

# Sistemas Tutores Afetivos: Estado da Arte e Desafios para o Reconhecimento de Emoções Relacionadas à Aprendizagem via Expressões Faciais

Ernani Gottardo<sup>1</sup>, Andrey Ricardo Pimentel<sup>2</sup>

<sup>1</sup>Instituto Federal de Educação Ciência e Tecnologia do RS (IFRS)  
Erechim – RS – Brasil

<sup>2</sup>Departamento de Informática - Universidade Federal do Paraná (UFPR)  
Curitiba – PR – Brasil.

ernani.gottardo@erechim.ifrs.edu.br, andrey@inf.ufpr.edu.br

**Abstract.** *Given the inextricable link between learning and emotions, a growing body of research has investigated how to provide functionality for recognizing and adapting to learners' affective reactions in educational software. In this paper, techniques for recognition of basic emotions from facial expressions are presented. The results show that for some basic emotions with most prominent expressions, such as joy, the results are considered promising. However, some learning-centered emotions such as sadness are not easily recognized via facial expressions, needing improvements and adaptations to currently existing methods.*

**Resumo.** *Considerando a intrínseca relação entre aprendizagem e emoções, pesquisas recentes têm investigado como prover funcionalidades de reconhecimento e adaptação às reações afetivas dos aprendizes em softwares educacionais. Nesse trabalho, técnicas de reconhecimento de emoções básicas via expressões faciais são apresentadas, juntamente com alguns experimentos. Os resultados demonstram que para algumas emoções básicas com expressões mais destacadas, como alegria, os resultados são considerados promissores. Entretanto, algumas emoções relacionadas à aprendizagem como tristeza não são facilmente reconhecidas via expressões faciais, necessitando de avanços e adaptações aos métodos existentes atualmente.*

## 1. Introdução

Um desafio frequentemente tratado em pesquisas na comunidade científica de informática em educação refere-se ao de desenvolvimento de softwares educacionais que ofereçam um suporte mais efetivo para os estudantes, principalmente agregando funcionalidades adaptativas às necessidades individuais. Alguns ambientes como os Sistemas Tutores Inteligentes (STI) implementam funcionalidades de adaptação, porém, essas funcionalidades normalmente baseiam-se apenas em informações cognitivas, não levando em conta as reações emocionais dos estudantes (Khan et al., 2010; Baker et al., 2012).

Considerando-se a estreita relação entre aprendizagem e emoções, nos últimos anos, a investigação do impacto no processo de aprendizado de fatores como motivação e estados afetivos dos estudantes têm emergido como uma linha de pesquisa promissora (Grafsgaard et al., 2013). Como exemplo, Jaques et al. (2012) destacam que seria muito importante para um sistema detectar quando um estudante está frustrado a fim de encorajá-lo a continuar estudando e a realizar as suas atividades

Entretanto, Picard et al. (2014) observam que estamos ainda nos primeiros estágios do entendimento da complexidade e da relação entre afeto, aprendizagem e *feedback* e que essas questões permanecem um desafio em pesquisas envolvendo ambientes computacionais de aprendizagem adaptativos.

Para que se possa realizar qualquer ação de adaptação às reações afetivas dos estudantes é necessário primeiramente reconhecê-las com exatidão. Nesse sentido, um conjunto de pesquisas tem sido desenvolvido, utilizando técnicas que incluem desde sensores de baixo nível como sinais fisiológicos (e.g. batimentos cardíacos, condutividade da pele) até propostas que utilizam apenas dados de interação com o ambiente.

Uma abordagem que se destaca no processo de reconhecimento das emoções é a utilização de expressões faciais (Sarrafzadeh et al., 2008; Grafsgaard et al., 2013). Isso se deve ao fato de que (a) existe uma forte relação entre características faciais e estados afetivos, (b) as expressões faciais não dependem do ambiente, e (c) não requerem dispositivos caros ou altamente intrusivos para seu reconhecimento.

Nesse trabalho serão apresentadas as etapas e demonstrados os resultados de experimentos com a utilização de métodos de inferência de emoções básicas utilizando expressões faciais. Esses resultados servirão como base para a análise e discussão dos principais desafios e necessidades de avanços nessa área.

Na descrição dos resultados deste trabalho, este documento está estruturado em cinco seções. Na primeira seção apresentou-se a introdução do trabalho, destacando o contexto em que o trabalho está inserido. Na Seção 2, são descritos alguns trabalhos correlatos. Na seção 3 e 4, apresentam-se, respectivamente, detalhes e análise dos resultados dos experimentos realizados. O artigo é finalizado com as conclusões e perspectivas de continuidade deste trabalho.

## **2. Trabalhos Relacionados**

Considerando a importância do desenvolvimento de funcionalidade adaptativas aos estados afetivos dos aprendizes, nos últimos anos observa-se o crescente desenvolvimento de pesquisas que aplicam conceitos oriundos da área conhecida como Computação Afetiva (em inglês *Affective Computing*) em ambientes de software educacionais (D'mello e Graesser, 2012).

Computação Afetiva é uma área multidisciplinar de pesquisa definida por Picard (1997) como “computação que está relacionada com, que surge de ou deliberadamente influencia emoções”(p. 3). A Computação Afetiva busca desenvolver métodos para dotar computadores de várias características sociais, tais como reconhecer emoções e responder apropriadamente a elas, expressar emoções, mostrar empatia, identificar personalidade, entre outros (Jaques et al., 2012).

Relativamente ao processo de reconhecimento de emoções, cabe ressaltar que emoções ou estados emocionais não são diretamente observáveis ou acessados. O que é revelado de forma voluntária ou involuntária pelas pessoas são padrões de expressões. Considerando esses padrões, pessoas ou sistemas podem aplicar técnicas para inferir ou estimar o estado emocional, considerando um determinado nível de erro ou incerteza (Picard, 1997).

A coleta dados para servir de base para a construção em ambientes adaptativos às reações afetivas dos estudantes têm sido realizada utilizando técnicas diversificadas que, frequentemente, requerem a utilização de dispositivos externos. Dentre esses dispositivos, pode-se citar sensor de postura e movimento (Grafsgaard et al., 2013), monitor de batimento cardíaco (Grafsgaard et al., 2013, Picard, 1997), sensores de gestos e expressões faciais (Alexander, 2007; Sarrafzadeh et al., 2008; Grafsgaard et al., 2013; D'mello e Graesser 2012), sensor de condutividade da pele e temperatura (Picard, 1997).

É importante observar que alguns dispositivos citados anteriormente (e.g. monitor cardíaco, sensor de condutividade de pele) podem ser considerados intrusivos, de alto custo ou inviáveis de utilização em larga escala. Além disso, muitos desses dispositivos não são uma solução satisfatória em cenários nos quais sensores externos podem distrair os usuários e interferir nas tarefas principais (D'mello et al, 2008).

Cada modalidade tem suas vantagens e desvantagens para sua utilização como fonte viável de inferência de emoções. Dentre os fatores que interferem na escolha destaca-se: i) a validade do sinal como uma forma natural de identificação de estados afetivos, ii) a confiabilidade do sinal em ambientes reais, iii) o tempo de obtenção e iv) custo e intrusividade do sinal (Calvo et al., 2010)

Nesse contexto, a utilização das expressões faciais para a inferência de estados afetivos é uma das técnicas que apresenta uma relação eficiente nos quatro aspectos citados anteriormente. Isso se deve ao fato de que a face desempenha um papel central no processo de comunicação e interação dos seres humanos, sendo uma das principais e mais aparentes formas de representar o que uma pessoa está tentando comunicar (Picard, 1997).

O método *Facial Action Coding System* (FACS) (Ekman, 1992) é a técnica mais utilizada em pesquisas que utilizam informações de expressões faciais (Picard, 2014). No FACS cada movimento independente da face é denominado *Action Unit*- AU (ex. movimento de sobrancelhas, pálpebras, lábios) que representam os movimentos dos músculos usados para gerar expressões. São definidos aproximadamente 58 AU's, cujas combinações resultam em um grande conjunto de expressões faciais possíveis. É importante observar que, segundo Ekman (1992) alguns desses músculos não podem ser manipulados conscientemente e somente são movimentados por emoções genuínas.

Dada sua abrangência, o métodos FACS tem sido útil para descoberta de movimentos faciais que são indicativos de estados afetivos (Bosch et al., 2014). Contudo, algumas limitações no modelo devem ser consideradas. Conforme observa (Picard, 1997) essa técnica falhará no reconhecimento do estado de alegria caso a pessoa que estiver experimentando alegria suprimir todas as expressões faciais. Outra limitação importante refere-se ao tempo necessário para identificação das expressões e também ao fato de que essa técnica foi desenvolvida originalmente para especialistas humanos codificar emoções baseadas em figuras estáticas.

A acurácia na inferência utilizando expressões faciais é variável entre os diversos estados afetivos. Conforme concluem (McDaniel2007 et al., 2007), encantamento é facilmente detectado por apresentar um grande conjunto de AUs. Por outro lado, frustração (e.g. meio sorriso) e aborrecimento mostraram-se difíceis de serem diferenciadas do estado ‘neutro’ por apresentarem poucas expressões destacadas.

Considerando as limitações do modelo FACS, pesquisas têm sido desenvolvidas para seu aprimoramento. O método denominado FACS+ (Essa et al., 1997) propõem a utilização de técnicas de processamento de fluxo óptico juntamente com um modelo dinâmico de movimento. Essa proposta também leva em consideração um componente temporal (não presente no modelo FACS) em que as ações faciais ocorrem em três fases distintas: aplicação, liberação e relaxamento.

Outras pesquisas promissoras na área de reconhecimento de expressões faciais têm demonstrado a viabilidade de distinguir sorrisos de frustração e satisfação (Hoque et al., 2012) ou expressão de aflição real e deliberada (Littlewort, et al., 2009).

Com o objetivo de analisar os resultados práticos do reconhecimento de emoções via expressões faciais, foi realizado um procedimento experimental que será descrito na sequência.

### **3.1 Descrição dos dados utilizado no experimento**

Para a realização de experimentos utilizando o método de inferência e emoções via expressões faciais foi utilizado nesse trabalho o banco de dados “Cohn-Kanade *Expression Database*” (Kanade et al., 2000). Esse banco de dados foi projetado pelo Grupo de Análise Afetiva da Universidade de Pittsburg com o objetivo de propiciar o desenvolvimento de pesquisas envolvendo análise e síntese de automática de imagens faciais.

As expressões faciais constantes no banco de dados foram coletadas a partir de 210 adultos com idades entre 18 e 50 anos. Nesse grupo, 69% eram do sexo feminino, 31% masculinos, 81% Euro-Americanos, 13% Afro-Americanos e 6% de outros grupos.

Para a coleta das imagens, os participantes foram orientados pelos pesquisadores a realizar uma série de sequência de expressões faciais. Cada sequência deveria começar com uma expressão neutra e proceder para um pico da expressão solicitada (e.g. raiva). O pico de cada expressão para cada sequência foi então codificada utilizando o método FACS (Ekman, 1992) e também classificado em uma das 8 categorias a seguir: neutro, raiva, desprezo, desgosto, medo, alegria, tristeza e surpresa. Essas categorias são baseadas nas expressões básicas definidas em Ekman (1992).

É importante destacar que a classificação das emoções nas categorias descritas foi realizada considerando o que foi requisitado ao participante e não necessariamente o que esse participante expressou efetivamente. Esse método pode diferir em aparência e em tempo comparado com comportamentos espontâneos (Kanade et al., 2000). Destaca-se também que na obtenção das imagens questões de iluminação e ângulo das imagens foram aspectos controlados, utilizando-se lâmpadas de alta intensidade e também coletando imagens em ângulos de rotação de 30 graus.

No total o banco de dados disponibiliza 486 imagens, podendo ser utilizado para treinamento e testes de algoritmos de aprendizagem de máquina por disponibilizar

imagens com seus respectivos metadados relativos às Unidades de Ações (*Action Units*) e rótulos classificando cada imagem.

### 3.2 Descrição da metodologia do experimento

Para a obtenção de um modelo de inferência de emoções a partir de uma base de dados rotulada, diversas tarefas intermediárias precisam ser realizadas. Nesse artigo utilizou-se com referência a metodologia proposta nos trabalhos de Bartlett et al.(2006) e Littlewort et al. (2004). As principais etapas dessa metodologia envolvem:

1. Detecção e extração da face a partir das imagens do *corpus* sendo utilizado
2. Extração de atributos da face detectada utilizando Filtros Gabor
3. Treinamento de classificadores com base nos atributos extraídos pelos Filtros Gabor para construção de um modelo de inferência de emoções

Para a execução das etapas descritas foram feitos acima tomou-se como base a implementação de Mella e Bellavista (2016). A partir dessa solução base alguns ajustes de parâmetros e adaptações na implementação foram realizados com vistas a adequar às necessidades desse trabalho. Além disso, é importante destacar que todas as etapas utilizam de forma intensa funções e procedimentos disponíveis na plataforma OPENCV (*Open Source Computer Vision*), que é uma solução aberta e largamente utilizada na comunidade científica em aplicações que envolvam processamento de imagens.

Na primeira etapa, a detecção da face nas imagens foi realizada utilizando o método de detecção de objetos usando a técnica conhecida como “*Haar Feature-based Cascade Classifier*” (Viola e Jones, 2001). Esse método utiliza uma abordagem de aprendizagem de máquina na qual classificadores são treinados utilizando um conjunto de imagens positivo e negativo (i.e imagens contendo faces e outras não). Aplicando-se esse método sobre o banco de dados descrito, novas imagens foram então geradas contendo apenas a região da face, que é a área de interesse (*ROI – Region of Interest*).

A plataforma OPENCV provê modelos prontos para a detecção de faces como, por exemplo, o modelo “*haarcascade\_frontal\_face\_defaul.xml*”. Entretanto existem outras opções como o modelo “*haarcascade\_frontalface\_cbcl.xml*” desenvolvido pelo *Center for Biological & Computational Learning* (CBCL) no MIT. Avaliando-se essas duas opções, no contexto deste trabalho decidiu-se pela primeira por apresentar melhores resultados na detecção de face com menos requisitos de qualidade de iluminação. Entende-se que em um contexto educacional real não se pode esperar ou garantir que as condições de iluminação sejam ideais. Entretanto, a segunda opção (CBCL), quando consegue detectar a face (i.e. boas condições de iluminação), apresenta maiores índices de acerto na classificação final das emoções.

Na segunda etapa, um conjunto de filtros Gabor disponíveis no OPENCV foram aplicados sobre as imagens resultantes da etapa 2 (região facial) para extração de características ou atributos das imagens. Esse tipo de filtro é frequentemente usado no processamento de imagens em tarefas como detecção de bordas, representação e discriminação de textura (Bartlett et al., 2006; Littlewort et al., 2004)

Finalmente, considerando os atributos obtidos pela aplicação dos filtros Gabor, técnicas de aprendizagem de máquinas supervisionadas podem ser utilizadas para o treinamento de classificadores e construção de modelos de inferência. Nesse estudo foram utilizados o classificador SVM (*Support Vector Machines*), considerando sua

eficácia e performance em cenários com um grande número de atributos, que nesse caso são gerados pelos filtros Gabor.

### 3.3 Aplicação da metodologia descrita e resultados do experimento

Seguindo-se as etapas da metodologia descrita anteriormente, um modelo de classificação para as 7 emoções consideradas (mais o estado neutro) foi gerado. Os principais parâmetros utilizados no estudo foram definidos experimentalmente, testando-se algumas opções e escolhendo a que obteve melhor resultado. O tamanho da imagem da face foi configurado como 48 x 48. Outros parâmetros relevantes referem-se aos filtros Gabor, para os quais foram utilizados os seguintes valores: Sigma=3, Lambda=5 e Theta=4.

Na Tabela 1 a seguir são descritos os resultados da aplicação do modelo de inferência de emoções utilizando como teste 189 das 486 imagens do banco de dados Cohn-Kanade, selecionadas aleatoriamente. Na tabela é apresentada a emoção rotulada no banco de dados, o número de imagens testada dessa emoção e o percentual de acerto do modelo de classificação. Na coluna observação são descritos os principais aspectos relativos aos erros de classificação.

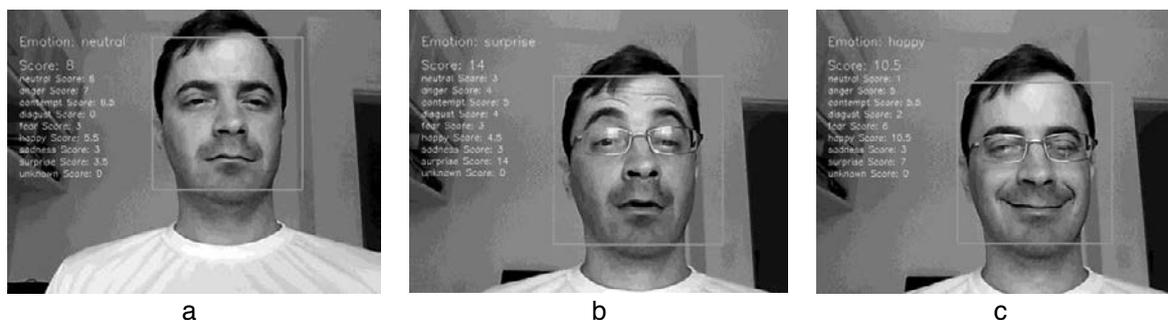
**Tabela 1 - Resultados da aplicação do modelo de inferência de emoções no banco de dados Cohn-Kanade**

Emoção	Número de Imagens Testadas	Percentual de acerto	Observações sobre os erros de classificação
Tristeza	11	27%	55% foram incorretamente classificadas como Neutro
Neutro	99	95%	O percentual de erro ficou igualmente dividido entre surpresa, desgosto, desprezo e raiva
Desgosto	17	94%	As classificações incorretas foram todas como Neutro (6%)
Raiva	4	100%	O único caso com nenhum erro de classificação
Surpresa	27	93%	As classificações incorretas foram todas como Neutro (7%)
Medo	6	50%	O percentual de erro (50%) ficou igualmente distribuído entre neutro, desprezo e felicidade
Desprezo	7	0%	Nesse caso, as emoções foram classificadas como: Neutro 72%, Tristeza 14% e Raiva 14%
Felicidade	18	94%	As classificações incorretas foram todas como Medo (6%)

Com o objetivo de realizar um teste em um cenário real, com características semelhantes a um estudante usando um software em um ambiente educacional, foi aplicado o modelo gerado para realizar a inferência utilizando vídeo do autor do

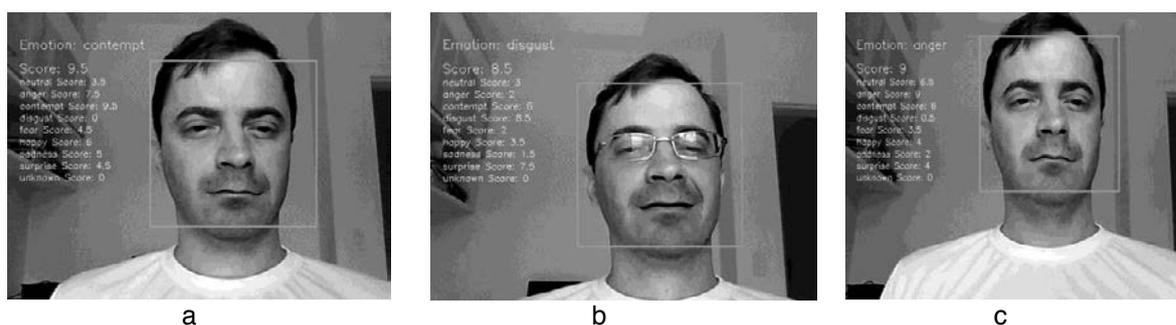
trabalho obtido por uma webcam VGA integrada do notebook modelo HP 430. Nesse vídeo foram simuladas algumas expressões para verificação do processo de inferência.

Na Figura 1 são exibidos três quadros desse vídeo em escala de cinza com o respectivo resultado do modelo de inferência descrito na legenda. Nesses quadros considera-se, na avaliação do autor, que houve acerto no processo de classificação que é impresso pelo modelo de inferência na parte superior esquerda da imagem.



**Figura 1 – Quadros do vídeo demonstrando experimento com o autor em que as imagens foram classificadas como: a) neutro, b) surpresa e c) felicidade**

Na Figura 2 foram selecionados outros três quadros do mesmo vídeo em que, na avaliação do autor, foi considerado que houve erro no processo de classificação.



**Figura 2 - Quadros do vídeo demonstrando experimento com o autor em que as imagens foram classificadas como: a) desprezo, b) desgosto e c) raiva**

#### 4. Análise dos Resultados

Analisando-se os resultados obtidos com o processo de teste com o banco de dados Cohn-Kanade apresentados na Tabela 1 pode-se perceber alguns comportamentos destacáveis. Primeiramente, observa-se que emoções como felicidade, raiva, surpresa e desgosto obtiveram índices elevados de acerto na classificação. Esses índices superam 90% de acurácia, e poderiam, se analisados de forma isolada, serem considerados excelentes.

Porém, índices extremamente menores foram obtidos com as emoções tristeza, medo, chegando a nenhum acerto para a emoção desprezo, demonstrando a dificuldade de inferência desse grupo de emoções, via expressões faciais. Uma observação específica poderia ser feita em relação à emoção Tristeza, que certamente representa uma importante situação de atenção, monitoramento e adaptação em um ambiente educacional. Porém, a classificação dessa emoção em 55% dos casos foi feita como Neutra.

Percebe-se que o grupo de emoções com acurácia superior apresenta expressões com características bastante específicas e altamente diferenciadas entre si e do estado neutro. Por outro lado, expressões como desgosto, tristeza e desprezo apresentam muitas semelhanças entre si e também com o estado neutro o que explica a maior dificuldade de classificação.

No experimento realizado com vídeos contendo expressões simuladas pelo próprio autor, conforme apresentado nas Figuras 1 e 2, os resultados foram analisados visualmente de forma manual pelo próprio autor. Através dessa análise, pode-se perceber que os resultados são muito similares aos obtidos no experimento com o banco de dados Cohn-Kanade, no qual as expressões mais destacadas (e.g raiva, surpresa felicidade) apresentam taxas de acertos superiores às demais.

Esse resultado tem sua relevância considerando-se que nesse segundo experimento as condições de captura das imagens são feitas em um ambiente natural, sem trabalhar aspectos como iluminação ou qualidade de equipamento. Outro aspecto a destacar é que com a aplicação da técnica de reconhecimento facial “*haarcascade\_frontal\_face\_defaul.xml*” disponível no OPENCV conseguiu-se altas taxas de reconhecimento facial no vídeo. Testes com o modelo “*haarcascade\_frontalface\_cbcl1.xml*” demonstraram que os índices de detecção da face foram extremamente menores, devido principalmente a iluminação.

É importante destacar que no experimento com o banco de dados Cohn-Kanade a escolha aleatória de um grupo de imagens representa uma limitação. Essa escolha poderia ser implementada em versões futuras utilizando-se métodos como validação cruzada, por exemplo. No experimento com vídeos do próprio autor, a principal limitação refere-se à falta de índices globais de acurácia e a análise manual feito pelo autor que poderiam ser obtidos classificando-se todos os quadros por um especialista.

## 5. Conclusões e Trabalhos Futuros

Mesmo com os avanços alcançados nos últimos anos, diversos desafios ainda precisam ser vencidos para a disseminação de aplicações baseadas em técnicas da computação afetiva em ambientes educacionais. Segundo Van Den Broek (2012), as principais barreiras para o avanço da computação afetiva são: definições conceituais relacionadas às emoções, operacionalização baseada nas definições conceituais e mapeamento dos sinais em emoções.

Analisando-se os resultados desse trabalho, mesmo considerando algumas limitações nos experimentos, pode-se concluir que, levando-se em conta o contexto educacional, diversas questões se impõem: i) as emoções experimentadas por estudantes são menos destacadas ou intensas, tornando mais difícil sua identificação ou distinção; ii) o conjunto de emoções básicas não representa completamente as emoções relacionadas à aprendizagem como por exemplo frustração e aborrecimento; iii) as condições de captura de imagens (e.g. iluminação, posição da face) em um ambiente real de aprendizagem não são as mesmas que a utilizada para construção de bancos de dados em laboratório.

Diante disso, conclui-se que muitas barreiras precisam ser vencidas para a evolução do reconhecimento de emoções via expressões faciais aplicadas ao domínio educacional. Uma das necessidades fundamentais refere-se à construção e disponibilização de bancos de dados que tenha foco no conjunto de emoções

relacionadas ao contexto educacional. É fundamental também que esses conjuntos de dados sejam construídos utilizando ambientes mais próximos ao natural, como por exemplo, com estudantes utilizando um software educacional, utilizando técnicas consolidadas para a anotação da base. Além disso, a inclusão de registros detalhados relativos à interação do estudante com o ambiente seria de extrema valia para a construção de modelos cognitivo-afetivos de inferência de emoções.

Finalmente, é importante observar que teorias e pesquisadores da área de emoção e cognição concordam largamente que as emoções nos seres humanos baseiam-se em dois componentes básicos e interligados: 1) emoções são físicas, enfatizando seu componente corporal; e 2) emoções são cognitivas, ressaltando seu componente mental (Picard, 1997). Dessa maneira, considerar esses dois aspectos em conjunto representa ainda um desafio e, ao mesmo tempo, um caminho promissor em investigações futuras com vistas ao aprimoramento de sistemas tutores que reconheçam e adaptem-se de maneira mais afetiva às reações afetivas dos aprendizes.

## Referências

- Alexander, S. T. V. (2007) “An affect-sensitive intelligent tutoring system with an animated pedagogical agent that adapts to student emotion like a human tutor”. Tese de Doutorado. Philosophy in Computer Science, Massey University, Albany, New Zealand.
- Baker, R.S.J.d., Gowda, S.M., Wixon, M., Kalka, J., Wagner, A.Z., Salvi, A., Alevan, V., Kusbit, G.W., Ocumpaugh, J., Rossi, L. (2012) “Towards Sensor-Free Affect Detection in Cognitive Tutor Algebra”. Proceedings of the 5th International Conference on Educational Data Mining. pp. 126–133.
- Bartlett, M. S., Littlewort, G., Frank, M., Lainscsek, C., Fasel, I., Movellan, J. (2006) . Fully automatic facial action recognition in spontaneous behavior. In in 7th International Conference on Automatic Face and Gesture Recognition, pages 223-230.
- Bosch, N., Chen, Y., and D’Mello, S. (2014). It’s written on your face: detecting affective states from facial expressions while learning computer programming. In Intelligent Tutoring Systems, pages 39–44. Springer.
- Calvo, R., D’Mello, S., et al. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *Affective Computing, IEEE Transactions on*, 1(1):18–37.
- D’mello, S.K., Craig, S.D., Witherspoon, A.W., Mcdaniel, B.T., Graesser, A.C. (2008) “Automatic Detection of Learner’s Affect from Conversational Cues”. *User Modeling and User-Adapted Interaction*, p. 45-80.
- D’mello, S.; Graesser, A. (2012). Dynamics of affective states during complex learning. *Learning and Instruction*, v. 22, n. 2, p. 145–157.
- Ekman, P. (1992). An argument for basic emotions. *Cognition & emotion* 6.3-4 169-200.
- Essa, I., Pentland, A. P., et al. (1997). Coding, analysis, interpretation, and recognition of facial expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 19(7):757–763.

- Grafsgaard, J. F., Wiggins, J. B., Boyer, K. E., Wiebe, E. N., Lester, J. C. (2013) “Embodied Affect in Tutorial Dialogue: Student Gesture and Posture”. *Artificial Intelligence in Education, Lecture Notes in Computer Science* vol. 7926, p. 1-10.
- Hoque, M. E., McDuff, D. J., and Picard, R. W. (2012). Exploring temporal patterns in classifying frustrated and delighted smiles. *Affective Computing, IEEE Transactions on*, 3(3):323–334.
- Jaques, P. A., Nunes, M. A. S. N., Isotani, S., Bittencourt, I. (2012) “Computação Afetiva aplicada à Educação: Dotando Sistemas Tutores Inteligentes de Habilidades Sociais”. *Anais do XXXII Congresso da Sociedade Brasileira de Computação, DesafIE 2012*.
- Kanade, T., Cohn, J. F., & Tian, Y. (2000). Comprehensive database for facial expression analysis. Paper presented at the Fourth IEEE International Conference on Automatic Face and Gesture Recognition.
- Khan, F. A., Graf, S., Weippl, E. R., Tjoa, A. M. (2010) “Identifying and Incorporating Affective States and Learning Styles in Web-based Learning Management Systems”. *Interaction Design and Architecture(s) Journal - IxD&A*, N. 9-10, p. 85-103.
- Littlewort, G., Bartlett, M. S., Fasel, I., Susskind, J., Movellan, J. (2004). Dynamics of facial expression extracted automatically from video. In *J. Image & Vision Computing*, pages 615-625.
- Littlewort, G. C., Bartlett, M. S., and Lee, K. (2009). Automatic coding of facial expressions displayed during posed and genuine pain. *Image and Vision Computing*, 27(12):1797–1803.
- McDaniel, B., D’Mello, S., King, B., Chipman, P., Tapp, K., and Graesser, A. (2007). Facial features for affective state detection in learning environments. In *Proceedings of the 29th Annual Cognitive Science Society*, pages 467–472. Cognitive Science Society Austin, TX.
- Mella, L., Bellavista, D. (2016). *Emotime: Recognizing emotional states in faces*. *Elaborazione Delle Immagini LM*. University of Bologna.
- Picard, R. W. (1997) *Affective computing*. Cambridge, Mass: MIT Press.
- Picard, R. W., Fedor, S., Ayzenberg, Y. (2014). Multiple arousal theory and daily-life electrodermal activity asymmetry. *Emotion Review*.
- Sarrafzadeh, A., Alexander, S., Dadgostar, F., Fan, C., Bigdeli, A. (2008) “ ‘How do you know that I don’t understand?’ A look at the future of intelligent tutoring systems”. *Computers in Human Behavior* 24, p. 1342–1363.
- Van Den Broek, E. L. (2012). Affective computing: a reverence for a century of research. In *Cognitive Behavioural Systems*, pages 434–448. Springer.
- Viola, P., Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (Vol. 1, pp. I-511). IEEE.