
Uma abordagem para indicar o estado emocional
de usuários em tempo de interação

Vinicius Pereira Gonçalves

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Vinícius Pereira Gonçalves

Uma abordagem para indicar o estado emocional de usuários em tempo de interação

Tese apresentada ao Instituto de Ciências Matemáticas e de Computação - ICMC-USP, como parte dos requisitos para obtenção do título de Doutor em Ciências - Ciências de Computação e Matemática Computacional. *EXEMPLAR DE DEFESA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientador: Prof. Dr. Jo Ueyama

USP – São Carlos
Maio 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)

G634u	<p>Gonçalves, Vinícius Pereira</p> <p>Uma abordagem para indicar o estado emocional de usuários em tempo de interação / Vinícius Pereira Gonçalves; orientador Jó Ueyama. - São Carlos - SP, 2016.</p> <p>115 p.</p> <p>Tese (Doutorado - Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, 2016.</p> <p>1. Emoções. 2. Avaliação. 3. Sensores. 4. Componentes. 5. Psicólogos. 6. Design Emocional. 7. Computação Afetiva. 8. Interação Humano-Computador. I. Ueyama, Jó, orient. II. Título.</p>
-------	---

Vinícius Pereira Gonçalves

**An approach to indicate the users' emotional state at
interaction time**

Doctoral dissertation submitted to the Instituto de
Ciências Matemáticas e de Computação - ICMC-
USP, in partial fulfillment of the requirements for the
degree of the Doctorate Program in Computer
Science and Computational Mathematics.
EXAMINATION BOARD PRESENTATION COPY

Concentration Area: Computer Science and
Computational Mathematics

Advisor: Prof. Dr. Jo Ueyama

**USP – São Carlos
May 2016**

AGRADECIMENTOS

A Deus, pela sua presença constante em minha vida. Obrigado Jesus!

A minha família e parentes pelo amor, carinho e apoio incondicionais. Em especial, aos meus pais Lucicléa e Clovis, a minha irmã Lucélia e minha sobrinha Juju que acabou de chegar ao mundo. Amo muito vocês!

Ao Prof. Jó Ueyama, por sua orientação presente e pelo perfil multidisciplinar.

À Prof.^a Vânia Neris, por sua orientação presente e pelo comprometimento.

Ao Jó e a Vânia pela amizade e por tudo mais.

À Prof.^a Thienne Johnson, por sua orientação e apoio no Doutorado Sanduíche.

Ao Gabriel Gian, Sibelius e Gabriel Martins meus amigos e alunos de IC, pelo apoio técnico e risadas nos momentos tensos.

Ao Eduardo, Victor, Alexandrina, Caio e Alan pela revisão do meu texto e apoio nas questões da psicologia e IA.

Ao Hospital Espírita de Marília pela parceria no Projeto de Extensão.

Aos amigos do LaSDPC e LIFeS, por compartilharem conhecimento e vivências.

As Reps Tacacrazy e Uva Preta que fizeram a minha vida em São Carlos mais divertida.

Aos amigos do caaso e federupas, paraenses e agregados, por participarem dos meus experimentos.

Aos amigos da 3AtoS e Fábrica de Artes, por terem incluído as artes cênicas na minha vida.

Aos amigos da IEQ-Botafogo que me acolheram como um filho em São Carlos.

Aos membros da Banca Examinadora, pelas contribuições no trabalho.

Aos Profissionais da USP, pela postura Ética.

À CAPES, CNPQ, FAPESP, ICMC/USP e UofA, pelo apoio financeiro.

*“As emoções são intermináveis.
Quanto mais as exprimimos,
mais maneiras temos de as exprimir.”
(Edward Morgan Forster)*

RESUMO

GONÇALVES, V. P.. **Uma abordagem para indicar o estado emocional de usuários em tempo de interação.** 2016. 115 f. Tese (Doutorado em em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

O estado emocional dos usuários influencia na tomada de decisão e é essencial para o conhecimento e explicação do comportamento dos usuários com aplicações computacionais. Sistemas computacionais fazem parte da vida cotidiana, influenciando o comportamento humano e estimulando mudanças nos estados emocionais. A avaliação das emoções dos usuários, durante a interação com aplicações computacionais, pode ajudar a fornecer interfaces flexíveis e melhores sistemas de recomendação. No entanto, as emoções são complexas e difíceis de identificar ou avaliar. Pesquisas anteriores demonstraram que o uso de sensores individuais, em um cenário do mundo real, não fornece uma avaliação emocional precisa. Acredita-se que somente uma visão holística pode permitir tirar conclusões significativas sobre o estado emocional dos usuários. Assim, nesta Tese, propõe-se uma abordagem chamada UserSense, que considera vários sensores para estimar os estados emocionais dos usuários em tempo de interação. A abordagem proposta, a partir de múltiplos sensores, considera várias entradas de usuários (como a fala, movimentos faciais, frequência cardíaca e atividades) e utiliza métodos de inteligência artificial para mapear essas diferentes respostas em um ou mais estados emocionais. Para a validação, os dados coletados pelos sensores durante a interação do usuário foram confrontados com os dos especialistas (psicólogos). A Teoria Componencial das Emoções e o Espaço Emocional Semântico de Scherer são utilizados na fundamentação teórica da abordagem UserSense que, baseada no *Middleware Adaptativo OpenCom*, incorpora funcionalidades para carregar e descarregar novos recursos, necessários a infraestrutura proposta. Os resultados experimentais mostram que a combinação de resultados gerados por vários sensores, fornece uma avaliação mais precisa dos estados emocionais do que considerar sensores individualmente.

Palavras-chave: Emoções, Avaliação, Sensores, Componentes, Psicólogos, Design Emocional, Computação Afetiva, Interação Humano-Computador.

ABSTRACT

GONÇALVES, V. P.. **Uma abordagem para indicar o estado emocional de usuários em tempo de interação**. 2016. 115 f. Tese (Doutorado em em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação (ICMC/USP), São Carlos – SP.

Users' emotional state influence decision making and is essential for the knowledge and explanation of users' behavior with computer applications. Computer systems are part of everyday life, influencing human behavior and stimulating changes in users' emotional states. The assessment of users' emotions during interaction with computer applications can help to provide tailorable interfaces and better recommendations systems. However, emotions are complex and difficult to identify or assess. Previous studies have shown that the use of single sensors, in a real-world scenario, does not provide accurate emotional assessment. We believe that only this holistic view can allow us to draw significant conclusions about the users' emotional states. Hence, in this Thesis we propose a approach called UserSense, that takes into account multiple sensors to estimate the users' emotional states at interaction time. The proposed multi-sensing approach considers several inputs from users (such as speech, facial movements, heart rate and activities), and uses artificial intelligent methods to map these different responses into one or more emotional states; for validation, the data collected by the sensors during the user interaction were confronted with specialists (psychologists). Componential Emotion Theory and Scherer's Emotional Semantic Space are used in the theoretical foundation of the UserSense approach, that takes into account Adaptive Middleware OpenCom to embed functionalities to load and unload new required features to proposed infrastructure. The experimental results show that the combination of outputs generated by multiple sensors provides a more accurate assessment of emotional states than considering sensors individually.

Key-words: Emotions, Assessment, Sensors, Components, Psychologists, Emotional Design, Affective Computing, Human-Computer Interaction.

LISTA DE ILUSTRAÇÕES

Figura 1 – Componentes emocionais de tempo real. As respostas emocionais de tempo real dos usuários são produzidas por quatro componentes. Esta pesquisa usa sensores para avaliar esses componentes.	29
Figura 2 – Taxonomia de métodos, técnicas e instrumentos para avaliação emocional. Adaptado de (XAVIER; NÉRIS, 2012).	31
Figura 3 – Octantes de classificação emocional adotados. As emoções são divididas em oito dimensões que refletem a experiência emocional. Adaptado de (SCHERER, 2005).	32
Figura 4 – Processo de aplicação da Abordagem Híbrida para a Avaliação da Experiência Emocional. Adaptado de (XAVIER; NÉRIS, 2014).	34
Figura 5 – O processo de estimativa das emoções do usuário em tempo de interação tem um mapeamento entre dados de sensoriamento e oitantes padrões (oitante padrão é construído para corresponder com a Figura 3). Modelo de classificação composto por 8 classificadores independentes são utilizados para obter valores emocionais intermediários para cada sensor e fornecer estados emocionais finais.	52
Figura 6 – Cenário de interação de componentes UserSense.	54
Figura 7 – Reconhecimento de Emoções Faciais.	55
Figura 8 – Detectando as emoções e processando as expressões faciais do usuário durante o jogo.	56
Figura 9 – Detectando as emoções e processando de dados de frequência cardíaca do usuário durante o jogo.	58
Figura 10 – Aplicação do Método de Decomposição Binária.	62
Figura 11 – Usuário interagindo durante o 1º experimento com o jogo Super Mario Bros. (a) usuário com o notebook, (b) tela do jogo, (c) captura pelo sistema de reconhecimento de emoções faciais.	64
Figura 12 – Cartão SAM(Self-Assessment Manikin) utilizado nos experimentos. Cartões SAM foram adotados para avaliar os Sentimentos Subjetivos ao final do jogo. Adaptado de (BRADLEY; LANG, 1994).	68
Figura 13 – Avaliação emocional do usuário antes e depois da interação com o jogo.	69

Figura 14 – Oitantes identificados para cada momento da interação, para um usuário. O eixo <i>Y</i> representa os oitantes e o eixo <i>X</i> representa o momento de interação (instância). Cada sensor identifica, em cada momento de interação, um conjunto de oitantes conforme mostra a figura. O estado emocional final para cada instância é dado pela eleição da maioria dos oitantes sugerida pelos sensores. Por exemplo, no instante 16, os oitantes sugeridos são 3, 4 e 8 - haja vista que, dois sensores sugerem esses oitantes, enquanto que os oitantes 2 e 5 são ignorados no momento, porque apenas um sensor os sugeriu.	70
Figura 15 – Resultados em tempo de interação (empregando sensores) e no final do jogo (empregando SAM). A avaliação estatística mostra diferenças significativas.	73
Figura 16 – Consumo de memória da UserSense.	74
Figura 17 – Acurácia da avaliação das emoções por meio de aprendizado de máquina e sensores individuais.	76
Figura 18 – Grafos gerados a partir da avaliação dos psicólogos.	78
Figura 20 – Similaridade entre os gráficos obtidos por cada sensor individualmente (S_I , S_{II} , S_{III}) e o grafo gerado pela avaliação dos especialistas.	79
Figura 21 – Similaridade entre os gráficos obtidos por cada sensor individualmente (S_I , S_{II} , S_{III}) ou pela composição dos sensores (S_{I+II} , S_{I+III} , S_{II+III} , $S_{I+II+III}$) e o grafo gerado pela avaliação dos especialistas.	79

LISTA DE TABELAS

Tabela 1 – Visão geral das principais pesquisas publicadas sobre a estimativa de emoções.	49
Tabela 2 – 41 atributos para descrever cada instância.	60
Tabela 3 – Resultados do experimento com 30 usuários. Pode-se verificar que os oitantes obtidos pelos sensores durante o tempo de interação diferem dos oitantes obtidos ao final do experimento (empregando SAM). Na segunda coluna, o valor percentual mostra o período de tempo que o usuário foi identificado no oitante específico.	72
Tabela 4 – Desempenho da Usersense no Motorola Moto X.	73

SUMÁRIO

1	INTRODUÇÃO	19
1.1	Contexto, Motivação e Problemática	19
1.2	Objetivos e Abordagem de Pesquisa	21
1.3	Hipóteses	23
1.4	Organização da Tese	24
2	EMOÇÕES NA INTERAÇÃO HUMANO-COMPUTADOR	25
2.1	Emoções	25
2.2	Teoria Componential das Emoções	28
2.3	Avaliação da Resposta Emocional	30
2.4	Considerações Finais	34
3	CAPTURA, ANÁLISE E CLASSIFICAÇÃO AUTOMATIZADAS DE REPOSTAS EMOCIONAIS	35
3.1	Conceituação	35
3.2	Trabalhos relacionados	37
3.3	Discussão	49
3.4	Considerações finais	50
4	ABORDAGEM PROPOSTA: USERSENSE	51
4.1	Visão Componentizada da UserSense	52
4.1.1	<i>UserSense no Contexto de Jogos Eletrônicos</i>	53
4.2	Sensores para Detecção de Emoções	54
4.2.1	<i>Sensor I (S_I): Reconhecimento de Emoções Faciais para Expressões Motoras</i>	55
4.2.2	<i>Sensor II (S_{II}): Sensores lógicos para Tendências Comportamentais</i>	56
4.2.3	<i>Sensor III (S_{III}): Reconhecimento de Fala para Avaliações Cognitivas</i>	56
4.2.4	<i>Sensor IV (S_{IV}): Frequência Cardíaca para Reações Fisiológicas</i>	57
4.3	Modelo de Classificação da UserSense	59
4.3.1	<i>Descrevendo as Instâncias da UserSense</i>	59
4.3.2	<i>Classificação Multirrótulo</i>	60
4.4	Considerações finais	61
5	AVALIAÇÃO EXPERIMENTAL DA USERSENSE	63

5.1	Planejamento dos Experimentos	63
5.2	Critérios de Inclusão de Usuários nos Experimentos	64
5.3	Critérios de Exclusão de Usuários dos Experimentos	65
5.4	Riscos em Participar dos Experimentos	65
5.5	Benefícios em Participar dos Experimentos	65
5.6	Tamanho da Amostra de Usuários	65
5.7	Considerações finais	66
6	AVALIAÇÃO DAS EMOÇÕES DURANTE E APÓS INTERAÇÃO .	67
6.1	Avaliando a Experiência Emocional do Usuário	67
6.2	Avaliação Não Paramétrica das Respostas Emocionais do Usuário .	70
6.3	Avaliação de Desempenho e Sobrecarga da UserSense	72
6.4	Considerações Finais	74
7	AVALIAÇÃO INTELIGENTE E POR SIMILARIDADE DE MÚLTIPLOS SENSORES	75
7.1	Combinando Aprendizado de Máquina e Múltiplos Sensores para Avaliar as Emoções do Usuário	75
7.2	Avaliação de Similaridade de Múltiplos Sensores	76
7.2.1	<i>Construção dos Grafos</i>	77
7.2.2	<i>Similaridade de Grafos</i>	78
7.2.3	<i>Análise dos Resultados</i>	78
7.3	Considerações finais	80
8	CONCLUSÕES	81
8.1	Síntese das Contribuições	81
8.2	Limitações e Trabalhos Futuros	83
8.3	Publicações e Trabalhos Desenvolvidos	84
8.4	Considerações Finais	85
	REFERÊNCIAS	87
APÊNDICE A	EMOÇÕES PRESENTES NA ESTRUTURA DO ESPAÇO EMOCIONAL SEMÂNTICO	97
APÊNDICE B	PARECER SUBSTANCIADO E APROVADO PELO COMITÊ DE ÉTICA DA FACULDADE DE MEDICINA DA USP	101
APÊNDICE C	TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO	105

APÊNDICE D	AUTORIZAÇÃO DE CAPTAÇÃO E EXIBIÇÃO DE IMAGEM, SOM E NOME	109
APÊNDICE E	QUESTIONÁRIO DE LEVANTAMENTO DE PERFIS DOS USUÁRIOS	113

INTRODUÇÃO

A emoção é um componente fundamental do ser humano, permitindo-o desfrutar de experiências e auxiliá-lo na tomada de decisões (SCHERER; CESCHI, 2000; BRAVE; NASS, 2003; ZHOU *et al.*, 2011; PETER; URBAN, 2012; RODRIGUES *et al.*, 2015; GONÇALVES *et al.*, 2016), papel importante na interação entre as pessoas (ZHOU *et al.*, 2011), por isso os aspectos emocionais vêm sendo estudados há muito tempo na área da psicologia (LICHTENSTEIN *et al.*, 2008).

Nos últimos anos, houve um aumento significativo de pesquisas sobre o tema das emoções na área de Interação Humano-Computador (IHC). Isso porque 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.

Concomitantemente ao crescimento do número de usuários de sistemas computacionais tem-se a possibilidade de capturar e analisar uma maior quantidade de dados desses indivíduos em contextos com maior variedade. Em um primeiro momento, a detecção da emoção pode ser utilizada para o design de novas maneiras de interação do usuário; posteriormente, pode caminhar para os dispositivos cientes de contexto (CAMPBELL; CHOUDHURY, 2012; RODRIGUES *et al.*, 2016). Nesse desenvolvimento, de acordo com o estado emocional do usuário em uma situação específica como de estresse ou mau humor, as Tecnologias da Informação e Comunicação (TICs) poderiam ser criadas com a finalidade de intervir na realização de uma tarefa, oferecendo sugestões (por exemplo, sugerir assistir um seriado para reduzir o estresse) ou mesmo tomando decisões em prol do usuário (por exemplo, pagar uma conta que irá vencer).

1.1 Contexto, Motivação e Problemática

A evolução das tecnologias, interfaces, artefatos e mesmo do usuário traz novos desafios aos pesquisadores de IHC uma vez que, a partir dela, percebe-se que há uma diversidade de perfis e novas formas de interação exigindo mais foco e atenção para os fatores humanos ao longo do desenvolvimento dessas tecnologias (GONCALVES; NERIS; UYAMA, 2011; CAMPBELL; CHOUDHURY, 2012; GONCALVES *et al.*, 2013a; SCHERER, 2013).

As TICs fazem parte do cotidiano, influenciam na estrutura das relações sociais e podem levar o usuário a sentir-se feliz ao receber uma sugestão de música de sites como o Pandora ¹ ou irritado por perder no jogo Super Mario Bros por Game Over (PEDERSEN; TOGELIUS; YANNAKAKIS, 2010). Esse fato evidencia que os sistemas computacionais não têm mais como único propósito ajudar os usuários a completar suas tarefas, porque a emoção e a cultura destacam-se como elementos importantes de estudo para a IHC.

(GELLER, 2014) afirma que o interesse científico na maneira como o humor e a emoção interferem na interação de usuários com sistemas computacionais parece ser recorrente nos dias atuais, no entanto o interesse de outras áreas sobre os aspectos emocionais dos indivíduos permeia os séculos. (TZU; PIN, 2015) refere-se à “arte de estudar o humor” como uma estratégia militar em seu clássico “A Arte da Guerra”. Outros livros notáveis incluindo “The expression of the emotions in man and animals” de (DARWIN, 1872) e “Essays on the Anatomy and Philosophy of Expression” (BELL, 1824) trazem indícios de que as emoções são psicofisiológicas e universais na natureza. Ainda nessa direção, (JAMES, 1884), em seu artigo “What is an Emotion?”, postulou que algumas emoções têm respostas distintas e, portanto, podem ser medidas, analisadas e classificadas. Aproximadamente um século depois, (PICARD, 1995) apresenta uma abordagem que considera a influência cultural e de aprendizado nas emoções, sendo o primeiro autor a afirmar categoricamente o valor das emoções na interação de usuários com sistemas computacionais.

De acordo com (BRAVE; NASS, 2003; GONÇALVES *et al.*, 2016), vários psicólogos argumentam que é impossível um indivíduo ter um pensamento ou executar uma ação sem envolver, mesmo que inconscientemente, o seu sistema emocional. Nesse sentido, considerar os aspectos emocionais no desenvolvimento de soluções de design, utilizar recursos computacionais para a captura de emoções e fazer uso de tais dados para atender as preferências, necessidades e intenções de uso do público alvo, enriquecem o processo de design, implementação e avaliação das interfaces de usuário (GONÇALVES *et al.*, 2012; RODRIGUES *et al.*, 2015; GONÇALVES *et al.*, 2016).

Diferentes abordagens têm sido utilizadas para avaliar a resposta emocional de usuários, isso porque apenas métodos qualitativos não são suficientes para a avaliação das respostas emocionais dos usuários, faz-se necessário investigar também abordagens que considerem dados quantitativos para a análise das respostas emocionais dos indivíduos (PETER; URBAN, 2012; RAMAKRISHNAN; EMARY, 2013; XAVIER; NERIS, 2014; RODRIGUES *et al.*, 2015; GONÇALVES *et al.*, 2016). Assim, um dos caminhos que se apresenta é fazer uso de instrumentos de autorrelato como Questionários, Think Aloud (SOMEREN *et al.*, 1994), The Self-Assessment Manikin (SAM) (BRADLEY; LANG, 1994) e Emocards (REIJNEVELD *et al.*, 2003). Ademais, outra alternativa para a avaliação emocional é o uso de sensores físicos e lógicos (por exemplo, verificando a frequência cardíaca e o tempo de interação do usuário) como forma de coleta de dados que podem ser modelados e analisados para se avaliar a resposta emocional do usuário.

Esta pesquisa de doutorado caminha em direção ao Design Emocional, haja vista que traz, para as decisões de design, os aspectos emocionais dos usuários (NORMAN, 2004). Isso porque nessa área de pesquisa, há a necessidade de considerar, durante o processo de design, além dos fatores de

¹ <http://www.pandora.com>

usabilidade e funcionalidade, também a experiência emocional ² dos usuários. Assim, para que os sistemas computacionais interativos possam reagir ao estado emocional de um indivíduo, é necessário uma infraestrutura capaz de capturar, analisar e classificar as repostas emocionais em tempo real.

No contexto do Design Emocional, (NORMAN, 2004) acredita que um design com boa usabilidade e utilidade não necessariamente seja o mais agradável de usar, e que um design atrativo não seja o mais eficiente. Desse modo, deve haver um equilíbrio entre a usabilidade, funcionalidade e a experiência emocional (HAYASHI; BARANAUSKAS, 2011; GONÇALVES *et al.*, 2015; RODRIGUES *et al.*, 2015). No mais, esse trabalho pretende ter uma visão holística da interação de usuários com sistemas computacionais, considerando sensores que possam abranger os quatro aspectos de tempo real que, segundo (SCHERER, 2005), compõem as emoções: avaliações cognitivas, tendências comportamentais, reações fisiológicas e expressões motoras, apresentados no Capítulo 2.

1.2 Objetivos e Abordagem de Pesquisa

O objetivo desta Tese é definir uma abordagem que auxilie a captura, análise e classificação de repostas emocionais dos usuários em tempo de interação (ou tempo de execução se for adotada uma visão técnica focando na perspectiva do sistema), visando permitir aplicações computacionais que possam se flexibilizar a partir de tal classificação (por exemplo, alterar elementos na interface de um jogo eletrônico para deixar o usuário menos estressado).

Para tanto, foi desenvolvido uma abordagem denominada de UserSense baseada no *Middleware* Adaptativo OpenCom (UEYAMA *et al.*, 2009; GONÇALVES *et al.*, 2015), que proporciona flexibilidade na análise das emoções. Para a coleta do estado emocional, a UserSense faz uso de sensores que capturam dados relacionados a diferentes componentes da emoção a partir da classificação proposta pela Estrutura do Espaço Emocional Semântico apresentada no Capítulo 2. Além da captura em tempo de interação, a UserSense é apoiada pelo *Middleware* Adaptativo OpenCom permitindo que componentes sejam adicionados e/ou excluídos a infraestrutura em tempo de interação. Por exemplo, caso o contexto de uso seja dispositivos móveis, pode-se adicionar ao UserSense o componente sensor de movimento, assim como, pode-se, também, adicionar o componente reconhecedor de face em um estudo de caso com videogame.

Em face de uma motivação, que enfatiza a interação do usuário com sistemas computacionais considerando aspectos emocionais, é interessante investigar mecanismos que têm sido aplicados para a captura, análise e classificação de repostas emocionais. Nesse sentido, foi elaborada e conduzida uma Revisão Sistemática na Literatura, seguindo o roteiro proposto por (KITCHENHAM, 2004), para extrair informações relacionadas à experiência emocional de usuários de sistemas computacionais. Atualmente, após uma revisão dos conceitos realizada por (KITCHENHAM *et al.*, 2010) para os termos Revisão Sistemática e Levantamento do Estado da Arte, convencionou-se nomear de Mapeamento Sistemático.

² O termo experiência emocional refere-se a seu significado mais genérico, utilizado para representar os aspectos emocionais referentes aos quatro domínios (Valência, Excitação, Sentimento de controle e Facilidade de conclusão do objetivo) responsáveis pela diferenciação entre emoções (SCHERER, 2005; GONÇALVES *et al.*, 2016).

Todos os passos envolvidos durante a realização da Revisão Sistemática foram documentados com o objetivo de garantir a validade dos resultados obtidos a partir dos dados analisados. Nesse sentido, as seguintes questões de pesquisa foram consideradas na seleção dos trabalhos relacionados.

- Quais são os métodos, algoritmos, instrumentos, ferramentas e técnicas utilizados para a captura, análise e classificação das emoções?
- Quais dos quatro aspectos de tempo real que compõem as emoções têm sido avaliados: avaliações cognitivas, tendências comportamentais, reações fisiológicas e expressões motoras?
- Quais dos quatro domínios responsáveis pela diferenciação das emoções têm sido considerados: Valência, Excitação, Sentimento de controle e Facilidade de conclusão do objetivo?
- Quais trabalhos consideram a análise empírica dos usuários e a avaliação analítica dos especialistas durante o processo de classificação das emoções?

Além das referências obtidas na Revisão Sistemática, novas referências foram obtidas por indicações de especialistas das áreas de IHC e Sensores e também por meio de novas buscas nas bases de dados utilizadas anteriormente durante a realização da Revisão Sistemática.

Os resultados da Revisão Sistemática apresentam indícios de que fazer com que sistemas computacionais interativos reajam ao estado emocional do usuário não é uma tarefa trivial, haja vista que se faz necessária a captura de respostas emocionais de maneira automatizada, em tempo real e por meio de diferentes fontes. Além da coleta de dados, recursos computacionais devem ser utilizados para analisar e classificar esses dados emocionais, com o objetivo de se concluir o estado emocional do usuário a partir do grau de Valência, Excitação, Sentimento de controle e Facilidade de conclusão do objetivo (domínios responsáveis pela diferenciação das emoções, de acordo com (SCHERER, 2005)).

A estratégia mais comum para a avaliação das emoções em tempo real é usar sensores. Todavia, a maioria dos trabalhos encontrados na literatura definem sensores apenas para um dos aspectos que compõem as emoções, por exemplo, reações fisiológicas. Ainda com relação a classificação das emoções, as pesquisas apresentam diferentes abordagens aplicadas apenas aos domínios de Valência e Excitação.

A UserSense busca contribuir com o Design Emocional, trazendo para as decisões de design os aspectos emocionais dos usuários. Dessa maneira, prover uma plataforma que apoie a análise das emoções e o impacto delas na experiência emocional do usuário, possibilitando, assim, uma visão holística dos indícios de respostas emocionais pode contribuir para uma redução dos problemas encontrados em análises emocionais unilaterais.

Para analisar a captura por diferentes tipos de sensores, aproximar-se do público alvo e entender melhor o processo de classificação da experiência emocional dos usuários, o autor desta Tese adotou o experimento exploratório de maneira iterativa e incremental como estratégia de pesquisa (YIN, 2003).

Nesse sentido, o OpenCom foi escolhido entre outros *Middleware* por ser genérico e flexível o suficiente a ponto de permitir, por exemplo, que de acordo com o dispositivo e contexto de uso, sensores possam ser acoplados e desacoplados na UserSense em tempo de interação. Pois, a cada novo experimento, sensores ruins foram substituídos ou aperfeiçoados; classificações semi-automatizadas passaram a ser

mais automatizadas e a fusão de sensores passou a ter uma melhor acurácia. Portanto, o intuito dos experimentos é verificar a viabilidade e refinar a solução proposta.

Para a coleta de dados, calibração dos sensores e parametrização do algoritmo de análise foi realizada a avaliação da interação de usuários com jogos eletrônicos altamente interativos, a partir de sensores como: câmera (reconhecimento de emoções expressas pela face), microfone (fala), frequência cardíaca (batimentos cardíacos) e sensores lógicos (sistemas de *software* que capturam dados sobre o desempenho do jogador na partida, por exemplo, se o usuário passou de fase ou venceu um obstáculo difícil).

Os experimentos contaram com a participação de estudantes de graduação e pós-graduação interagindo com *laptops*. Para a construção do modelo computacional usado pela UserSense, essa pesquisa contou com o auxílio de psicólogos especialistas na análise de emoções para verificar a viabilidade da proposta e auxiliar nas etapas de construção do modelo de aprendizado de máquina.

Ressalta-se que os procedimentos metodológicos foram desenvolvidos e são explicados em mais detalhes, na medida em que os capítulos são apresentados.

1.3 Hipóteses

Para a realização dos objetivos supramencionados, a hipótese principal - Tese - abordada nesta pesquisa é que para que os sistemas computacionais interativos possam reagir, ou seja, adaptar-se ao estado emocional de um indivíduo, é necessário criar infraestruturas de *software* e *hardware* (por exemplo, com auxílio de sensores e métodos de classificação) que apoiem a captura, análise e classificação de indícios de respostas emocionais de usuários visando permitir que aplicações computacionais possam se flexibilizar a partir de tal classificação.

As hipóteses secundárias são:

- Apesar da diversidade de métodos e maneiras de se avaliar a resposta emocional do usuário, a análise em tempo real e por meio de diferentes fontes traz desafios e uma riqueza de detalhes que destoam das abordagens unilaterais ou parciais, por exemplo, quando o usuário declara que está calmo durante um jogo e um sensor de frequência cardíaca verifica uma alta variação dos batimentos durante a interação.
- Acredita-se que as respostas emocionais de um ser humano não possam ser explicadas apenas pela soma dos seus componentes ou parte deles. Na visão holística defendida nesta Tese, o organismo humano como um todo determina como se comportam as partes e considera as suas inter-relações.
- O uso de abordagens, que fazem a análise em tempo real, agrega mais informações do que avaliações feitas apenas ao final da interação. Acredita-se que há uma diferença significativa entre a coleta de repostas emocionais em tempo de interação e aquelas realizadas após a interação.
- As emoções são capazes de influenciar na tomada de decisão dos usuários, sendo que, sistemas capazes de provocar emoções positivas são frequentemente mais utilizados. Nesse sentido, o Design

Emocional traz experiências emocionais dos usuários para o processo de interação, para que as TICs possam inspirar maior confiança e se tornem mais fáceis de serem aprendidas e utilizadas.

1.4 Organização da Tese

Esta Tese de Doutorado está dividida em oito capítulos. O Capítulo 1 contextualiza a proposta do projeto de pesquisa, apresentando a motivação, as hipóteses e os objetivos pretendidos.

O Capítulo 2 aborda os principais conceitos e teorias envolvidos nesta pesquisa, que serão utilizados para fundamentar a avaliação da experiência emocional do usuário em tempo de interação.

O Capítulo 3 faz um levantamento de trabalhos relacionados e aborda as principais técnicas e teorias utilizadas na captura, análise e classificação automatizadas da experiência emocional de usuários.

O Capítulo 4 apresenta a proposta defendida nesta Tese de Doutorado, trazendo detalhes sobre a abordagem UserSense: a arquitetura componentizada, os sensores utilizados nesta pesquisa, o mapeamento e análise dos estados emocionais, por fim, seus modelos de classificação.

O Capítulo 5 mostra detalhes sobre o planejamento e execução dos experimentos exploratórios, assim como as questões éticas envolvidas nesta pesquisa.

O Capítulo 6 trata de duas avaliações relacionadas a abordagem UserSense. A primeira refere-se a avaliação de emoções durante e após a interação de usuários com sistemas computacionais. Já a segunda é uma avaliação de desempenho e sobrecarga da UserSense em dispositivos móveis.

O Capítulo 7 é uma continuidade dos processos avaliativos. A primeira descreve a aplicação dos modelos de aprendizado de máquina, utilizados nesta Tese, para avaliar as emoções do usuário de maneira automatizada. A segunda mostra a aplicação da avaliação de similaridade de grafos para verificar se o uso de múltiplos sensores é mais eficiente do que usar sensores isoladamente.

Por fim, no Capítulo 8 é possível encontrar as conclusões desta tese, entre elas, as principais contribuições, as publicações obtidas, bem como suas limitações e as propostas de trabalhos futuros.

EMOÇÕES NA INTERAÇÃO HUMANO-COMPUTADOR

Durante décadas, pesquisadores de aspectos emocionais tentam definir um espaço dimensional que faça sentido para a diversidade da experiência emocional dos usuários. De acordo com (SCHERER, K. R., 2001), a emoção pode ser definida como um processo de inter-relacionadas mudanças sincronizadas nas relações de todos ou a maioria dos diferentes componentes do organismo, em resposta à avaliação de um estímulo de evento externo ou interno considerado relevante para as principais preocupações do organismo.

Diante do contexto dessa pesquisa em que o pesquisador considera as emoções a partir de uma visão holística, em que nada pode ser explicado pelo ordenamento ou conjunto de componentes, mas pelas relações que eles possuem uns com os outros e com o próprio todo, é adotado nesta pesquisa a Teoria Componencial das Emoções. Essa abordagem baseada em componentes fornece uma maneira abrangente para entender os diferentes aspectos das emoções e suas inter-relações, como destacado a Seção 2.2.

Apresentados os aspectos que compõem o estado emocional, é necessário definir uma abordagem que apoie a classificação das respostas emocionais dos usuários. Assim, este trabalho utiliza uma estrutura que ajuda a diferenciar estados emocionais de maneira sistemática e inter-relacionada, como mostrado na Seção 2.3

2.1 Emoções

Neste doutorado, considera-se que as emoções devem ser pensadas a partir de uma visão holística, em que nada pode ser explicado pela mera ordenação ou disposição de componentes, mas antes, pelas relações que eles mantêm entre si e com o próprio todo - uma oposição à filosofia reducionista. Acredita-se que não há uma definição que seja certa ou errada. Os pesquisadores que estudam as emoções e seus aspectos devem adotar as definições que melhor atendam suas necessidades e que estejam de acordo com seus preceitos.

Não existe uma definição consensual de emoção, mesmo entre os pesquisadores de psicologia,

e muito menos por meio das inúmeras disciplinas que estudam aspectos emocionais. Entretanto, a fim de contribuir com a visão holística, algumas definições consideradas pertinentes e adequadas serão aqui apresentadas.

Questões relacionadas ao papel das emoções na subjetividade humana tiveram início no mundo oriental com as discussões de Tzu nos anos de 470 a 399 a.C. (BERCHT, 2001) e, na civilização ocidental, com Sócrates (séc. VI a.C.). Platão (427-347 a.C.) fez uma primeira tentativa de estudo dos aspectos emocionais ao estruturar a alma: cognição, emoção e motivação (SCHERER, 1995).

(DARWIN, 1872), o pai da Teoria da Evolução, fez uso de seus estudos para explicar a presença de emoções nas pessoas e nos animais. Todavia, Darwin concentrou-se em discutir não a origem ou o papel das emoções nos seres humanos, mas o porquê de sua expressão da forma que se conhece. Darwin defende que as expressões fisiológicas das emoções foram selecionadas segundo a sua teoria da evolução, de forma que a sudorese excessiva e aumento dos batimentos cardíacos, por exemplo, observadas na resposta emocional de tensão, preparam o animal para a ação que pode vir a tomar.

(PINTO, 2001) defende que a emoção é uma experiência subjetiva que envolve a mente, o corpo e o organismo humano como um todo. Segundo o autor, é uma reação complexa desencadeada por um estímulo ou pensamento e envolve reações orgânicas e sensações pessoais. É uma resposta que envolve diferentes componentes, nomeadamente uma reação observável, uma excitação fisiológica, uma interpretação cognitiva, uma expressão motora e uma experiência subjetiva (ELLSWORTH; SCHERER, 2003; NIEDENTHAL; KRAUTH-GRUBER; RIC, 2006; SCHERER, 2005).

(DAMÁSIO, 2003) define emoções como uma variação psíquica e física desencadeada por um estímulo, sendo que as emoções são um meio natural de avaliar o ambiente ao redor e de reagir de maneira adaptativa.

Segundo (HÖÖK *et al.*, 2008), nas últimas décadas, houve um aumento do número de pesquisas sobre o papel da emoção aplicada a diversas áreas do conhecimento, como à psicologia (c.f., (ELLSWORTH; SCHERER, 2003)), à neurologia (c.f., (LEDOUX, 1998)), à medicina (por exemplo, (DAMASIO; MARG, 1995)) e à sociologia (c.f., (KATZ, 2001)). Esse investimento em novas pesquisas sobre a emoção estimulou inúmeras ideias entre os pesquisadores de IHC, sendo o trabalho de (PICARD, 1995) um dos pioneiros na área. Todavia, como em qualquer movimento dentro de IHC, há diferentes perspectivas teóricas sobre o tema (BOEHNER *et al.*, 2005).

A reação complementar e/ou contrária aos modelos cognitivistas de (PICARD, 1995; SCOLARI S.H.P.; MARAR, 2005) surgem, por exemplo, em pesquisas desenvolvidas por (SCHERER, 2005; BOEHNER *et al.*, 2005; HÖÖK *et al.*, 2008; GAVER, 2009; GONÇALVES *et al.*, 2016; FONTAINE; SCHERER; SORIANO, 2013). A reação, ao invés de considerar uma estrutura predominantemente cognitivista do Design Emocional, baseia-se na fenomenologia e analisa a emoção construída na interação entre as pessoas e entre pessoas e os sistemas computacionais.

Na IHC, entende-se a importância de considerar as emoções dos usuários explicitamente nos processos de avaliação e nas decisões de design. Entretanto, a discussão continua em pelo menos alguns dos elementos centrais defendidos por (FRIJDA, N. H. *et al.*, 2007; BRAVE; NASS, 2003; NORMAN, 2004; FONTAINE; SCHERER; SORIANO, 2013; FRIJDA; SCHERER *et al.*, 2009; SCHERER, 2005),

conforme descrito na sequência.

As emoções são extraídas quando algo relevante acontece com o organismo, tendo uma relação direta com as suas preferências, necessidades, intenções de uso, objetivos, valores e bem-estar geral. A relevância é determinada pela avaliação de eventos a partir de critérios, em particular a novidade ou imprevisibilidade de um estímulo ou evento, sua agradabilidade intrínseca ou desagradado, e sua consistência motivacional (SCHERER, K. R., 2001). Portanto, cabe ao organismo analisar os eventos favoráveis para satisfazer uma necessidade, definir uma preferência, alcançar um objetivo, ou defender um valor; ou até mesmo a obstrução para alcançar qualquer um desses.

Encontrar uma definição para a emoção, o afeto, o humor e o sentimento, não é uma tarefa simples, por se tratar de um assunto complexo e que vem gerando discussões entre pesquisadores da área da psicologia. O que se sabe é que não existe uma definição, ou certa, ou errada, ou única para cada um dos termos (LAURANS; DESMET; HEKKERT, 2009).

No trabalho realizado por (KLEINGINNA; KLEINGINNA, 1981), por exemplo, é possível encontrar uma lista com mais de cem definições para o termo emoção (HAYASHI; BARANAUSKAS, 2011; LAURANS; DESMET; HEKKERT, 2009; SCHERER, 2005). Em IHC, os termos afeto e emoção são frequentemente utilizados como sinônimos (BENTLEY; JOHNSTON; BAGGO, 2005). (NORMAN, 2004) escreve em seu trabalho o termo “sistema de afeto e emoção” e refere-se também ao termo “sistema emocional (afetivo)”, o que para (BENTLEY; JOHNSTON; BAGGO, 2005) sugere que ambos os termos estariam sendo utilizados como sinônimos. No entanto, (BENTLEY; JOHNSTON; BAGGO, 2005) acredita que fazer uma distinção sobre esses termos, poderia contribuir para o estudo do afeto e da emoção no contexto de IHC.

As emoções exercem controle sobre os episódios de ação de prontidão, no sentido de afirmar (nem sempre com sucesso) prioridades no controle do comportamento e experiência (FRIJDA, N. H. et al., 2007). Nesse contexto, (NORMAN, 2004) defende que o afeto pode ser definido como um termo geral para o sistema de julgamento, seja consciente ou subconsciente, e a emoção entendida como a experiência consciente do afeto.

O sentimento que uma pessoa sente e que não se sabe qual o motivo ou a causa é o afeto. Todavia, o sentimento consciente que se sabe a causa e o motivo é a emoção. Nessa visão, o afeto e a emoção se influenciam. O afeto que inclui a emoção é um sistema de julgamento sobre o que é bom ou ruim e do que é seguro ou perigoso (NORMAN, 2004; GONÇALVES *et al.*, 2016; RODRIGUES *et al.*, 2016).

O humor, assim como as emoções, é capaz de influenciar o comportamento (XAVIER; NÉRIS, 2012). O termo humor pode ser utilizado para descrever um estado provocado por um evento direto ou indireto de maneira não intencional (GONÇALVES *et al.*, 2016; RODRIGUES *et al.*, 2015). Sendo que, o humor representa uma experiência mais global, por exemplo, uma pessoa pode dizer que está triste por não conseguir realizar uma tarefa (trata-se de uma emoção uma vez que o motivo é conhecido) ou, que se sente deprimida sem ter um motivo ou causa específica (trata-se de humor visto que essa situação refere-se a um estado de ânimo) (BRAVE; NASS, 2003). Assim, uma pessoa com bom humor tende a ter uma visão positiva sobre qualquer coisa, como aguardar um download de um vídeo, sem se estressar, em um dia que a conexão com a Internet está ruim. Por outro lado, uma pessoa de mau humor tende a ter uma visão negativa, como ficar furiosa por não conseguir carregar a página do Facebook. Em adição, as

emoções frequentemente causam ou contribuem para o humor (BRAVE; NASS, 2003).

Como supra mencionado, outro termo importante, geralmente confundido com a emoção, é o sentimento. Para (SCHERER, 2005), o sentimento é definido como um dos componentes da emoção. O sentimento não representa o estado emocional de um indivíduo, mas sim, uma atribuição verbal que um indivíduo faz sobre um episódio emocional ¹ (BRAVE; NASS, 2003). Segundo (GONÇALVES *et al.*, 2015), é a interpretação consciente que o indivíduo realiza sobre o que ele sente no momento, e que é transmitida de uma maneira verbal, por exemplo, quando uma pessoa diz que se sente “feliz” por ter baixado um filme ou que está “frustrada” por não saber usar o site de compra de passagem aérea.

Ainda falando das questões emocionais, essa área de pesquisa envolve vários subsistemas do organismo que tendem a experiências em tempo real e a um certo grau de episódios emocionais, chegando ao ponto de se tornar altamente sincronizados (SCHERER, 2005). As emoções podem também preparar o organismo para lidar com eventos importantes em suas vidas e, assim, ter um estímulo motivacional forte, produzindo estados de ação de prontidão (FRIJDA, N. H. *et al.*, 2007).

Nesta Tese adota-se a visão de (SCHERER, 2005; FONTAINE *et al.*, 2002; ELLSWORTH; SCHERER, 2003; NIEDENTHAL; KRAUTH-GRUBER; RIC, 2006) para as emoções. Nesse sentido, o autor defende que a emoção é um episódio de alterações inter-relacionadas e sincronizadas nos estados de todos ou quase todos os quatro componentes (subsistemas) do organismo, em resposta à avaliação de um estímulo de evento externo ou interno que é considerado relevante para as preocupações principais do organismo. Os quatro componentes mencionados da emoção são: (a) avaliações cognitivas (avaliação de objetos e eventos), (b) reações fisiológicas (sensações corporais), (c) expressões motoras (face, voz, gestos), e (d) tendências comportamentais (preparação e direcionamento de ações).

2.2 Teoria Componencial das Emoções

A emoção é uma reação complexa que envolve todo o organismo do indivíduo, tendo uma relação direta com as suas necessidades, metas, valores e bem-estar em geral. Nesse sentido, diferentes componentes são estudados em pesquisas sobre a emoção e trazem à tona a discussão sobre o que são emoções e o que é necessário ser estudado quando se quer delimitar o significado da emoção (FONTAINE *et al.*, 2002; SCHERER, 2005). Alguns pesquisadores se concentram nos antecedentes das emoções e a avaliação de situações (FRIJDA, 1987; PARKINSON, 1995; SCHERER, K. R., 2001), em termos de bem-estar de um indivíduo (ELLSWORTH; SCHERER, 2003). Outros pesquisadores estudam os padrões de reação psicofisiológica (EKMAN; FRIESEN, 1969; 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 experiências subjetivas (YIK; RUSSELL; BARRETT, 1999) de emoção, bem como inibição e controle sobre as emoções (FONTAINE *et al.*, 2002).

Diante do exposto, a Teoria Componencial das Emoções, descrita por (ELLSWORTH; SCHE-

¹ O estado emocional refere-se ao exato momento em que uma emoção foi gerada (SCHERER, 2005). Diferentemente do episódio emocional que refere-se usualmente à experiência que o indivíduo consegue recordar quando é questionado (FRIJDA *et al.*, 2000).

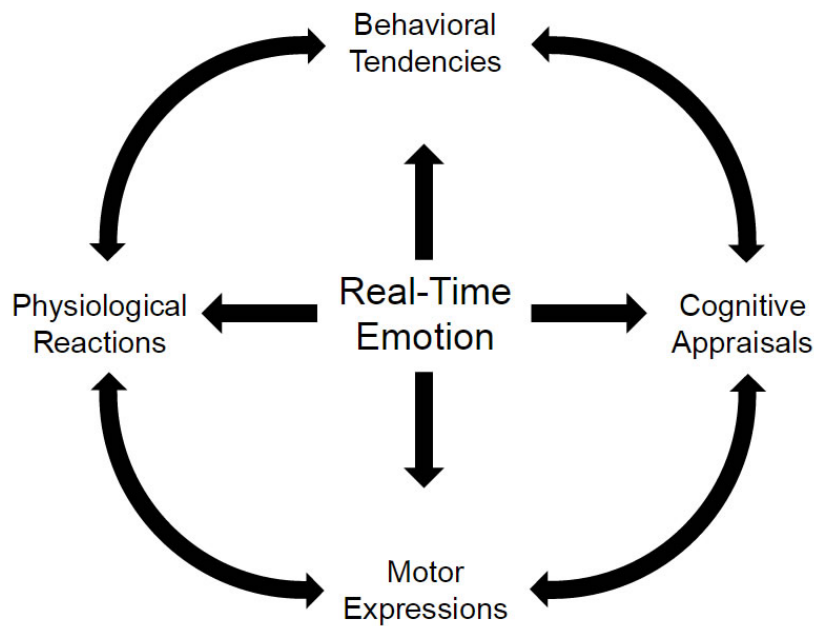


Figura 1 – Componentes emocionais de tempo real. As respostas emocionais de tempo real dos usuários são produzidas por quatro componentes. Esta pesquisa usa sensores para avaliar esses componentes.

RER, 2003; NIEDENTHAL; KRAUTH-GRUBER; RIC, 2006; SCHERER, 2005), entende as emoções como sendo níveis variáveis de mudanças inter-relacionadas entre um conjunto de quatro componentes: avaliações cognitivas, reações fisiológicas, expressões motoras e tendências comportamentais.

Vale mencionar que esta pesquisa apresenta uma abordagem para a análise dos componentes emocionais de tempo real, portanto, não foi abordada a análise de episódio emocional (Sentimento Subjetivo), de maneira prioritária. De acordo com (SCHERER, 2005), não há método objetivo capaz de avaliar os Sentimentos Subjetivos, pois a única maneira de obter essas informações é perguntando ao próprio indivíduo (*feedback*).

Os pesquisadores (MAHLKE; MINGE, 2008) destacam que uma abordagem baseada nesses componentes oferece uma maneira abrangente para entender os diferentes aspectos das emoções em todo tipos de situações. Assim, diferentes aspectos podem ser utilizados para a análise emocional, sem, no entanto, deixar de considerar a relação entre os diferentes componentes (SCHERER, K. R., 2001). A seguir, são apresentados e definidos cada um dos quatro componentes da emoção de tempo real definidos pelos autores (ver Figura 1).

As *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 a se adaptar a uma determinada situação (EKMAN; FRIESEN, 1969). Exemplos desse componente incluem: o tempo necessário para completar uma tarefa, o número de erros, a precisão de 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).

As *Avaliações Cognitivas* estão ligadas à interpretação de uma situação e contribuem para o desenvolvimento das emoções. O indivíduo está constantemente avaliando o mundo ao seu redor,

procurando perceber as qualidades afetivas de eventos, objetos e recursos (FONTAINE *et al.*, 2002). Fazer isso implica em utilizar recursos sensoriais, perceptivos e cognitivos, incluindo o pensamento racional. Contexto e avaliação estão ligados, acentuando as funções de adaptação das emoções, auxiliando o indivíduo a preparar a Tendência Comportamental adequada aos incidentes com os possíveis resultados relevantes. Uma mesma situação pode induzir diferentes emoções, dependendo da interpretação feita pelo usuário (MAHLKE; MINGE, 2008; SCHERER, 2005).

As *Expressões Motoras*, também conhecidas como reações expressivas, são responsáveis por comunicar as Tendências Comportamentais (MAHLKE; MINGE, 2008; SCHERER, 2005). Esse componente das emoções implica em alterações nas expressões faciais, vocais e gestuais, que acompanham a experiência emocional do usuário. Nesse contexto, o rosto e a voz sofrem mudanças com o grau de Excitação (LANG *et al.*, 1993). Em termos de respostas emocionais, tem-se, por exemplo, um olhar de ódio, franzir a testa, comprimir os lábios ou até mesmo um sorriso. As Expressões Motoras estão relacionadas às características do discurso da fala, como a velocidade, a intensidade, a melodia e o som. Verbalmente, um “fino” tom de voz poderá ser produzido uma vez que certas partes da área vocal, como faringe, cavidade nasal e laringe, fiquem secas e enrijecidas (LEVENTHAL, 1984).

E por fim, as *Reações Fisiológicas*, também são conhecidas como Mudanças Psicofisiológicas (FONTAINE *et al.*, 2007), Neurofisiológicas (HAYASHI; BARANAUSKAS, 2011) ou Ativação Fisiológica (LAURANS; DESMET; HEKKERT, 2009). Essas são aspectos emocionais que permitem ao usuário revelar espontaneamente e inconscientemente suas emoções. De acordo com (STEMMLER, 2003; SCHERER, 2005; MAHLKE; MINGE, 2008), esse componente regula o Sistema Nervoso Central (SNC), que recebe informações dos cinco sentidos e direciona as ordens ao sistema motor; ao Sistema Neuroendócrino (SNE), que é responsável pela regularização das mudanças internas do corpo e ao Sistema Nervoso Autônomo (SNA), este controla as funções do corpo, respiração, digestão, circulação, entre outros.

A Teoria Componencial das Emoções, com todas as teorias e componentes mencionados acima, oferece uma abordagem teórica e empírica para compreensão de como mensurar as emoções seguindo a visão holística defendida nesta Tese.

2.3 Avaliação da Resposta Emocional

De acordo com a posição desses autores (ELLSWORTH; SCHERER, 2003; NIEDENTHAL; KRAUTH-GRUBER; RIC, 2006; SCHERER, 2005), a mesma defendida nesta Tese, a avaliação emocional deve analisar os quatro componentes da emoção: (a) avaliações cognitivas (avaliação de objetos e eventos), (b) reações fisiológicas (sensações corporais), (c) expressões motoras (face, voz, gestos), e (d) tendências comportamentais (preparação e direcionamento de ações).

Considerando os componentes responsáveis pelas emoções, são inúmeros os métodos, técnicas, ferramentas e instrumentos utilizados para avaliação das emoções. A Figura 2 apresenta uma taxonomia que classifica tais emoções em verbal ou não verbal, que tratam de algumas medidas para avaliação das repostas emocionais disponíveis na literatura. As medidas verbais são aquelas em que o usuário explicitamente fala o que está sentindo, ou ainda, utilizam palavras escritas para verbalizar as emoções

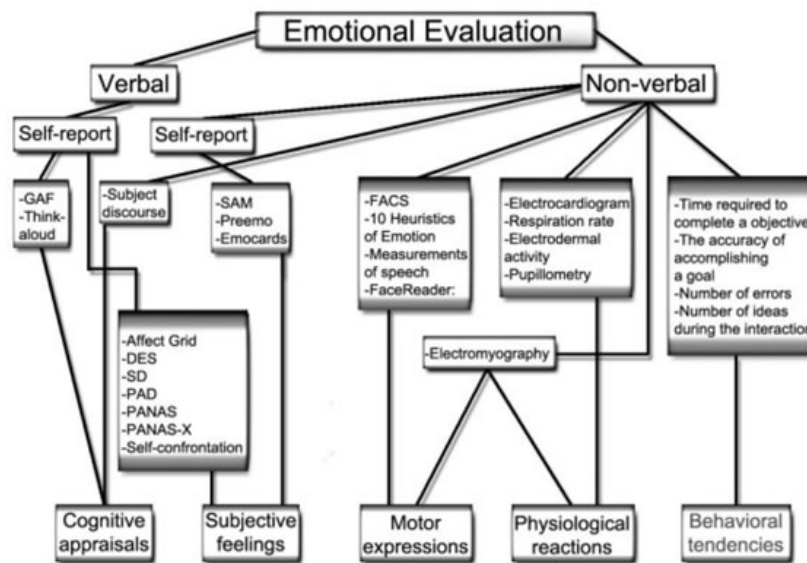


Figura 2 – Taxonomia de métodos, técnicas e instrumentos para avaliação emocional. Adaptado de (XAVIER; NÉRIS, 2012).

(DESMET, 2004); já as medidas não verbais são consideradas discretas, independente da cultura e linguagem; entretanto, elas também podem ser subjetivas.

Seguindo a taxonomia apresentada na Figura 2, medidas não verbais que podem ser utilizadas para a avaliação das expressões motoras incluindo o Facial Action Coding System (FACS) (HAGER; EKMAN; FRIESEN, 2002) e as Dez Heurísticas da Emoção (LERA; GARRETA-DOMINGO, 2007). A avaliação heurística das emoções é um processo custoso, não automatizado, que depende do especialista e não permite que uma aplicação computacional reaja ao seu resultado em tempo de interação. Em contrapartida, o FACS fornece o reconhecimento de emoções faciais em tempo real.

Concernente as tendências comportamentais, por exemplo - dois modelos, usados na IHC, podem auxiliar na avaliação de indicadores de desempenho: GOMS e o Keystroke Level (PREECE; ROGERS; SHARP, 2005). O GOMS é um modelo de engenharia oneroso, que requer extrema atenção do especialista para modelar todas as operações necessárias para que o usuário realize uma tarefa de um *software* (PREECE; ROGERS; SHARP, 2005).

(SCHERER, 2005) propôs uma abordagem de recurso de design a partir da qual as emoções são definidas na medida em que: (a) concentram-se em eventos específicos, (b) envolvem a avaliação de características intrínsecas dos objetos ou eventos, bem como de sua consistência, motivo e a presença de condições favoráveis a motivos específicos, (c) afetam a maioria ou todos os subsistemas (componentes) corporais que podem tornar-se, até certo ponto, sincronizados e (d) estão sujeitas as mudanças rápidas.

No contexto de IHC, a classificação dos diferentes termos emocionais é importante para estudar o impacto das emoções na experiência emocional do usuário (GONÇALVES *et al.*, 2015; RODRIGUES *et al.*, 2015; GONÇALVES *et al.*, 2016).

Além de definir as medidas utilizadas para a coleta das emoções, é importante identificar uma maneira de classificá-las, ou seja, relacionar as diferentes emoções existentes. Sabe-se que as emoções são distintas, no entanto, assim como (RUSSELL, 1980) sugere, elas estão relacionadas de maneira

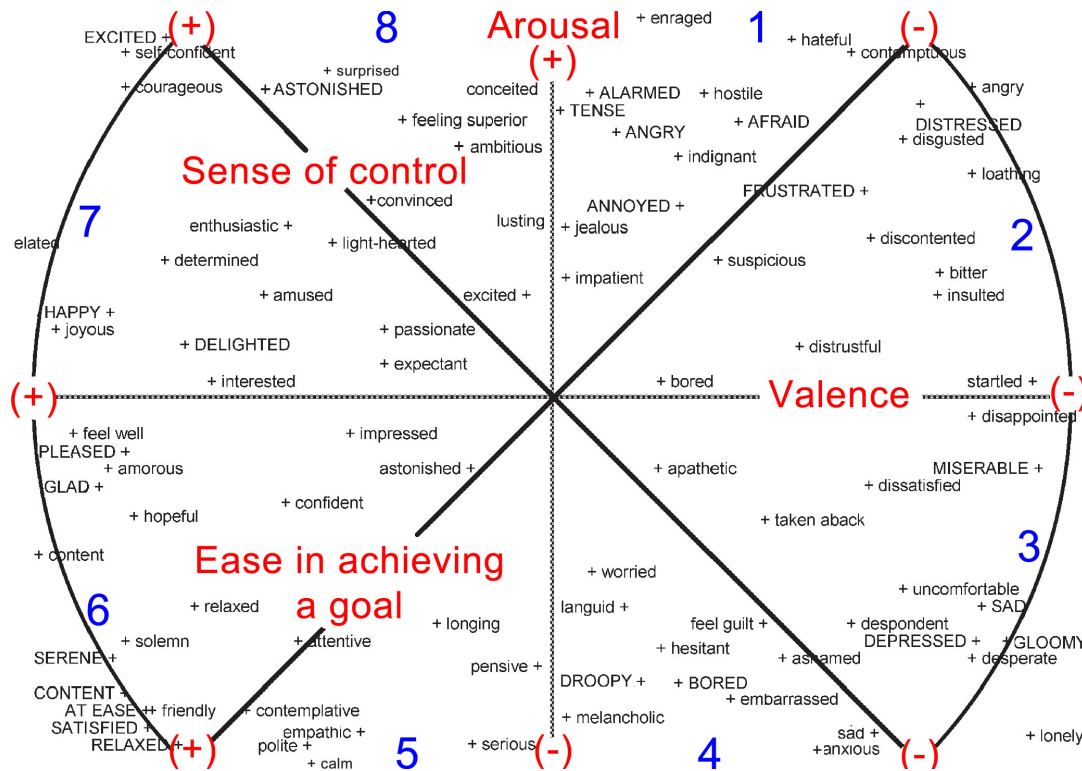


Figura 3 – Octantes de classificação emocional adotados. As emoções são divididas em oito dimensões que refletem a experiência emocional. Adaptado de (SCHERER, 2005).

sistemática.

No contexto da IHC, a classificação de diferentes estados emocionais precisa ser priorizada para compreender o impacto das emoções sobre a experiência emocional. Uma maneira de classificar as emoções surgiu com o trabalho desenvolvido por (SCHERER, 2005). Nele, ele realiza uma extensão do modelo bidimensional (Valência e Excitação) de (RUSSELL, 2003), por meio da inclusão destes dois domínios que possuem um forte impacto na diferenciação das emoções, que são: “goal conduciveness” (avalia a facilidade de atingir um ou vários objetivos/necessidades, a partir de uma consequência de ações ou eventos) e “coping potential” (avalia o sentimento de controle do organismo sobre um evento). (SCHERER, 2005) realiza também a inclusão de oitenta termos emocionais adicionais que podem ser visualizados na Figura 3. Essa classificação é denominada “Estrutura do Espaço Emocional Semântico”. O Apêndice A apresenta uma tabela de emoções provenientes de uma versão traduzida dessa estrutura. Os domínios de experiência emocional no contexto do IHC podem ser resumido como:

- Excitação (*Arousal*): domínio fortemente relacionado ao impacto de excitação e motivação do usuário;
- Sentimento de controle (*Sense of control*): reflete o sentimento de controle do usuário sobre a interação/dispositivo;
- Valência (*Valence*): representa o grau de satisfação, prazer e desprazer;
- Facilidade de conclusão do objetivo (*Ease in achieving a goal*): reflete a facilidade e a dificuldade de alcançar um objetivo.

Segundo (XAVIER; NÉRIS, 2012), por meio da utilização da Estrutura do Espaço Emocional Semântico, é possível identificar o impacto que uma emoção provocada reflete nos quatro domínios da experiência emocional (Valência, Excitação, Sentimento de controle e Facilidade de conclusão do objetivo). Como exemplo dessa identificação, pode-se citar que a emoção “tense” está diretamente ligada a um alto grau de excitação, à obstrução da realização da tarefa, uma satisfação intermediária, e sugere um alto sentimento de controle do usuário.

Tendo em vista os aspectos multidimensionais da Estrutura do Espaço Emocional Semântico de (SCHERER, 2005), será feito uso dessa abordagem ao considerar a visão holística que a proposta reflete sobre os componentes emocionais do indivíduo, a saber: avaliações cognitivas, mudanças psicofisiológicas, expressões motoras e tendências comportamentais (ELLSWORTH; SCHERER, 2003; NIEDENTHAL; KRAUTH-GRUBER; RIC, 2006; SCHERER, 2005).

Diante do exposto, nesta Tese adota-se a posição de que a experiência emocional de um usuário pode ser avaliada a partir da relação entre os quatro domínios descritos anteriormente e as respostas emocionais coletadas por meio de sensores (SCHERER; CESCHI, 2000; GONÇALVES *et al.*, 2016; GONCALVES *et al.*, 2013b), conforme descrito a seguir.

Adaptando a Abordagem Híbrida para o contexto desta Tese, na primeira etapa (Figura 4), defini-se que os sensores podem ser usados para avaliar os componentes da emoção (SCHERER, 2005). Quatro desses cinco componentes (tendências comportamentais, reações fisiológicas, expressões motoras e avaliações cognitivas) podem dar indícios sobre o estado emocional de um usuário durante uma interação, sem interrompê-la. Por outro lado, o componente sentimento subjetivo permite a avaliação de um episódio emocional, em outras palavras, uma avaliação subjetiva de um estado emocional já ocorrido (*feedback*). Desse modo, sugere-se que os sensores podem ser adotados para avaliação dos quatro primeiros componentes de emoção, em tempo de interação e a avaliação dos sentimentos subjetivos pode ser utilizado para confrontar os resultados obtidos com os outros quatro componentes (conforme pode ser observado no estudo apresentado no Capítulo 6).

Na segunda etapa, são analisados os resultados das medições efetuadas durante a interação (Figura 4). Os resultados de cada medição são generalizados em termos de positivo, neutro ou negativo. Cada resultado positivo ou negativo está relacionado com o respectivo hemisfério (um dos quatro oitantes), quer pertencendo ao domínio positivo ou negativo da Estrutura do Espaço Emocional Semântico. Por exemplo, os designers muitas vezes querem saber se o usuário vai se sentir feliz e ter uma elevada sensação de controle durante a interação, e, assim, também ficar feliz com o dispositivo, para que a tarefa possa ser realizada sem qualquer dificuldade. No entanto, dependendo de cada caso, o designer pode desejar que o usuário fique nervoso e/ou animado durante a interação, ou mesmo que o usuário deva sentir-se calmo e relaxado.

Nessa abordagem, mesmo que algumas das medidas avaliadas sejam precisas e apenas um oitante seja identificado, como resultado, é necessário generalizar as medições de modo que a relação entre os resultados das medições generalizadas e os hemisférios de cada domínio da experiência emocional possa permitir que cada oitante tenha a mesma probabilidade de ser escolhido. Assim, isto irá tornar possível identificar o(s) oitante(s) resultante(s) através de um processo que envolve a incidência dos oitantes (respostas) que são escolhidos no mesmo passo.

No terceiro passo (Figura 4), deve-se incrementar os oitantes na base dos resultados das medições avaliadas para alcançar a incidência dos oitantes. Por exemplo, se uma medição está relacionada com o domínio da excitação, tem um resultado positivo e a abordagem também escolhe os oitantes 1,2,7 e 8 com o hemisfério positivo para o domínio da excitação, é preciso aumentar a incidência de cada um dos quatro oitantes em uma única unidade.

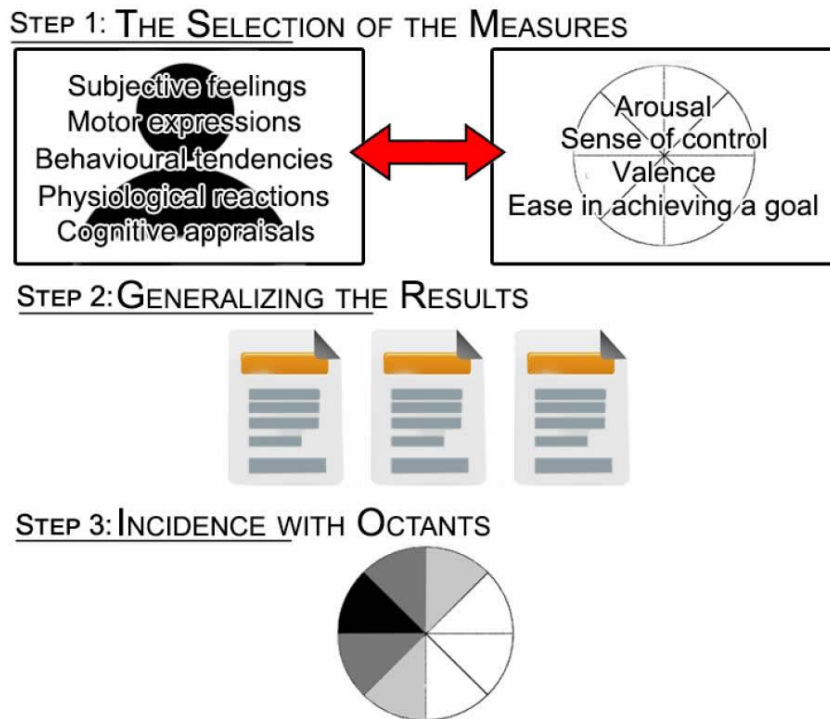


Figura 4 – Processo de aplicação da Abordagem Híbrida para a Avaliação da Experiência Emocional. Adaptado de (XAVIER; NERIS, 2014).

2.4 Considerações Finais

Esta Tese defende que, para a avaliação da resposta emocional de um indivíduo, é necessário considerar: (a) as contínuas mudanças nos processos de avaliação em todos os níveis de processamento do sistema nervoso central, (b) os padrões de respostas geradas nos sistemas nervoso neuroendócrino, autonômico e somático, (c) as mudanças motivacionais produzidas pelo resultado da avaliação, de tendências de ações particulares e (d) os padrões de expressões facial e vocal, bem como os movimentos do corpo.

Nessa pesquisa, considera-se que a experiência emocional de um usuário pode ser avaliada com base na relação entre os quatro domínios da Estrutura do Espaço Emocional Semântico e as respostas emocionais que são coletadas por sensores em tempo real. No mais, considerando os aspectos multidimensionais da Estrutura do Espaço Emocional Semântico, é possível determinar indícios de uma emoção provocada quando refletida nos quatro domínios da experiência emocional. Por exemplo, pode-se mencionar que o sentimento de frustração está diretamente ligado ao que está sendo obstruído na realização de uma tarefa, uma sensação de controle intermediário e um alto grau de excitação apontam para uma experiência negativa.

CAPTURA, ANÁLISE E CLASSIFICAÇÃO AUTOMATIZADAS DE REPOSTAS EMOCIONAIS

Esta pesquisa enfatiza a interação do usuário com sistemas computacionais considerando aspectos emocionais de tempo real. Frente a esse contexto, foi elaborada e conduzida uma Revisão Sistemática que investigou, na literatura, os mecanismos aplicados para a captura, análise e classificação de respostas emocionais. Os resultados dessa Revisão Sistemática foram publicados e estão disponíveis em (GONCALVES *et al.*, 2013b). No mais, esta seção traz itens de interesse observados a partir desse levantamento bibliográfico, assim como a atualização dessa revisão.

Para estimar um estado emocional, uma visão parcial ou unilateral dos indícios emocionais apresentados por um usuário podem não revelar a riqueza de detalhes que uma visão abrangente apresenta. Como não há maneira para classificar uma resposta emocional como certa ou errada, sugere-se a estratégia de considerar diferentes fontes de dados. Acredita-se que somente uma visão holística pode permitir que se tire conclusões mais significativas sobre o estado emocional dos usuários.

3.1 Conceituação

Segundo (DESMET, 2004), 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”. Diante do exposto, as expressões motoras estão relacionadas com as expressões faciais, gestos corporais e características do discurso da fala como: a velocidade, a intensidade, a melodia e o som.

Além das expressões motoras, as reações fisiológicas permitem avaliar respostas emocionais dos usuários em um contexto experimental, uma vez que os usuários revelam espontaneamente e inconscientemente suas emoções. Em (LANE *et al.*, 2011), é possível verificar que a coleta de dados foi realizada para verificar se o usuário está tendo um bom relacionamento social; se ele está tendo tempo necessário para

descanso por meio do microfone; e se está praticando exercícios físicos por meio do acelerômetro e do GPS.

De acordo com (MAHLKE; MINGE, 2008), estudos sugerem que existe uma relação entre a dilatação da pupila e o prazer de um estímulo, ou seja, os estímulos agradáveis são suscetíveis à geração de pupilas mais dilatadas. O largo espectro de aplicações e o constante aumento da capacidade de processamento dos computadores digitais têm motivado pesquisadores a tentar reproduzir tal habilidade em máquinas e programá-las para executar tarefas cada vez mais complexas (FILHO *et al.*, 2014). De fato, procedimentos de classificação têm ajudado na análise de repostas emocionais: médicos no diagnóstico da depressão, usuários que buscam mudança de comportamento em sua rotina, pilotos em manobras de aterrissagem, etc.

Pesquisas em ciência da computação têm, nas últimas décadas, gerado uma ampla quantidade de métodos que “aprendem” a partir de exemplos, e que podem ser utilizados para extrair dados empíricos para a classificação a partir de padrões (DUDA; HART *et al.*, 1973; CLARK; NIBLETT, 1989; CRISTIANINI; SHAW-TAYLOR, 2000; FILHO *et al.*, 2014). Tais técnicas de classificação estão, cada vez mais, sendo aplicadas para a análise de respostas emocionais (ØHRN; ROWLAND, 2000; CHANEL *et al.*, 2009; ZHOU *et al.*, 2011; RAMAKRISHNAN; EMARY, 2013; GONÇALVES *et al.*, 2016; GONCALVES *et al.*, 2013b).

(DUDA; HART *et al.*, 1973) definem um problema de classificação, como encontrar uma função $f(x)$ de uma base de dados composta pelo par (x_i, y_i) , na qual y_i PERTENCE A $1, \dots, k$ e k é o número de classes do problema. Ou seja, um procedimento de classificação é algum método formal capaz de decidir, com base em informações fornecidas, a que grupo ou população um determinado objeto pertence. De uma maneira geral, este projeto está interessado em buscar e testar métodos que, a partir de diferentes fontes de dados, possa analisar e classificar, em tempo de interação, o estado emocional de um usuário interagindo com um sistema computacional. Abaixo, segue uma súmula conceitual de alguns algoritmos de classificação utilizados no reconhecimento de emoções.

kNN – k-nearest neighbor classification - Classificação dos k-vizinhos mais próximos. O kNN, usado nos trabalhos de (BAILENSEN *et al.*, 2008; YUAN *et al.*, 2010), é uma abordagem de classificação não-paramétrica, por exemplo, não assume nenhuma distribuição de probabilidade à priori para os dados. 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 *et al.*, 2011).

SVM – Support Vector Machine. O algoritmo SVM, usado por (KATSIS *et al.*, 2008; RAMAKRISHNAN; EMARY, 2013; PETER; URBAN, 2012; LITTLEWORT *et al.*, 2011) para o reconhecimento de emoções, é um dos métodos mais robustos e acurados entre os algoritmos conhecidos. Segundo (VAPNIK, 2013; FACELI *et al.*, 2011), o objetivo do SVM é encontrar a melhor função para distinguir duas classes dos dados de treinamento. A garantia da melhor função é atingida porque o SVM maximiza a margem entre as classes, oferecendo assim a melhor habilidade de generalização. De acordo com (VAPNIK, 2013), uma margem pode ser interpretada como uma quantidade de espaço ou distância entre as classes, conforme definido pelo hiperplano.

Fuzzy Logic - Lógica Fuzzy. A Lógica Fuzzy, usada nos trabalhos de (KATSIS *et al.*, 2008; MANDRYK; ATKINS, 2007; PETER; URBAN, 2012), é capaz de processar dados incompletos e prover

soluções aproximadas para problemas que outros métodos encontram dificuldades em solucionar (CLARK; NIBLETT, 1989).

Decision Tree – Árvore de Decisão. A Árvore de Decisão, aplicada nos trabalhos de (BAILEN-SON *et al.*, 2008; PETER; URBAN, 2012; FRANTZIDIS *et al.*, 2010), utiliza uma árvore para realizar a classificação ou a estimação de um valor de um determinado teste. Um dos algoritmos mais famosos para gerar essas árvores é o C4.5. O C4.5 cresce a árvore utilizando uma abordagem dividir-conquistar, a árvore é ramificada utilizando o atributo na obtenção do melhor ganho de informação (QUINLAN, 1986).

Bayesian Networks – Redes Bayesianas. Rede Bayesiana é um modelo de grafo probabilístico aplicado nos trabalhos de (PETER; URBAN, 2012; RAMAKRISHNAN; EMARY, 2013; CHANEL *et al.*, 2009) e representa um conjunto de variáveis aleatórias e suas dependências condicionais como um grafo acíclico direto (FACELI *et al.*, 2011). Por exemplo, uma rede bayesiana pode representar a relação entre os sintomas e as doenças. Dados os sintomas, podemos inferir a probabilidade de várias doenças.

A área de aprendizado de máquina visa construir modelos que aprendam com dados. Dentro do aprendizado de máquina, se encontram os algoritmos de classificação, que são métodos de aprendizado supervisionado. No aprendizado supervisionado, os conjuntos de exemplos são rotulados, ou seja, sabe-se a qual classe pertence cada instância (QUINLAN, 1986; FACELI *et al.*, 2011; VAPNIK, 2013). No contexto de emoções, por exemplo, pode-se considerar que cada instância é uma interação com um sistema computacional, e assim tem-se a emoção correspondente para cada uma dessas interações. Os algoritmos de classificação tentam generalizar uma função ou mapeamento da entrada para a saída (rótulos) que posteriormente será utilizado para prever a saída de exemplos (ou instâncias) não vistos.

O reconhecimento de emoções é normalmente abordado de dois modos: (a) o primeiro trata de um problema de classificação, no qual o objetivo é treinar um modelo que retorna como saída de uma emoção (Felicidade, Tristeza, por exemplo), a entrada do modelo utiliza-se de um vetor de *features*, por exemplo, características da face ou da voz, a posição da boca ou a posição das sobrancelhas; e o rótulo de tal vetor de *features*, a emoção presente nas características foi determinado por um especialista; (b) o segundo modela o problema utilizando regressão.

Regressão é um modelo que prediz um valor real como saída. Classificação é um modelo que prediz uma classe de um conjunto finito como saída. Portanto, o modelo de regressão é treinado utilizando um vetor de *features* mais o rótulo dos mesmos (nesse caso, o rótulo não é uma emoção - Felicidade, Tristeza-, mas sim valores de Valência e Excitação). Como resultado, o modelo de regressão ao receber um vetor de *features* desconhecidos prediz os valores de Valência e Excitação para tal vetor.

3.2 Trabalhos relacionados

Considerando que esta pesquisa traz uma abordagem para apoiar a captura, análise e classificação de indícios de respostas emocionais de usuários em tempo de interação, a seguir, serão apresentados alguns trabalhos apontados pela literatura que estão relacionados a esta Tese.

- *Speech emotion recognition approaches in human computer interaction* (RAMAKRISHNAN; EMARY, 2013)

O principal objetivo do trabalho desenvolvido por (RAMAKRISHNAN; EMARY, 2013) foi apresentar e avaliar alguns recursos empregados no reconhecimento de emoções em discurso, do inglês *Speech Emotion Recognition* (SER), assim como descrever diversas áreas de interesse onde esses recursos podem ser aplicados. Para as avaliações, foram utilizadas duas bases de dados: Berlin emotional database e Danish emotional database. Tais bases de dados possuem gravações em idiomas distintos, respectivamente alemão e dinamarquês, de discursos realizados por profissionais de maneira atuada (representação cênica) e captadas por meio do microfone.

Os *software* openSMILE e Praat foram utilizados para extrair características relevantes desses discursos. Essas particularidades pertencem a dois grupos principais: acústicas e linguísticas. No primeiro grupo estão contidas propriedades referentes a descritores de baixo nível (da sigla em inglês LLD), características relacionadas às pausas e aspectos Zipf. Por outro lado, o segundo grupo abrange particularidades contidas nas alterações gramaticais. Existem diferentes abordagens para as características linguísticas (segundo grupo), tais como: Árvores Semânticas (*Semantic Trees*), Análise Semântica Latente (*Latent Semantic Analysis*), Aprendizado Baseada em Transformação (*Transformation-Based Learning*), Redes Bayesianas (*Bayesian Networks*) e Modelagem do Espaço Vetorial (*Vector Space Modeling*).

Posteriormente, (RAMAKRISHNAN; EMARY, 2013) fazem a extração das características com o auxílio dos algoritmos classificadores – SVM e Modelo Oculto de Markov (*Hidden Markov Models* - HMM) – esses algoritmos foram implementados e executados no *software* MATLAB. Para discriminar o estado emocional, são utilizados os domínios de Valência (negativo-positivo) e Excitação (baixa-alta), considerando as características de Pitch e MFCC (BUSSO; LEE; NARAYANAN, 2009; SOBOL-SHIKLER; ROBINSON, 2010; ERRO *et al.*, 2010).

De acordo com (RAMAKRISHNAN; EMARY, 2013), os resultados obtidos mostram que as características de frequência e MFCC são eficientes para distinguir entre as emoções de alta Valência, por exemplo: raiva, medo e alegria; e as emoções de baixa Valência, por exemplo: tristeza. Esses recursos também são efetivos para classificar as emoções de excitação semelhante, por exemplo a raiva vs. alegria. Segundo (RAMAKRISHNAN; EMARY, 2013), entre os classificadores, o SVM proporciona uma melhor taxa de reconhecimento sobre HMM.

Diante dos resultados, também são identificadas variações causadas pelas diferenças nos sistemas vocais de cada ser humano em termos de fatores como tamanho e forma. Isso faz com que existam variações de características emocionais nos discursos de indivíduos jovens e indivíduos de meia idade e, também entre masculino e feminino. Além disso, na interação diária entre homem e computador, as emoções são totalmente espontâneas com variações elevadas. Isso torna extremamente difícil e custosa a tarefa de identificá-las em discursos.

Esta pesquisa de doutorado, em contraste ao trabalho desenvolvido por (RAMAKRISHNAN; EMARY, 2013), não fez uso apenas de sensores que captam o discurso e o analisa, mas considera sensores de diferentes atributos (por exemplo, captura e análise da face, variabilidade da frequência cardíaca, etc...). Assim, acredita-se que é possível aumentar a acurácia no reconhecimento de emoções na

interação humano-computador. Nesse sentido, além de considerar os domínios de valência e excitação nesta pesquisa de doutorado, foram abordados mais dois domínios, a saber, sentimento de controle e facilidade de conclusão do objetivo.

- ***Emotion in human-computer interaction (PETER; URBAN, 2012)***

(PETER; URBAN, 2012) apresenta uma abordagem geral de reconhecimento de emoções utilizando sensores fisiológicos, utilizando tal abordagem eles desenvolveram o EREC (Sistema de Reconhecimento de Emoções) e mostram dois exemplos de aplicações afetivas utilizando o EREC.

A abordagem geral apresentada consiste em quatro etapas: Monitoramento, Pré-processamento dos dados, Extração de Características e Classificação. O Monitoramento coleta Reações Fisiológicas do usuário. O pré-processamento visa validar os sinais capturados, realizando suavização, normalização, sincronização e discretização dos mesmos. Em seguida, características são extraídas dos dados processados para realizar uma boa estimativa do estado emocional do usuário. A última etapa é a classificação, na qual diferentes abordagens podem ser utilizadas, por exemplo, árvores de decisão para processamento rápido ou SVM para resultados mais acurados (PETER; URBAN, 2012).

Ao utilizar tal abordagem, o sistema EREC é proposto. Esse sistema coleta dados de quatro tipo de sensores: atividade eletrodérmica, temperatura do ar atmosférico, variabilidade da frequência cardíaca. O pré-processamento de dados, a extração de características e a classificação utilizam-se do *framework* OmniRoute (MADER *et al.*, 2004).

(PETER; URBAN, 2012) trazem duas aplicações utilizando o sistema EREC: uma ferramenta de avaliação de usabilidade emocional e um ambiente de e-learning emocional. A primeira aplicação busca integrar a emoção reconhecida no processo de avaliação de usabilidade com o 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 negativas para o aprendizado provendo outras formas de interação com o conteúdo a ser aprendido.

Num cenário real, um único sensor não é capaz de prover, permanentemente, informação afetiva do usuário na vida cotidiana. Sensores visuais podem perder o foco no globo ocular; sensores fisiológicos podem não ser usados corretamente afetando a confiabilidade dos mesmos. Portanto, essa pesquisa visa utilizar uma abordagem de sensores multimodal tanto para aumentar a acurácia do reconhecimento da emoção como reduzir situações em que os sensores podem apresentar dados errôneos.

- ***The computer expression recognition toolbox (CERT) (LITTLEWORT *et al.*, 2011)***

Esse artigo apresenta o CERT (Ferramenta computacional para reconhecer expressões). O CERT captura automaticamente a intensidade de 19 ações faciais do FACS (Sistema de codificação único de ações faciais) e de 6 expressões faciais protótipas. Segundo (LITTLEWORT *et al.*, 2011), ele é capaz de reconhecer 10 características da face, assim como a orientação 3-D da cabeça. CERT alcançou em média com uma acurácia de 90.1% no banco de dados de expressões faciais posadas (*Extended Cohn-Kanade*), e 80% no banco de dados de expressões faciais espontâneas.

O pipeline do CERT apresenta cinco etapas. A primeira etapa é a detecção de face, o detector de face foi treinado utilizando uma extensão da abordagem de (VIOLA; JONES, 2004). A segunda etapa é a detecção de características faciais, um conjunto de 10 características faciais é detectado, a saber – o canto interno e externo dos olhos, a ponta do nariz, o canto interno e externo da boca, e o centro da boca, cada um utiliza um detector de característica específico (LITTLEWORT *et al.*, 2011). A terceira etapa é o registro da face que consiste em prever a localização da face baseado no conjunto de 10 características faciais detectadas. A quarta etapa é o processo de extração de características, de acordo com (LITTLEWORT *et al.*, 2011) um filtro é aplicado na face e o resultado é armazenado em um vetor de características.

A quinta etapa é o reconhecimento das ações unitárias da face (AU), o vetor de características da etapa anterior é utilizado como entrada de uma SVM para cada AU, a saída é a intensidade de cada AU. A última etapa consiste na dinâmica e intensidade das expressões faciais.

(LITTLEWORT *et al.*, 2011) apresenta três extensões do CERT: detector de sorriso, estimação da orientação 3-D da face e um reconhecedor básico de emoções. Além disso, dois exemplos de aplicações são apresentados: detecção automática de fadiga no motorista e sistema de aprendizado automático.

Esse trabalho mostra-se muito importante para esta pesquisa, porém o uso de uma abordagem de sensores multimodal, por exemplo, o uso de captura por vídeo em conjunto com o áudio e outros sensores, pode levar a um aumento da acurácia na avaliação da experiência emocional do usuário.

- ***Open source smartphone libraries for computational social science (LATHIA *et al.*, 2013)***

A pesquisa desenvolvida por (LATHIA *et al.*, 2013) teve por objetivo simplificar o desenvolvimento de aplicações que coletam dados de sensores por *smartphones*. Para isso, foi desenvolvido um conjunto de bibliotecas open-source que coleta, armazena, transfere e consulta dados de sensores de *smartphones*.

(LATHIA *et al.*, 2013) desenvolveu três bibliotecas: a primeira é responsável pela coleta dos dados dos sensores; a segunda gerencia como os dados coletados devem ser armazenados, consultados ou transmitidos; a terceira permite adicionar notificações baseado em eventos dos sensores ou de maneira programada pelo estudo do pesquisador.

Essas bibliotecas foram utilizadas para construir o aplicativo *Emotion Sense*. O *Emotion Sense* utiliza o método de amostragem de experiência em conjunto com a coleta de dados de sensores de *smartphones* e fornece feedback sobre como humor do usuário está relacionado ao dados dos sensores coletados pelo seu *smartphones*.

Além disso, essas bibliotecas foram oferecidas como uma ferramenta opcional para alunos de pós-graduação para uso em projeto final do curso de *Mobile Sensing*. As experiências reportadas pelos alunos que utilizaram essas bibliotecas foram, em geral, positivas e a quantidade de código necessária foi ligeiramente menor com o uso da biblioteca.

Em resumo, essas bibliotecas facilitam o desenvolvimento de aplicações que necessitam coletar dados dos sensores de *smartphones*, tornando o desenvolvimento de tais aplicações mais rápido e uniforme.

Diante disso, esta pesquisa de doutorado se beneficiaria com o uso de tais bibliotecas na etapa de concepção da UserSense para um contexto mobile.

- ***Short-term emotion assessment in a recall paradigm (CHANEL et al., 2009)***

(CHANEL et al., 2009) avalia em seu trabalho as emoções humanas de 10 usuários por meio do eletroencefalograma para a avaliação cognitiva; e atividade eletrodérmica, variabilidade da frequência cardíaca e taxa da frequência respiratória para a análise das reações fisiológicas. Fazer uso de medidas que analisam apenas dois componentes da emoção pode acabar comprometendo a visão holística que esta pesquisa de doutorado apresenta. Diferente desta pesquisa de doutorado, que apresenta a coleta em tempo de interação, (CHANEL et al., 2009) faz sua análise com base em episódios emocionais e os classifica a partir dos domínios da Valência e Excitação, sendo mais específico: excitação positiva, excitação negativa e neutro. Portanto, (CHANEL et al., 2009) desconsidera em sua pesquisa os domínios de Sentimento de controle e Facilidade de conclusão do objetivo propostos por (SCHERER, 2005).

Dois métodos de análise discriminante, a saber, a Análise Discriminante Linear (*Linear Discriminant Analysis - LDA*) e da Análise Discriminante Quadrática (*Quadratic Discriminant Analysis - QDA*) são usados nesse trabalho, ambos, baseados em inferência bayesiana. Para esse estudo (CHANEL et al., 2009), definiu-se três classes: “calmo”, “positivo” e “negativo”, e ao final fez a classificação (fusão) das fontes coletadas com o auxílio de uma Máquina de Vector de Relevância (*Relevance Vector Machine - RVM*), que diferentemente do SVM, fornece uma classificação probabilística.

Em estatística, os procedimentos que estão entre os mais antigos e tradicionais na literatura de classificação de padrões são a análise discriminante linear, a análise discriminante quadrática e a regressão logística. Esses métodos, cujos fundamentos estão estabelecidos na literatura há bastante tempo, têm baixo custo computacional. No entanto, outros métodos como redes neurais artificiais, árvores de classificação e regressão e análise discriminante de núcleo têm recebido bastante atenção dos pesquisadores em aprendizado de máquina e apresentado resultados satisfatórios (LACHENBRUCH; GOLDSTEIN, 1979; MARDIA; KENT; BIBBY, 1979; JOHNSON; WICHERN et al., 1992; ANDERSON-COOK, 2004).

- ***Comparative analysis of emotion estimation methods based on physiological measurements for real-time applications (KUKOLJA et al., 2014)***

Esse artigo apresenta um estudo comparativo baseado em sete métodos de aprendizado de máquina comumente utilizados para estimar emoções com base em reações fisiológicas. (KUKOLJA et al., 2014) realizou um experimento envolvendo 14 usuários que tiveram suas emoções estimuladas pelas imagens da base de dados do IAPS. (KUKOLJA et al., 2014) usou duas abordagens diferentes para a seleção referente estado emocional: (1) avaliações subjetivas - uma abordagem baseada em julgamentos de auto-relato pessoal sobre o estado emocional provocado e (2) as anotações de estímulos - uma abordagem utilizando o cumprimento das normas sociais aceitas.

Para a análise das características emocionais, (KUKOLJA et al., 2014) considerou os aspectos de reações fisiológicas a partir da atividade eletrodérmica (AED), eletroencefalograma (EEG), taxa da frequência respiratória e temperatura da pele. Os sentimentos considerados nesse estudo, realizado por (KUKOLJA et al., 2014), foram tristeza, felicidade, surpresa, tristeza, raiva, nojo e medo.

- ***A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies (MANDRYK; ATKINS, 2007)***

Em (MANDRYK; ATKINS, 2007), é apresentado um método baseado em Lógica Fuzzy para modelar emoções de maneira contínua usando dados fisiológicos. Este método é dividido em duas etapas, no qual: 1) os dados coletados são normalizados e calculados os valores de Excitação e Valência, e 2) posteriormente, essas informações foram utilizadas para definir a emoção do usuário em: entediado, desafiado, alta excitação, frustrado ou alegre. Para isso, foram realizados testes experimentais com vinte e quatro participantes do sexo masculino, com idades entre 18 e 27 anos e com experiência em tecnologia.

Os participantes deveriam disputar um período do game NHL 2003 da empresa EA Sports em um console modelo PS2 da fabricante Sony sob três condições (ambiente de disputa): 1) contra um amigo co-localizado, 2) contra uma pessoa desconhecida co-localizada, e 3) contra o próprio console. No mais, é importante ressaltar que, entre cada condição, o participante permanece por cinco minutos em um período de descanso. Segundo (MANDRYK; ATKINS, 2007), isso permite que as informações fisiológicas retornem aos níveis iniciais antes de cada condição de jogo.

Para a avaliação dos Sentimentos Subjetivos, os participantes também preencheram, durante os períodos de descanso, relatórios com auto-relatos das emoções sentidas e um relatório após o término das três condições. Além disso, foram registradas as Reações Fisiológicas e Expressões Motoras dos participantes para uma posterior avaliação objetiva. Todavia, o método proposto utilizará exclusivamente informações fisiológicas para classificar as emoções dos participantes. As informações consideradas para a análise nos experimentos são: atividade eletrodérmica, eletrocardiograma (ECG), eletromiografia do músculo zigomático maior, EMG do músculo corrugador do supercílio e variabilidade da frequência cardíaca.

Inicialmente, o modelo proposto transforma os dados fisiológicos em valores de Excitação e Valência. Para isso, recebe quatro entradas (AED, EMG, VFC) com seus valores normalizados (0-100) e retorna duas saídas (Excitação e Valência) em percentual. De acordo com (MANDRYK; ATKINS, 2007), essa etapa utiliza vinte e duas funções de pertinência triangulares (*triangular membership functions*) e foram distribuídas uniformemente em todo o espectro (LEFTERI; ROBERT, 1997).

Na segunda etapa do método proposto, as informações de Excitação e Valência (entradas com valores percentuais) são utilizadas para modelar as emoções de entediado, desafiado, alta excitação, frustrado e alegre. Um total de sessenta e sete funções de pertinência trapezoidais (*trapezoidal membership functions*) foram geradas para mapear os valores dos dados de entrada nas cinco emoções definidas como saída. Os resultados obtidos por (MANDRYK; ATKINS, 2007) demonstram que o modelo, baseado em Lógica Fuzzy proposto, apresenta um comportamento menos reacionário e mais moderado quando comparado a abordagem manual.

Porém, o modelo proposto obteve baixa acurácia para modelar a emoção de Excitação, cerca de 50%. Isso acontece porque assumiu-se que a excitação de um jogador aumentaria com o desafio, no entanto, alguns participantes declararam que utilizam uma estratégia (individual) de tentar relaxar quando são desafiados, com objetivo de melhorar seu desempenho. Em outro caso, utilizando as respostas do questionário pós-experimento, o modelo proposto chegou a obter uma acurácia de 17% quando questiona

sobre a condição mais desafiadora.

Como já citado, esta pesquisa de doutorado objetiva aumentar a acurácia do reconhecimento de emoções na interação humano-computador incorporando sensores de diferentes atributos, por meio da sua visão holística. Assim, além de considerar os domínios de Valência e Excitação foram abordados mais dois domínios, Sentimento de Controle e Facilidade de Conclusão do Objetivo.

- *Affect prediction from physiological measures via visual stimuli (ZHOU et al., 2011)*

(ZHOU et al., 2011) faz um estudo que tem como objetivo prever diferentes estados emocionais de usuários por meio de reações fisiológicas comparando três tipos de modelos computacionais. Um experimento foi desenvolvido para provocar respostas emocionais com fotos de emoções padronizadas quando algumas reações fisiológicas foram medidas. Três algoritmos de classificação – Regras de Decisão (*Decision Rules*), kNN e Árvore de Decomposição (*Decomposition Tree*) – foram aplicados com base na técnica de Conjuntos Ásperos (*Rough Set*), para a construção de modelos de previsão das características fisiológicas extraídas.

A técnica de Conjuntos Ásperos (ØHRN; ROWLAND, 2000) foi desenvolvida na Polônia no início de 1980 e preocupa-se com a análise classificatória de informações imprecisas, incertas ou incompletas expressas em termos de dados adquiridos com a experiência. Basicamente, essa teoria lida com a aproximação de conjuntos que são difíceis de descrever com a informação disponível.

A pesquisa desenvolvida por (ZHOU et al., 2011) traz apenas sensores que analisam as reações fisiológicas e não considera os outros componentes da emoção apontados por (ELLSWORTH; SCHERER, 2003; NIEDENTHAL; KRAUTH-GRUBER; RIC, 2006; SCHERER, 2005). Para esse estudo, (ZHOU et al., 2011) faz uso de um sistema de detecção fisiológica chamado E-Prime (Pittsburgh, EUA) para coletar a atividade eletrodérmica, taxa da frequência respiratória e dados de eletroencefalograma (EEG). Outra ferramenta utilizada foi o sistema ProComp Infinity (Montreal, Canadá) para coletar sinais de eletromiografia (EMG) facial.

Nessa pesquisa, (ZHOU et al., 2011) utiliza a fala do usuário para a definição das classes (saída), por exemplo, quando um indivíduo diz: “Ao olhar para esse imagem eu fico triste”. Para os dados de entrada utilizados nos atributos com a finalidade de treinar os métodos de classificação foram considerados os dados brutos capturados durante o experimento. Além disso, foram criados três tipos de modelos de previsão, a saber, específicos de gênero (masculino versus feminino), específicos para cultura (chineses versus indianos versus ocidentais), e os modelos gerais (participantes com diferentes gêneros e culturas como amostras), e as comparações diretas foram feitas entre esses modelos. Segundo (ZHOU et al., 2011), os resultados sugerem que os modelos específicos têm desempenho melhor do que os modelos gerais.

Diferentemente desta pesquisa de doutorado, que traz uma visão mais abrangente para a análise das emoções, (ZHOU et al., 2011) apresenta em sua pesquisa experimentos que fazem a captura apenas de reações fisiológicas, desconsiderando outros componentes da emoção como as expressões motoras que analisam a voz.

- ***A GMM based 2-stage architecture for multi-subject emotion recognition using physiological responses (YUAN et al., 2010)***

(YUAN et al., 2010) propõe a arquitetura de um sistema de reconhecimento de emoções, com o intuito de melhorar o desempenho ao lidar com o contexto multi-assunto (*multi-subject context*). (YUAN et al., 2010) faz uma análise de reações fisiológicas baseada em dois estágios do sistema de reconhecimento de emoções usando para a 1ª fase, de captura, o Modelo de Mistura Gaussiana (*Gaussian Mixture Model - GMM*) e para a 2ª fase de classificação, Busca Sequencial Flutuante para Frente (*Sequential Floating Forward Search - SFFS*) com kNN. Na primeira fase, os dados de cada assunto são treinados em modelos Modelos de Mistura Gaussiana separados. O sistema ProComp Infiniti (Montreal, Canadá) foi usado para a coleta dos dados.

(YUAN et al., 2010) faz a captura das reações fisiológicas por meio de Atividade eletrodérmica, eletrocardiograma (ECG), eletromiografia do músculo zigomático maior, EMG do músculo corrugador do supercílio, taxa da frequência respiratória, volume de sangue do pulso (VSP).

Vinte oito imagens do IAPS foram selecionadas para provocar emoções nos usuários. Inicialmente, os participantes foram convidados a classificar as imagens em termos de seu nível de Excitação e Valência em uma escala de 1-9 e verbalmente falar uma única palavra de emoção que melhor descrevesse seus sentimentos depois de ver a imagem.

(YUAN et al., 2010) faz uso de um método simples de busca para localizar os pontos de pico dos batimentos cardíacos. Considerando os dois estágios para o reconhecimento de emoções propostos, (YUAN et al., 2010) adota o conjunto de recursos estatísticos de domínio de tempo propostos por (PICARD; VYZAS; HEALEY, 2001), porque, de acordo com (YUAN et al., 2010), essas características têm aparecido em vários estudos anteriores e mostrado eficiência na classificação em estados emocionais (GU et al., 2008; GU et al., 2009; HAAG et al., 2004; MERA; ICHIMURA, 2004). Diante do exposto, seis características foram extraídas de cada reação fisiológica capturada.

A pesquisa desenvolvida por (YUAN et al., 2010) faz a captura predominante de reações fisiológicas desconsiderando outros componentes da emoção que foram abordados nesta pesquisa de doutorado, como por exemplo, tendências comportamentais. Além disso, para o contexto desta Tese fez-se a captura e a classificação em tempo de interação do usuário com um sistema computacional.

- ***EMG and GSR signals for evaluating user's perception of different types of ceramic flooring (LAPARRA-HERNÁNDEZ et al., 2009)***

(LAPARRA-HERNÁNDEZ et al., 2009) apresenta os resultados de um projeto de pesquisa sobre a avaliação de três tipos de reações fisiológicas: eletromiografia do músculo zigomático maior, EMG do músculo corrugador do supercílio e atividade eletrodérmica. A captura foi feita com o auxílio do Varioport System (Becker MEDITEC, Karlsruhe, Alemanha).

O músculo zigomático maior é um músculo que se estende do osso zigomático até o ângulo da boca (MERA; ICHIMURA, 2004). Na contração, puxa o ângulo da boca súpero-lateralmente, dando, por exemplo, a expressão facial de sorriso.

O supercílio corrugador é um pequeno, estreito, músculo piramidal próximo ao olho. Ele está localizado na extremidade medial da sobrancelha, sob as frontalis e um pouco acima do músculo orbicular dos olhos (MERA; ICHIMURA, 2004). Segundo (LERA; GARRETA-DOMINGO, 2007) esse músculo leva o usuário a franzir a testa. Sobrancelhas franzidas podem ser um sinal de uma necessidade de se concentrar, descontentamento ou de percepção de falta de clareza.

As reações fisiológicas foram registradas enquanto três imagens de calibração (face neutra, rosto sorridente e rosto carrancudo) e oito imagens de pisos cerâmicos foram sendo projetados. As imagens foram retiradas de uma base de dados (MARTINEZ; BENAVENTE, 1998). Oito pessoas, quatro mulheres e quatro homens, com idades entre 24 e 30 anos, participaram do estudo.

(LAPARRA-HERNÁNDEZ *et al.*, 2009) faz uso de dois conjuntos de imagens: o primeiro foi usado para calibrar a resposta dos usuários e consistia na fotografia de uma mulher jovem em três momentos - com a face neutra, Valência positiva (rosto sorridente) e Valência negativa (franzindo a testa). O segundo conjunto de imagens foi escolhido para avaliar a resposta dos usuários a diferentes tipos de piso cerâmico. A cada tipo havia sido incorporada uma imagem que mostrava uma sala com uma mulher idosa. O único elemento de mudança em todas as imagens eram os tipos de pisos.

Quando as reações fisiológicas foram processadas, cada tipo de imagem (calibragem e piso de cerâmica) foi analisada individualmente. Os valores de eletromiografia e atividade eletrodérmica foram normalizados para permitir a comparação dos valores de cada tipo de piso para cada assunto.

A pesquisa de doutorado apresentada aqui, também traz uma abordagem de captura em tempo real, no entanto percebe-se que o estudo apresentado por (LAPARRA-HERNÁNDEZ *et al.*, 2009) está relacionado apenas a análise de reações fisiológicas durante a observação de imagens e não com a interação entre usuários e um sistema computacional. Além disso, (LAPARRA-HERNÁNDEZ *et al.*, 2009) considera em sua análise apenas o domínio da Valência, fato que pode acabar comprometendo a visão holística que poderia ter sido dada ao seu estudo.

- ***On the classification of emotional biosignals evoked while viewing affective pictures: an integrated data-mining-based approach for healthcare applications (FRANTZIDIS *et al.*, 2010)***

(FRANTZIDIS *et al.*, 2010) propõe um arquitetura para a análise de reações fisiológicas emocionais provocadas a partir da visualização de uma coleção de imagens padronizadas do IAPS. Esse projeto visa o monitoramento da atividade eletrodérmica (EDA) e o eletroencefalograma de pacientes idosos e/ou doentes crônicos. O IAPS é um procedimento para indução de emoções desenvolvido por (LANG; BRADLEY; CUTHBERT, 2008), que foi classificado por um grande número de participantes em termos de Valência e Excitação.

Para esse estudo (FRANTZIDIS *et al.*, 2010) faz a classificação das emoções como base nos domínios da Valência e Excitação, por meio de uma abordagem de classificação de dados, a saber, o algoritmo de Árvore de Decisão (*Decision Tree*).

A valência e o gênero dos participantes serviram como atributos de entrada para um classificador de distância de Mahalanobis (*Mahalanobis distance classifier*), que classifica os dados para valores altos e baixos de Valência. A análise foi feita a partir de episódios emocionais e as características diferentes

provenientes da análise foram utilizadas pelo classificador de distância de Mahalanobis com o intuito de detectar com mais precisão a dimensão da excitação.

A distância de Mahalanobis é baseada nas correlações entre variáveis com as quais distintos padrões podem ser identificados e analisados. Esse método é útil para determinar a similaridade entre uma amostra desconhecida e uma conhecida (CRISTIANINI; SHAWE-TAYLOR, 2000).

Segundo (FRANTZIDIS *et al.*, 2010), algoritmo de árvore de decisão foi adotada, uma vez que é uma técnica de aprendizado eficiente para a representação da regra de classificação. Além disso, o algoritmo encontra características mais robustas para a divisão inicial dos conjuntos de dados reais.

A precisão da árvore de decisão foi medida por meio de uma validação cruzada (*cross validation*). Isto implica que os dados de treinamento foram divididos em 11 subgrupos. Um deles foi considerado como o conjunto de teste e os restantes serviram para o processo de formação. Os resultados foram descritos na linguagem *Extensible Markup Language* (XML).

Diferentemente desta pesquisa de doutorado, (FRANTZIDIS *et al.*, 2010) tem como resultado de sua classificação apenas valores baixos e altos para valência e excitação não considerando outros domínios da emoção. Ao seu estudo, além da ineficiência do trato desses domínios, também acrescenta-se a ausência da análise feita a partir de outras fontes de captura, como por exemplo, o reconhecimento de emoções faciais e variabilidade da frequência cardíaca, com o objetivo de dar uma abrangência maior a sua análise.

- ***Toward emotion recognition in car-racing drivers: a biosignal processing approach* (KATSIS *et al.*, 2008)**

(KATSIS *et al.*, 2008) apresenta em seu trabalho uma metodologia e um sistema para apoiar a avaliação do estado emocional de motoristas de carro de corrida, no qual um teste de viabilidade do sistema foi feito a partir de dados obtidos de dez indivíduos em um ambiente de corrida simulado. A abordagem proposta realiza a avaliação dos estados emocionais utilizando eletromiografia facial, eletrocardiograma, taxa da frequência respiratória e atividade eletrodérmica. O sistema consiste das seguintes fases: 1) Módulo de multisensorial, responsável pela captura, pré-processamento e de transmissão sem fio dos dados. 2) Módulo de computação centralizada que extrai as características especiais dos dados capturados e apresenta o estado emocional do sujeito, com base em um conjunto de dados que contém as características extraídas, juntamente com a anotação dos médicos especialistas. 3) A interface do sistema.

A pesquisa de doutorado apresentada nesta monografia procura ter uma visão holística, que considera diferentes componentes emocionais de tempo real (avaliação cognitiva, reação fisiológica, expressão motora e tendência comportamental) para analisar as emoções e as classifica segundos os domínios da Valência, Excitação, Sentimento de controle e Facilidade de conclusão do objetivo (SCHERER, 2005).

Diante desse contexto e considerando a Estrutura do Espaço Emocional Semântico de (SCHERER, 2005), (KATSIS *et al.*, 2008) traz uma proposta com apenas 4 classes emocionais, a saber, alta tensão (presente no oitante 1), baixa tensão (presente no oitante 5), decepção (presente no oitante 3) e euforia (presente no oitante 6). Considerar apenas reações fisiológicas e quatro tipos de emoções pode reduzir a abrangência que uma avaliação de repostas emocionais requer.

No contexto dessa pesquisa, (KATSIS *et al.*, 2008) faz uso da SVM e de um Sistema Adaptativo de Inferência Neuro-Fuzzy (*Adaptive Neuro-Fuzzy Inference System* - ANFIS) para a classificação. Uma tarefa de classificação baseada em SVM envolve treinamento e testes, que consistem em uma série de instâncias de dados. Cada instância no conjunto de treinamento contém uma etiqueta de classe e vários atributos. O objetivo da SVM é produzir um modelo que prevê a classe de instâncias de dados no conjunto de testes em que apenas os atributos são dados (DUDA; HART *et al.*, 1973).

Técnicas de Neuro-Fuzzy têm surgido a partir da fusão de Redes Neurais Artificiais e Sistemas de Inferência Fuzzy e formam uma estrutura popular para resolver problemas do mundo real. Um Sistema Neuro-Fuzzy é baseada em um Sistema Fuzzy, treinado por um algoritmo de classificação derivado da teoria de Redes Neurais (DUDA; HART *et al.*, 1973; CLARK; NIBLETT, 1989). ANFIS (JANG; SUN; MIZUTANI, 1997) é uma arquitetura que é praticamente uma evolução de um Sistema Fuzzy Sugeno (*Sugenotype Fuzzy System*) (SUGENO, 1985).

As taxas globais de classificação obtidas com a utilização de validação cruzada dez vezes são 79,3% e 76,7% para o SVM e o ANFIS, respectivamente. Na validação cruzada k vezes estratificada, as dobras são criadas de um modo que contenham, aproximadamente, a mesma proporção de rótulos de previsão como o conjunto de dados original. Estudos mostraram que a validação cruzada estratificada tende a gerar resultados com menor viés e variância quando comparada com a validação cruzada regular k vezes (KOHAVI *et al.*, 1995).

Nesse trabalho, para estimar o desempenho dos classificadores, uma abordagem de dez vezes a validação cruzada estratificada é usada. Estudos empíricos mostram que dez parece ser um número ideal de dobras que otimiza o tempo que leva para completar o teste, minimizando o viés e a variância associada ao processo de validação (BREIMAN *et al.*, 1984).

- ***Identifying emotion by keystroke dynamics and text pattern analysis* (NAHIN *et al.*, 2014)**

A plataforma desenvolvida por (NAHIN *et al.*, 2014) detecta as emoções por meio de uma análise dos padrões de digitação dos usuários e o tipo de textos (palavras e frases) digitados por eles. A ferramenta Weka foi usada por (NAHIN *et al.*, 2014) para classificar o conjunto de dados coletados com o auxílio de um algoritmo de aprendizado de máquina supervisionado.

Para a análise dos padrões de digitação dos usuários foram verificados os textos fixos (definidos pelos autores) e livres (não definidos pelos autores). Nesse estudo, o autor constatou que textos fixos apresentaram melhores resultados na indicação de emoções específicas.

Depois de analisar os padrões de digitação, (NAHIN *et al.*, 2014) analisou padrões de textos digitados. Palavras consideradas não-emocionais foram identificadas e removidas da avaliação. O Modelo Vetorial em Sistemas de Recuperação da Informação (*Vector Space Model*) foi utilizado para classificar as frases. Os resultados do trabalho mostram respostas interessantes para os sete tipos de emoções.

(NAHIN *et al.*, 2014) realiza apenas a análise das tendências comportamentais, assim, ele pode ter uma visão parcial do estado emocional dos usuários. Outros aspectos, tais como a fala e as expressões faciais, consideradas na visão abrangente defendida nesta Tese, ficaram de fora do estudo de (NAHIN *et*

al., 2014). Acredita-se que o uso de vários sensores possa ajudar a trazer uma maior acurácia e abrangência na avaliação das emoções.

Além disso, (NAHIN *et al.*, 2014) traz, em seu trabalho, apenas a análise de sete tipo emoções: raiva, nojo, culpa, medo, alegria, tristeza e vergonha. Fato que pode acabar comprometendo a visão abrangente que se busca nesta Tese com o auxílio da Estrutura do Espaço Emocional Semântico.

- ***Real-time classification of evoked emotions using facial feature tracking and physiological responses (BAILENSON et al., 2008)***

(BAILENSON *et al.*, 2008) propõe e apresenta um modelo baseado em algoritmos de classificação – Regressão Linear (*Linear Regression*) e Rede Neural Artificial (*Artificial Neural Network*) – para análise em tempo real de emoções do usuário. A análise foi feita a partir de vídeos de 41 indivíduos assistindo filmes emocionalmente evocativos. Um dos sensores realizava o reconhecimento de emoções faciais (diversão, neutro e tristeza) e outro a análise da variabilidade da frequência cardíaca (VFC) e da atividade eletrodérmica (AED). Cinco especialistas fizeram a análise dos vídeos, *frame por frame*, com o objetivo de definir as classes (saídas) para os algoritmos de classificação, por exemplo, analisando que o usuário estava triste.

(BAILENSON *et al.*, 2008) apresenta algoritmos baseados em pontos extraídos da face dos sujeitos, bem como suas reações fisiológicas. Além disso, essa pesquisa trabalha apenas com o eixo da Valência, mais especificamente, com as emoções: diversão, neutro e tristeza. Para esta pesquisa de doutorado, trabalhou-se com os quatro domínios da Estrutura do Espaço Emocional Semântico de (SCHERER, 2005). O objetivo desta Tese é fazer a captura em tempo real de usuários interagindo com sistemas computacionais, considerando outros componentes da emoção e não apenas expressões motoras e reações fisiológicas.

- ***MoodScope: Building a mood sensor from smartphone usage patterns (LIKAMWA et al., 2013)***

Esse artigo apresenta uma abordagem que infere o humor do usuário utilizando o padrão de uso do *smartphones*, chamado MoodScope. (LIKAMWA *et al.*, 2013) realizou um estudo com 32 usuários durante dois meses, atingindo uma acurácia de 66% no início com um modelo estatístico genérico, e 93% após os 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 ele o consulte em outras situações, por exemplo, local e tempo, pode melhorar a consciência do usuário sobre suas mudanças de humor. Outra aplicação seria o compartilhamento automático de humor para seus amigos ou a identificação do humor durante uma conversa no telefone. Além disso, aplicativos sensíveis ao humor do usuário podem prover serviços personalizados para ele.

O MoodScope diferencia-se de outras aplicações de detecção de humor por não utilizar dados de sensores físicos do *smartphones*, nem de áudio e vídeo. (LIKAMWA *et al.*, 2013) utiliza 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

Tabela 1 – Visão geral das principais pesquisas publicadas sobre a estimativa de emoções.

Trabalhos	Características							
	Tendên. compor.	Avalia. cognit.	Express. motor.	Reaç. fisiol.	Valênc.	Excitaç.	Sentim. control.	Facil. conclu. objet.
(RAMAKRISHNAN; EMARY, 2013)	Não	Sim	Não	Não	Sim	Sim	Não	Não
(PETER; URBAN, 2012)	Não	Não	Não	Sim	Não	Não	Não	Não
(LITTLEWORT <i>et al.</i> , 2011)	Não	Não	Sim	Não	Não	Não	Não	Não
(LATHIA <i>et al.</i> , 2013)	Sim	Não	Não	Não	Não	Não	Não	Não
(CHANEL <i>et al.</i> , 2009)	Não	Sim	Não	Sim	Sim	Sim	Não	Não
(KUKOLJA <i>et al.</i> , 2014)	Não	Sim	Não	Sim	Sim	Sim	Não	Não
(MANDRYK; ATKINS, 2007)	Não	Não	Não	Sim	Sim	Sim	Não	Não
(ZHOU <i>et al.</i> , 2011)	Não	Sim	Não	Sim	Não	Não	Não	Não
(YUAN <i>et al.</i> , 2010)	Não	Não	Não	Sim	Sim	Sim	Não	Não
(LAPARRA-HERNÁNDEZ <i>et al.</i> , 2009)	Não	Não	Não	Sim	Sim	Não	Não	Não
(FRANTZIDIS <i>et al.</i> , 2010)	Não	Sim	Não	Sim	Sim	Sim	Não	Não
(KATSIS <i>et al.</i> , 2008)	Não	Não	Não	Sim	Não	Não	Não	Não
(NAHIN <i>et al.</i> , 2014)	Sim	Não	Não	Sim	Sim	Sim	Não	Não
(BAILENSEN <i>et al.</i> , 2008)	Não	Não	Sim	Sim	Não	Não	Não	Não
(LIKAMWA <i>et al.</i> , 2013)	Sim	Não	Não	Não	Não	Não	Não	Não
Esta Tese	Sim	Sim	Sim	Sim	Sim	Sim	Sim	Sim

que coleta os dados em plano de fundo; e a segunda que é utilizada para o usuário reportar o seu humor. O treinamento de modelos personalizados para cada usuário é realizado na nuvem para não consumir recursos do *smartphones* dos usuários.

A pesquisa desenvolvida por (LIKAMWA *et al.*, 2013) tem uma abordagem distinta desta pesquisa de doutorado, pois não utiliza dados de sensores físicos, nem imagem nem áudio. Outro fator interessante é que não há uma captura de dados em tempo real, haja vista que o usuário deve reportar seu humor quatro vezes ao dia, o que pode desestimular o usuário a usá-lo. Uma possível alternativa para a pesquisa de (LIKAMWA *et al.*, 2013) seria uma abordagem multimodal utilizando dados de sensores físicos e/ou lógicos para classificar o humor do usuário em situações em que o reconhecer de humor que usa somente o padrão de uso do *smartphones* não tiver certeza do humor do usuário.

3.3 Discussão

A Tabela 1 fornece um resumo detalhado dos estudos de investigação na literatura, predominantemente preocupada com a captura das respostas emocionais relacionados às reações fisiológicas, um fato que pode comprometer a riqueza de detalhes que uma análise emocional deve ter. O autor deste estudo busca uma visão holística da experiência emocional do usuário com o objetivo de proporcionar uma visão mais abrangente e maior grau de precisão para a captura, análise e classificação das emoções.

Com base na pesquisa bibliográfica, pode ser visto que não existe um consenso de opinião no que refere-se ao que é responsável pela definição das classes a serem usadas no treinamento dos algoritmos de

classificação. Alguns estudos, como (BAILENSEN *et al.*, 2008; KATSIS *et al.*, 2008; RAMAKRISHNAN; EMARY, 2013) apontam que podem ser os especialistas (por exemplo, os psicólogos), enquanto outros pesquisadores (ZHOU *et al.*, 2011; CHANEL *et al.*, 2009; YUAN *et al.*, 2010; LIKAMWA *et al.*, 2013) argumentam que pode ser o próprio usuário a realizar esse definição. Diante do exposto, a presente pesquisa se propõe a avaliar a experiência emocional do usuário por meio de experimentos exploratórios e através de uma abordagem analítica multimodal de sensores. Além disso, com base nos dados coletados de usuários interagindo com sistemas computacionais e análise de especialistas (psicólogos), duas abordagens são adotadas, a saber: (i) Estrutura do Espaço Emocional Semântico e (ii) Aprendizado de Máquina.

Um grande número de estudos na literatura sobre classificação de emoções estão preocupados apenas com os domínios Valência e Excitação, fato que pode prejudicar a precisão de uma avaliação da experiência emocional do indivíduo. Avaliar a experiência emocional em termos de Valência e Excitação em um sistema computacional é um fator-chave nas decisões de design, uma vez que é possível identificar, respectivamente, o grau de satisfação e motivação do usuário durante a interação. De qualquer modo, também devem ser levados em conta os outros domínios da experiência emocional, como "Sentimento de controle" e "Facilidade de conclusão do objetivo", enquanto as emoções estão sendo avaliadas. Isso abre a perspectiva de uma associação de emoções que envolve os quatro domínios da experiência emocional e que podem permitir que os aplicativos computacionais se flexibilizem na adoção desta abordagem.

3.4 Considerações finais

Na visão holística defendida nesta Tese, o organismo humano, considerado em sua totalidade, determina como as diferentes partes comportam-se e a maneira como elas estão interligadas. Como argumentou (PETER; URBAN, 2012; KUKOLJA *et al.*, 2014; LI *et al.*, 2015; GONÇALVES *et al.*, 2016; WRZUS; MEHL, 2015; WANG *et al.*, 2014), um único sensor em um cenário do mundo real não é capaz de fornecer informações afetivas sobre o usuário em uma base diária. Sensores visuais podem perder o foco no globo ocular, enquanto sensores fisiológicos podem não ser usados corretamente afetando a confiabilidade dos mesmos.

Em vista disso, existe claramente uma necessidade de se aproximar do público alvo, de modo que se possa ter uma melhor compreensão de como a experiência emocional dos usuários pode ser classificada com base em uma abordagem multimodal. No próximo Capítulo, serão apresentados detalhes sobre a abordagem para captura, análise e classificação automatizadas da experiência emocional de usuários em tempo de interação, proposta nesta Tese.

ABORDAGEM PROPOSTA: USERSENSE

A pesquisa desenvolvida neste doutorado propõe e apresenta uma abordagem chamada UserSense, baseada no *Middleware* Adaptativo OpenCom que utiliza múltiplos sensores para avaliar as emoções do usuário em tempo de interação. A estrutura proposta é ilustrada na Figura 5. Em primeiro lugar, os sensores são usados para medir emoções em tempo real. Em seguida, um modelo de classificação de aprendizado de máquina é usado para combinar as saídas dos sensores, e fornecer o estado emocional do usuário nesse momento como resultado final.

Como este trabalho está interessado em prever o estado emocional do usuário em um determinado momento, pode-se dizer que cada instância do conjunto de dados corresponde a medição das emoções de um usuário em um determinado momento. A terminologia “instância” será usada em todo o texto. Nesse problema de classificação, não é feita distinção entre indivíduos e não se considera dependência entre momentos diferentes do mesmo usuário. Os instantes de todos os usuários são colocados no mesmo conjunto de dados, ou seja, as instâncias são tratadas independentes de usuário e tempo.

O rosto do usuário é analisado por um sistema que foi implementado utilizando os principais recursos advindos do FaceTracker (SARAGIH; LUCEY; COHN, 2011). A captura dessas expressões faciais é realizada com o auxílio de uma câmera durante a interação com sistemas computacionais. Mais detalhes são explicados na Seção 4.2.1.

O sensor lógico (*logical sensor*) mostrado na Figura 5 considera respostas emocionais do que o usuário está fazendo para delinear suas emoções. Por exemplo, ao jogar um jogo eletrônico, a Tendência Comportamental verifica se o utilizador está ganhando mais do que perdendo em um dado intervalo de tempo. Através desses dados, a UserSense pode "estimar" se o usuário está feliz ou frustrado. Mais detalhes são explicados na Seção 4.2.2.

O discurso do usuário foi capturado por um microfone e um sensor que inspeciona cada palavra emitida por ele. Esse sensor conta com a ajuda do *software* da Google para identificar cada palavra pronunciada. Este processo está detalhado na Seção 4.2.3. Além disso, o método Think Aloud foi utilizado para apoiar a Avaliação Cognitiva do usuário.

Ainda no contexto dessa pesquisa, os valores de frequência cardíaca são lidos por um *smartwatch* da Motorola - Moto 360. Ele pode transmitir o valor de frequência cardíaca para um *smartphone* ou para

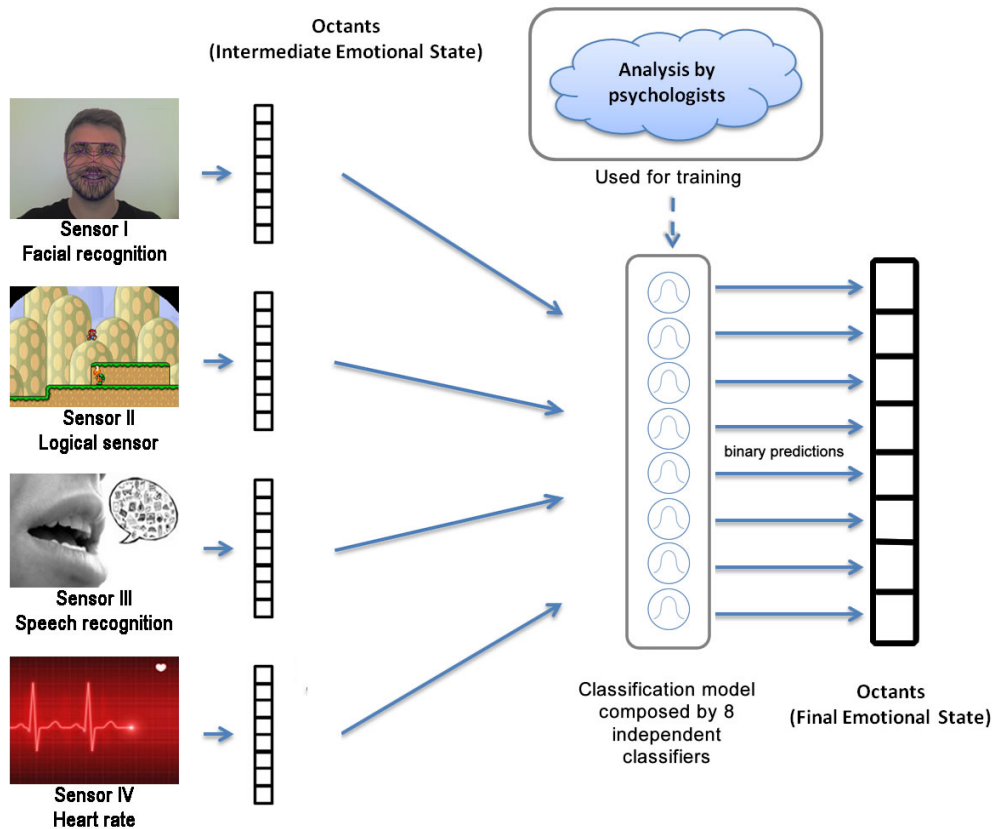


Figura 5 – O processo de estimativa das emoções do usuário em tempo de interação tem um mapeamento entre dados de sensoriamento e oitantes padrões (oitante padrão é construído para corresponder com a Figura 3). Modelo de classificação composto por 8 classificadores independentes são utilizados para obter valores emocionais intermediários para cada sensor e fornecer estados emocionais finais.

um PC usando tecnologia sem fio, como *Bluetooth* ou *Wi-Fi*. Mecanismo detalhado na Seção 4.2.4.

Como está sendo realizado o aprendizado supervisionado, em que o método de aprendizado máquina usa instâncias rotuladas para aprender modelos, as instâncias dos dados precisam ser rotuladas. Nesse contexto, um rótulo será o estado emocional do usuário naquele momento; mais especificamente, um rótulo irá corresponder a um ou mais oitantes que correspondem ao estado emocional do usuário. Os rótulos das instâncias são utilizadas pelo método de aprendizado durante a fase de treinamento.

Na classificação, o rótulo das instâncias de treinamento é geralmente definido por um especialista do domínio. Por exemplo, o julgamento de um biólogo seria confiável se tivéssemos a tarefa de rotular casos biológicos, tais como proteínas. No caso de classificação de emoções, usa-se a análise de um ou mais psicólogos - se há mais de um psicólogo envolvido na análise, uma estratégia utilizada para definir o rótulo é usar a(s) emoção(ões) mais frequentemente escolhidas.

4.1 Visão Componentizada da UserSense

Esta pesquisa discute o uso de uma abordagem genérica e adaptativa para diferentes plataformas como PC's, dispositivos móveis e integrados. Nesse sentido, (UEYAMA *et al.*, 2009) afirmam que a reconfiguração em tempo de execução é uma característica fundamental para lidar com a heterogeneidade de *hardware* inerente as diversas plataformas.

Nesse contexto, considerando inúmeras combinações de sensores em diferentes dispositivos, que podem sensorar um usuário para uma posterior inferência emocional, foi desenvolvido uma camada de *software* denominada UserSense baseada no modelo de componentes do OpenCom (UEYAMA *et al.*, 2009).

A abordagem UserSense apresenta características garantidas pelo uso do *middleware* flexível OpenCom como a generalidade, a arquitetura flexível, extensível e independente de linguagem. Tal arquitetura baseia-se em um microkernel, em que as funcionalidades são tanto incrementadas quanto solicitadas e descartadas assim que deixam de ser requisitadas. Assim, um componente é criado para cada sensor a ser utilizado, incluídos assim que acionados. Da mesma maneira, cria-se um componente para a fusão do processamento de dados coletados pelos sensores.

Os componentes de processamento presentes nessa abordagem podem ser compartilhados entre dois outros componentes que representam um mesmo tipo de aspecto emocional, como as tendências comportamentais. Nesse contexto, um usuário monitorado por sensores lógicos, por exemplo, pode deixar de interagir com um jogo eletrônico e começar a interagir com o *instagram*. Nessa situação, ambos os componentes, que representam os sensores lógicos de cada aplicação, seriam processados usando o componente “*processa tend. comp.*¹” que armazena os dados em um repositório.

Tais componentes de processamento também podem ser compartilhados com componentes que representam dados de diferentes dispositivos. Conforme foi utilizado no experimento exploratório com jogos eletrônicos, os dados são enviados a UserSense quando essa é alocada em um servidor.

4.1.1 UserSense no Contexto de Jogos Eletrônicos

A UserSense é uma abordagem flexível que apoia a captura, análise e classificação de repostas emocionais em tempo real. Essa característica flexível permite a utilização de um número variável de sensores conforme a disponibilidade e a relevância do uso de cada um deles em um determinado contexto. Por exemplo, um sensor pode estar indisponível por conta de pouca bateria em um dispositivo móvel, ou o processamento de imagens de uma câmera em um ambiente escuro pode ser pouco relevante.

No ambiente utilizado para o experimento exploratório com usuários, durante uma interação com um jogo eletrônico, foram introduzidos 4 tipos de sensores para o monitoramento: sensor de reconhecimento de emoções faciais, sensor de frequência cardíaca, sensor de reconhecimento de fala e sensores lógicos.

Tanto o sensor de face quanto o de fala e o lógico foram obtidos através de um *laptop*, já o sensor de frequência cardíaca através de um *smartwatch* que enviava os dados para um celular como descrito no Capítulo 4. A UserSense foi alocada em um servidor onde recebia os dados dos sensores utilizados nos experimentos exploratórios.

A Figura 5 ilustra a arquitetura de componentes da UserSense. Os componentes representados pela cor verde (parte superior da Figura), por exemplo, *sensor face* e *sensor fala* referem-se a etapa de captura da abordagem proposta, por meio de sensores. Os componentes representados pela cor salmão (parte inferior da Figura), por exemplo *processa expres. mot.* e *processa aval. cogn.* fazem referência a

¹ *processa tend. comp.* é a abreviação de processamento de tendências comportamentais

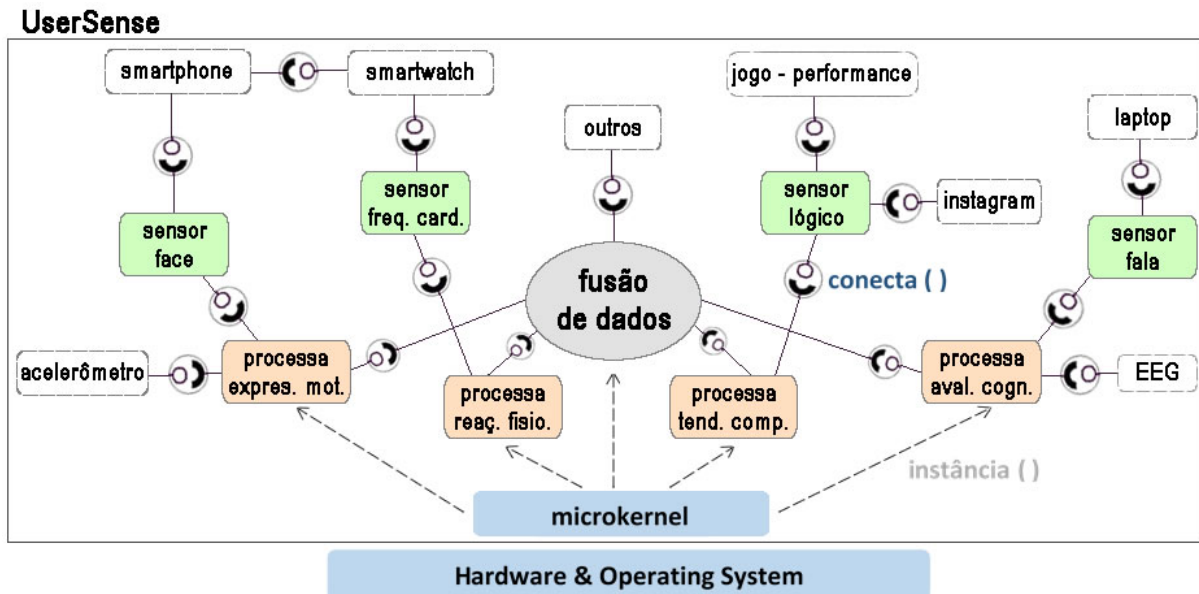


Figura 6 – Cenário de interação de componentes UserSense.

etapa de análise da UserSense, considerando a Estrutura do Espaço Emocional Semântico de Scherer para o mapeamento dos estados emocionais. Por fim, a etapa de classificação das emoções é representada pelo componente *fusão de dados* na cor cinza (no centro da Figura).

O componente *fusão de dados*, criado para a classificação dos oito oitantes da Estrutura do Espaço Emocional Semântico, é um modelo de classificação que utiliza métodos de aprendizado de máquina para estimar o estado emocional dos usuários, tendo como entrada os dados que vêm dos sensores e como saída a avaliação dos psicólogos.

No mais, alguns componentes adicionais representados pelas caixas brancas e linhas tracejadas, como por exemplo o acelerômetro, foram incluídos para demonstrar a adaptabilidade da UserSense a diferentes contextos e plataformas. Componentes adicionais são componentes incluídos apenas quando necessário ou que não foram implementados para esse contexto, mas é interessante incluí-los na Figura para exemplificar que um receptáculo pode alocar diferentes componentes.

4.2 Sensores para Detecção de Emoções

No contexto desta pesquisa, sensores físicos são pequenos dispositivos de baixo custo capazes de perceber alguns fenômenos em seus arredores, captar dados e transmiti-los para um *data sink* por meio da comunicação com outros dispositivos, usando um cabo ou meio sem fio (por exemplo, em uma rede de sensores sem fio (FILHO *et al.*, 2014)). Um medidor de frequência cardíaca, por exemplo, é um sensor físico de interesse neste projeto. A mesma ideia pode ser aplicada a pequenos sistemas de *software* capazes de perceber fenômenos, captar dados e transmiti-los, por exemplo, a taxa de erros do usuário durante a execução de uma tarefa. Nesse caso, eles são chamados sensores lógicos (ROSALES *et al.*, 2011). A seguir, cada sensor individual, utilizado neste estudo, será explicado.

4.2.1 Sensor 1 (S_I): Reconhecimento de Emoções Faciais para Expressões Motoras

As Expressões Motoras estão relacionados às expressões faciais, gestos físicos e aspectos de conteúdo lexical, tais como: velocidade, intensidade, melodia e som. De acordo com Desmet (DESMET, 2004), cada emoção está ligada a um padrão fixo de Expressão Motora. Para ilustrá-lo, autor apresenta o 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”.

O sistema de Reconhecimento de Emoções Faciais, usado nesse estudo e desenvolvido por Santos et al. (2013), foi concebido em dois módulos distintos: Extração Automática de Características Faciais (EACF) e Reconhecimento de Expressões Faciais (REF). As imagens recebidas a partir da câmera são analisadas pelo módulo EACF que extrai as características faciais. O EACF envia esses dados para o módulo de REF que, por sua vez, determina as características expressivas. A Figura 7 mostra o processo de reconhecimento.

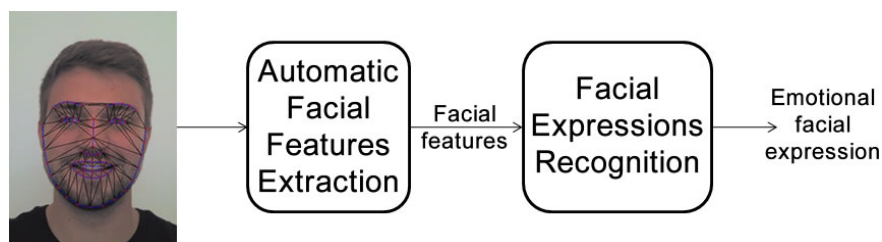


Figura 7 – Reconhecimento de Emoções Faciais.

O módulo EACF processa a imagem, procura por um rosto humano e o analisa de modo a produzir uma série de características que representem respostas emocionais dos usuários. A matriz é baseada em pontos e texturas que descrevem as características da expressão facial em uma imagem. Para isso, foi utilizado o sistema idealizado por (SARAGIH; LUCEY; COHN, 2011).

O módulo de REF é responsável por receber o conjunto de características produzidas e efetuar um mapeamento adequado entre as características obtidas a partir da face humana e o estado emocional correspondente. Esse mapeamento é realizado por uma Rede Neural Artificial (RNA), que permite que uma emoção humana seja reconhecida.

Essa RNA é composta por uma camada de entrada, uma camada oculta e uma camada de saída. As entradas da RNA são os pontos extraídos pelo módulo EACF. A saída é a emoção correspondente expressa pelo usuário, que pode ser felicidade, raiva, tristeza, surpresa e calma.

Nessa tese, adaptou-se o sistema Reconhecimento de Emoções Faciais para obter os oitantes a partir desse sensor. Assim, propôs-se o seguinte procedimento de mapeamento para a Estrutura do Espaço Emocional Semântico mostrada na Figura 3: Felicidade (oitantes 6,7), Surpresa (oitante 8), Calma (oitante 5), Raiva (oitantes 1,2) e Tristeza (oitantes 3,4).

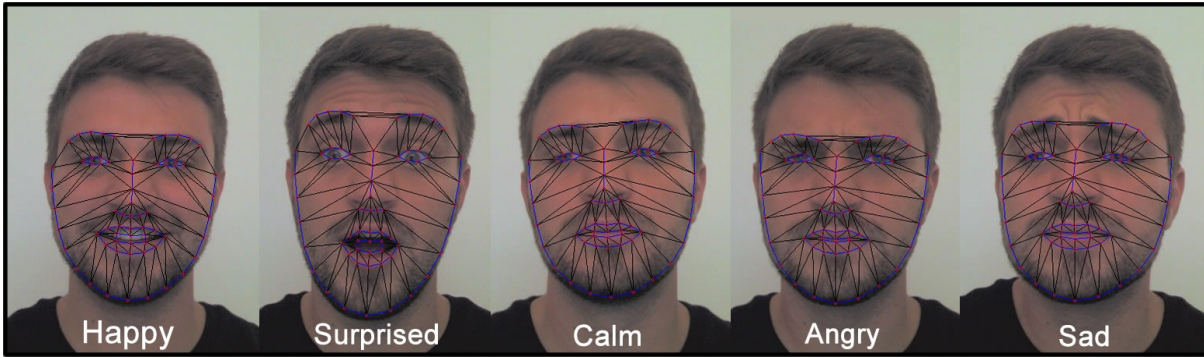


Figura 8 – Detectando as emoções e processando as expressões faciais do usuário durante o jogo.

4.2.2 *Sensor II (S_{II}): Sensores lógicos para Tendências Comportamentais*

De acordo com (MAHLKE; MINGE, 2008), as tendências comportamentais são aspectos emocionais que utilizam medidas não verbais, em geral são avaliadas por indicadores de desempenho tais como o tempo necessário para completar uma tarefa, a precisão de se alcançar um objetivo, o número de erros, o número de pedidos por ajuda, bem como o número de ideias criativas durante a interação.

A UserSense traz doze tipos de sensores lógicos que capturam o desempenho de cada jogador, mapeados para a Estrutura do Espaço Emocional Semântico, e aplicam o coeficiente Kappa ² para a opinião dos especialistas. Os sensores lógicos junto com seus oitantes estão mapeadas a seguir: (1) GET_MUSHROOM: apanhar o cogumelo (oitantes 1,6,7,8); (2) GET_FLOWER: apanhar a flor que dá poder (oitantes 1,6,7,8); (3) WIN: concluir a fase (oitantes 5,6,7,8); (4) KILL_ENEMY_FIREBALL: matar um inimigo com uma bola de fogo (oitantes 4,5,6,7); (5) KILL_ENEMY_SHELL: matar um inimigo com um casco de tartaruga (oitantes 1,6,7,8); (6) KILL_ENEMY_STOMP: matar um inimigo pulando em cima dele (oitantes 1,6,7,8); (7) DIE_ROLE: morrer caindo em um buraco (oitantes 1,2,3,8); (8) DIE_OPPONENT: ser morto por um oponente (oitantes 2,3,4,5); (9) DIE_TIME: não concluir a fase à tempo (oitantes 1,2,3,8); (10) GET_HURT: ser ferido por um inimigo (oitantes 2,3,4,5); (11) NEW_LIFE: ganhar uma vida extra (oitantes 5,6,7,8); (12) FIRE: disparar uma bola de fogo (oitantes 4,5,6,7).

4.2.3 *Sensor III (S_{III}): Reconhecimento de Fala para Avaliações Cognitivas*

As avaliações cognitivas, ligadas à interpretação de uma situação, contribuem para o desenvolvimento das emoções (análise subjetiva do discurso) (FONTAINE *et al.*, 2002). De acordo com Ericsson e Simon (ERICSSON; SIMON, 1993), o método Think Aloud desenha os pensamentos na memória de curto prazo. Haja vista que todos os processos cognitivos viajam através de memória de curto prazo e os pensamentos conscientes do contexto podem ser relatados no momento em que são processados.

A coleta de dados em tempo real, a partir da memória de curto prazo, é preferível porque os

² O coeficiente Kappa é uma abordagem estatística que mede o grau em que os especialistas concordam em uma dada classificação (COHEN *et al.*, 1960). Nesse trabalho, o Coeficiente Kappa obtido foi de 0,80, quando se compara a opinião de três especialistas.

pensamentos gerados a partir da memória de longo prazo de indivíduos são frequentemente manchados pela percepção. Ericsson and Simon (ERICSSON; SIMON, 1993) afirmam que, uma vez que a informação entra a memória de longo prazo, os indivíduos de maneira incorreta podem descrever os processos que efetivamente estão sendo utilizados. Verbalizações que ocorrem simultaneamente com os processos cognitivos são em grande parte independente da interpretação por parte do sujeito (SOMEREN *et al.*, 1994).

A abordagem UserSense é apoiada pelo método Think Aloud, no qual pesquisadores sondam temas correlatos ao contexto de uso, haja vista que os usuários podem facilmente se distrair durante as atividades do experimento (ERICSSON; SIMON, 1993). Caso o silêncio permaneça por vários segundos, os pesquisadores simplesmente sondam os usuários para continuar a falar. Além disso, é interessante estimular os usuários a pensar em voz alta, mas não deve-se desvirtuar a análise adicionando ideias externas aos processos internos de cognição.

Para o desenvolvimento deste trabalho, um sensor de reconhecimento de fala foi utilizado para analisar emoções declaradas pelo usuário e por meio do aplicação do método Think Aloud apoiar a avaliação cognitiva. A UserSense oferece um procedimento de mapeamento entre a emoção falada e os oito oitantes da Estrutura do Espaço Emocional Semântico, (Figura 3), como: Tenso (oitante 1), Frustrado (oitante 2), Insatisfeito (oitante 3), Ansioso (oitante 4), Neutro (oitante 5), Contente (oitante 6), Entusiasmado (oitante 7) e Excitado (oitante 8).

Para sensor de reconhecimento de fala foi implementada uma aplicação *web* que visa transformar a fala do usuário em texto para fazer a captura de emoções e enviá-las para o UserSense. A aplicação utiliza o microfone do computador para a captura da fala do usuário e a envia para a Cloud Speech API³ disponibilizada pela empresa Google que utiliza uma técnica de aprendizado de máquina para fazer o reconhecimento da fala. O texto inferido por meio de um modelo de rede neural artificial é retornado para a aplicação que agora possui um reconhecimento aproximado da fala do usuário em texto. Após receber os dados, a aplicação faz uma busca no texto apenas por palavras que expressam emoções, e que no final serão enviadas para o servidor com o UserSense, juntamente com o instante no qual elas foram ditas.

4.2.4 **Sensor IV (S_{IV}): Frequência Cardíaca para Reações Fisiológicas**

Na busca de um sensor cardíaco acessível e de funções personalizáveis foi escolhido a plataforma Android para desenvolver um aplicativo capaz de monitorar a frequência cardíaca de um usuário e, posteriormente, avaliá-la. Para essa coleta de dados, utiliza-se os aparelhos Motorola Moto X (*smartphone*) e Moto 360 (*smartwatch*).

A ideia básica foi ter um relógio registrando os dados de um usuário e transmitindo para o servidor, para isso o relógio é pareado com um *smartphone* Android através de uma camada de dados *bluetooth* e o *smartphone* comunica com o *back-end* via conexão *socket* – servidor do Lapsd.

O aplicativo é dividido em 2 partes: a primeira atua somente no relógio, responsável por compilar os dados do sensor cardíaco e enviar ao celular via *bluetooth* e a segunda recebe os dados enviados pelo relógio e os envia ao servidor.

³ <https://cloud.google.com/speech>



Figura 9 – Detectando as emoções e processando de dados de frequência cardíaca do usuário durante o jogo.

Inicialmente, determinou-se que o relógio buscaria dados sobre a frequência cardíaca do usuário a cada 8 segundos e imediatamente enviaria ao *back-end* com o intuito de armazená-los, mas devido ao gasto excessivo de bateria dos dispositivos, mudou-se o intervalo de transmissão de dados relógio/celular para a cada 30 segundos.

O sensor do *smartwatch* funciona por meio da fotopletismografia (PPG), obtida por meio de um oxímetro de pulso, que ilumina a pele do usuário e faz a medição das mudanças na absorção de luz. A variação do volume sanguíneo causado no local faz com que a absorção de luz mude, então o sensor pode identificar cada ciclo cardíaco para, então, fazer o cálculo da frequência cardíaca (CARVALHO *et al.*, 2012).

No desenvolvimento do aplicativo foi observada uma limitação técnica ao utilizar medições frequentes pois o relógio desligava o sensor cardíaco e retornava sempre para mesmo valor. Para contornar esse problema, utilizou-se o procedimento de deixar o aplicativo no modo *STAY_AWAKE*, no qual se consegue prioridade de recursos e a tela do aparelho não apaga. Dessa maneira, os pesquisadores podem acompanhar a frequência cardíaca por meio dos dados mostrados na tela do relógio no pulso do usuário (ver Figura 9).

Após o servidor receber pacotes contendo os dados coletados, a verificação é feita de modo a classificá-los em três categorias: Neutro, Negativo ou Positivo. A partir do trabalho desenvolvido por (RICHTER *et al.*, 2011), propôs-se o seguinte procedimento de mapeamento da frequência cardíaca para a Estrutura do Espaço Emocional Semântico: caso no período de 10 segundos, haja um aumento da frequência cardíaca de 3.1 ± 1.9 bpm, considera-se que a valência é negativa; caso ocorra uma diminuição de 2.3 ± 2.1 bpm, considera-se que a valência é positiva e em outros casos, ela é considerada neutra. Essas classificações retornavam o servidor que as salvava em um arquivo específico referente a cada usuário.

4.3 Modelo de Classificação da UserSense

Como a UserSense conta com vários sensores para aumentar a confiabilidade na análise das emoções do usuário, usa-se um modelo de classificação para combinar as saídas dos sensores de classificação para emoções do usuário. Nesse momento do trabalho, será descrita a maneira como tal modelo pode ser obtido utilizando métodos de classificação de aprendizado de máquina.

A classificação é um dos problemas mais importantes em aprendizado de máquina. Dado um conjunto de dados composto de n pares (x_i, y_i) , em que cada x_i é um item de dados (instância) e y_i representa sua classe (rótulo), um algoritmo de classificação deve encontrar uma função, por uma fase de treinamento ou de ajuste, que mapeia cada item de dados à sua classe correta.

No contexto do reconhecimento de emoções em tempo real, considera-se como instância de dados as medidas dadas pelos sensores em um determinado instante; a Seção 4.3.1 apresenta a maneira de usar essas medições para descrever cada instância. O rótulo de cada uma delas é o estado emocional do usuário naquele momento. Como este trabalho usa a opinião de psicólogos para rotular as instâncias, um rótulo corresponderá a um ou mais oitantes que correspondem ao estado emocional do usuário. A Seção 4.3.2 explica como utilizar os rótulos da instância para gerar o modelo de classificação. Conforme discutir-se-á a seguir em mais detalhes, foram gerados oito modelos de classificação, um para cada oitante.

4.3.1 Descrevendo as Instâncias da UserSense

Uma estratégia comum para descrever as instâncias é definir um conjunto de atributos e listar os valores do atributo para cada instância. Por exemplo, os alunos de uma sala de aula são descritos em termos dos seguintes atributos: sexo, idade, altura e peso. Nesse exemplo, uma instância (ou seja, um estudante) poderia ser {masculino,12,1.50m,45kg}.

No contexto dessa tese, o intuito é classificar o estado emocional de uma pessoa em um determinado momento. Assim, cada instância será definida pelas medições dos sensores para uma determinada pessoa em um determinado momento. Como as saídas dos sensores são todas as informações que tem-se sobre a instância (desconsiderando todas as informações sobre o próprio usuário), os atributos serão definidos em função das saídas dos sensores.

Existem muitas maneiras pelas quais o conjunto de atributos podem ser definidos. Nesta pesquisa, a escolha de atributos foi motivada pela escolha de métodos de classificação (ou seja, SVMs e Árvores de Decisão). No entanto, um outro conjunto de atributos pode ser usado na UserSense.

No total, usou-se 41 atributos para descrever cada instância. Em seguida, eles serão descritos. Note que não há redundância na informação representada pelos diferentes tipos de atributos, deixando o método de aprendizado decidir quais atributos são relevantes, ou não, para a tarefa de aprendizado.

- O primeiro conjunto de atributos indica os oitantes que foram estimados por cada sensor. Para representar essa informação tem-se um atributo binário $O_{i,j}$ para cada oitante i (com i de 1..8) e sensor j (com j de 1..4), onde $O_{i,j}$ é igual a 1 se o sensor j previu o oitante i e 0 caso contrário. Como tem-se quatro sensores e um atributo binário para cada sensor, a isso acrescenta-se 32 atributos.

Tabela 2 – 41 atributos para descrever cada instância.

Nome	Atributos							
$O_{i,j}$	$O_{1,1} = 0$	$O_{2,1} = 0$	$O_{3,1} = 1$	$O_{4,1} = 1$	$O_{5,1} = 0$	$O_{6,1} = 0$	$O_{7,1} = 0$	$O_{8,1} = 1$
	$O_{1,2} = 0$	$O_{2,2} = 1$	$O_{3,2} = 1$	$O_{4,2} = 1$	$O_{5,2} = 1$	$O_{6,2} = 0$	$O_{7,2} = 0$	$O_{8,2} = 0$
	$O_{1,3} = 0$	$O_{2,3} = 0$	$O_{3,3} = 1$	$O_{4,3} = 1$	$O_{5,3} = 0$	$O_{6,3} = 0$	$O_{7,3} = 0$	$O_{8,3} = 0$
$Count_i$	$Count_1 = 0$	$Count_2 = 1$	$Count_3 = 3$	$Count_4 = 3$	$Count_5 = 1$	$Count_6 = 0$	$Count_7 = 0$	$Count_8 = 1$
MF_i	$MF_1 = 0$	$MF_2 = 0$	$MF_3 = 1$	$MF_4 = 1$	$MF_5 = 0$	$MF_6 = 0$	$MF_7 = 0$	$MF_8 = 0$
$CountMF$	$CountMF = 2$							

- O segundo conjunto de atributos indica quantos sensores preveem cada oitante. Para representar essa informação, tem-se um atributo numérico $Count_i$ para cada oitante i (com i de 1..8). Como tem-se um atributo $Count$ para cada oitante, tem-se um total de oito atributos.
- O terceiro conjunto de atributos indica quais oitantes foram previstos na maioria das vezes quando se considera todos os sensores. Para representar isso, tem-se um atributo binário MF_i para cada oitante i (com i de 1..8). Como tem-se um atributo MF para cada oitante, tem-se um total de oito atributos.
- Tem-se um atributo numérico $CountMF$ que conta quantos oitantes apareceram mais vezes.

Com o intuito de esclarecer melhor a abordagem proposta, será descrito com mais detalhes os atributos supramencionados relacionando-os com os dados das instâncias. Para isso, considere uma situação em que o estado emocional de um usuário foi identificado por meio de três sensores: o sensor de reconhecimento facial indica os oitantes 3, 4 e 8; enquanto o sensor lógico seleciona os oitantes 2, 3, 4 e 5; e o sensor de fala sugere os oitantes 3 e 4.

A Tabela 2 ilustra como as saídas dos sensores são convertidas para os valores dos 41 atributos que são considerados nesta abordagem. No contexto desta Tese, usa-se o rótulo “1” para representar o sensor de reconhecimento facial, enquanto rótulo “2” para representar o sensor lógico e o rótulo “3” para representar o sensor de reconhecimento de fala.

Como a classificação é uma tarefa de aprendizado supervisionado, cada instância x_i deve ser associada a um rótulo y_i . Nesse contexto, como as instâncias foram rotuladas com base nas análises de alguns psicólogos (nos dados utilizados nos experimentos, os rótulos foram definidos por três psicólogos), usou-se a emoção(ões) frequentemente escolhidas como o rótulo. Essa estratégia direciona-se para um problema de multirrótulo, como descrito a seguir.

4.3.2 Classificação Multirrótulo

Em um problema de classificação multirrótulo, cada exemplo pode pertencer simultaneamente a mais de uma classe. A classificação das emoções é um problema multirrótulo, uma vez que uma ou mais emoções podem ser atribuídas a um usuário num mesmo dado momento.

Muitas abordagens têm sido utilizadas na literatura para lidar com problemas multirrótulo. Neste trabalho, usa-se o Método de Decomposição Binária (CARVALHO; FREITAS, 2009), que consiste em transformar um problema originalmente multirrótulo em um certo número de problemas binários. E cada

problema binário será gerado de tal maneira que um classificador pode ser treinado para responder a seguinte pergunta “Será que esta instância pertence a uma determinada classe ou não?”. Na verdade, haverá uma classificação binária para cada classe, que irá distinguir essa classe de todas as outras.

Ao combinar os resultados dos classificadores individuais, pode-se obter uma previsão definitiva multirrótulo. Por exemplo, para o caso de três classes, o problema multirrótulo será transformado em três problemas binários, e um classificador será treinado para resolver cada um deles. O classificador 1 responde a seguinte pergunta “Será que esta instância pertence à classe 1?”. Os classificadores 2 e 3 respondem a mesma pergunta para suas respectivas classes. Para uma dada instância x , o classificador 1 prediz x como pertencente à classe 1, o classificador 2 prediz x como não pertencente à classe 2, e o classificador 3 prediz x como pertencente à classe 3, logo a previsão multirrótulo para x será as classes 1 e 3.

No contexto deste trabalho, foi gerado um classificador para cada oitante. A motivação para a escolha dessa abordagem é o fato de que cada problema binário resultante pode ser tratado com um modelo tradicional de aprendizado de máquina padrão, por exemplo SVM e Árvores de Decisão padrão, que são os métodos utilizados neste trabalho.

Na abordagem UserSense, tem-se um problema binário para cada classe, haja vista que pode-se transformar os dados de treinamento multirrótulo em dados de treinamento binário para essa classe. Nos dados de treinamento resultante, considera-se como exemplos positivos todos os casos em que a previsão multirrótulo contém essa classe. Caso contrário, considera-se as instâncias como negativas.

Há múltiplas maneiras de implementar a UserSense. Na implementação apresentada nesta Tese, cria-se um classificador binário para cada um dos oito domínios que agrupam os estados emocionais, ou seja, tem-se oito modelos para prever emoções. Cada um desses modelos tem como entrada os dados provenientes de quatro diferentes sensores (fala, face, lógico e frequência cardíaca) e uma previsão final pode ser obtida por um comitê (previsão de diferentes modelos). Avaliou-se essa implementação com duas opções diferentes e independentes de métodos de classificação de aprendizado de máquina: SVMs e Árvores de Decisão.

Como ilustração, considere r um exemplo de aplicação que tem apenas três instâncias de formação: x_1 , x_2 e x_3 . Suponha que: a instância x_1 pertence aos oitantes 1, 2, e 3; x_2 pertence aos oitantes 1 e 2; e x_3 pertence ao oitante 4. Conforme pode ser visto na Figura 10, o problema binário para o oitante 1, x_1 e x_2 será considerado como positivo, enquanto x_3 será considerado como negativo. O mesmo é válido para o oitante 2. Para o oitante 3, x_1 será considerado como positivo, enquanto x_2 e x_3 são considerados como negativos. Para o oitante 4, x_3 será considerado como positivo, e x_1 e x_2 são considerados como negativos. Para todos os oitantes restantes, todas as instâncias são consideradas negativas.

4.4 Considerações finais

Este Capítulo apresentou a UserSense. Essa abordagem explora o conceito de componentes que são carregados, instanciados e destruídos quando solicitados. Uma abordagem baseada em componentes que contribui com um modelo genérico para a estruturação e extensibilidade de *software* baseada em componentes. No *Middleware* Adaptativo OpenCom, uma abordagem apoiada por componentes e fortemente

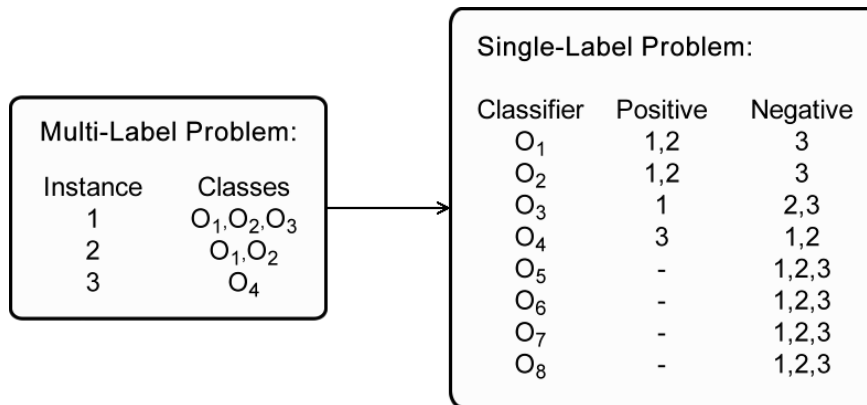


Figura 10 – Aplicação do Método de Decomposição Binária.

acoplada ao conjunto de componentes que, de acordo com (COULSON *et al.*, 2008), cooperam para a solução de problemas. Este Capítulo apresentou também cada um dos sensores utilizados no contexto desta pesquisa e a relação deles com a Teoria Componential das Emoções (ver Figura 1). Ele também mapeou as respostas emocionais trazidas pelos sensores com Estrutura do Espaço Emocional Semântico mostrada na Figura 3.

No mais, a abordagem UserSense apresenta um modelo de classificação para cada oitante da Estrutura do Espaço Emocional Semântico, ou seja, tem-se 8 modelos para fazer a predição das emoções. Esses modelos têm como entrada o que vem dos 4 sensores. Para a UserSense foram adotados dois algoritmos de classificação de aprendizado de máquina (SVM e Árvores de Decisão) que podem ser usados para gerar esses modelos de classificação. Portanto, pode-se usar um deles ou o outro para gerar os modelos (são duas implementações diferentes do que esta sendo proposto). No próximo Capítulo, serão apresentados detalhes sobre os experimentos exploratórios realizados nesta pesquisa, assim como a aprovação desta pesquisa pelo Comitê de Ética em Pesquisa da Faculdade de Medicina da Universidade de São Paulo (CAAE-45081415.0.0000.0065).

AVALIAÇÃO EXPERIMENTAL DA USERSENSE

Com o objetivo de aproximar-se e entender mais sobre as emoções declaradas e expressas pelos usuários durante a interação com sistemas computacionais interativos, foram realizados alguns experimentos com usuários com níveis de escolaridade variando de Ensino Médio à Doutorado e com diferentes experiências com os jogos. Além disso, 30% dos participantes disseram que jogar jogos de vídeo é a sua principal atividade de lazer, 80% eram do sexo masculino, 77% têm uma experiência com jogo de mais de dois anos, 37% deles jogam 1 a 3 vezes por semana, 83% declaram jogar no computador, 33% gastam dinheiro em jogos e, 70% dos participantes jogam jogos online. O estudo foi aprovado pelo Comitê de Ética em Pesquisa da Faculdade de Medicina da Universidade de São Paulo (CAAE-45081415.0.0000.0065), ver Apêndice B.

5.1 Planejamento dos Experimentos

Para o planejamento dos experimentos exploratórios, foram definidos os objetivos, a hipótese, a metodologia, o material e os dispositivos utilizados para verificação da hipótese.

Hipótese: para que os sistemas computacionais interativos possam reagir ao estado emocional de um indivíduo, pode-se se criar uma infraestrutura de *software* e *hardware* com auxílio de sensores e métodos de classificação, que apoie a captura, análise e classificação de indícios de respostas emocionais de usuários visando permitir que aplicações computacionais se flexibilizem a partir de tal classificação.

Objetivo: coletar e analisar respostas emocionais de usuários interagindo com um jogo eletrônico em um computador (Ver Figura 11), considerando dados relacionados com respostas emocionais obtidos durante a interação.

Metodologia Aplicada: com o intuito de analisar e coletar respostas emocionais do usuário na interação com sistemas computacionais interativos, um grupo de usuários foi convidado a participar de um experimento de uso do computador. O objetivo da atividade era que esses usuários jogassem o jogo Super Mario Brose (PEDERSEN; TOGELIUS; YANNAKAKIS, 2010). Cada usuário jogou por 10 minutos e

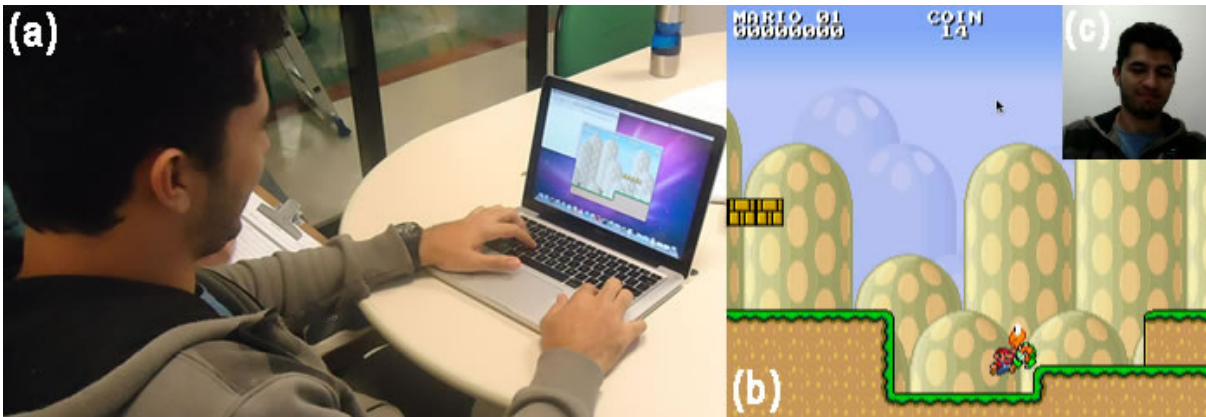


Figura 11 – Usuário interagindo durante o 1º experimento com o jogo Super Mario Bros. (a) usuário com o notebook, (b) tela do jogo, (c) captura pelo sistema de reconhecimento de emoções faciais.

durante a interação aplicou-se o método Think Aloud (Someren, 1994). No mais, Reconhecimento de Emoções Faciais, Reconhecimento de Fala, Sensores Lógicos e Frequência Cardíaca foram utilizados para a captura de dados em tempo de interação.

Material de Apoio: para a condução dos experimentos foram preparados: um Termo de Consentimento Livre e Esclarecido, um Termo de Autorização de Captação e Exibição de Imagem, Som e Nome e um Questionário de Levantamento de Perfil. O Termo de Consentimento Livre e Esclarecido elucida, aos participantes, o objetivo da pesquisa, a voluntariedade na participação e o caráter científico da mesma (ver Apêndice C). O Termo de Autorização de Captação e Exibição de Imagem, Som e Nome esclarece aos participantes que há captura de imagens e voz e que os dados coletados são utilizados para fins científicos (ver Apêndice D). O Questionário de Levantamento de Perfil possuía questões de cunho social e cultural que permitiram traçar o perfil desses usuários (ver Apêndice E).

Durante a condução dos experimentos exploratórios foi apresentada aos usuários uma tabela de emoções proveniente de uma versão traduzida da Estrutura do Espaço Emocional Semântico de Scherer, para auxiliar na aplicação do método Think Aloud (SOMEREN *et al.*, 1994) (ver Apêndice A). Para essa pesquisa, a captura de imagem e som foi essencial para aplicação desse método.

Dispositivos Utilizados: os usuários realizaram a interação individualmente e no decorrer do experimento jogaram o jogo Super Mario Bros em um computador Macbook Pro. Durante os 10 minutos de interação, o usuário não foi interrompido. A captura de imagens e voz ocorreu pela câmera do próprio computador; e os valores de frequência cardíaca foram lidos por um *smartwatch* da Motorola - Moto 360.

5.2 Critérios de Inclusão de Usuários nos Experimentos

Os sujeitos da pesquisa foram adolescentes e adultos com faixa etária entre 17 e 45 anos, homens e mulheres, com experiência ou não em jogos, independentemente de classe social, cor e escolaridade, provenientes de amostragens da população universitária do estado de São Paulo. Os usuários que participaram da pesquisa não foram pagos, mas convidados por listas de e-mail da universidade. A cada novo experimento, diferentes usuários puderam participar da pesquisa.

Os sujeitos realizaram a interação individualmente. Não houve comparação ou exposição de

qualquer usuário, pois essa pesquisa está interessada em coletar/calibrar sensores e parametrizar o *software* tentando fazer com que esse consiga coletar e analisar emoções de diferentes usuários.

5.3 Critérios de Exclusão de Usuários dos Experimentos

Não participaram como sujeitos da pesquisa pessoas que não estejam na faixa etária indicada de 17 e 45 anos.

5.4 Riscos em Participar dos Experimentos

A desmotivação está entre os riscos da participação dos sujeitos na pesquisa, pois nem todos os usuários poderiam ter afinidade pelo jogo. Os participantes poderiam sentir timidez, pois conforme está previsto no método Think aloud e considerando o contexto desta pesquisa, o usuário deveria relatar suas emoções durante o experimento exploratório.

O usuário estava interagindo com sistemas computacionais, por exemplo, jogos eletrônicos altamente interativos. Portanto, poderia ocorrer a alteração no seu estado emocional.

O questionário também poderia ser considerado um risco, que foi contornado uma vez que o objetivo dos questionários é entender um pouco sobre o perfil dos usuários. Com essa finalidade, as questões foram elaboradas para mitigar ao mínimo possível de desconforto e constrangimento. Além disso, vale destacar que todos os protocolos éticos foram cumpridos, os dados coletados são confidenciais e os usuários poderiam interromper a participação no experimento exploratório a qualquer momento.

A escolha e definição dos jogos é feita pensando no público alvo com o intuito de permitir o acesso ao maior número possível de usuários.

5.5 Benefícios em Participar dos Experimentos

Os benefícios relacionados à participação dos sujeitos da pesquisa incluem a análise dos seus estados emocionais, a descontração e diversão que muitas vezes o jogo proporciona. O jogo como atividade prazerosa, de acordo com (PEDERSEN; TOGELIUS; YANNAKAKIS, 2010), proporciona a oportunidade de uma participação efetiva que na sua essência vai determinar uma tomada de decisão.

Além disso, a partir dos experimentos foi possível aprimorar a abordagem para estimar o estado emocional de indivíduos na interação com sistemas computacionais. Os sujeitos da pesquisa puderam contribuir também com a inclusão social e digital, haja vista que o estado emocional pode ser usado como artifício para melhorar a forma como as soluções computacionais são apresentadas à população.

5.6 Tamanho da Amostra de Usuários

Foram realizados 5 experimentos, com a participação média de 30 pessoas em cada.

5.7 Considerações finais

A partir dos dados obtidos na interação com o jogo Super Mario Bros, foi possível conduzir a avaliação emocional do usuário em tempo de real. Considerando a Estrutura do Espaço Emocional Semântico e Abordagem Híbrida para a Avaliação da Experiência Emocional, foram definidas medidas capazes de avaliar os quatro domínios da experiência emocional do usuário (Valência, Excitação, Sentimento de controle e Facilidade de conclusão do objetivo) definidos por (SCHERER, 2005). O próximo Capítulo traz uma avaliação das emoções durante e após a interação de usuários com um jogo eletrônico. Além disso, o Capítulo apresenta uma avaliação de desempenho e sobrecarga da abordagem UserSense em dispositivos móveis.

AVALIAÇÃO DAS EMOÇÕES DURANTE E APÓS INTERAÇÃO

De acordo com (SCHERER, 2005), não há nenhum método objetivo capaz de avaliar Sentimentos Subjetivos, porque a única maneira de obter essa informação é perguntar ao próprio usuário. Em outras palavras, este fator emocional só permite avaliar um episódio emocional, todavia torna-se impossível realizar uma avaliação objetiva do estado emocional (durante a interação e sem interrupção do usuário).

A Experiência Subjetiva (ou Sentimento Subjetivo) é responsável pela avaliação de um episódio emocional, ou seja, da experiência subjetiva emocional consciente que se relaciona com a capacidade do indivíduo para regular as emoções (DESMET, 2004). Como (RUSSELL; WEISS; MENDELSON, 1989) propôs, tais experiências são consideradas percepções: sentir medo é perceber-se com medo. Mais tecnicamente, (YIK; RUSSELL; BARRETT, 1999) caracterizou esse sentimento como uma meta-experiência, pois os dados brutos em que ele se baseia inclui 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).

As medidas disponíveis na literatura para avaliar os Sentimentos Subjetivos são geralmente de auto-avaliação e não-verbais, como SAM (BRADLEY; LANG, 1994), Emocards (REIJNEVELD *et al.*, 2003) e PrEmo (DESMET, 2004). Uma medida verbal é o Affect Grid (RUSSELL; WEISS; MENDELSON, 1989). SAM é um instrumento que utiliza pictogramas em vez de palavras para representar as respostas emocionais (veja Figura 12). O usuário deve escolher, entre os pictogramas, aqueles que melhor representam suas emoções (DESMET, 2004).

6.1 Avaliando a Experiência Emocional do Usuário

No contexto desta pesquisa, o SAM foi escolhido porque este instrumento de auto-avaliação pode avaliar os domínios que são diretamente correspondentes aos domínios da Estrutura do Espaço Emocional

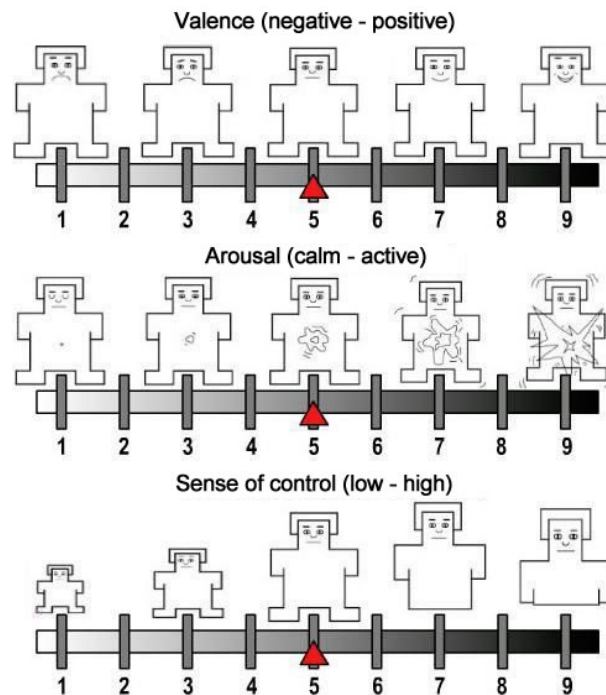


Figura 12 – Cartão SAM(Self-Assessment Manikin) utilizado nos experimentos. Cartões SAM foram adotados para avaliar os Sentimentos Subjetivos ao final do jogo. Adaptado de (BRADLEY; LANG, 1994).

Semântico definidos por (SCHERER, 2005).

Esse tipo de escala de avaliação é composta por três conjuntos de pictogramas que representam os domínios de Valência (satisfeito - insatisfeito), Excitação (motivado - relaxado) e Sentimento de controle (completamente no controle - fora de controle) (MAHLKE; MINGE, 2008). Cada domínio é representado por uma escala de valores que varia de 1-9. O usuário é instruído a escolher um dos cinco pictogramas ou intermediário entre os dois - o que melhor representa o seu estado emocional, no que diz respeito ao domínio que está sendo avaliado. Os valores dos círculos selecionados são verificados para obter o resultado - o domínio é avaliado como positivo para valores entre 1 e 4; o domínio é avaliado como negativo para valores entre 6 e 9; e o valor 5 avalia o domínio como neutro.

Apesar do foco desta Tese não ser a avaliação de episódios emocionais, julgou-se interessante realizar um estudo que analisa as respostas emocionais dos usuários durante e após a interação. Pois, acredita-se que métodos de autoavaliação para a análise das emoções, aplicados ao final da interação, trazem resultados distintos de técnicas que verificam as emoções em tempo de interação. Acredita-se que o uso de abordagens que fazem a análise em tempo real traz uma riqueza de detalhes que destoam de avaliações feitas apenas ao final da interação. Logo, os dados capturados pelos sensores durante a interação do usuário foram confrontados com a auto-avaliação emocional que utiliza o instrumento SAM ao final da interação.

A partir dos dados obtidos na interação com o jogo Super Mario Bros, foi possível conduzir a avaliação emocional do usuário em tempo de real e após interação. Considerando a abordagem UserSense, foram definidas medidas/sensores capazes de avaliar os quatro domínios da experiência emocional do usuário (Valência, Excitação, Sentimento de controle e Facilidade de conclusão do objetivo) definidos por (SCHERER, 2005).

Durante esse experimento que capturou respostas emocionais de usuários interagindo com um jogo, foram registradas as falas dos jogadores por meio de um sensor de Reconhecimento de Fala, que apoiou a aplicação do método Think Aloud. Tá método dos usuários auxiliou na Avaliação Cognitiva em tempo de interação, por exemplo, quando o jogador disse “Estou Ansiosa, deu vontade de fumar.”, corresponde a estar no oitante 4; outro usuário, por exemplo, disse “Que tenso, deu até calor agora!”, esse sentimento corresponde a estar no oitante 1 da Estrutura do Espaço Emocional Semântico.

A Figura 13 traz o exemplo de um usuário que, por meio da técnica Think Aloud, disse em um determinado momento da interação “Eu morri porque o tempo acabou! Estou triste.”. No mais, considerando todas as falas desse usuário percebe-se que o mesmo se encontra nos oitantes 3 e 4. O sensor lógico apontou que nesse momento do jogo o usuário havia morrido porque o tempo da fase havia acabado DIE_TIME, semelhantemente, considerando todos os sensores lógicos acionados, esse usuário se encontra nos oitantes 1, 2, 3 e 8. Por fim, o Reconhecimento de Emoções Faciais disse que, na maior parte do tempo de interação, o usuário estava Triste e, portanto, se encontrava nos oitantes 3 e 8. Todavia, ao final da interação quando foi realizar a autoavaliação com o SAM, o jogador disse que estava Satisfeito (oitantes 5, 6, 7 e 8), Motivado (oitantes 1, 2, 7 e 8) e Totalmente no Controle da situação (oitantes 1, 6, 7 e 8).

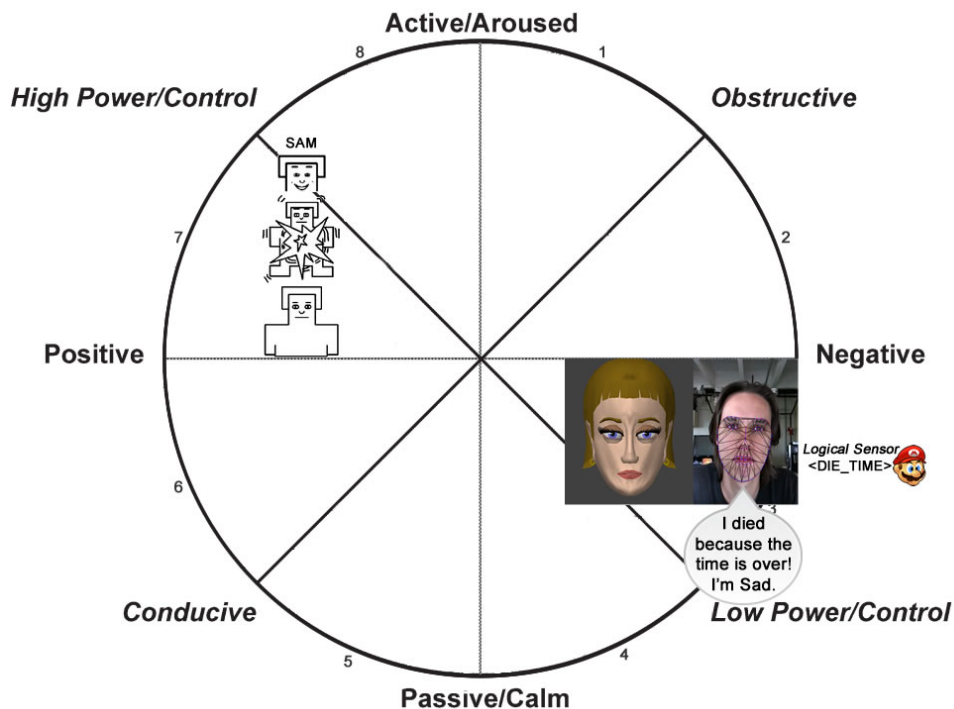


Figura 13 – Avaliação emocional do usuário antes e depois da interação com o jogo.

Diante do cenário desse experimento, foi possível verificar que para a análise da resposta emocional em tempo de interação, o usuário teve maior incidência no oitante 3. Entretanto, no que diz respeito a autoavaliação realizada após a interação, o usuário teve maior incidência nos oitantes 7 e 8. A partir do exemplo demonstrado na Figura 13, foi possível perceber indícios de que nem sempre aquilo que o usuário declara em uma avaliação ao final da interação corresponde ao que ele demonstrou durante a interação.

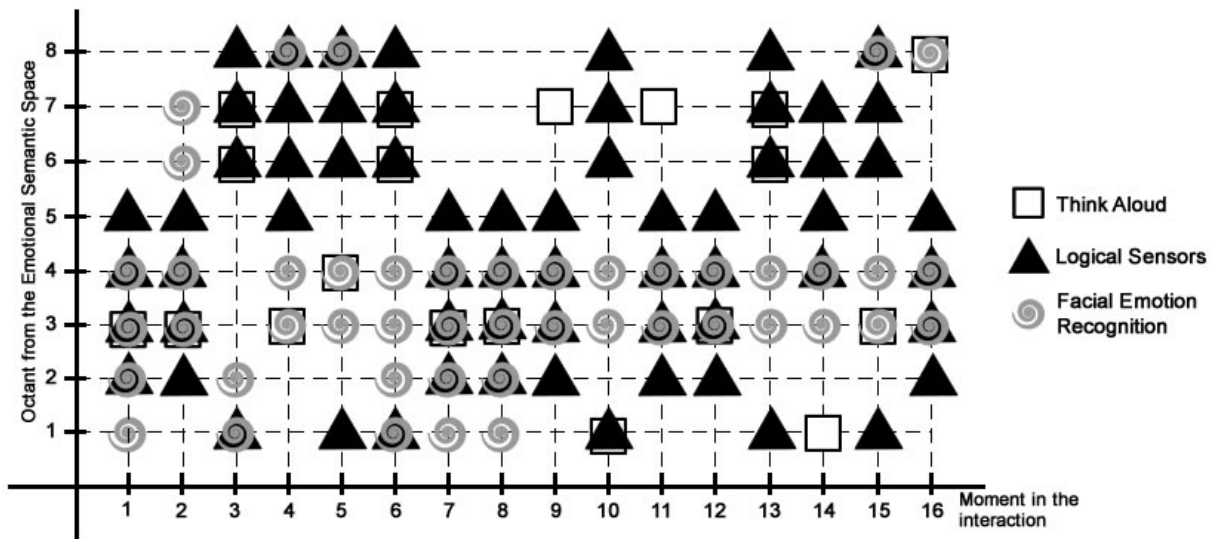


Figura 14 – Oitantes identificados para cada momento da interação, para um usuário. O eixo Y representa os oitantes e o eixo X representa o momento de interação (instância). Cada sensor identifica, em cada momento de interação, um conjunto de oitantes conforme mostra a figura. O estado emocional final para cada instância é dado pela eleição da maioria dos oitantes sugerida pelos sensores. Por exemplo, no instante 16, os oitantes sugeridos são 3, 4 e 8 - haja vista que, dois sensores sugerem esses oitantes, enquanto que os oitantes 2 e 5 são ignorados no momento, porque apenas um sensor os sugeriu.

6.2 Avaliação Não Paramétrica das Respostas Emocionais do Usuário

Tendo executado o experimento da coleta de respostas emocionais de usuários na interação com um jogo, foi possível perceber a riqueza de detalhes e a distinção de uma abordagem que avalia as emoções em tempo de interação se comparada a métodos e técnicas que fazem a análise da experiência emocional considerando um episódio emocional do indivíduo.

As avaliações cognitivas estão ligadas à interpretação (avaliação) de uma situação e contribuem para o desenvolvimento das emoções (SCHERER, 2005). Uma mesma situação pode induzir diferentes emoções, dependendo da interpretação feita pelo usuário. Nesse, sentido o uso o método Think aloud trouxe resultados para o experimento que sugerem uma relação direta entre o que o usuário declarou estar sentindo com os dados capturados pelos Sensores Lógicos. Por exemplo, quando o 1º usuário diz “Isso me deixa feliz, matar vários de uma só vez com o casco. Eu estava tentando isso desde o início” e o Sensor Lógico acionado foi KILL_ENEMY_SHELL, assim o gráfico da Figura 14 mostra a incidência dos oitantes para sensores, em que é possível observar que o usuário 1 passou 35% do tempo de interação no oitante 3 (ver Tabela 3). Em outro momento observado, por exemplo, o 10º usuário declarou “Apareceu uma flor no jogo. Eu estou feliz!” e o sensor lógico aponta GET_FLOWER.

A partir da aplicação da abordagem UserSense foi possível ter indícios da experiência emocional de cada usuário durante cada instante da interação, assim como considerar a autoavaliação do episódio emocional do jogador, segundo os quatro domínios da Estrutura do Espaço Emocional Semântico. Por exemplo, o 22º usuário, durante a interação com o Jogo Super Mario Bros, pode ter experimentado emoções como frustração, ansiedade e apreensão, haja vista que o sensor lógico registrou nesse instante

do jogo o GET_HURT, em paralelo o Reconhecimento de Emoções Faciais detectou que o usuário estava com Raiva. Ainda analisando o 22º usuário, é possível observar na Tabela 3 em que 42% do tempo do tempo de interação, esse jogador estava tendo uma experiência emocional correspondente ao oitante 2, que está associado a emoções como raiva, frustração, angústia e descontentamento. Todavia, quando esse participante faz a autoavaliação do episódio emocional do jogo, ele declara estar feliz, contente, satisfeito e confiante (oitante 6).

A partir da Tabela 3 é possível observar a compilação dos dados da aplicação da abordagem UserSense apresentada no Capítulo 5. Na primeira coluna, tem-se um conjunto de 30 usuários que participaram do experimento de interação com o jogo. Na segunda coluna, tem-se os oitantes predominadas de cada usuário durante os 10 minutos de interação - considerando os sensores: Reconhecimento de Fala (Think Aloud), Reconhecimento de Emoções Faciais e Sensores Lógicos. Por fim, a terceira coluna traz oitantes predominantes de cada usuário para a autoavaliação do episódio emocional de jogar.

Diferenciar as emoções, assim como a adoção de um modelo para a descrição emocional, é uma etapa relevante para o estudo das emoções. No contexto de IHC, os sentimentos são importantes, pois eles são persistentes; uma pessoa pode acreditar que continua feliz mesmo que algo ruim tenha acontecido e modificado seu estado emocional (FRIJDA; SCHERER *et al.*, 2009; GONÇALVES *et al.*, 2016). Por exemplo, o 24º usuário declara “Minha nossa senhora!!! Que susto.” e o Reconhecimento de Emoções Faciais aponta que o jogador mostra indícios de estar com Raiva. Em adição, os sentimentos são responsáveis por influenciar as pessoas a buscarem ou evitarem objetos ou situações de acordo com seus interesses (RODRIGUES *et al.*, 2016).

Com o intuito de realizar uma validação estatística (correlação e comparação) sobre os dados coletados para a Figura 15, realizou-se inicialmente uma verificação da adequação a normalidade dos dados. Isso é, analisar se a distribuição é ou não uma Gaussiana. Para tanto, utilizou-se o *Shapiro-Wilk Normality Test*, o qual apresentou os seguintes p – valores para o conjunto de sensores e SAM: 0,1115 e $2,569e - 05$. Considerando 95% de confiança, o conjunto sensores (Reconhecimento de Fala/Think Aloud, Reconhecimento de Emoções Faciais e Sensores Lógicos) possui a distribuição da normalidade aceita, enquanto o conjunto SAM não possui. Nesse sentido, o *Wilcoxon Rank Sum Test* é o método não paramétrico mais indicado para realizar a comparação entre os conjuntos, apresentando o seguinte p – valor : $4,347e - 05$.

Na Figura 15, é apresentado o resultado da análise não paramétrica entre o conjunto de sensores e SAM em função dos oitantes. A comparação entre as avaliações das respostas emocionais em tempo de interação (por exemplo, sensores, oitante 3 como resultante) e depois da interação (por exemplo, SAM, oitante 7 como resultante) mostra que os resultados retornados são distintos nessa análise aplicada. Nota-se que é possível afirmar que os métodos considerados são estatisticamente diferentes (p – valor : $4,347e - 05$, considerando 95%). Portanto, a avaliação de respostas emocionais em tempo de interação apresenta resultados distintos de uma análise que ocorre ao final da tarefa.

Tabela 3 – Resultados do experimento com 30 usuários. Pode-se verificar que os oitantes obtidos pelos sensores durante o tempo de interação diferem dos oitantes obtidos ao final do experimento (empregando SAM). Na segunda coluna, o valor percentual mostra o período de tempo que o usuário foi identificado no oitante específico.

Usuário	Oitantes obtidos pelos sensores durante o jogo	Oitantes obtidos pelo SAM no final do jogo
1	3 (35%)	1, 7, 8
2	3 (33%)	3, 4
3	1, 3, 4, 8 (20%)	7, 8
4	3, 4, 6 (23%)	2, 5, 7, 8
5	3 (23%)	2, 5, 7, 8
6	1 (25%)	5
7	2, 3, 6 (30%)	7, 8
8	5 (29%)	5, 7, 8
9	5 (24%)	2, 5, 7, 8
10	3 (24%)	7, 8
11	6 (67%)	7, 8
12	5 (38%)	5, 7, 8
13	3, 4 (31%)	5, 7, 8
14	2, 3 (25%)	7, 8
15	3 (24%)	7, 8
16	2 (33%)	2, 5, 7, 8
17	3 (22%)	5, 7, 8
18	1, 2 (47%)	7, 8
19	4 (21%)	7, 8
20	6 (27%)	2, 5, 7, 8
21	4 (26%)	5, 7, 8
22	2 (42%)	6
23	4, 5 (23%)	7, 8
24	4, 6, 7 (20%)	2, 5, 7, 8
25	4, 6 (25%)	5
26	1 (29%)	2, 5, 7, 8
27	7 (44%)	7, 8
28	3 (29%)	7, 8
29	8 (30%)	7, 8
30	2 (32%)	5, 7, 8

6.3 Avaliação de Desempenho e Sobrecarga da UserSense

Uma avaliação independente foi realizada para avaliar a Usersense e a viabilidade do seu uso em dispositivos móveis. Nessa avaliação, foram medidos o desempenho e sobrecarga incorrida pelo Usersense em um *smartphone* Motorola Moto X com SO Android na versão 4.4.4 KitKat. O telefone celular possui um processador Quad-Core 2.5 GHZ e tem 2GB de memória RAM, juntamente com 32GB de armazenamento (microSD).

Esse teste tem por objetivo analisar o impacto global ao implantar a UserSense em um dispositivo

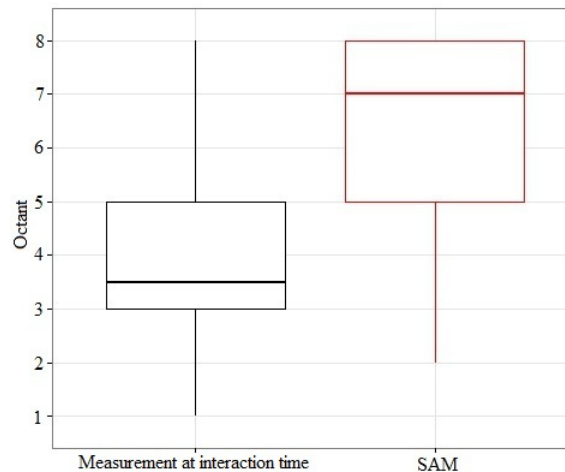


Figura 15 – Resultados em tempo de interação (empregando sensores) e no final do jogo (empregando SAM). A avaliação estatística mostra diferenças significativas.

com recursos escassos. Para isso, processar as mesmas medidas realizadas acima, ou seja, calcular a sobrecarga tomada para carregar e instanciar componentes UserSense e compará-los com os dos objetos Java.

A Tabela 4 mostra uma comparação entre a utilização de componentes UserSense e objeto Java. O tempo medido para carregar um único componente nulo UserSense é de $12\mu s$. Já um único objeto nulo Java, ou seja, sem implementação do método, necessita de $58\mu s$. A grande sobrecarga incorrida pelo Java é atribuído principalmente ao tempo necessário para ler o objeto a partir do disco.

Tabela 4 – Desempenho da Usersense no Motorola Moto X.

Operação	UserSense	Objeto Java
Tempo de Carregamento	$12\mu s$	$58\mu s$
Tempo Instanciação	$452\mu s$	$10\mu s$

No processo de instanciação, um componente nulo UserSense levou $452\mu s$ em relação a $10\mu s$ levados para instanciar um objeto nulo Java. As diferenças na presente medição podem ser atribuídas ao maior tamanho do arquivo, exigido pelos componentes UserSense ao armazenar a informação para os meta dados. Além disso, essas diferenças também podem ser atribuídas ao mecanismo de instanciação complexo, empregado pela UserSense, que inclui o tempo necessário para armazenar a informação do componente, suas interfaces e seus receptáculos no registro.

Todas as informações dos componentes, interfaces e receptáculos manuseadas pelo kernel UserSense são mantidas no registro, em virtude disso, novas instanciações de componentes, interfaces e receptáculos previamente instanciados levam um tempo desprezível em sua instanciação. Nota-se que a sobrecarga da UserSense em relação ao Java não é muito diferente quando a soma da carga, os tempos de instanciação e o mecanismo de registro da UserSense são considerados.

O gráfico da Figura 16 mostra um comparativo entre o uso de memória de componentes UserSense e objeto Java. O uso de memória de um objeto Java nulo é de 39 bytes e de um componente da UserSense sem receptáculo e sem interface é 52 bytes. Assim, verifica-se que o uso de memória deles são praticamente iguais, mesmo que o componente da UserSense seja mais complexo do que um objeto Java nulo. A partir

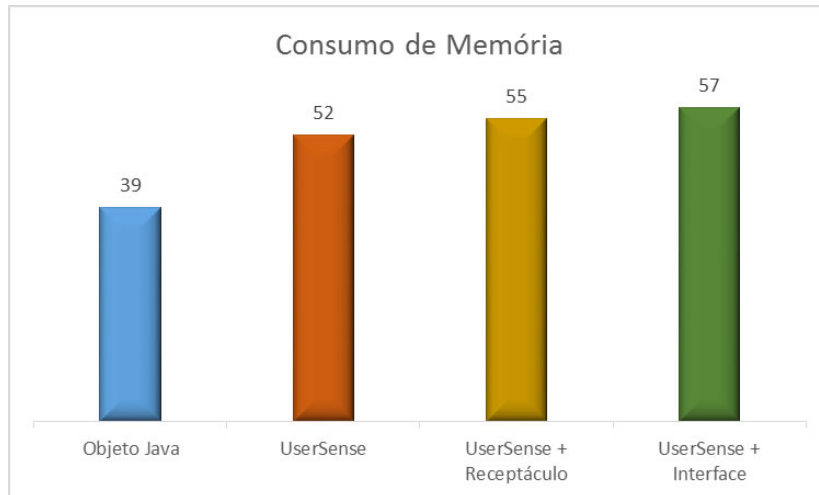


Figura 16 – Consumo de memória da UserSense.

desse gráfico verifica-se que um receptáculo usa aproximadamente 3 bytes e uma interface 5 bytes. Logo, é possível inferir que um componente UserSense com receptáculo ou com interface não consome muita memória se comparado a um objeto Java nulo.

Com o objetivo de atender a diversidade de sensores em uma interação do usuário com um PC, um dispositivo móvel ou integrado, este trabalho desenvolveu a Usersense que, baseada no modelo de componentes do OpenCom (UEYAMA *et al.*, 2009) possibilita o processamento flexível de dados de sensores, permitindo que a aplicação se adeque ao número de sensores disponível e relevantes em tempo real.

6.4 Considerações Finais

Os resultados desse estudo mostram que a análise em tempo real oferece informações mais ricas do que as avaliações que são feitas apenas no final da interação. Tal avaliação do episódio emocional permite um viés indesejável quando se avalia as emoções do usuário. Por exemplo, o usuário pode ser indevidamente influenciado pelos últimos momentos de interação com o jogo. Ou o usuário pode apenas querer mostrar o seu prazer em participar de um experimento exploratório.

Os resultados apresentados aqui não implicam que as emoções coletadas após a interação não podem ser utilizadas para a avaliação de uma solução de *software*. Este estudo defende o uso de uma ampla gama de métodos para a captura de respostas emocionais dos usuários durante o tempo de interação. No próximo Capítulo, serão apresentados detalhes sobre as técnicas de aprendizado de máquina, que também são empregadas nesta Tese, para que se possa inferir as emoções do usuário de maneira mais autônoma e inteligente. Isso envolve a análise dos psicólogos (durante o treinamento) para a UserSense, que por sua vez é capaz de aprender e usar esse conhecimento automaticamente ao avaliar as emoções dos usuários.

AVALIAÇÃO INTELIGENTE E POR SIMILARIDADE DE MÚLTIPLOS SENSORES

Essa seção apresenta a avaliação realizada, empregando a abordagem proposta e apresentada no Capítulo 4. Destacam-se ainda os benefícios da análise de emoções por meio de vários sensores. Posteriormente, explora-se a análise de similaridade para fundamentar as conclusões, particularmente relacionadas às vantagens da utilização de vários sensores em relação ao uso de um único sensor.

7.1 Combinando Aprendizado de Máquina e Múltiplos Sensores para Avaliar as Emoções do Usuário

Nesta seção, serão mostrados os resultados da UserSense, considerando os dados coletados, como já foi mostrado no Capítulo 5. A UserSense utilizou dois algoritmos de aprendizado de máquina distintos - Árvores de Decisão com algoritmo C4.5 (QUINLAN, 1986) e SVMs (VAPNIK, 2013) com parâmetros *default* do ambiente Weka (HALL *et al.*, 2009) - que são métodos de classificação sólidos e, muitas vezes, considerados quando é necessária uma classificação (CLARK; NIBLETT, 1989; CRISTIANINI; SHAWE-TAYLOR, 2000). Como afirmado anteriormente, a seleção dos métodos inteligentes nesta Tese é apenas uma opção de implementação da abordagem UserSense.

No entanto, outros métodos de aprendizado de máquina, também, poderiam ser explorados nessa avaliação, tais como Redes Neurais Artificiais e Redes Bayesianas. Além disso, é importante destacar que nesses experimentos a Árvore de Decisão e a SVM são usadas separadamente e os resultados são apresentados para cada uma delas, individualmente.

Nesse contexto, tem-se muitas instâncias e não se sabe a emoção dos usuários para cada instância. Assim, usa-se o processo de aprendizado supervisionada, que fornece o objeto de entrada e o valor de saída desejado e, em seguida, o algoritmo realiza o treinamento e as respostas são dadas pelos especialistas do domínio (psicólogos).

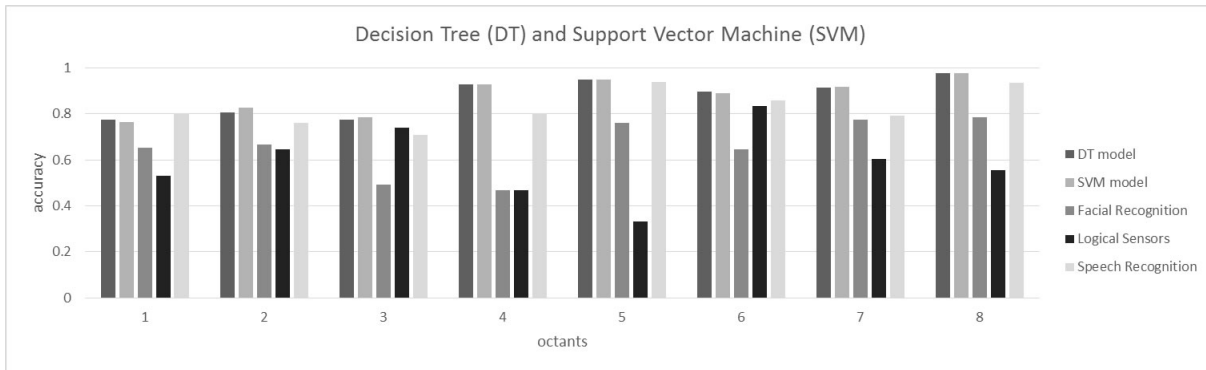


Figura 17 – Acurácia da avaliação das emoções por meio de aprendizado de máquina e sensores individuais.

Usa-se nesta Tese a opinião da maioria dos 3 psicólogos para criar um conjunto de dados de treinamento e ajudar a construir o modelo de classificação. Para um conjunto de vários usuários e dados de treinamento, os 3 psicólogos dizem suas opiniões a respeito da emoção que o usuário estava experimentando em cada conjunto de dados em um determinado momento. Para cada conjunto de dados, se, pelo menos, 2 psicólogos escolheram o mesmo oitante, esse oitante é selecionado como classe do psicólogo no modelo de treinamento.

Para realizar os experimentos, foi utilizada uma técnica de Validação Cruzada com 5 partições e as instâncias foram misturadas independentemente de usuário, para gerar a partição. Optou-se por ter 5 partições em vez de 10, que é mais comumente adotado na literatura sobre o aprendizado de máquina (FACELI *et al.*, 2011), por causa do tamanho do conjunto de dados. Como tem-se um conjunto de dados com tamanho limitado (com 345 casos no total), a utilização de uma Validação Cruzada com 5 partições pode levar a um conjunto de dados de teste mais representativo. Tudo isso, considerando o conjunto de testes para cada execução da validação cruzada.

O gráfico da Figura 17 apresenta a acurácia média da validação cruzada para a UserSense com Árvore de Decisão e SVM, para cada um dos oitantes. Além disso, comparou-se os resultados com os dados diretamente por cada sensor individualmente.

Os resultados mostram que a abordagem UserSense (com as técnicas de aprendizado de máquina selecionada) obtiveram a maior acurácia para sete dos oito oitantes. O oitante 1 foi o único no qual o modelo foi superado por um sensor individual (ou seja, o reconhecimento de fala). No entanto, a diferença na acurácia entre esses sensores e as duas versões do modelo para a UserSense é bem menor em comparação com a diferença de acurácia para os outros oitantes. Isso demonstra um alto grau de consistência dos resultados encontrados.

7.2 Avaliação de Similaridade de Múltiplos Sensores

Para confirmar que o uso de múltiplos sensores na verificação das emoções do usuário é mais eficiente do que o uso de sensores individuais, comparou-se a avaliação de especialistas (especialistas no domínio, psicólogos) com a avaliação por sensores individuais e com a de vários sensores. Para esse fim, fez-se uso de grafos de similaridade. Os grafos são estruturas relacionais que permitem modelar as

relações entre pares de objetos, entidades ou exemplos. De um modo geral, os objetos são representados pelas arestas e as relações entre eles são representados por arestas (NEWMAN, 2003; NEWMAN, 2010).

Um conjunto de dados $X = \{X_1, \dots, X_n\}$ com $X \in \mathbb{R}^m$ pode ser descrito por um grafo $G = (V, E)$, onde V é um conjunto de n vértices e cada vértice $v_i \in V$ corresponde a um ponto $X_i \in X$. Além disso, X_i tem $X_{i,j}$ atributos para cada oitante i (veja mais detalhes no Capítulo 4). Um grafo G também consiste em um conjunto E de m arestas representando similaridade entre vértices.

7.2.1 Construção dos Grafos

No contexto desta pesquisa, para construir os grafos, usa-se os Mk -Vizinhos Mais Próximos ($MkNN$) (CHEN; FANG; SAAD, 2009), com k no intervalo de $k = \{0, \dots, 15\}$. A maioria dos métodos de construção de grafos é baseada em k -Vizinhos Mais Próximos (kNN), que conecta cada ponto de X_i aos seus k -Vizinhos Mais Próximos, isso é, cada vértice está associado a um conjunto de k vértices mais próximos de acordo com uma medida de similaridade, por exemplo a distância euclidiana.

Além do kNN , existe também os grafos kNN Mutual ($MkNN$) em que há uma conexão entre dois vértices somente se a regra da vizinhança é cumprida por ambos os vértices (JEBARA; WANG; CHANG, 2009; MAIER; LUXBURG; HEIN, 2009; CHEN; FANG; SAAD, 2009). O kNN tem um grau de complexidade $O(dn^2)$, onde d é o número de dimensões de dados e n o número de pontos de dados. Similar $MkNN$ é quadrática também, mas dada uma estrutura de dados especiais é possível computar o grafo $MkNN$ em quase grau linear (CHEN; FANG; SAAD, 2009).

Nesta Tese, usa-se o método $MkNN$ e o parâmetro k foi fixado em um intervalo de $k = \{0, \dots, 15\}$. Considerando 8 oitantes O_t e as instâncias $X_i = \{X_{i,1}, X_{i,2}, X_{i,3}, X_{i,4}, X_{i,5}, X_{i,6}, X_{i,7}, X_{i,8}\}$, onde $X_{i,j}$ é 1 se os oitantes t pertencem a ele e 0 contrário. Por exemplo, suponha que o objeto X_i pertence aos oitantes 1 e 2, então $X_i = \{1, 0, 1, 0, 0, 0, 0, 0\}$, ou $X_i = \{0.5, 0, 0.5, 0, 0, 0, 0, 0\}$ quando utilizar valores normalizados (veja mais detalhes no Capítulo 4). Além disso, para representar a composição de sensores, esta Tese combinou e normalizou as instâncias. Por exemplo, considerando $X_i^{S_I} = \{0.5, 0, 0.5, 0, 0, 0, 0, 0\}$ é inferida a partir do Sensor I e $X_i^{S_{II}} = \{0, 0, 0, 0, 0, 0, 0.5, 0.5\}$ é inferida a partir do Sensor II. A composição dos Sensores I e II é $X_i^{S_I+S_{II}} = \{0.25, 0, 0.25, 0, 0, 0, 0.25, 0.25\}$. Em seguida, a distância euclidiana (ED) (Eq. 7.1) foi usada para calcular a similaridade entre X_i e X_j .

$$ED(X_i^{S_I}, X_j^{S_{II}}) = \sqrt{\sum_{t=1}^8 (X_{it}^{S_I} - X_{jt}^{S_{II}})^2} \quad (7.1)$$

Com base nesse conceito, o grafo de similaridade foi construído para o Sensor I (S_I), Sensor II (S_{II}), Sensor III (S_{III}), Sensor I e II (S_{I+II}), Sensor I e III (S_{I+III}), Sensor II e III (S_{II+III}), Sensor I, II e III ($S_{I+II+III}$). Nessa análise de similaridade, também foram incluídos os grafos produzidos para a avaliação dos especialistas do domínio. A Figura 18 ilustra alguns grafos que foram gerados a partir da avaliação dos psicólogos, onde é possível perceber que quanto maior o k , maior é o número de conexões. Assim, comparou-se a similaridade entre cada um grafo obtido por meio dos sensores (ou conjunto de sensores) com o grafo construído para a avaliação dos especialistas.

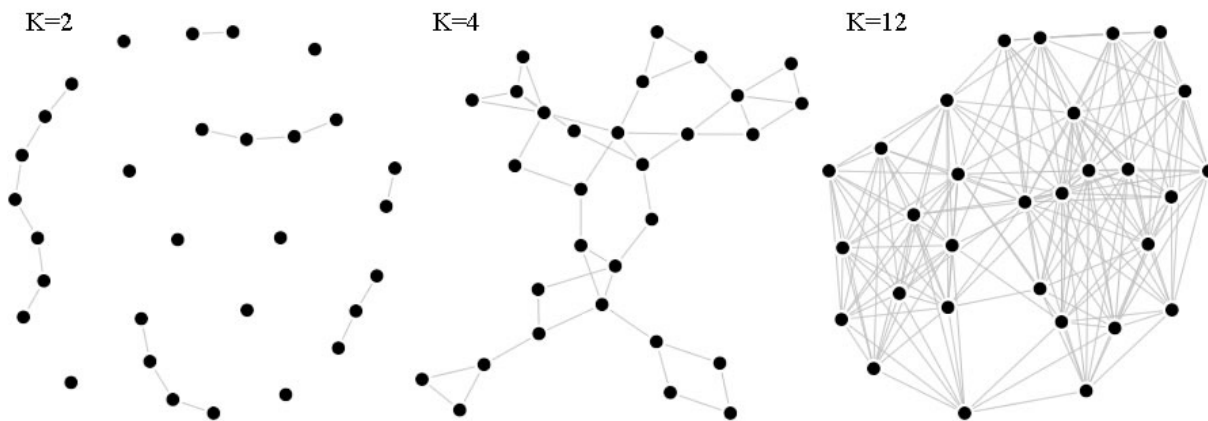


Figura 18 – Grafos gerados a partir da avaliação dos psicólogos.

7.2.2 Similaridade de Grafos

Na sequência da fase de construção dos grafos, pode-se agora comparar os grafos dos sensores com os grafos dos especialistas, por meio de uma análise de similaridade utilizando o Graph Edit Distance (GED). GED é a ferramenta mais tradicional e amplamente adotada para medir a similaridade de grafos (GAO *et al.*, 2010).

A ideia básica do GED é quantificar a dissimilaridade de dois grafos pelo número mínimo de operações (adição ou exclusão de vértices e arestas) para converter um dado grafo em outro. Quanto mais dissimilares dois gráficos possuem, mais operações devem ser executadas (GAO *et al.*, 2010). Por exemplo, considerando a Figura 19, para converter G_1 em G_2 , ele será obrigado a excluir uma aresta; adicionar um vértice; e incluir uma aresta. Dado que os custos das operações de edição (exclusão/inserção) são 1 e os custos das operações de substituição são 0, então, isso resulta em $GED(G_1, G_2) = 3$.

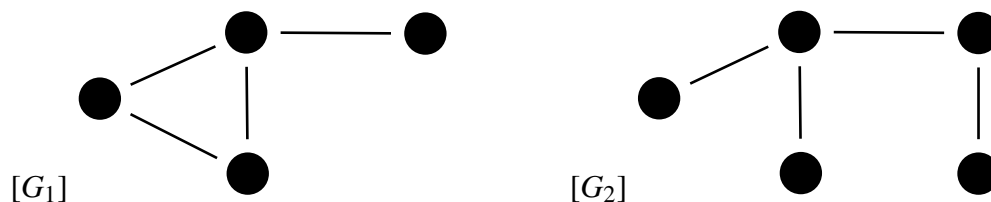


Figura 19 – A distância de edição de grafos entre G_1 e G_2 é 3 operações de edição, ou seja, elimina-se uma aresta; insere-se um vértice e insere-se uma aresta.

7.2.3 Análise dos Resultados

É importante ressaltar que esta Tese usa o conjunto de dados apresentado no Capítulo 5 para realizar essa análise de similaridade. As Figuras 20 e 21 comparam a similaridade entre cada um dos sensores e a composição de múltiplos sensores com os resultados dos especialistas. Isto significa que os grafos obtidos por cada sensor ou composição deles é comparado com o grafo obtido por especialistas que dependem da GED.

O gráfico da Figura 20 mostra o desempenho de cada sensor individualmente (S_I , S_{II} , S_{III}). Do mesmo modo, o gráfico da Figura 21 mostra o desempenho de cada sensor individual. Ele também mostra

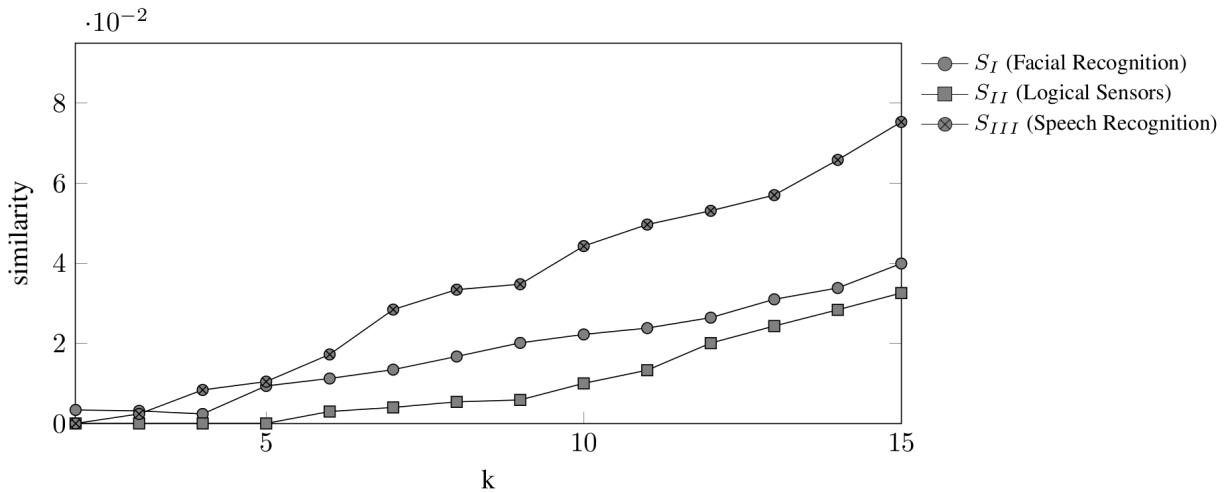


Figura 20 – Similaridade entre os gráficos obtidos por cada sensor individualmente (S_I , S_{II} , S_{III}) e o grafo gerado pela avaliação dos especialistas.

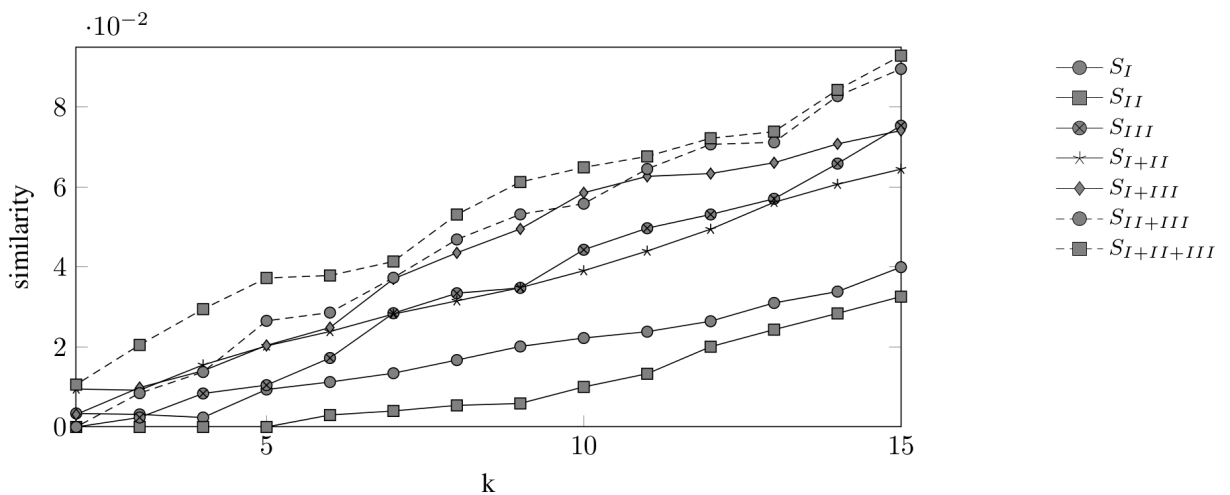


Figura 21 – Similaridade entre os gráficos obtidos por cada sensor individualmente (S_I , S_{II} , S_{III}) ou pela composição dos sensores (S_{I+II} , S_{I+III} , S_{II+III} , $S_{I+II+III}$) e o grafo gerado pela avaliação dos especialistas.

a composição de dois sensores (S_{I+II} , S_{I+III} , S_{II+III}) e todos os sensores ($S_{I+II+III}$). Na Figura 20, observa-se que o sensor S_{III} obteve o melhor desempenho independentemente da estrutura do grafo (k -Vizinhos Mais Próximos). Ao contrário disso, os sensores S_I e S_{II} trouxeram resultados que são próximos uns dos outros.

A partir do gráfico da Figura 21, é possível notar que a composição de todos os sensores obtém um melhor desempenho (similaridade) do que a de apenas um único sensor. Particularmente, a composição dos três sensores ($S_{I+II+III}$) obtém o melhor desempenho do que qualquer outra estrutura de grafo.

Nota-se que os resultados de vários sensores, particularmente a composição de todos os sensores, é o mais próximo da análise feita pelos especialistas no domínio (psicólogos). Assim, pode-se concluir que a utilização de vários sensores é mais eficaz do que a utilização de um único sensor para verificar as emoções do usuário.

7.3 Considerações finais

A partir da condução dos experimentos foi possível constatar que os diferentes valores de entrada dos múltiplos sensores trazem uma riqueza de detalhes que não é possível em outras abordagens que utilizam apenas um sensor.

Por exemplo, em um experimento exploratório em que quatro sensores são utilizados e o usuário joga um jogo eletrônico, um sensor lógico (que pode determinar tendências comportamentais) verifica que o jogador perdeu duas vezes num intervalo de tempo de 40 segundos; por causa disso, o jogador declara estar frustrado (avaliação cognitiva). O sensor de frequência cardíaca (que analisa reações fisiológicas) mostra um aumento de 3.1bpm. O sensor que reconhece emoções faciais (que determina expressões motoras), por sua vez, sugere que o usuário está triste. De posse desses dados, há indícios de que, nesta situação, o jogador está com uma Valência negativa, e experimenta uma Excitação média, com um baixo Sentimento de controle e Dificuldade de conclusão do objetivo. Essa conclusão não poderia ser obtida caso fosse utilizada apenas uma avaliação para um único aspecto emocional com um único tipo de sensor.

Segundo (SCHERER, 2005; GONÇALVES *et al.*, 2016; NORMAN, 2004; FONTAINE; SCHERER; SORIANO, 2013; RODRIGUES *et al.*, 2016; FRIJDA; SCHERER *et al.*, 2009), para uma avaliação emocional ideal é preciso avaliar: (1) as contínuas mudanças nos processos de avaliação cognitiva, (2) os padrões de respostas geradas nos sistemas nervoso neuroendócrino, autonômico e somático, (3) as mudanças motivacionais produzidas pelo resultado da avaliação de tendências de ação particulares, (4) os padrões de expressões facial e vocal, bem como os movimentos do corpo e, (5) a natureza do estado dos sentimentos subjetivos experimentados que refletem todas essas mudanças dos componentes. Dessa maneira, acredita-se que as avaliações mais abrangentes de respostas emocionais pode ser realizada por uma combinação de instrumentos (por exemplo, diferentes sensores) associadas a diferentes componentes emocionais.

CONCLUSÕES

Na área de IHC, inúmeros são os métodos, técnicas e instrumentos que apoiam a avaliação de respostas emocionais. Todavia, como a maioria dessas medidas não é aplicada de maneira automatizada, apesar de serem muito eficientes, essas abordagens não ocorrem em tempo de interação.

Logo, para que os sistemas computacionais interativos possam reagir ao estado emocional de um indivíduo é necessário se criar uma infraestrutura (por exemplo, o auxílio de sensores e métodos de classificação) que apoie a captura, análise e classificação de indícios de respostas emocionais de usuários, visando permitir a flexibilização das aplicações computacionais a partir de tal classificação.

8.1 Síntese das Contribuições

Nesta pesquisa, é apresentada uma abordagem para o reconhecimento de emoções baseada nos componentes emocionais em tempo real. Para a avaliação de cada um desses componentes, utilizou-se sensores e algoritmos de aprendizado de máquina. As saídas desses sensores são combinadas e consideradas na sua classificação, os quatro domínios responsáveis pela diferenciação das emoções, a saber, Valência, Excitação, Sentimento de controle e Facilidade de conclusão do objetivo. Entre as principais contribuições deste projeto de doutorado, pode-se ressaltar:

- Esta pesquisa permitiu o desenvolvimento e a validação de uma abordagem apoiada na captura, análise e classificação de indícios de respostas emocionais de usuários em tempo real. A utilização da abordagem UserSense, proposta nesta Tese, oferece uma maneira flexível de selecionar quais sensores e algoritmos de aprendizado de máquina podem ser utilizados na avaliação dos diferentes componentes da emoção e permite, conseqüentemente, analisar e classificar os resultados obtidos a partir desses sensores. Tal abordagem apresenta um modelo de classificação para cada oitante da Estrutura do Espaço Emocional Semântico, ou seja, tem-se 8 modelos para fazer a predição das emoções que usam os dados provenientes dos sensores como entrada e consideram a avaliação dos especialistas do domínio na definição das classes;

- A escolha dos sensores não foi feita de maneira arbitrária. Ao contrário, a seleção está diretamente ligada aos quatro componentes da emoção de tempo real e permite avaliar os diferentes aspectos emocionais, a saber, avaliações cognitivas (reconhecimento de fala), tendências comportamentais (sensores lógicos), reações fisiológicas (frequência cardíaca) e expressões motoras (reconhecimento de face). O uso de sensores possibilita avaliar a experiência emocional do usuário a partir de uma visão holística, que visa reduzir o viés das avaliações unilaterais. A partir da análise de similaridade apresentada nesta Tese, foi possível indicar que a utilização de mais sensores na avaliação das emoções é mais eficiente do que usar sensores isoladamente. Por esse motivo, a abordagem UserSense traz uma avaliação mais abrangente das respostas emocionais provocadas nos usuários durante a interação com os sistemas computacionais;
- Durante o desenvolvimento desta pesquisa, ocorreu a aproximação com usuários reais para verificar a viabilidade da proposta e aprimorar a abordagem de avaliação das emoções de maneira automatizada. Nesse sentido, a participação dos usuários contribuiu com a inclusão social e digital, pois o estudo das emoções é uma poderosa ferramenta para melhorar a maneira como as soluções computacionais são apresentadas à diversidade populacional. No mais, esta pesquisa também contou com o apoio de psicólogos que participaram diretamente da condução dos estudos de caso e validação da proposta. Diante do exposto, esta Tese contribui para o “Acesso Participativo e Universal do Cidadão Brasileiro ao Conhecimento” (BARANAUSKAS; SOUZA, 2006), o quarto Desafio da Pesquisa em Computação no Brasil;
- A abordagem proposta considera um microkernel, cujo tempo de execução da UserSense é mantido em seu mínimo para garantir a implantação em diversos ambientes, tais como PCs, *smartphones*, sensores e *tablets*. O microkernel UserSense incorpora apenas funções para carregar e descarregar novos recursos necessários. Por exemplo, no momento ainda não tem-se o “Eletroencefalograma”, no entanto, ele poderá ser implantado assim que estiver disponível e for necessário para o contexto de uso. Da mesma maneira, se o usuário não estiver usando um *smartwatch* então o *software* do sensor de frequência cardíaca, ele não será carregado até que seja necessário. Assim, a possibilidade de poupar recursos torna-se uma vantagem chave, particularmente em dispositivos de recursos limitados, tais como *smartphones* e sensores;
- A abordagem UserSense auxilia na captura e classificação de dados para que os sistemas computacionais interativos possam reagir ao estado emocional de um indivíduo. Portanto, a avaliação emocional dos usuários é disponibilizada para que as aplicações cientes de contexto possam, por exemplo, influenciar na tomada de decisão e na mudança de comportamento dos usuários. Pois, uma abordagem de captura por sensores em tempo de interação pode ser uma maneira de auxiliar nessa interação flexível. A UserSense permite avaliar o impacto das emoções na experiência emocional dos usuários de maneira automatizada. Contribuindo, então, com o estado da arte da área de Design Emocional;
- Apesar do objetivo principal desta pesquisa não ser a avaliação dos episódios emocionais (componente do sentimento subjetivo), foi conduzido um estudo secundário sobre o tema. Ele mostrou que há uma riqueza de detalhes e distinção de informações provenientes de abordagens que avaliam

as emoções em tempo de interação se comparadas a métodos e técnicas que fazem a análise da experiência emocional considerando um episódio emocional do indivíduo (*feedback*).

8.2 Limitações e Trabalhos Futuros

A abordagem UserSense auxilia na classificação da experiência emocional do usuário e, consequentemente, deixa as aplicações computacionais cientes do estado emocional do usuário. A flexibilização dos sistemas será considerada nos trabalhos futuros. No mais, essas informações emocionais poderão ser utilizadas em um processo colaborativo de mapeamento flexível de Personas, por exemplo idosos, voltado para aplicações de Saúde em Casas Inteligentes através de tecnologias de Internet das Coisas.

Devido à complexidade do estudo e a dificuldade de se avaliar as emoções, esta pesquisa limitou-se apenas à aplicação de um tipo de estudo de caso (jogos eletrônicos) para demonstrar a viabilidade da abordagem e aprimoramento da captura, análise e classificação automatizadas da experiência emocional de usuários em tempo real. Assim, entre os trabalhos futuros a serem desenvolvidos, destaca-se a realização de um estudo para comparar a aplicação e avaliação da abordagem proposta em diferentes contextos de avaliação emocional disponíveis na literatura.

Os algoritmos de classificação foram utilizados com seus parâmetros padrão. Um processo de ajuste de parâmetros (*tunning*) para as especificidades do problema tratado poderia promover modelos com resultados ainda melhores do que os obtidos. Como trabalho futuro, pretende-se ajustar os parâmetros dos algoritmos de classificação.

Observou-se também que, durante os experimentos, alguns resultados obtidos através de um sensor, por exemplo, o reconhecimento de fala, e outro sensor, por exemplo, o reconhecimento de face foram divergentes. Isso sugere que a avaliação dos domínios da experiência emocional, por ambos os sensores, é importante para a obtenção de uma visão mais abrangente da avaliação, uma vez que não há como classificar uma resposta emocional como certa ou errada.

Os usuários nem sempre conseguem expressar verbalmente, nomear, ou escolher algo que represente suas emoções. Nesse sentido, em alguns momentos da avaliação, o sensor de reconhecimento de fala que auxiliou a avaliação cognitiva por meio do método Think Aloud poderia ter sido substituído por um EEG. Em trabalhos futuros, pretende-se incluir novos sensores em diferentes contextos, pois, ter uma abordagem flexível (como é o caso da UserSense) que fornece uma maneira alternativa de se avaliar as emoções e seu impacto na experiência emocional, possibilita uma avaliação mais abrangente, amenizando problemas encontrados por esse tipo de avaliação.

Outrossim, realizar estudos de caso com usuários, muitas vezes é uma tarefa hercúlea, pois nem sempre se consegue muitos voluntários durante a etapa de recrutamento. Esse fato pode limitar a quantidade de instâncias para os experimentos e, por conseguinte, dificultar a avaliação emocional ao trazer resultados pouco relevantes.

8.3 Publicações e Trabalhos Desenvolvidos

Este Doutorado além de trazer uma proposta interdisciplinar, contou com o apoio de um Grupo de Pesquisa em Psicologia da Universidade Federal de São Carlos (UFSCar) que apoiou a condução dos estudos e experimentos com usuários. Nesse sentido, considerando o objetivo desta pesquisa e o processo de formação no Doutorado algumas publicações e trabalhos foram desenvolvidos.

Abaixo seguem as publicações e trabalhos desenvolvidos durante o período de pesquisa no Doutorado. Os itens 1 e 2 referem-se a dois artigos de Journal, que apresentam resultados dos estudos e experimentos sobre a abordagem para captura, análise e classificação automatizadas da experiência emocional de usuários apresentada nesta Tese.

O item 3 refere-se a um artigo de Journal, fruto da pesquisa desenvolvida durante o período de Doutorado Sanduíche na Universidade do Arizona (Estados Unidos), sob a orientação da Prof^ª. Dr^ª. Thienne Johnson.

O item 4 refere-se a um artigo de Journal, que apresenta resultados da pesquisa desenvolvida no projeto de extensão descrito no item 9. Como se tratava de uma pesquisa envolvendo diferentes projetos de doutorado, optou-se por colocar os autores em ordem alfabética.

O item 5 refere-se a um artigo de Journal, que traz resultados de uma pesquisa sobre sensores, desenvolvida no grupo de pesquisa Wireless Networking and Energy Systems (WINES) coordenado pelo Prof. Dr. Jó Ueyama.

O item 6 refere-se a um capítulo de livro, fruto da pesquisa de Mestrado desenvolvida pelo autor desta monografia, acrescida de aspectos emocionais abordados nesta tese.

O item 7 refere-se a um artigo premiado como Best Paper, que apresenta resultados da pesquisa desenvolvida no projeto de extensão descrito no item 9. Como tratava-se de uma pesquisa envolvendo diferentes projetos de doutorado, optou-se por colocar os autores em ordem alfabética.

O item 8 refere-se a um artigo, que traz os resultados de uma revisão sistemática sobre o tema de pesquisa desta Tese.

Por fim, o item 9 refere-se a um Projeto de Extensão desenvolvido durante o doutorado, por meio de uma parceria entre a USP, UFSCar e o Hospital Espírita de Marília. Na sequência, são apresentadas as referências desses trabalhos:

1. **Journal 1:** GONÇALVES, V.P.; et al. (2016) Assessing Users' Emotion at Interaction Time: A Multimodal Approach with Multiple Sensors. *Soft Computing Journal (SOCO)*, v. 1, p.1-15.
2. **Journal 2:** GONÇALVES, V.P.; et al. (2016) Enhancing Intelligence in Multimodal Emotion Assessments. *Applied Intelligence Journal (APIN)*. Submitted.
3. **Journal 3:** GONÇALVES, V.P.; et al. (2015) Providing Adaptive Smartphone Interfaces Targeted at Elderly People: An Approach that Takes Into Account Diversity Among the Elderly. *Universal Access in the Information Society Journal (UAIS)*. v. 1, p.1-21.

4. **Journal 4:** RODRIGUES, K.R.H.; GONÇALVES, V.P.; et al. (2015) Personas-Driven Design for Mental Health Therapeutic Applications. SBC Journal on 3D Interactive Systems (JIS), v. 6, p.18-34.
5. **Journal 5:** ROCHA FILHO, G.P; GONÇALVES, V.P.; et al. (2014) NodePM: A Remote Monitoring Alert System for Energy Consumption Using Probabilistic Techniques. Sensors Journal. v. 14, n. 1, p.848-867.
6. **Capítulo de Livro:** RODRIGUES, K.R.H.; GONÇALVES, V.P.; et al. (2016) Tecnologias de Informação e Comunicação Inclusivas e o Envelhecer (Cap.). In: Promoção da Saúde para um Envelhecimento Saudável. Coleção Envelhecimento: Saberes e Vivências. Campos, A.C.V; Berlezi, E.M.; Corrêa, A.H.M. (Org.). 1ed. Ijuí-RS: Unijuí, v. 4, p.265-288.
7. **Paper Premiado como Best Paper:** RODRIGUES, K.R.H.; GONÇALVES, V.P.; et al. (2014) Enriquecimento de Personas para Apoio ao Design de Aplicações Terapêuticas para a Saúde Mental. In: XIII Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais. Foz do Iguaçu. Anais do IHC 2014. v. 1. p.51-60.
8. **Paper 2:** GONÇALVES, V.P.; et al. (2013) Emoções na Interação Humano-Computador: Um Estudo Considerando Sensores. In: XII Simpósio de Fatores Humanos em Sistemas Computacionais. Manaus. Anais do IHC 2013. p.252-255.
9. **Projeto de Extensão:** participação na elaboração e execução de um Projeto de Extensão em parceria com o Hospital Espírita de Marília: “Design Participativo de Aplicações Terapêuticas com Jovens Dependentes Químicos e Pacientes Depressivos”.

8.4 Considerações Finais

Esta Tese defende o uso de uma ampla gama de sensores para avaliar as emoções dos usuários. Acredita-se que somente uma visão holística das emoções pode permitir a obtenção de conclusões significativas sobre os estados emocionais dos usuários. Este trabalho proporcionou uma experiência real, analisando emoções por meio do estudo de caso com usuários interagindo com um jogo eletrônico. Os dados relativos às respostas emocionais - como a fala, o desempenho, as expressões faciais e a frequência cardíaca - foram coletados por sensores durante a interação do usuário e estes foram usados como entrada para construir um modelo de aprendizado de máquina.

Os resultados da avaliação da abordagem UserSense têm demonstrado que a adoção de múltiplos sensores para análise das emoções do usuário é mais eficiente do que a utilização de um sensor individual. Também foi possível constatar a mesma conclusão a partir da análise de similaridade de grafos. A abordagem apresentada nesta Tese permite o desenvolvimento de aplicações inovadoras e pode ser usada como base em projetos de pesquisa nas áreas de IHC, Sensores e Psicologia. É possível também que a UserSense sirva de ponto de partida para novos projetos que melhorem e estendam o suporte do Design Emocional à interação multimodal. Este trabalho traz avanços significativos e contribuições para aplicações computacionais, enfatizando os aspectos emocionais nas decisões de design através de uma visão abrangente da experiência emocional do usuário.

REFERÊNCIAS

- ANDERSON-COOK, C. M. An introduction to multivariate statistical analysis. **Journal of the American Statistical Association**, American Statistical Association, v. 99, n. 467, p. 907–909, 2004. Citado na página 41.
- BAIENSON, 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 36, 37, 48, 49 e 50.
- BARANAUSKAS, M. C. C.; SOUZA, C. d. Desafio 4: Acesso participativo e universal do cidadão brasileiro ao conhecimento. **Computação Brasil**, v. 7, n. 23, p. 7, 2006. Citado na página 82.
- BELL, C. **Essays on the Anatomy and Philosophy of Expression**. [S.l.]: J. Murray, 1824. Citado na página 20.
- BENTLEY, T.; JOHNSTON, L.; BAGGO, K. von. Evaluation using cued-recall debrief to elicit information about a user's affective experiences. In: COMPUTER-HUMAN INTERACTION SPECIAL INTEREST GROUP (CHISIG) OF AUSTRALIA. **Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future**. [S.l.], 2005. p. 1–10. Citado na página 27.
- BERCHT, M. Em direção a agentes pedagógicos com dimensões afetivas. **Instituto de Informática. UFRGS. Tese de Doutorado. Porto Alegre**, 2001. Citado na página 26.
- BOEHNER, K.; DEPAULA, R.; DOURISH, P.; SENEGERS, P. Affect: from information to interaction. In: ACM. **Proceedings of the 4th decennial conference on Critical computing: between sense and sensibility**. [S.l.], 2005. p. 59–68. Citado na página 26.
- BRADLEY, M. M.; LANG, P. J. Measuring emotion: The self-assessment manikin and the semantic differential. **Journal of Behavior Therapy & Experimental Psychiatry**, v. 25, n. 1, p. 49–59, 1994. Citado 4 vezes nas páginas 11, 20, 67 e 68.
- BRAVE, S.; NASS, C. The human-computer interaction handbook. In: JACKO, J. A.; SEARS, A. (Ed.). Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 2003. cap. Emotion in Human-computer Interaction, p. 81–96. ISBN 0-8058-3838-4. Citado 5 vezes nas páginas 19, 20, 26, 27 e 28.
- BREIMAN, L.; FRIEDMAN, J.; STONE, C. J.; OLSHEN, R. A. **Classification and regression trees**. [S.l.]: CRC press, 1984. Citado na página 47.
- BUSSO, C.; LEE, S.; NARAYANAN, S. Analysis of emotionally salient aspects of fundamental frequency for emotion detection. **Audio, Speech, and Language Processing, IEEE Transactions on**, IEEE, v. 17, n. 4, p. 582–596, 2009. Citado na página 38.
- CAMPBELL, A.; CHOUDHURY, T. From smart to cognitive phones. **IEEE Pervasive Computing**, IEEE, n. 3, p. 7–11, 2012. Citado na página 19.
- CARVALHO, A. C. de; FREITAS, A. A. A tutorial on multi-label classification techniques. In: **Foundations of Computational Intelligence Volume 5**. [S.l.]: Springer, 2009. p. 177–195. Citado na página 60.

- CARVALHO, S.; LEITE, J.; GALDO-ÁLVAREZ, S.; GONÇALVES, Ó. F. The emotional movie database (emdb): a self-report and psychophysiological study. **Applied psychophysiology and biofeedback**, Springer, v. 37, n. 4, p. 279–294, 2012. Citado na página 58.
- 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 5 vezes nas páginas 36, 37, 41, 49 e 50.
- CHEN, J.; FANG, H.-r.; SAAD, Y. Fast approximate knn graph construction for high dimensional data via recursive lanczos bisection. **J. Mach. Learn. Res.**, JMLR.org, v. 10, p. 1989–2012, dez. 2009. ISSN 1532-4435. Citado na página 77.
- CLARK, P.; NIBLETT, T. The cn2 induction algorithm. **Machine learning**, Springer, v. 3, n. 4, p. 261–283, 1989. Citado 4 vezes nas páginas 36, 37, 47 e 75.
- COHEN, J. *et al.* A coefficient of agreement for nominal scales. **Educational and psychological measurement**, Durham, v. 20, n. 1, p. 37–46, 1960. Citado na página 56.
- COULSON, G.; BLAIR, G.; GRACE, P.; TAIANI, F.; JOOLIA, A.; LEE, K.; UEYAMA, J.; SIVAHARAN, T. A generic component model for building systems software. **ACM Transactions on Computer Systems (TOCS)**, ACM, v. 26, n. 1, p. 1, 2008. Citado na página 62.
- CRISTIANINI, N.; SHAWE-TAYLOR, J. **An introduction to support vector machines and other kernel-based learning methods**. [S.l.]: Cambridge university press, 2000. Citado 3 vezes nas páginas 36, 46 e 75.
- DAMÁSIO, A. Ao encontro de espinosa, trad. **PEA., Pub. Europa-América, Lisboa**, 2003. Citado na página 26.
- DAMASIO, A. R.; MARG, E. Descartes' error: Emotion, reason, and the human brain. **Optometry and Vision Science**, [Baltimore, MD: Published for the Academy by Williams & Wilkins, c1989-, v. 72, n. 11, p. 847–847, 1995. Citado na página 26.
- DARWIN, C. The expression of the emotions in man and animals. **London, UK: John Marry**, 1872. Citado 2 vezes nas páginas 20 e 26.
- DESMET, P. Funology. In: BLYTHE, M. A.; OVERBEEKE, K.; MONK, A. F.; WRIGHT, P. C. (Ed.). Norwell, MA, USA: Kluwer Academic Publishers, 2004. cap. Measuring Emotion: Development and Application of an Instrument to Measure Emotional Responses to Products, p. 111–123. ISBN 1-4020-2966-7. Citado 4 vezes nas páginas 31, 35, 55 e 67.
- DUDA, R. O.; HART, P. E. *et al.* **Pattern classification and scene analysis**. [S.l.]: Wiley New York, 1973. v. 3. Citado 2 vezes nas páginas 36 e 47.
- EKMAN, P.; FRIESEN, W. V. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. **Semiotica**, v. 1, n. 1, p. 49–98, 1969. Citado 2 vezes nas páginas 28 e 29.
- ELLSWORTH, P. C.; SCHERER, K. R. Appraisal processes in emotion. **Handbook of affective sciences**, v. 572, p. V595, 2003. Citado 6 vezes nas páginas 26, 28, 29, 30, 33 e 43.
- ERICSSON, K.; SIMON, H. **Protocol analysis; Verbal reports as data (revised edition)**. **Bradfordbooks**. [S.l.]: MIT Press, Cambridge, MA, 1993. Citado 2 vezes nas páginas 56 e 57.
- ERRO, D.; NAVAS, E.; HERNÁEZ, I.; SARATXAGA, I. Emotion conversion based on prosodic unit selection. **Audio, Speech, and Language Processing, IEEE Transactions on**, IEEE, v. 18, n. 5, p. 974–983, 2010. Citado na página 38.

FACELI, K.; LORENA, A.; GAMA, J.; CARVALHO, A. Inteligência artificial: Uma abordagem de aprendizagem de máquina, chapter classificação multirrótulo. **LCT. Citado nas**, p. 25, 2011. Citado 3 vezes nas páginas 36, 37 e 76.

FILHO, G. P. R.; UEYAMA, J.; VILLAS, L. A.; PINTO, A. R.; GONÇALVES, V. P.; PESSIN, G.; PAZZI, R. W.; BRAUN, T. Nodepm: A remote monitoring alert system for energy consumption using probabilistic techniques. **Sensors**, v. 14, n. 1, p. 848–867, 2014. ISSN 1424-8220. Citado 2 vezes nas páginas 36 e 54.

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 4 vezes nas páginas 28, 29, 30 e 56.

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 30.

FONTAINE, J. R.; SCHERER, K. R.; SORIANO, C. **Components of emotional meaning: A source-book**. [S.l.]: OUP Oxford, 2013. Citado 2 vezes nas páginas 26 e 80.

FRANTZIDIS, C. A.; BRATSAS, C.; KLADOS, M. A.; KONSTANTINIDIS, E.; LITHARI, C. D.; VIVAS, A. B.; PAPADELIS, C. L.; KALDOUDI, E.; PAPPAS, C.; BAMIDIS, P. D. On the classification of emotional biosignals evoked while viewing affective pictures: an integrated data-mining-based approach for healthcare applications. **Information Technology in Biomedicine, IEEE Transactions on**, IEEE, v. 14, n. 2, p. 309–318, 2010. Citado 4 vezes nas páginas 37, 45, 46 e 49.

FRIJDA, N. H. Emotion, cognitive structure, and action tendency. **Cognition and emotion**, Taylor & Francis, v. 1, n. 2, p. 115–143, 1987. Citado na página 28.

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 28.

FRIJDA, N. H. *et al.* The psychologists' point of view. **Handbook of emotions**, Guilford New York, v. 2, p. 59–74, 2000. Citado na página 28.

FRIJDA, N. H.; SCHERER, K. R. *et al.* Emotion definitions (psychological perspectives). Oxford University Press, 2009. Citado 3 vezes nas páginas 26, 71 e 80.

FRIJDA, N. H. *et al.* What might emotions be? comments on the comments. **Information sur les Sciences Sociales**, v. 46, p. 433–443, 2007. Citado 3 vezes nas páginas 26, 27 e 28.

GAO, X.; XIAO, B.; TAO, D.; LI, X. A survey of graph edit distance. **Pattern Analysis and Applications**, Springer-Verlag, v. 13, n. 1, p. 113–129, 2010. Citado na página 78.

GAVER, W. Designing for emotion (among other things). **Philosophical Transactions of the Royal Society of London B: Biological Sciences**, The Royal Society, v. 364, n. 1535, p. 3597–3604, 2009. Citado na página 26.

GELLER, T. How do you feel?: Your computer knows. In: **Communications of the ACM**. [S.l.]: ACM, 2014. v. 57, n. 1, p. 24–26. ISSN 10.1145/2555809. Citado na página 20.

GONÇALVES, V. P.; GIANCRISTOFARO, G. T.; FILHO, P. G.; JOHNSON, T.; CARVALHO, V.; PESSIN, G.; NERIS, V. P. de A.; UEYAMA, J. Assessing users' emotion at interaction time: a multimodal approach with multiple sensors. **Soft Computing**, Springer, p. 1–15, 2016. Citado 11 vezes nas páginas 19, 20, 21, 26, 27, 31, 33, 36, 50, 71 e 80.

- GONÇALVES, V. P.; NERIS, V. P.; UEYAMA, J.; SERAPHINI, S. An analitic approach to evaluate flexible mobile user interfaces for the elderly. In: **Proceedings of the 14th International Conference on Enterprise Information System**. [S.l.: s.n.], 2012. v. 3, p. 91–96. Citado na página 20.
- GONCALVES, V. P.; NERIS, V. P.; UEYAMA, J.; SERAPHINI, S.; DIAS, T. C.; FILHO, G. P. R. Senior citizens in interaction with mobile phones: A flexible middleware approach to support the diversity. In: **The 11th International Conference on Software Engineering Research and Practice**. [S.l.: s.n.], 2013. p. 315–321. Citado na página 19.
- GONCALVES, V. P.; NERIS, V. P. A.; UEYAMA, J.; GIANCRISTOFARO, G. T.; FILHO, G. P. R.; SERAPHINI, S. Emoções na interação humano-computador: um estudo considerando sensores. In: **Proceedings of the 12th Brazilian Symposium on Human Factors in Computing Systems**. Porto Alegre, Brazil, Brazil: Brazilian Computer Society, 2013. (IHC '13), p. 252–255. ISBN 978-85-7669-278-2. Citado 3 vezes nas páginas 33, 35 e 36.
- GONCALVES, V. P.; NERIS, V. P. D. A.; UEYAMA, J. Interacao de idosos com celulares: Flexibilidade para atender a diversidade. In: **Proceedings of the 10th Brazilian Symposium on on Human Factors in Computing Systems and the 5th Latin American Conference on Human-Computer Interaction**. Porto Alegre, Brazil, Brazil: Brazilian Computer Society, 2011. (IHC+CLIHC '11), p. 343–352. ISBN 978-85-7669-257-7. Citado na página 19.
- GONÇALVES, V. P.; NERIS, V. P. de A.; SERAPHINI, S.; DIAS, T. C.; PESSIN, G.; JOHNSON, T.; UEYAMA, J. Providing adaptive smartphone interfaces targeted at elderly people: an approach that takes into account diversity among the elderly. **Universal Access in the Information Society**, Springer, p. 1–21, 2015. Citado 3 vezes nas páginas 21, 28 e 31.
- GU, Y.; TAN, S.; WONG, K.; HO, M.; QU, L. Using ga-based feature selection for emotion recognition from physiological signals. In: IEEE. **Intelligent Signal Processing and Communications Systems, 2008. ISPACS 2008. International Symposium on**. [S.l.], 2009. p. 1–4. Citado na página 44.
- GU, Y.; TAN, S. L.; WONG, K. J.; HO, M. H. R.; QU, L. Emotion-aware technologies for consumer electronics. In: IEEE. **Consumer Electronics, 2008. ISCE 2008. IEEE International Symposium on**. [S.l.], 2008. p. 1–4. Citado na página 44.
- HAAG, A.; GORONZY, S.; SCHAICH, P.; WILLIAMS, J. Emotion recognition using bio-sensors: First steps towards an automatic system. In: SPRINGER. **ADS**. [S.l.], 2004. p. 36–48. Citado na página 44.
- HAGER, J. C.; EKMAN, P.; FRIESEN, W. V. Facial action coding system. **Salt Lake City, UT: A Human Face**, 2002. Citado na página 31.
- HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The weka data mining software: an update. **ACM SIGKDD explorations newsletter**, ACM, v. 11, n. 1, p. 10–18, 2009. Citado na página 75.
- HAYASHI, E.; BARANAUSKAS, M. C. C. 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 3 vezes nas páginas 21, 27 e 30.
- HÖÖK, K.; STÅHL, A.; SUNDSTRÖM, P.; LAAKSOLAAHTI, J. Interactional empowerment. In: ACM. **Proceedings of the SIGCHI Conference on Human Factors in Computing Systems**. [S.l.], 2008. p. 647–656. Citado na página 26.
- JAMES, W. Ii.—what is an emotion? **Mind**, Mind Assoc, n. 34, p. 188–205, 1884. Citado na página 20.
- JANG, J.-S. R.; SUN, C.-T.; MIZUTANI, E. Neuro-fuzzy and soft computing: a computational approach to learning and machine intelligence. prentice hall, 1997. Citado na página 47.

- JEBARA, T.; WANG, J.; CHANG, S.-F. Graph construction and b-matching for semi-supervised learning. In: **Proceedings of the 26th Annual International Conference on Machine Learning**. New York, NY, USA: ACM, 2009. (ICML '09), p. 441–448. ISBN 978-1-60558-516-1. Citado na página 77.
- JOHNSON, R. A.; WICHERN, D. W. *et al.* **Applied multivariate statistical analysis**. [S.l.]: Prentice hall Englewood Cliffs, NJ, 1992. v. 4. Citado na página 41.
- KATSIKIS, 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 5 vezes nas páginas 36, 46, 47, 49 e 50.
- KATZ, J. **How emotions work**. [S.l.]: University of Chicago Press, 2001. Citado na página 26.
- KITCHENHAM, B. Procedures for performing systematic reviews. **Keele, UK, Keele University**, v. 33, n. 2004, p. 1–26, 2004. Citado na página 21.
- KITCHENHAM, B.; PRETORIUS, R.; BUDGEN, D.; BRERETON, O. P.; TURNER, M.; NIAZI, M.; LINKMAN, S. Systematic literature reviews in software engineering—a tertiary study. **Information and Software Technology**, Elsevier, v. 52, n. 8, p. 792–805, 2010. Citado na página 21.
- KLEINGINNA, P. R.; KLEINGINNA, A. M. A categorized list of emotion definitions, with suggestions for a consensual definition. **Motivation and emotion**, Springer, v. 5, n. 4, p. 345–379, 1981. Citado na página 27.
- KOHAVI, R. *et al.* A study of cross-validation and bootstrap for accuracy estimation and model selection. In: **Ijcai**. [S.l.: s.n.], 1995. v. 14, n. 2, p. 1137–1145. Citado na página 47.
- KUKOLJA, D.; POPOVIĆ, S.; HORVAT, M.; KOVAČ, B.; ČOSIĆ, K. Comparative analysis of emotion estimation methods based on physiological measurements for real-time applications. **International journal of human-computer studies**, Elsevier, v. 72, n. 10, p. 717–727, 2014. Citado 3 vezes nas páginas 41, 49 e 50.
- LACHENBRUCH, P. A.; GOLDSTEIN, M. Discriminant analysis. **Biometrics**, JSTOR, p. 69–85, 1979. Citado na página 41.
- LANE, N. D.; MOHAMMOD, M.; LIN, M.; YANG, X.; LU, H.; ALI, S.; DORYAB, A.; BERKE, E.; CHOUDHURY, T.; CAMPBELL, A. T. Bewell: A smartphone application to monitor, model and promote wellbeing. In: **in Intl. ICST Conf. on Pervasive Computing Technologies for Healthcare**. [S.l.: s.n.], 2011. Citado na página 35.
- LANG, P. J.; BRADLEY, M. M.; CUTHBERT, B. N. International affective picture system (iaps): Affective ratings of pictures and instruction manual. **Technical report A-8**, 2008. Citado na página 45.
- LANG, P. J.; GREENWALD, M. K.; BRADLEY, M. M.; HAMM, A. O. Looking at pictures: Affective, facial, visceral, and behavioral reactions. **Psychophysiology**, CAMBRIDGE UNIVERSITY PRESS, v. 30, p. 261–261, 1993. Citado na página 30.
- LAPARRA-HERNÁNDEZ, J.; BELDA-LOIS, J.; MEDINA, E.; CAMPOS, N.; POVEDA, R. Emg and gsr signals for evaluating user's perception of different types of ceramic flooring. **International Journal of Industrial Ergonomics**, Elsevier, v. 39, n. 2, p. 326–332, 2009. Citado 3 vezes nas páginas 44, 45 e 49.
- LATHIA, N.; RACHURI, K.; MASCOLO, C.; ROUSSOS, G. Open source smartphone libraries for computational social science. In: **ACM. Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication**. [S.l.], 2013. p. 911–920. Citado 2 vezes nas páginas 40 e 49.

- LAURANS, G.; DESMET, P. M.; HEKKERT, P. P. Assessing emotion in interaction: Some problems and a new approach. In: UNIVERSITE DE TECHNOLOGIE DE COMPIEGNE. **Proceedings of the 4th International Conference on Designing Pleasurable Products and Interfaces, DPPI'09, Compiègne, Oct. 2009**. [S.l.], 2009. Citado 2 vezes nas páginas 27 e 30.
- LEDOUX, J. **The emotional brain: The mysterious underpinnings of emotional life**. [S.l.]: Simon and Schuster, 1998. Citado na página 26.
- LEFTERI, H. T.; ROBERT, E. U. Fuzzy and neural approaches in engineering. **A Wiley-Interscience Publication**, 1997. Citado na página 42.
- LERA, E. de; GARRETA-DOMINGO, M. Ten emotion heuristics: guidelines for assessing the user's affective dimension easily and cost-effectively. In: BRITISH COMPUTER SOCIETY. **Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI... but not as we know it-Volume 2**. [S.l.], 2007. p. 163–166. Citado 2 vezes nas páginas 31 e 45.
- LEVENTHAL, H. A perceptual-motor theory of emotion. **Advances in experimental social psychology**, v. 17, p. 117–182, 1984. Citado na página 30.
- LI, T.; AN, C.; CAMPBELL, A. T.; ZHOU, X. Hilight: Hiding bits in pixel translucency changes. **ACM SIGMOBILE Mobile Computing and Communications Review**, ACM, v. 18, n. 3, p. 62–70, 2015. Citado na página 50.
- 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 19.
- 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 3 vezes nas páginas 48, 49 e 50.
- 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 4 vezes nas páginas 36, 39, 40 e 49.
- MADER, S.; PETER, C.; GÖCKE, R.; SCHULTZ, R.; VOSKAMP, J.; URBAN, B. A freely configurable, multi-modal sensor system for affective computing. In: **Affective Dialogue Systems**. [S.l.]: Springer, 2004. p. 313–318. Citado na página 39.
- 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 5 vezes nas páginas 29, 30, 36, 56 e 68.
- MAIER, M.; LUXBURG, U. von; HEIN, M. Influence of graph construction on graph-based clustering measures. In: MAX-PLANCK-GESELLSCHAFT. Red Hook, NY, USA: Curran, 2009. p. 1025–1032. ISBN 978-1-605-60949-2. Citado na página 77.
- MANDRYK, R. L.; ATKINS, M. S. A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies. **International Journal of Human-Computer Studies**, Elsevier, v. 65, n. 4, p. 329–347, 2007. Citado 3 vezes nas páginas 36, 42 e 49.
- MARDIA, K.; KENT, J.; BIBBY, J. Multivariate analysis academic press inc. **London) LTD**, 1979. Citado na página 41.

MARTINEZ, A.; BENAVENTE, R. The ar face database cvc. **Copyright of Informatica (03505596) is the property of Slovene Society Informatika and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use**, 1998. Citado na página 45.

MERA, K.; ICHIMURA, T. Emotion analyzing method using physiological state. In: SPRINGER. **Knowledge-Based Intelligent Information and Engineering Systems**. [S.l.], 2004. p. 195–201. Citado 2 vezes nas páginas 44 e 45.

NAHIN, A. N. H.; ALAM, J. M.; MAHMUD, H.; HASAN, K. Identifying emotion by keystroke dynamics and text pattern analysis. **Behaviour & Information Technology**, Taylor & Francis, v. 33, n. 9, p. 987–996, 2014. Citado 3 vezes nas páginas 47, 48 e 49.

NEWMAN, M. **Networks: An Introduction**. New York, NY, USA: Oxford University Press, Inc., 2010. Citado na página 77.

NEWMAN, M. E. J. The structure and function of complex networks. **SIAMR**, JSTOR, v. 45, n. 2, p. 167–256, 2003. Citado na página 77.

NIEDENTHAL, P. M.; KRAUTH-GRUBER, S.; RIC, F. **Psychology of emotion: Interpersonal, experiential, and cognitive approaches**. [S.l.]: Psychology Press, 2006. Citado 6 vezes nas páginas 26, 28, 29, 30, 33 e 43.

NORMAN, D. A. Emotion design: Why we love (or hate) everyday things. Basic Books, 2004. Citado 5 vezes nas páginas 20, 21, 26, 27 e 80.

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 2 vezes nas páginas 36 e 43.

PARKINSON, B. Ideas and realities of emotion routledge. **London, United Kingdom**, 1995. Citado na página 28.

PEDERSEN, C.; TOGELIUS, J.; YANNAKAKIS, G. N. Modeling player experience for content creation. **IEEE Trans. Comput. Intellig. and AI in Games**, v. 2, n. 1, p. 54–67, 2010. Citado 3 vezes nas páginas 20, 63 e 65.

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 7 vezes nas páginas 19, 20, 36, 37, 39, 49 e 50.

PICARD, R. W. **Affective computing MIT Media Laboratory Perceptual Computing Section Technical Report No. 321**. [S.l.]: Cambridge, 1995. Citado 2 vezes nas páginas 20 e 26.

PICARD, R. W.; VYZAS, E.; HEALEY, J. Toward machine emotional intelligence: Analysis of affective physiological state. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 23, n. 10, p. 1175–1191, 2001. Citado na página 44.

PINTO, A. Psicologia geral. **Lisboa: Universidade Aberta**, n. 227, p. 340, 2001. Citado na página 26.

PREECE, J.; ROGERS, Y.; SHARP, H. **Design de interação**. [S.l.]: Bookman, 2005. Citado na página 31.

QUINLAN, J. R. Induction of decision trees. **Machine learning**, Springer, v. 1, n. 1, p. 81–106, 1986. Citado 2 vezes nas páginas 37 e 75.

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 6 vezes nas páginas 20, 36, 37, 38, 49 e 50.

REIJNEVELD, K.; LOOZE, M. de; KRAUSE, F.; DESMET, P. Measuring the emotions elicited by office chairs. In: **Proceedings of the 2003 International Conference on Designing Pleasurable Products and Interfaces**. New York, NY, USA: ACM, 2003. (DPPI '03), p. 6–10. ISBN 1-58113-652-8. Citado 2 vezes nas páginas 20 e 67.

RICHTER, S.; DETER, H.-C.; RUDAT, M.; SCHÄCHINGER, H.; ZIMMERMANN-VIEHOFF, F.; WEBER, C. Anger and cardiovascular startle reactivity in normotensive young males. **International Journal of Psychophysiology**, Elsevier, v. 79, n. 3, p. 364–370, 2011. Citado na página 58.

RODRIGUES, K.; GARCIA, F.; BOCANEGRA, L.; GONÇALVES, V.; CARVALHO, V.; NERIS, V. Personas-driven design for mental health therapeutic applications. **SBC**, v. 6, n. 1, p. 18, 2015. Citado 5 vezes nas páginas 19, 20, 21, 27 e 31.

RODRIGUES, K.; GONÇALVES, V.; ; NERIS, V. Tecnologias de informação e comunicação inclusivas e o envelhecer. In: **Promoção da Saúde para um Envelhecimento Saudável**. [S.l.]: Unijuí, 2016. v. 4, p. 265–288. Citado 4 vezes nas páginas 19, 27, 71 e 80.

ROSALES, G. C. M.; ARAUJO, R. Borges de; OTSUKA, J. L.; ROCHA, R. V. da. Using logical sensors network to the accurate monitoring of the learning process in distance education courses. In: **IEEE. Advanced Learning Technologies (ICALT), 2011 11th IEEE International Conference on**. [S.l.], 2011. p. 573–575. Citado na página 54.

RUSSELL, J. A circumplex model of affect. **Journal of personality and social psychology**, v. 39, n. 6, p. 1161–1178, 1980. ISSN 0022-3514. Citado na página 31.

RUSSELL, J. A. Core affect and the psychological construction of emotion. **Psychological Review**, p. 145–172, 2003. Citado na página 32.

RUSSELL, J. A.; WEISS, A.; MENDELSON, G. A. Affect grid: a single-item scale of pleasure and arousal. **Journal of Personality and Social psychology**, American Psychological Association, v. 57, n. 3, p. 493, 1989. Citado na página 67.

SARAGIH, J. M.; LUCEY, S.; COHN, J. F. Deformable model fitting by regularized landmark mean-shift. **Int. J. Comput. Vision**, Kluwer Academic Publishers, Hingham, MA, USA, v. 91, n. 2, p. 200–215, jan. 2011. ISSN 0920-5691. Citado 2 vezes nas páginas 51 e 55.

SCHERER, K. What are emotions? and how can they be measured? **Social Science Information**, jan. 2005. Citado 20 vezes nas páginas 11, 21, 22, 26, 27, 28, 29, 30, 31, 32, 33, 41, 43, 46, 48, 66, 67, 68, 70 e 80.

SCHERER, K. R. Plato's legacy: Relationships between cognition, emotion, and motivation. **Geneva Studies in Emotion and Communication**, v. 9, n. 1, p. 1–7, 1995. Citado na página 26.

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 4 vezes nas páginas 25, 27, 28 e 29.

SCHERER, K. R. The nature and dynamics of relevance and valence appraisals: Theoretical advances and recent evidence. **Emotion Review**, SAGE Publications, v. 5, n. 2, p. 150–162, 2013. Citado na página 19.

- SCHERER, K. R.; CESCHI, G. Criteria for emotion recognition from verbal and nonverbal expression: Studying baggage loss in the airport. **Personality and Social Psychology Bulletin**, Sage Publications, v. 26, n. 3, p. 327–339, 2000. Citado 2 vezes nas páginas 19 e 33.
- SCOLARI S.H.P.; MARAR, J. H. R. Reflections on the emotions intensity variables in design. In: **Proceedings of the 5th International Congress Of Design Research**, [S.l.: s.n.], 2005. p. 586–592. Citado na página 26.
- SOBOL-SHIKLER, T.; ROBINSON, P. Classification of complex information: Inference of co-occurring affective states from their expressions in speech. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 32, n. 7, p. 1284–1297, 2010. Citado na página 38.
- SOMEREN, M. W. V.; BARNARD, Y. F.; SANDBERG, J. A. *et al.* **The think aloud method: A practical guide to modelling cognitive processes**. [S.l.]: Academic Press London, 1994. v. 2. Citado 3 vezes nas páginas 20, 57 e 64.
- STEMMLER, G. Methodological considerations in the psychophysiological study of emotion. **Handbook of affective sciences**, p. 225–255, 2003. Citado 2 vezes nas páginas 28 e 30.
- SUGENO, M. **Industrial applications of fuzzy control**. [S.l.]: Elsevier Science Inc., 1985. Citado na página 47.
- TZU, S.; PIN, S. **A arte da guerra**. [S.l.]: WWF Martins Fontes, 2015. Citado na página 20.
- 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 4 vezes nas páginas 21, 52, 53 e 74.
- VAPNIK, V. **The nature of statistical learning theory**. [S.l.]: Springer Science & Business Media, 2013. Citado 3 vezes nas páginas 36, 37 e 75.
- VIOLA, P.; JONES, M. J. Robust real-time face detection. **International journal of computer vision**, Springer, v. 57, n. 2, p. 137–154, 2004. Citado na página 40.
- WANG, R.; CHEN, F.; CHEN, Z.; LI, T.; HARARI, G.; TIGNOR, S.; ZHOU, X.; BEN-ZEEV, D.; CAMPBELL, A. T. Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: ACM. **Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing**. [S.l.], 2014. p. 3–14. Citado na página 50.
- WRZUS, C.; MEHL, M. R. Lab and/or field? measuring personality processes and their social consequences. **European Journal of Personality**, Wiley Online Library, v. 29, n. 2, p. 250–271, 2015. Citado na página 50.
- XAVIER, R. A. C.; NÉRIS, V. P. de A. A hybrid evaluation approach for the emotional state of information systems users. In: **ICEIS (3)**. [S.l.: s.n.], 2012. p. 45–53. Citado 4 vezes nas páginas 11, 27, 31 e 33.
- XAVIER, R. A. C.; NERIS, V. P. de A. Measuring the emotional experience of users through a hybrid semantic approach. In: SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. **Proceedings of the 13th Brazilian Symposium on Human Factors in Computing Systems**. [S.l.], 2014. p. 226–235. Citado 3 vezes nas páginas 11, 20 e 34.
- 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 28 e 67.

YIN, R. K. Case study research design and methods third edition. **Applied social research methods series**, Sage Publications Inc, v. 5, 2003. Citado na página [22](#).

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 4 vezes nas páginas [36](#), [44](#), [49](#) e [50](#).

ZHOU, F.; QU, X.; HELANDER, M. G.; JIAO, J. R. Affect prediction from physiological measures via visual stimuli. **International Journal of Human-Computer Studies**, Academic Press, Inc., Duluth, MN, USA, v. 69, n. 12, p. 801–819, dez. 2011. ISSN 1071-5819. Citado 5 vezes nas páginas [19](#), [36](#), [43](#), [49](#) e [50](#).

EMOÇÕES PRESENTES NA ESTRUTURA DO ESPAÇO EMOCIONAL SEMÂNTICO

Durante a condução dos experimentos exploratórios foi apresentada aos usuários uma tabela de emoções proveniente de uma versão traduzida da Estrutura do Espaço Emocional Semântico de Scherer, para auxiliar na aplicação do método Think Aloud, como pode ser verificado a seguir.

Emoções presentes na Estrutura do Espaço Emocional Semântico

Quadrante 1	Quadrante 2	Quadrante 3	Quadrante 4
Raiva	Raiva	Triste	Triste
Alarmado (a)	Amargurado (a)	Decepcionado (a)	Ansioso (a)
Assustado (a)	Angustiado (a)	Desanimado (a)	Apreensivo (a)
Estressado (a)	Desconfiado (a)	Desapontado (a)	Cabisbaixo (a)
Furioso (a)	Descontente	Desconfortável	Cansado (a)
Impaciente	Entediado (a)	Desesperado (a)	Chateado (a)
Indignado (a)	Frustrado (a)	Insatisfeito (a)	Comovido (a)
Irritado (a)	Insultado (a)	Melancólico (a)	Desanimado (a)
Assustado (a)			Desmotivado (a)
Medo			Entediado (a)
Tenso (a)			Preocupado (a)
			Desanimado (a)
Quadrante 5	Quadrante 6	Quadrante 7	Quadrante 8
Neutro (a)	Feliz	Feliz	Surpreso (a)
Pensativo (a)	Aliviado (a)	Afoito (a)	Aventureiro (a)
Contemplativo (a)	Amigável	Alegre	Compenetrado (a)
Educado (a)	Bem	Animado (a)	Convencido (a)
Empático (a)	Calmo (a)	Apaixonado (a)	Empolgado (a)
Pacífico (a)	Confiante	Autoconfiante	Excitado (a)
Sonolento (a)	Contente	Bravo (a)	Motivado (a)
Tranquilo (a)	Descontraído (a)	Corajoso (a)	Persuasivo (a)
Normal	Gentil	Despreocupado (a)	Desejar
	Impressionado (a)	Encantado (a)	Prepotente
	Relaxado (a)	Entusiasmante	Superior
	Satisfeito (a)	Esperançoso (a)	Triunfante
	Tranquilo (a)	Eufórico (a)	Vaidoso (a)
		Extasiado (a)	

PARECER SUBSTANCIADO E APROVADO PELO COMITÊ DE ÉTICA DA FACULDADE DE MEDICINA DA USP

Os detalhes da condução dos experimentos exploratórios com usuários presentes nesta Tese de doutorado foram submetidos ao Comitê de Ética em Pesquisa envolvendo Seres Humanos - Faculdade de Medicina da Universidade de São Paulo (FMUSP). O parecer substanciado e aprovado pelo Comitê pode ser verificado a seguir.

FACULDADE DE MEDICINA DA
UNIVERSIDADE DE SÃO
PAULO - FMUSP



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Um framework para captura, análise e classificação automatizadas da experiência emocional de usuários em tempo de interação

Pesquisador: Vinícius Pereira Gonçalves

Área Temática:

Versão: 2

CAAE: 45081415.0.0000.0065

Instituição Proponente: UNIVERSIDADE DE SAO PAULO

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 1.103.031

Data da Relatoria: 03/06/2015

Apresentação do Projeto:

Projeto de pós-graduação intitulado "Um framework para captura, análise e classificação automatizadas da experiência emocional de usuários em tempo de interação".

Objetivo da Pesquisa:

Desenvolver/estabelecer um sistema de captura, análise e classificação automatizadas da experiência emocional de usuários em tempo de interação com sistemas eletrônicos diversos.

Avaliação dos Riscos e Benefícios:

Sem risco, pois avaliará voluntários em atividades rotineiras. Benefícios informativos.

Comentários e Considerações sobre a Pesquisa:

O projeto está muito bem apresentado, tanto na forma como no conteúdo, como manifestado no parecer anterior.

Não havia menção ao número total de participantes na pesquisa. Porém, isso foi corrigido no projeto e demais documentos de apresentação do mesmo.

Considerações sobre os Termos de apresentação obrigatória:

De acordo.

Endereço: DOUTOR ARNALDO 251 21º andar sala 36

Bairro: PACAEMBU

CEP: 01.246-903

UF: SP

Município: SAO PAULO

Telefone: (11)3893-4401

E-mail: cep.fm@usp.br

FACULDADE DE MEDICINA DA
UNIVERSIDADE DE SÃO
PAULO - FMUSP



Continuação do Parecer: 1.103.031

Recomendações:

Nenhuma.

Conclusões ou Pendências e Lista de Inadequações:

Nenhuma.

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Considerações Finais a critério do CEP:

SAO PAULO, 11 de Junho de 2015

Assinado por:
Roger Chammas
(Coordenador)

Endereço: DOUTOR ARNALDO 251 21º andar sala 36

Bairro: PACAEMBU

CEP: 01.246-903

UF: SP

Município: SAO PAULO

Telefone: (11)3893-4401

E-mail: cep.fm@usp.br

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

O Termo de Consentimento Livre e Esclarecido foi utilizado durante a condução dos experimentos exploratórios e elucidada aos participantes o objetivo da pesquisa, a voluntariedade na participação e o caráter científico da mesma, como pode ser verificado a seguir.

Termo de Consentimento Livre e Esclarecido

1. Você está sendo convidado para participar da pesquisa “Um framework para captura, análise e classificação automatizadas da experiência emocional de usuários em tempo de interação”.
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, com a instituição.
5. Essa pesquisa tem por objetivo captar e analisar respostas emocionais de usuários interagindo com o jogo Super Mário Bros, com o auxílio de quatro sensores: face, fala, frequência cardíaca e lógico (desempenho).
6. Sua participação nesta pesquisa consistirá em jogar o jogo Super Mário Bros em um Notebook e preencher questionários para posteriores avaliações.
7. A sua participação na pesquisa pode envolver algum desconforto relacionado ao tempo despendido com a realização de questionários e entrevistas, sendo que faremos o possível para minimizar possíveis desconfortos. Em relação ao conteúdo dos questionários e entrevistas serão planejados a evitar possíveis constrangimentos ou desconfortos, e caso ocorram você pode se recusar a responder ou mesmo interromper a sua participação a qualquer momento, sem qualquer prejuízo em sua relação com a instituição ou com o pesquisador.
8. Os benefícios relacionados a sua participação estão em contribuir com a pesquisa, para que se saiba mais sobre emoções na interação com dispositivos computacionais.
9. As informações obtidas por meio dessa pesquisa serão confidenciais e asseguramos o sigilo sobre sua participação.
10. Os dados não serão divulgados de forma a possibilitar sua identificação.
11. 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).
12. Você receberá uma cópia deste termo onde consta o telefone e o endereço do pesquisador principal, podendo tirar suas dúvidas sobre o projeto e sua participação, agora ou a qualquer momento.

Vinicius Pereira Gonçalves
Instituto de Ciências Matemáticas e de Computação
Universidade de São Paulo
13566-590 São Carlos-SP
Tel.: 16-33739638

Endereço e telefone do Pesquisador Principal: Rua Maestro João Seppe, 303, Jd. Paraíso
13561-180 São Carlos-SP
Tel.: 16-981273468

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

São Carlos, 09 de Novembro de 2015

Sujeito da pesquisa

AUTORIZAÇÃO DE CAPTAÇÃO E EXIBIÇÃO DE IMAGEM, SOM E NOME

O Termo de Autorização de Captação e Exibição de Imagem, Som e Nome foi utilizado durante a condução dos experimentos exploratórios e esclarece aos participantes que há captura de imagens e voz e que os dados coletados são utilizados para fins científicos, como pode ser verificado a seguir.

AUTORIZAÇÃO DE CAPTAÇÃO E EXIBIÇÃO DE IMAGEM, SOM E NOME

Eu, _____, (nacionalidade) _____, (estado civil) _____ portador da Cédula de Identidade RG nº _____ (ou Registro Nacional de Estrangeiro _____), inscrito no CPF/MF sob o número _____ autorizo a captação, utilização e exibição de minha voz e imagem pela UNIVERSIDADE DE SÃO PAULO – USP diretamente ou através do Instituto de Ciências de Computação e Matemática Computacional – 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, sejam 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 e voz, 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 ou 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 ou voz objetos da presente autorização, mesmo nas hipóteses de comercialização das obras, em parceira ou não com outras pessoas jurídicas.

3. Declaro estar ciente de que as imagens e voz captadas nesta 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 reclamado 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, 09 de Novembro de 2015.

(Assinatura)

(Nome por extenso)

QUESTIONÁRIO DE LEVANTAMENTO DE PERFIS DOS USUÁRIOS

O Questionário de Levantamento de Perfil foi utilizado durante a condução dos experimentos exploratórios e possuía questões de cunho social e cultural que permitiram traçar o perfil desses usuários, como pode ser verificado a seguir.

Questionário - Perfil do Participante

Nº

- 1. Sexo:**
 - Masculino
 - Feminino
- 2. Idade:** ____ anos
- 3. Qual o seu grau de escolaridade?**
 - Nenhuma escolaridade
 - Ensino Fundamental Incompleto
 - Ensino Fundamental
 - Ensino Médio Incompleto
 - Ensino Médio
 - Superior Incompleto
 - Superior
 - Mestrado Incompleto
 - Mestrado
 - Doutorado Incompleto
 - Doutorado
 - Pós-doutorado Incompleto
 - Pós-doutorado
- 4. Jogar é a sua principal atividade de lazer?**
 - Não
 - Sim
- 5. Qual a motivação para você jogar?**
 - Não Joga
 - Desafio
 - Lazer
 - Socializar
 - Vício
 - Outro _____
- 6. Você joga a quanto tempo?**
 - Não Jogo
 - 1 à 6 meses
 - 7 meses à 1 ano
 - 1 à 2 anos
 - Mais de 2 anos
- 7. Você joga com que frequência?**
 - Não Jogo
 - Menos de 1 vezes por semana
 - 1 à 3 vezes por semana
 - 4 à 7 vezes por semana
 - Mais de 7 vezes por semana
- 8. Você costuma jogar com qual tipo de dispositivo?**
 - Não Jogo
 - Computador
 - Celular
 - Tablet
 - Vídeo Game
 - Outro: _____
- 9. Quais recursos você costuma usar para jogar?**
 - Não Jogo
 - Internet
 - Mouse e Teclado
 - Touch Screen
 - Controle Remoto com sensor de movimento (ex: Nintendo Wii)
 - Controle Remoto sem sensor de movimento (ex: Joystick)
 - Kinect
 - Outro: _____
- 10. Você costuma gastar dinheiro com jogo?**
 - Não
 - Sim

Se sim, qual a média por mês?
