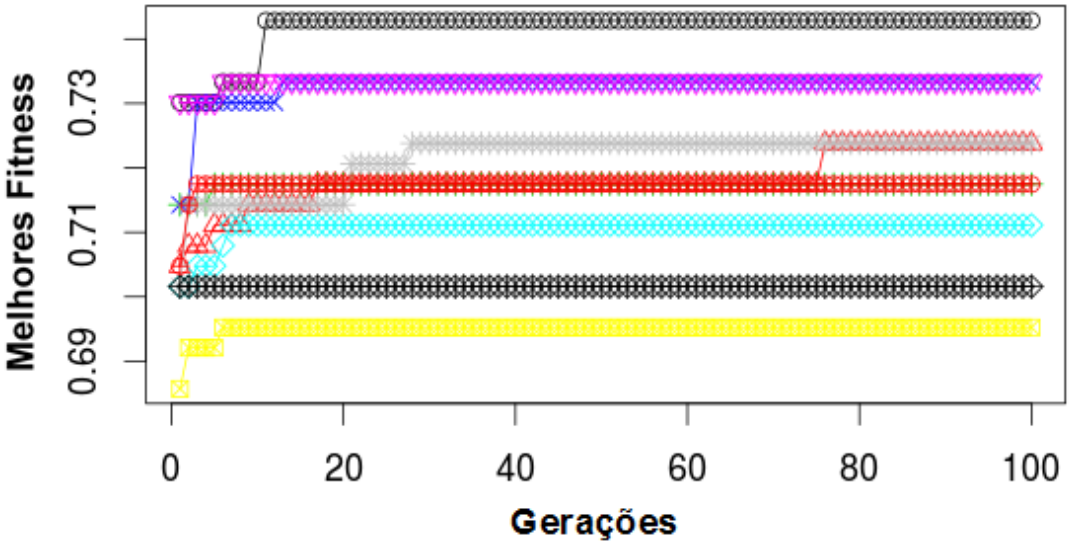

Figura 17 – Configuração: 50 indivíduos e 100 Interações.

Analisando as Figuras 15, 16 e 17, é possível observar de maneira empírica que, a configuração I12G25 não permitiu uma convergência mais consistente para um resultado não-ótimo. Enquanto as configurações I25G50 e I50G100 (Figuras 16 e 17) transmitem uma segurança maior ao indicarem os pesos otimizados ao Comitê, pois exibem um período maior de evolução sem avanço no melhor *Fitness*. Tal período é superior a 30 gerações em 80% dos casos na configuração I25G50 e superior a 70 gerações em 90% dos casos na configuração I50G100. Contudo, os pesos indicados pelas três configurações do AG foram utilizados na fase de avaliação da acurácia do Comitê.

A Figura 18 exibe a acurácia alcançada pelo Comitê com os pesos de ambas as abordagens (matemática e otimizada) durante a fase de avaliação do *cross-validation*. Apesar da observação anterior sobre a evolução do melhor *Fitness* obtido com a configuração I12G25, vale ressaltar que os pesos indicados por tal configuração alcançaram níveis de acurácia próximos aos das outras configurações (I25G50 e I50G100). Além disso, os níveis de acurácia obtidos com os pesos otimizados são realmente próximos aos resultados alcançados pelos pesos proporcionais à precisão individual de cada classificador que compõe o Comitê.

Diversos métodos estatísticos foram utilizados para realizar uma análise mais confiável dos dados. Inicialmente, utilizou-se o *Shapiro-Wilk* para avaliar se os grupos de resultados eram adequados à distribuição normal. Os p-valores obtidos são descritos em detalhe na Tabela 7, os quais confirmam a hipótese de adequação à distribuição normal de todos os grupos de resultados.

Tal resultado direciona a análise para o uso do teste-T (método paramétrico que permite uma confiabilidade maior na análise de dados com adequação à distribuição normal), a fim de checar se existe diferença estatisticamente significativa entre os resultados exibidos na Figura 18. Esse teste indicou que essa diferença não existe em nenhum dos grupos de resultados. Ainda, o p-valor apontado por tal teste é de 1 em todas as comparações pareadas, o que demonstra uma

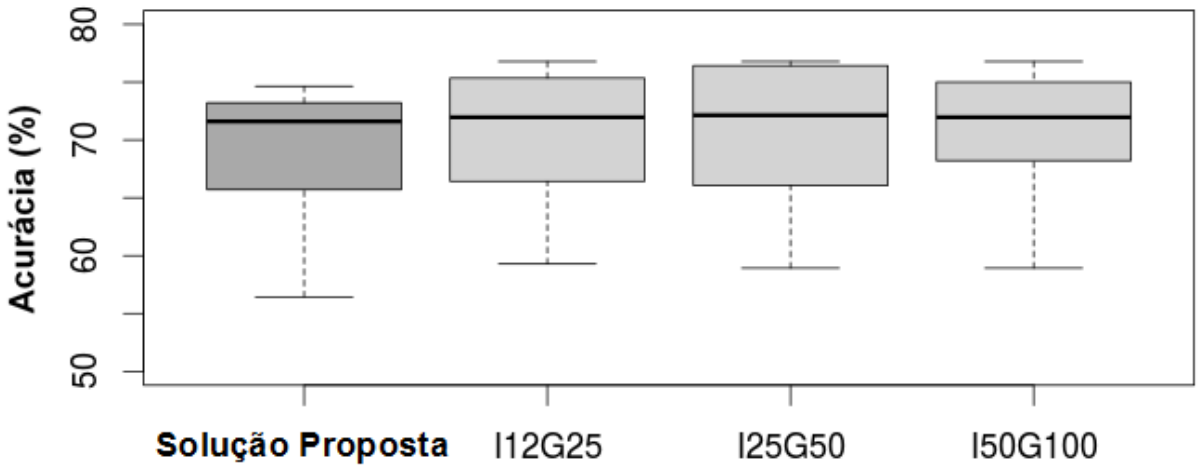

Figura 18 – Acurácia alcançada pelo Comitê com o uso de pesos das diferentes abordagens avaliadas.

Tabela 7 – O método de *Shapiro-Wilk* aprova a hipótese de adequação dos conjuntos de amostras à distribuição normal com intervalo de confiança de 95%. Valores-*p* superiores a 0,05 indicam compatibilidade com a distribuição normal.

Método Shapiro-Wilk	
Configurações	**p-value**
Solução Proposta	0.052
I12G25	0.087
I25G50	0.210
I50G100	0.192

semelhança muito forte entre os resultados.

Por fim, pode-se observar que o uso da abordagem evolutiva para otimizar os pesos do Comitê não proporcionou maior acurácia comparado à abordagem matemática utilizada previamente. Portanto, o modelo do Comitê proposto com pesos calculados a partir da acurácia individual de cada classificador se mostrou igualmente eficiente para ser ampliado a outras emoções.

5.4 Avaliação do Desempenho Computacional do *FlexPersuade*

Outro fator importante analisado foi referente ao desempenho computacional da metodologia proposta. É possível saber o comportamento e investigar maneiras eficientes de realizar novas implementações, prevenindo possíveis "gargalos". A arquitetura utilizada para testes foi a ***S**mart **A**rchitecture for In-**H**ome **H**ealth**c**are* (SAHHc) (MANO *et al.*, 2016). A SAHHc se baseia na utilização da *Internet of Things* (IoT) para o monitoramento inteligente e individualizado de pacientes idosos em suas casas. As principais características da SAHHc são a escalabilidade de

seus componentes e a capacidade de inferir as emoções ou os sentimentos das experiências dos pacientes (como a dor ou o início da depressão), aliados ao acompanhamento individualizado.

A SAHHc é composta por dois elementos: (i) Sensores e (ii) Nó Decisor. Os sensores são dispositivos distribuídos no ambiente, que podem ser encontrados em grande número e o seu objetivo é recolher informações (por exemplo, capturar imagens) sobre a saúde do paciente e enviá-las ao Nó Decisor. O Nó Decisor processa as informações recebidas e toma decisões que visam a preservar a saúde dos pacientes e/ou a dar-lhes um tratamento adequado. Inicialmente, o Nó Decisor é definido como um único elemento (uma máquina) e também trabalha como gerente de rede, mas suas funções podem ser divididas entre vários tomadores de decisão para fins de escalabilidade. Além disso, o Nó Decisor serve como uma interface entre o paciente e o cuidador.

A Figura 19 apresenta os elementos configuráveis (níveis) da arquitetura SAHHc, dividida em três níveis. O Nível 0 (dispositivos embarcados) realiza o reconhecimento da face, o pré-processamento e a identificação dos pacientes monitorados, bem como as suas expressões faciais. O Nível 1 (servidor local) realiza o monitoramento da casa e da saúde dos pacientes, além do controle de todo o sistema. Portanto, ele precisa ter mais poder de processamento e disponibilidade de armazenamento. Já no Nível 2 (servidor remoto), o processamento da imagem na nuvem é realizado quando o servidor local está sobrecarregado e desequilibrado. Então, ele fornece um esquema de balanceamento de carga e garante que a arquitetura SAHHc seja escalável.

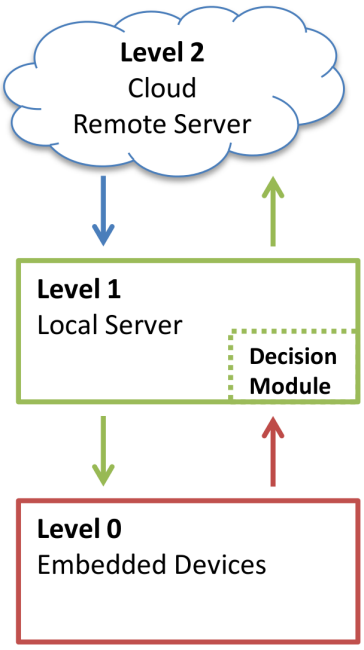

Figura 19 – Elementos configuráveis (níveis) da arquitetura SAHHc (MANO *et al.*, 2016).

5.4.1 Desempenho em Rede de Sensores

Para os testes da proposta deste trabalho em uma Rede de Sensores, foram definidos dois cenários de rede diferentes: (i) Nó Sensor e (ii) Nó Decisor Central. No primeiro cenário, no qual a identificação e a classificação das emoções são executadas nos nós sensores, os resultados permitem saber o intervalo mínimo em que cada imagem é capturada, visto que antes de capturá-la é necessário o processamento e a transmissão dos resultados ao Nó Decisor. No entanto, essa abordagem restringe o desempenho de todo processo, pois os dispositivos possuem recursos limitados. No segundo cenário, os nós sensores são responsáveis somente pela captura e pela transmissão das imagens ao Nó Decisor, o qual, por sua vez, é responsável pelo processamento de todas as imagens e pela tomada de decisões apropriadas. Embora o Nó Decisor tenha mais recursos computacionais disponíveis, uma vez que existem vários nós sensores, ele fará o processamento competitivo e pode ser sobrecarregado, caso a carga de trabalho seja muito pesada.

Para que as comparações de desempenho entre os dois cenários pudesse ser mais justa, assumiu-se que as imagens já estivessem armazenadas nos elementos que empregam a metodologia. Dessa maneira, os tempos de operação para cada experiência foram determinados para medir o nível de carga e permitir a sua comparação. Assim, 30 repetições foram realizadas para cada configuração em intervalos de 10 segundos, a fim de conseguir uma melhor análise estatística dos resultados.

Os experimentos foram planejados com o intuito de simular situações reais para representar o comportamento da arquitetura. Assim, no primeiro cenário (*OnBoard* - **Nível 0**), onde os nós sensores processam as imagens capturadas, não há competição para o processamento. Assume-se que os nós não sejam empregados em qualquer outra aplicação, podendo haver um número variado de nós sensores presentes na arquitetura ou ainda que o tempo de execução seja igual para todos os nós (uma vez que todos os nós sensores têm o mesmo microprocessador). Por outro lado, no segundo cenário, pode haver competição para o processamento, com um nível máximo igual ao número de nós sensores que compõem a arquitetura. No servidor local, o elemento responsável pelas decisões processadas por todas as imagens captadas pelos nós de sensores - **Nível 1**, No pior cenário possível, onde as imagens são recebidas pelo Nó Decisor ao mesmo tempo e as várias instâncias da metodologia proposta são iniciadas imediatamente, foram definidos 1, 2, 4 e 8 como os número de imagens simultâneas a serem processadas.

A Figura 20 exibe os tempos de execução obtidos para cada configuração experimental. Embora a configuração *OnBoard* não tenha concorrência para o processamento, essa abordagem apresentou o maior tempo de execução entre os experimentos. Isso se deve ao fato dos recursos computacionais serem limitados nos sensores (no experimento, foi considerado uma arquitetura com o modelo de Raspberry Pi B para os nós sensores). O Raspberry Pi modelo B tem processador ARM® (1 core - 700 Mhz do *Clock*), 512 MB de RAM e cartão SD de 8 GB. O Nó Decisor está hospedado em um servidor local equipado com um processador Intel® i5 (4 núcleos [dois

físicos e dois por Hyper-Threading] - 3 GHz do *Clock*), 16 GB de RAM e um disco de 1 TB.

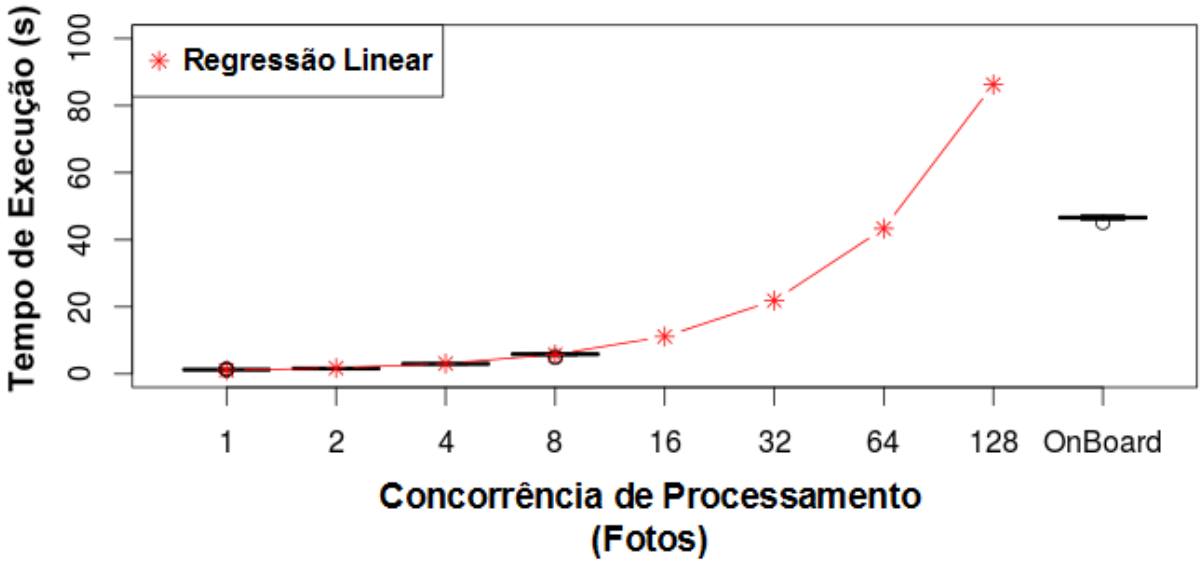

Figura 20 – Avaliação de desempenho: *OnBoard* versus Nó Decisor.

A Figura 20 apresenta os resultados referentes ao nível de concorrência em cada experimento. Nota-se que o tempo de execução é proporcional ao aumento da concorrência. No entanto, apesar do experimento apresentar 8 processos simultâneos (duas vezes o número de núcleos disponíveis no Nó Decisor), o tempo de execução observado no Nó Decisor é cerca de 5 vezes menor que o tempo de execução do experimento *OnBoard* (Nível 0).

Com base nesses resultados, foi realizada uma regressão linear para inferir o nível de concorrência necessária para o tempo de execução, a fim de verificar se o tempo de execução do Nó Decisor é maior que o do *OnBoard*, o qual sugere uma saturação da arquitetura. Pode-se pressupor que as características do aumento de tempo de execução são mantidas enquanto os cálculos são feitos para novos níveis de competição. Esses resultados são indicados por asteriscos na Figura 20 com uma linha reta (em vermelho) entre os pontos.

Assim, verificou-se que, quando existe uma competição envolvendo 64 imagens, o tempo de execução é próximo daquele apresentado pelo *OnBoard*, porém um pouco mais baixo. Quando o nível de competição é igual a 128 imagens, pôde-se observar que a saturação da arquitetura apresentou tempo de execução cerca de duas vezes mais elevado que no cenário *OnBoard*. Assim, assume-se que a execução da metodologia proposta no elemento Nó Decisor torna-se viável com uma arquitetura que possua até 64 elementos sensores. Caso seja necessário um maior número desses elementos, uma abordagem que envolve o uso de cenários *OnBoard* é mais apropriada.

Foi realizada uma avaliação estatística com o objetivo de verificar se os tempos de execução de 1 a 8 imagens diferiam ou não dos tempos de execução para a abordagem *OnBoard*. O teste de *Wilcoxon* foi utilizado e mostrou que os conjuntos são distintos. Esses resultados demonstram que a aplicação da metodologia proposta no elemento Nó Decisor produziu um desempenho satisfatório para as pequenas e médias implantações. No entanto, com as implemen-

tações de um grande número de elementos sensores é mais adequada a abordagem *OnBoard*, como um meio de evitar uma possível saturação da arquitetura.

5.4.2 Análise dos efeitos de transferência de imagens a serem processadas na nuvem

Uma estratégia alternativa que pode adotar a SAHHc é utilizar o conceito de Computação na Nuvem para executar a identificação e a classificação da emoção do usuário (**Nível 2**). Nesse caso, a arquitetura não é afetada pela sobrecarga ao processar as imagens captadas pelos elementos sensores. Por outro lado, a transferência de imagens para o processamento na nuvem pode causar congestionamento na transmissão e, portanto, aumentar o tempo necessário para o envio de mensagens.

O conceito de computação da nuvem, do inglês *Cloud Computing*, é de um sistema que tem recursos computacionais dinâmicos e pode, assim, manter as características necessárias de um dado serviço que pode ser contraído. Em vista disto, assume-se que o tempo de processamento pode ser ajustado de acordo com a aplicação atual. No entanto, o tempo utilizado para enviar imagens não faz parte das características dos serviços contratados, pois depende de recursos que não são controlados pelo provedor de *Cloud Computing*.

O tempo de transmissão é parte do tempo total, contado a partir do tempo de aquisição da imagem até o fim do procedimento metodológico proposto, e resulta em um aumento no intervalo mínimo entre os processos de monitorização. Com isto em mente, avaliamos o impacto das transmissões simultâneas que resultaram em um aumento do intervalo mínimo entre os processos de monitoramento. Realizou-se um experimento para medir a transmissão simultânea de imagens obtidas por câmeras internas, através da ferramenta de comunicação *Secure File Transfer Protocol* (SFTP). A ligação para a comunicação foi estabelecida entre o Nó Decisor e um servidor virtual remoto implementado na nuvem.

O servidor real é um Dell PowerEdge R720, Intel® Xeon 32 núcleos (dois processadores físicos com oito núcleos e HyperThreading (2x)) - 2 GHz de relógio, 256 GB de RAM e 6 HDs SAS hot-plug 3,5 de 600 GB em RAID 5. A máquina virtual foi hospedada nesse servidor real e foram usados 4 VCPUs, 16GB de RAM e 20 GB de disco.

O processo de transmissão de imagem segue o mesmo esquema de paralelismo que foi conduzido nos experimentos anteriores, isto é, 1, 2, 4, 8 imagens são enviadas através da rede, simultaneamente. O tempo decorrido para o envio dessa quantidade de imagens em paralelo encontra-se na Figura 21.

De acordo com a Figura 21, o tempo de transferência aumenta linearmente conforme o paralelismo aumenta. Em vista disto, uma regressão linear foi utilizada para estimar o tempo de transferência para 16, 32, 64 e 128 imagens. Pode-se dizer que o tempo total de processamento não excedeu o tempo de processamento no modo *OnBoard*. Este foi, no caso de: (a)

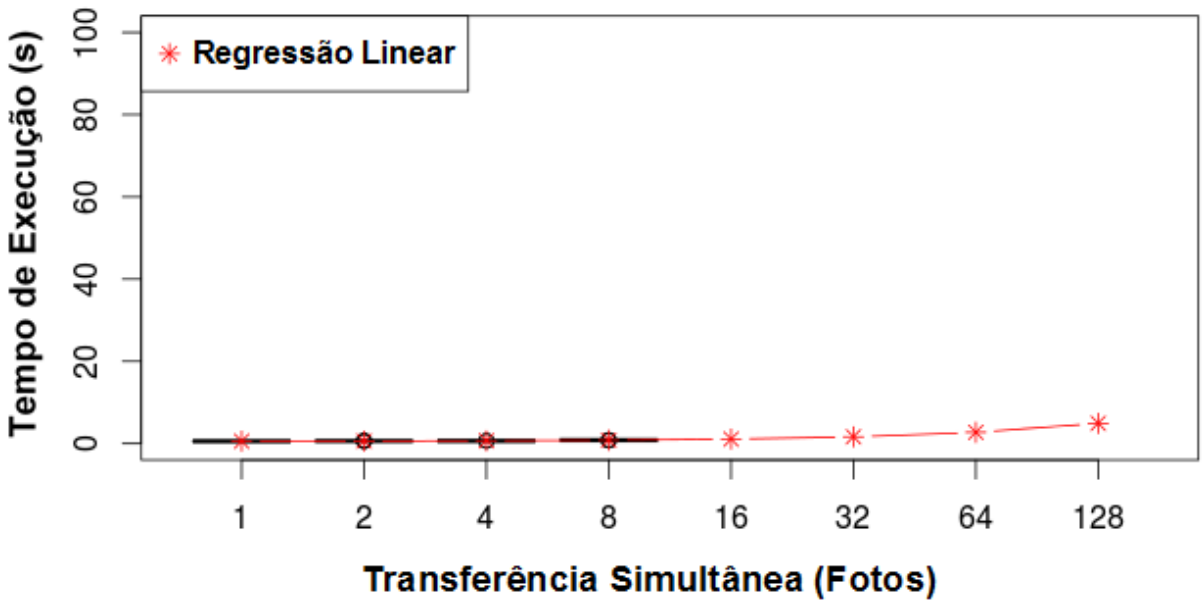

Figura 21 – Avaliação de desempenho do processamento na nuvem.

uma infraestrutura de nuvem virtual utilizada para obter um tempo de resposta similar para o processamento de 32 imagens simultaneamente (Figura 20), ou (b) utilizar um servidor com configuração semelhante para o elemento Nó Decisor. Isso nos ajuda a evitar a saturação da arquitetura, que é implementada para o processamento de 128 fotos. A razão para isso é que o servidor irá adaptar seus recursos de computação para manter o tempo de resposta estabelecido.

Foi realizada uma avaliação estatística com o objetivo de verificar se os tempos de transmissão de imagens do Nó Decisor diferia de acordo com a diferença no número de imagens. Os resultados do teste estatístico demonstrou que, para além da transmissão simultânea de imagens 2 e 4, as outras imagens de transmissão apresentam diferenças significativas. Esses resultados mostram que o tempo de transferência aumenta de forma significativa quando comparado ao crescimento do paralelismo, o que sugere que o ponto de saturação também pode ser obtido. No entanto, essa abordagem proporciona uma maior robustez para a execução da SAHHc, uma vez que permite o tratamento de um maior número de imagens sem exceder o tempo da abordagem *OnBoard*.

5.5 Avaliação da Interação com o Usuário (Player de música

Visando verificar a proposta de interação do *FlexPersuade*, descrita no Capítulo 4, um experimento comparativo foi executado com um conjunto de 30 indivíduos, com idade entre 18 e 80 anos. Essa seção descreve o planejamento, a execução e os resultados da observação da interação do usuário com o *FlexPersuade*. Vale salientar que esse experimento, junto com um aluno de Doutorado, Vinícius Pereira Gonçalves orientado pelo Prof. Jó Ueyama, foi submetido

e autorizado pelo Comitê de Ética em pesquisa envolvendo Seres Humanos da Faculdade de Medicina da Universidade de São Paulo (CAAE-45081415.0.0000.0065).

5.5.1 Planejamento

No início da avaliação da interação do *FlexPersuade* com usuários de diversas idades, foi realizado um planejamento a fim de verificar a satisfação/bem-estar do usuário diante a interação com o player de música.

- *Hipótese:* tendo como base a diversidade e o aumento da utilização de dispositivos computacionais, acredita-se que o *FlexPersuade* é uma solução que contempla situações específicas (emoção), no que diz respeito à interação do usuário com dispositivos computacionais.

- *Objetivo:* observar e analisar a satisfação na interação dos usuários com dispositivos computacionais de acordo com as suas emoções.

- *Metodologia:* com o intuito de analisar a interação do usuário com sistemas flexíveis de dispositivos computacionais, um grupo de 30 usuários, com idade entre 18 e 80 anos, foi convidado a participar de uma prática de uso do Notebook. Esses usuários foram divididos em dois grupos. O grupo 1 realizou o experimento com a atuação do *FlexPersuade* na interação do usuário com o player de música, enquanto o grupo 2 realizou a interação sem a interferência do *FlexPersuade*, ou seja, as músicas foram tocadas de modo totalmente aleatório. A Tabela 8 apresenta a forma que os dois grupos foram divididos.

Tabela 8 – Divisão dos grupos para os experimentos.

	18 a 30 anos	**31 a 50 anos**	**mais de 50 anos**
Grupo 1	5	5	5
Grupo 2	5	5	5

  Os experimentos foram divididos dessa maneira, de modo a verificar a satisfação do usuário na interação com o *FlexPersuade*. O objetivo da atividade era que esses usuários utilizassem o Notebook conforme a sua preferência, enquanto ouviam músicas no sistema desenvolvido. A partir da observação, verificou-se que o *FlexPersuade* conseguiu interagir com as emoções dos usuários durante a interação com os dispositivos computacionais.

- *Material de Apoio:* para a condução do experimento exploratório, foram preparados um Termo de Consentimento Livre e Esclarecido, uma Autorização de Captação e Exibição de Imagem e um Formulário de Observação do Participante. O Termo de Consentimento Livre e Esclarecido (vide Apêndice A) elucida aos participantes o objetivo da pesquisa, a voluntariedade na participação e o caráter científico da mesma. Autorização de Captação

Tabela 9 – Matriz de confusão referente à classificação do Comitê das imagens dos usuários submetidos ao teste.

	A	B	C	D	E	F	G
Alegria = A	**30**	0	0	0	0	0	0
Aversão = B	2	**21**	0	0	4	0	3
Medo = C	0	0	**20**	4	1	1	2
Neutro = D	0	0	2	**25**	0	0	3
Raiva = E	0	6	0	1	**19**	0	4
Surpresa = F	0	0	1	0	0	**29**	0
Triste = G	0	1	1	4	2	0	**22**

e Exibição de Imagem (vide Apêndice B) concede a exploração das imagens adquiridas realizadas no decorrer dos testes. O Formulário de Observação do Participante (vide Apêndice C) foi elaborado para auxiliar no experimento exploratório durante a interação do usuário com o dispositivo computacional.

- *Dispositivo utilizado:* os usuários foram organizados de maneira que realizassem o experimento de forma individualizada. Foi utilizado um NoteBook ASUS® Intel® Core i5, Ram 4 GB e HD 500 GB.

5.5.2 Resultados da Observação

É importante destacar que os experimentos foram avaliados em dois momentos distintos. A primeira avaliação teve como foco a análise do desempenho referente à acurácia alcançada na classificação com imagens diferentes das dos bancos utilizados para o treinamento do modelo baseado em EC. Já a segunda, verificou a satisfação dos usuários referente a soluções flexíveis através do player de música.

A Tabela 9, referente à primeira avaliação, exibe a Matriz de Confusão alcançada pelo Comitê de Classificação com pesos proporcionais à precisão individual de cada classificador para a classificação das imagens dos usuários. Foram utilizadas 30 imagens de cada emoção por participante, totalizando 210 imagens. Vale salientar que tais imagens não foram utilizadas para o treinamento dos modelos, ou seja, as imagens eram desconhecidas, e alcançaram níveis de acurácia de 79,04%, próximos aos níveis dos testes iniciais, com acurácia de 80,53% (Seção 5.2).

Na segunda avaliação, foi possível verificar a satisfação do grupo, representada no gráfico da Figura 22 com os dados do Formulário de Observação do Participante. Os usuários foram submetidos ao teste por 1 hora e, a cada 15 minutos, lhes era perguntado a sua satisfação, de 0 a 10, com a utilização do sistema. Pôde-se, então, observar que a satisfação dos usuários que interagiram com o *FlexPersuade* (grupo 1) foi superior a dos usuários que não interagiram com o player de música por meio do *FlexPersuade* (grupo 2).

A partir dos experimentos realizados com os usuários, foi possível inferir que o uso do

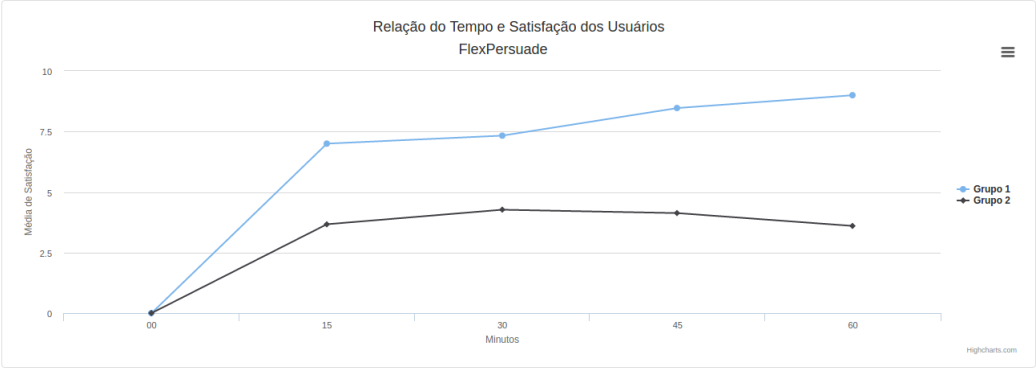

Figura 22 – Nível de satisfação de interação do usuário.

FlexPersuade atende grande parte da intenção de identificar o estado emocional e sugerir músicas de acordo com a emoção do usuário.

CAPÍTULO

6

CONCLUSÕES

6.1 Conlcusões e Trabalhos Futuros

Na área de IHC inúmeros são os métodos, técnicas e instrumentos que apoiam a avaliação de respostas emocionais, devido à oportunidade de sistemas computacionais reagirem ao estado emocional de um indivíduo. Sendo assim, é necessário criar um modelo (por exemplo, com o auxílio de sensores e de métodos de classificação) que suporte a captura, a análise e a classificação de indícios de respostas emocionais de usuários, permitindo que aplicações computacionais se flexibilizem a partir de tal classificação.

Diante disso, discute-se o uso de imagens de faces para a identificação e a classificação das emoções dos usuários de dispositivos computacionais. Este trabalho estabelece um modelo que utiliza o conceito de EC para a classificação das emoções com base na face dos usuários. Assim, o erro cometido por cada classificador individual pode ser reduzido através da combinação de todos os outros componentes que constituem o EC.

Os resultados mostraram que o Comitê proposto apresentou baixa dispersão em sua replicação e a mediana da acurácia foi superior aos resultados apresentados pelos respectivos algoritmos que constituem o modelo proposto (usados individualmente) para a classificação da emoção do usuário quando comparado com abordagens clássicas. Os resultados indicam que as emoções alegria, surpresa e raiva apresentam características que as tornam mais fáceis de serem identificadas, enquanto as outras emoções podem ser confundidas pelas características faciais, pois são mais difíceis de se identificar. No entanto, apesar dessa diferença na classificação de cada emoção, o modelo proposto apresenta menor custo e melhor sensibilidade na classificação das emoções comparado a outros modelos utilizados individualmente.

Pode-se observar que a abordagem utilizada inicialmente para a atribuição de pesos aos algoritmos que compõem o EC, (valores calculados a partir da acurácia individual de cada classificador) se mostrou igualmente eficiente quando comparada a uma abordagem evolutiva

para definir os pesos do Comitê. Observou-se também a satisfação dos usuários na utilização do *FlexPersuade* por meio player de música.

Vale ressaltar, portanto, que apesar dos bons resultados obtidos, acreditamos ser possível aumentar a acurácia de classificação do modelo desenvolvido. Assim, como trabalhos futuros, exploraremos: (i) outras abordagens para a identificação da emoção, por exemplo, o Componente de Tendências Comportamentais - Rede Social; (ii) os resultados da combinação de múltiplos sensores para a identificação da emoção e; (iii) estratégias de persuasão que melhor satisfazem as preferências dos usuários.

6.2 Atividades Acadêmicas e Complementares

6.2.1 Disciplinas Cursadas

Este discente, durante o ano de 2014, cursou sete disciplinas: Introdução à Inteligência Artificial, Tópicos em Computação e Matemática Computacional I, Introdução ao Aprendizado de Máquina, Tópicos em Computação Paralela, Teoria da Computação, Metodologia de Pesquisa Científica em Computação e Preparação Pedagógica (conceitos A, A, B, A, C, A e A, respectivamente), obtendo os créditos necessários no Programa de Mestrado no Instituto de Ciências Matemáticas e de Computação (ICMC-USP). No segundo semestre de 2015, cursou as disciplinas: Tópicos em Computação e Matemática Computacional II, Visualização Computacional e Tópicos em Computação Bioinspirada (conceitos A, B e A, respectivamente). Atualmente, cursa as disciplinas: Princípios de Interação Humano-Robô e Sistemas Operacionais, com reaproveitamento de créditos para o Programa de Doutorado do ICMC-USP.

6.2.2 Produção Científica

6.2.2.1 Periódico

- **Mano, L. Y.**; Meneguete, R. I.; Nakamura, L. H. V.; Libralon, G. L.; Faiçal, B. S.; Filho, G. P. R.; Giancristofaro, G. T.; Pessin, G; Gomes, P. H.; Ueyama; J. "Exploiting IoT technologies for enhancing Health Smart Homes through patient identification and emotion recognition". Computer Communications. **(DOI:10.1016/j.comcom.2016.03.010 - Impact Factor: 1.695)**.

- Faiçal, B.; Freitas, H.; Gomes, P.; **Mano, L.**; et al. "An adaptive approach for UAV-based pesticide spraying in dynamic environments." Computers and Eletronics in Agriculture. Elsevier, 2016. **(Fase de Rebuttal - Impact Factor: 1.761)**

- Faiçal, B. S.; **Mano, L. Y.**; Bayer, G.; Pessin, G.; Gomes, P. H.; Ueyama; J. "Enhancing Reliability in Wireless Sensor Networks for Adaptive River Monitoring Systems: Reflec-

tions on our Long-term Deployment in Brazil". Journal of Environmental Informatics. **(Submetido - Impact Factor: 2.792)**.

6.2.2.2 Conferência

- **Mano, L. Y.**; Giancristofaro, G. T.; Faiçal, B. S.; Libralon, G. L.; Pessin, G; Gomes, P. H.; Ueyama; J. "Exploiting the Use of Ensemble Classifiers to Enhance the Precision of User's Emotion Classification". 16th Engineering Applications of Neural Networks Conference, 16th EANN 2015, Grécia. **(DOI: 10.1145/2797143.2797165)**.

REFERÊNCIAS

BAILENSON, J. N.; PONTIKAKIS, E. D.; MAUSS, I. B.; GROSS, J. J.; JABON, M. E.; HUTCHERSON, C. A.; NASS, C.; JOHN, O. Real-time classification of evoked emotions using facial feature tracking and physiological responses. **International journal of human-computer studies**, Elsevier, v. 66, n. 5, p. 303–317, 2008. Citado 5 vezes nas páginas 21, 30, 35, 38 e 39.

BOUCKAERT, R. R.; FRANK, E.; HALL, M.; KIRKBY, R.; REUTEMANN, P.; SEEWALD, A.; SCUSE, D. **WEKA Manual for Version 3-7-8**. [S.l.]: January, 2013. Citado na página 50.

CAMPBELL, A.; CHOUDHURY, T. From smart to cognitive phones. **Pervasive Computing, IEEE**, v. 11, n. 3, 2012. Citado na página 42.

CANUTO, A. M. d. P. **Combining neural networks and fuzzy logic for applications in character recognition.** Tese (Doutorado) — University of Kent at Canterbury, 2001. Citado 2 vezes nas páginas 23 e 31.

CHANEL, G.; KIERKELS, J. J.; SOLEYMANI, M.; PUN, T. Short-term emotion assessment in a recall paradigm. **International Journal of Human-Computer Studies**, Elsevier, v. 67, n. 8, p. 607–627, 2009. Citado 3 vezes nas páginas 23, 30 e 43.

CLARK, P.; NIBLETT, T. The cn2 induction algorithm. **Machine learning**, Springer, v. 3, n. 4, p. 261–283, 1989. Citado 2 vezes nas páginas 30 e 32.

CRISTIANINI, N.; SHAWE-TAYLOR, J. **An introduction to support vector machines and other kernel-based learning methods**. [S.l.]: Cambridge university press, 2000. Citado 2 vezes nas páginas 30 e 32.

DARWIN, C.; WALLACE, A. On the tendency of species to form varieties; and on the perpetuation of varieties and species by natural means of selection. **Journal of the proceedings of the Linnean Society of London. Zoology**, Wiley Online Library, v. 3, n. 9, p. 45–62, 1858. Citado na página 54.

DESMET, P. A multilayered model of product emotions. **The Design Journal**, Bloomsbury Journals (formerly Berg Journals), v. 6, n. 2, p. 4–13, 2003. Citado 2 vezes nas páginas 26 e 29.

DUDA, P. E.; RICHARD, O. **Hart, Pattern Classification and Scene Analysis**. [S.l.]: John Wiley and Sons, New York, 1973. Citado na página 30.

DUDA, R. O.; HART, P. E.; STORK, D. G. **Pattern classification**. [S.l.]: John Wiley & Sons, 2012. Citado 2 vezes nas páginas 23 e 31.

EBNER, N. C.; RIEDIGER, M.; LINDENBERGER, U. Faces—a database of facial expressions in young, middle-aged, and older women and men: Development and validation. **Behavior research methods**, 2010. Citado na página 49.

EKMAN, P. **Cross-cultural studies of facial expression**. [S.l.]: New York: Academic, 1973. Citado na página 28.

____. **Darwin and facial expression: A century of research in review**. [S.l.]: Ishk, 2006. Citado 5 vezes nas páginas 15, 26, 29, 39 e 42.

EKMAN, P.; FRIESEN, W. V. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. **Nonverbal communication, interaction, and gesture**, p. 57–106, 1981. Citado na página 25.

ELLSWORTH, P. C.; SCHERER, K. R. Appraisal processes in emotion. **Handbook of affective sciences**, v. 572, p. V595, 2003. Citado na página 25.

ESLAMBOLCHILAR, P.; WILSON, M.; OAKLEY, I.; DEY, A. Pinc: persuasion, influence, nudge & coercion through mobile devices. In: ACM. **CHI'11 Extended Abstracts on Human Factors in Computing Systems**. [S.l.], 2011. p. 13–16. Citado na página 21.

EXAME. **O cansaço pode matar a inovação nas empresas**. 2013. <http://exame.abril.com.br/revista-voce-rh/edicoes/28/noticias/por-que-paramos-de-pensar>. Acessado em Janeiro de 2014. Disponível em: <http://exame.abril.com.br/revista-voce-rh/edicoes/28/noticias/por-que-paramos-de-pensar>. Citado na página 22.

FACELI, K. **Inteligência artificial: uma abordagem de aprendizado de máquina**. [S.l.]: Grupo Gen-LTC, 2011. Citado 3 vezes nas páginas 30, 31 e 54.

FOGG, B. J. Persuasive technology: using computers to change what we think and do. **Ubiquity**, ACM, v. 2002, n. December, p. 5, 2002. Citado na página 22.

FONTAINE, J. R.; POORTINGA, Y. H.; SETIADI, B.; MARKAM, S. S. Cognitive structure of emotion terms in indonesia and the netherlands. **Cognition & Emotion**, Taylor & Francis, v. 16, n. 1, p. 61–86, 2002. Citado 2 vezes nas páginas 25 e 26.

FONTAINE, J. R.; SCHERER, K. R.; ROESCH, E. B.; ELLSWORTH, P. C. The world of emotions is not two-dimensional. **Psychological science**, SAGE Publications, v. 18, n. 12, p. 1050–1057, 2007. Citado na página 26.

FRIJDA, N. H. **The emotions**. [S.l.]: Cambridge University Press, 1986. 255–31 p. Citado na página 25.

FRIJDA, N. H.; KUIPERS, P.; SCHURE, E. T. Relations among emotion, appraisal, and emotional action readiness. **Journal of personality and social psychology**, American Psychological Association, v. 57, n. 2, p. 212, 1989. Citado na página 25.

GOLDBERG, R. B.; BARKER, S. J.; PEREZ-GRAU, L. Regulation of gene expression during plant embryogenesis. **Cell**, Elsevier, v. 56, n. 2, p. 149–160, 1989. Citado na página 54.

GONÇALVES, V. P.; NERIS, V.; MORANDINI, M.; NAKAGAWA, E. Y.; UEYAMA, J. Uma revisão sistemática sobre métodos de avaliação de usabilidade aplicados em software de telefones celulares. In: BRAZILIAN COMPUTER SOCIETY. **Proceedings of the 10th Brazilian Symposium on on Human Factors in Computing Systems and the 5th Latin American Conference on Human-Computer Interaction**. [S.l.], 2011. p. 197–201. Citado 3 vezes nas páginas 21, 26 e 44.

GONÇALVES, V. P.; NERIS, V.; UEYAMA, J.; GIANCRISTOFARO, G. T.; SERAPHINI, S. *et al.* Emoções na interação humano-computador: um estudo considerando sensores. In: BRAZILIAN COMPUTER SOCIETY. **Proceedings of the 12th Brazilian Symposium on**

Human Factors in Computing Systems. [S.l.], 2013. p. 252–255. Citado 5 vezes nas páginas 21, 23, 29, 30 e 43.

GONÇALVES, V. P.; SERAPHINI, S.; NERIS, V. P.; UEYAMA, J. Flexinterface: a framework to provide flexible mobile phone user interfaces. In: **15th International Conference on Enterprise Information Systems (ICEIS 2013)**. [S.l.: s.n.], 2013. Citado na página 26.

GONÇALVES, V. P.; UEYAMA, J. Um estudo sobre o design, a implementação ea avaliação de interfaces flexíveis para idosos em telefones celulares. 2012. Citado 2 vezes nas páginas 23 e 44.

HAYASHI, E. C.; BARANAUSKAS, M. Towards a framework for the affective and emotional faces of usability. In: IEEE. **Information Society (i-Society), 2011 International Conference on**. [S.l.], 2011. p. 72–79. Citado na página 26.

KATSIS, C. D.; KATERTSIDIS, N.; GANIATSAS, G.; FOTIADIS, D. I. Toward emotion recognition in car-racing drivers: A biosignal processing approach. **Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on**, IEEE, v. 38, n. 3, p. 502–512, 2008. Citado na página 30.

KLEIN, J.; MOON, Y.; PICARD, R. W. This computer responds to user frustration:: Theory, design, and results. **Interacting with computers**, Elsevier, v. 14, n. 2, p. 119–140, 2002. Citado 2 vezes nas páginas 21 e 36.

LANG, P. J.; GREENWALD, M. K.; BRADLEY, M. M.; HAMM, A. O. Looking at pictures: Affective, facial, visceral, and behavioral reactions. **Psychophysiology**, Wiley Online Library, v. 30, n. 3, p. 261–273, 1993. Citado na página 27.

LANGNER, O.; DOTSCH, R.; BIJLSTRA, G. *et al.* Presentation and validation of the radboud faces database. **Cognition and Emotion**, 2010. Citado na página 49.

LAURANS, G.; DESMET, P. M.; HEKKERT, P. P. Assessing emotion in interaction: some problems and a new approach. In: **International Conference on designing pleasurable products and interfaces**. [S.l.: s.n.], 2009. p. 13–16. Citado na página 26.

LEVENTHAL, H. A perceptual-motor theory of emotion. **Advances in experimental social psychology**, v. 17, p. 117–182, 1984. Citado na página 27.

LIBRALON, G.; ROMERO, R. Mapping of facial elements for emotion analysis. In: **Proceedings of the Brazilian Conf. on Intelligent Systems**. [S.l.: s.n.], 2014. Citado 3 vezes nas páginas 29, 30 e 42.

LICHTENSTEIN, A.; OEHME, A.; KUPSCHICK, S.; JÜRGENSOHN, T. Comparing two emotion models for deriving affective states from physiological data. In: **Affect and Emotion in Human-Computer Interaction**. [S.l.]: Springer, 2008. p. 35–50. Citado na página 21.

LIKAMWA, R.; LIU, Y.; LANE, N. D.; ZHONG, L. Moodscope: Building a mood sensor from smartphone usage patterns. In: ACM. **Proceeding of the 11th annual international conference on Mobile systems, applications, and services**. [S.l.], 2013. p. 389–402. Citado 4 vezes nas páginas 21, 37, 38 e 39.

LITTLEWORT, G.; WHITEHILL, J.; WU, T.; FASEL, I.; FRANK, M.; MOVELLAN, J.; BARTLETT, M. The computer expression recognition toolbox (cert). In: IEEE. **Automatic Face & Gesture Recognition and Workshops (FG 2011), 2011 IEEE International Conference on**. [S.l.], 2011. p. 298–305. Citado na página 30.

LUCEY, P.; COHN, J. F. *et al.* The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In: **Computer Vision and Pattern Recognition Workshops, IEEE Computer Society Conf. on**. [S.l.: s.n.], 2010. Citado na página 49.

MAHLKE, S.; MINGE, M. Consideration of multiple components of emotions in human-technology interaction. In: **Affect and emotion in human-computer interaction**. [S.l.]: Springer, 2008. p. 51–62. Citado 6 vezes nas páginas 22, 25, 26, 27, 28 e 29.

MANO, L. Y.; FAIÇAL, B. S.; NAKAMURA, L. H.; GOMES, P. H.; LIBRALON, G. L.; MENEGUETE, R. I.; FILHO, P. G.; GIANCRISTOFARO, G. T.; PESSIN, G.; KRISHNAMACHARI, B. *et al.* Exploiting iot technologies for enhancing health smart homes through patient identification and emotion recognition. **Computer Communications**, Elsevier, 2016. Citado 3 vezes nas páginas 15, 57 e 58.

MANO, L. Y.; GIANCRISTOFARO, G. T.; FAIÇAL, B. S.; LIBRALON, G. L.; PESSIN, G.; GOMES, P. H. Exploiting the use of ensemble classifiers to enhance the precision of user's emotion classification. In: ACM. **Proceedings of the 16th International Conference on Engineering Applications of Neural Networks (INNS)**. [S.l.], 2015. p. 5. Citado 3 vezes nas páginas 15, 42 e 43.

MARTINEZ, A.; DU, S. A model of the perception of facial expressions of emotion by humans: Research overview and perspectives. **Journal of Machine Learning Research**, 2012. Citado na página 29.

MENA-CHALCO, J.; CESAR-JR, R.; VELHO, L. Banco de dados de faces 3d: Impa-face3d. **IMPA-RJ, Tech. Rep**, 2008. Citado na página 49.

MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. **Sistemas Inteligentes-Fundamentos e Aplicações**, v. 1, p. 1, 2003. Citado na página 31.

NERIS, V. P. de A.; BARANAUSKAS, M. C. C. Making interactive systems more flexible: an approach based on users' participation and norms. In: BRAZILIAN COMPUTER SOCIETY. **Proceedings of the IX Symposium on Human Factors in Computing Systems**. [S.l.], 2010. p. 101–110. Citado na página 23.

NIEDENTHAL, P.; KRAUTH-GRUBER, S.; RIC, F. Psychology of emotion. **Psychology of Emotion**, Psychology Press New York, 2006. Citado na página 25.

ØHRN, A.; ROWLAND, T. Rough sets: a knowledge discovery technique for multifactorial medical outcomes. **American journal of physical medicine & rehabilitation**, LWW, v. 79, n. 1, p. 100–108, 2000. Citado 3 vezes nas páginas 23, 30 e 43.

OINAS-KUKKONEN, H. A foundation for the study of behavior change support systems. **Personal and ubiquitous computing**, Springer-Verlag, v. 17, n. 6, p. 1223–1235, 2013. Citado na página 21.

PARKINSON, B. **Ideas and realities of emotion**. [S.l.]: Psychology Press, 1995. Citado na página 25.

PETER, C.; URBAN, B. Emotion in human-computer interaction. In: **Expanding the Frontiers of Visual Analytics and Visualization**. [S.l.]: Springer, 2012. p. 239–262. Citado 3 vezes nas páginas 21, 30 e 36.

PICARD, R. **Affective computing. 1997**. [S.l.]: MIT Press, 1997. Citado na página 21.

PICARD, R. W. **What does it mean for a computer to "have" emotions? Chapter to appear in: Emotions in Humans and Artifacts ed. by R. Trappl, P. Petta and S. Payr**. 2004. Citado 2 vezes nas páginas 21 e 38.

RAMAKRISHNAN, S.; EMARY, I. M. E. Speech emotion recognition approaches in human computer interaction. **Telecommunication Systems**, Springer, v. 52, n. 3, p. 1467–1478, 2013. Citado 3 vezes nas páginas 23, 30 e 43.

RUSSELL, J. A. A circumplex model of affect. **Journal of personality and social psychology**, American Psychological Association, v. 39, n. 6, p. 1161, 1980. Citado 6 vezes nas páginas 15, 27, 28, 35, 36 e 39.

SAIDIN, A. Z.; MACAULAY, C.; HINE, N. Persuasion knowledge toolkit: requirements gathering with designer. In: BRITISH COMPUTER SOCIETY. **Proceedings of the 25th BCS Conference on Human-Computer Interaction**. [S.l.], 2011. p. 503–508. Citado na página 22.

SANTOS, A. d. M. **Uma análise da aplicação do modelo de Rede Neural RePART em Comitês de classificadores**. Dissertação (Mestrado) — Universidade Federal do Rio Grande do Norte, 2008. Citado na página 31.

SANTOS V.C.; COCA, S. L. G. R. R. F. Imitation of facial expressions for learning emotions in social robotics. **Simpósio Brasileiro de Automação Inteligente**, Anais do SBAI-DINCON, XI, 2013. Citado 4 vezes nas páginas 15, 32, 33 e 44.

SARAGIH, J. M.; LUCEY, S.; COHN, J. F. Deformable model fitting by regularized landmark mean-shift. **International Journal of Computer Vision**, Springer, v. 91, n. 2, p. 200–215, 2011. Citado 3 vezes nas páginas 33, 38 e 42.

SCHERER, K. R. Appraisal considered as a process of multilevel sequential checking. **Appraisal processes in emotion: Theory, methods, research**, v. 92, p. 120, 2001. Citado 2 vezes nas páginas 25 e 26.

____. What are emotions? and how can they be measured? **Social science information**, Sage Publications, v. 44, n. 4, p. 695–729, 2005. Citado 10 vezes nas páginas 22, 25, 26, 27, 28, 29, 32, 36, 39 e 41.

SCHULLER, B.; REITER, S.; MULLER, R.; AL-HAMES, M.; LANG, M.; RIGOLL, G. Speaker independent speech emotion recognition by ensemble classification. In: IEEE. **Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on**. [S.l.], 2005. p. 864–867. Citado 4 vezes nas páginas 15, 23, 31 e 32.

SHARKEY, A. J. Multi-net systems. In: **Combining artificial neural nets**. [S.l.]: Springer, 1999. p. 1–30. Citado na página 31.

STEMMLER, G. Methodological considerations in the psychophysiological study of emotion. **Handbook of affective sciences**, p. 225–255, 2003. Citado 2 vezes nas páginas 25 e 26.

TECH, P. **Sensor ligado ao Kinect decifra ate estado emocional**. 2013. <http://planetech.uol.com.br/2013/04/04/sensor-ligado-ao-kinect-decifra-ate-estado-emocional>. Acessado em Novembro de 2013. Disponível em: <http://planetech.uol.com.br/2013/04/04/sensor-ligado-ao-kinect-decifra-ate-estado-emocional/>. Citado 2 vezes nas páginas 21 e 37.

TIAN, Y. L.; KANADE, T.; F., C. J. Handbook of face recognition. **Springer**, 2005. Citado na página 32.

UEYAMA, J. **A Run Time Component Model for Building System Software**. Tese (Doutorado) — Lancaster University, 2005. Citado na página 44.

UEYAMA, J.; PINTO, V. P.; MADEIRA, E. R.; GRACE, P.; JONHSON, T. M.; CAMARGO, R. Y. Exploiting a generic approach for constructing mobile device applications. In: ACM. **Proceedings of the Fourth International ICST Conference on COMmunication System softWAre and middlewaRE**. [S.l.], 2009. p. 12. Citado na página 44.

UEYAMA, J.; SCHMID, S.; COULSON, G.; BLAIR, G. S.; GOMES, A. T.; JOOLIA, A.; LEE, K. A re-configurable component model for programmable nodes. In: IEEE. **Distributed Computing Systems Workshops, 2004. Proceedings. 24th International Conference on**. [S.l.], 2004. p. 375–380. Citado na página 44.

UEYAMA, J.; VILLAS, L. A.; PINTO, A. R.; GONÇALVES, V. P.; PESSIN, G.; PAZZI, R. W.; BRAUN, T. *et al.* Nodepm: A remote monitoring alert system for energy consumption using probabilistic techniques. **Sensors**, Multidisciplinary Digital Publishing Institute, v. 14, n. 1, p. 848–867, 2014. Citado na página 30.

WESTERGREN, T. **PANDORA**. 2013. <http://www.pandora.com/restricted>. Acessado em Outubro de 2013. Disponível em: <http://www.pandora.com/restricted>. Citado na página 37.

XAVIER, R. A. C.; GARCIA, F. E.; NERIS, V. P. de A. Decisões de design de interfaces ruins eo impacto delas na interação: um estudo preliminar considerando o estado emocional de idosos. In: BRAZILIAN COMPUTER SOCIETY. **Proceedings of the 11th Brazilian Symposium on Human Factors in Computing Systems**. [S.l.], 2012. p. 127–136. Citado 3 vezes nas páginas 22, 27 e 28.

YIK, M. S.; RUSSELL, J. A.; BARRETT, L. F. Structure of self-reported current affect: Integration and beyond. **Journal of personality and social psychology**, American Psychological Association, v. 77, n. 3, p. 600, 1999. Citado 2 vezes nas páginas 25 e 26.

YUAN, G.; LIM, T. S.; JUAN, W. K.; RINGO, H. M.-H.; LI, Q. A gmm based 2-stage architecture for multi-subject emotion recognition using physiological responses. In: ACM. **Proceedings of the 1st Augmented Human International Conference**. [S.l.], 2010. p. 3. Citado na página 30.

ZHOU, F.; QU, X.; HELANDER, M. G.; JIAO, J. R. Affect prediction from physiological measures via visual stimuli. **International Journal of Human-Computer Studies**, Elsevier, v. 69, n. 12, p. 801–819, 2011. Citado 7 vezes nas páginas 21, 23, 30, 36, 38, 39 e 43.

ANEXO

A

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

1. Você está sendo convidado para participar da pesquisa “FlexPersuade - Explorando uma abordagem flexível em *softwares* de persuasão: um estudo de caso com *players* de música”.
2. Você foi selecionado para ser voluntário e sua participação não é obrigatória.
3. A qualquer momento você pode desistir de participar e retirar seu consentimento.
4. Sua recusa não trará nenhum prejuízo em sua relação com o pesquisador ou a instituição.
5. Essa pesquisa tem por objetivo captar e analisar respostas emocionais por meio da face, com o auxílio da câmera de dispositivos computacionais.
6. Sua participação nesta pesquisa consistirá em capturar imagens da sua face em um Notebook para posteriores avaliações.
7. Os benefícios relacionados a sua participação estão em contribuir com a pesquisa, para que se saiba mais sobre emoções na interação com dispositivos computacionais.
8. As informações obtidas por meio dessa pesquisa serão confidenciais e assegurados o sigilo sobre sua participação.
9. Os dados não serão divulgados de forma a possibilitar sua identificação.
10. Esse experimento foi submetido e autorizado pelo Comitê de Ética em pesquisa envolvendo Seres Humanos da Faculdade de Medicina da Universidade de São Paulo (CAAE-45081415.0.0000.0065).
11. Você receberá uma cópia desse termo onde consta o telefone e o endereço do pesquisador principal, podendo tirar dúvidas sobre o projeto e sua participação, agora ou a qualquer momento.

— — — — — — — — — — — —

Leandro Yukio Mano Alves
Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo
13566-590 São Carlos-SP
Tel.: 16-3373-9638

Endereço e telefone do Pesquisador Principal: Rua Maestro João Seppe, 303, Jd. Paraíso
13561-180 São Carlos-SP
Tel.: 16-98127-3468

Declaro que entendi os objetivos, riscos e benefícios de minha participação na pesquisa e concordo em participar.

São Carlos, 01 de Dezembro de 2015

— — — — — — — — — — — —

Sujeito da Pesquisa

ANEXO

B

AUTORIZAÇÃO DE CAPTAÇÃO E EXIBIÇÃO DE IMAGEM

Eu,_____, (nacionalidade) _____, (estado civil) _____portador da Cédula de Identidade RG nº _____, inscrito no CPF/MF sob o número _____autorizo a captação, utilização e exibição de minha imagem pela Universidade de São Paulo – USP diretamente ou através do Instituto de Ciências Matemáticas e de Computação – ICMC, diretamente ou através do Departamento de Computação – DC ou outra entidade vinculada ou contratada, a serem utilizadas em obras audiovisuais a serem produzidas para fins institucionais, didáticos e/ou científicos, seja, essas destinadas à divulgação ao público em geral e/ou apenas para uso interno desta instituição.

1. A presente autorização, concedida a título gratuito, confere à USP, diretamente ou através do Laboratório de Sistemas Distribuídos e Programação Concorrente – LaSDPC – ICMC, ou outra entidade vinculada ou contratada, o direito de utilizar minhas imagens, nas obras para veiculação interna na USP, bem como em eventos externos, no Brasil e no exterior por mídia escrita, eletrônica ou digital, tais como Revistas, Manuais, Portais de Internet, folders, atividades de caráter didático oi científico, trabalhos científicos, programas de Rádio e TV, publicações em geral, entre outros, a critério exclusivo da USP, desde que não haja desvirtuamento da sua finalidade.

2. As obras poderão ser distribuídas pelo Laboratório LaSDPC – ICMC – USP, ou por outra entidade vinculada ou contratada, de forma gratuita ou comercial, sendo certo que nada será devido pelo uso das imagens, objetos da presente autorização, mesmo nas hipóteses de comercialização das obras, em parceria ou não com outras pessoas jurídicas.

3. Declaro estar ciente de que as imagens captadas nessa ocasião farão parte de um “banco de imagens” pertencente ao Laboratório LaSDPC – ICMC – USP e que poderão ser utilizadas

a qualquer tempo e de acordo com os critérios da mesma.

4. Por esta ser a expressão da minha vontade declaro que autorizo o uso acima descrito sem que nada haja a ser declamado a título de direitos conexos à minha imagem ou a qualquer outro título, e assino a presente autorização em (02) duas vias de igual teor e forma.

São Carlos, 01 de Dezembro de 2015

— — — — — — — — — — — —

(Assinatura)

— — — — — — — — — — — —

(Nome por Extenso)

ANEXO

C

FORMULÁRIO DE OBSERVAÇÃO DO PARTICIPANTE

Nome: _____, idade: _____,

Avalie, de 0 à 10, onde 0 corresponde a muito ruim e 10 a muito bom, sua satisfação de interação com o dispositivo computacional, por meio do *player* de música. Considerar se as músicas estão sendo selecionadas de acordo com seu estado emocional.

- 15 minutos:
  Nota:
- 30 minutos:
  Nota:
- 45 minutos:
  Nota:
- 60 minutos:
  Nota:

São Carlos, 01 de Dezembro de 2015

— — — — — — — — — — — —

(Assinatura)