Reconhecimento dinâmico de emoções através de expressões faciais utilizando árvore de decisão

Adilmar C. Dantas[1], **Sara L. de Melo**[1], **Fábio F. Moura** [2], **Márcia A. Fernandes**[1]

[1]Departamento de Ciência da Computação
Universidade Federal de Uberlândia
Uberlândia – Minas Gerais – Brasil

[2]Departamento de Engenharia e Computação
Instituto Federal de Minas Gerais
Bambuí – Minas Gerais – Brasil
email{akanehar,saraluziamelo,fabioferreirademoura}@gmail.com, marcia@ufu.br

Abstract. *The classification emotions through facial expressions in computing environments has been widely investigated because it allows you to enlarge and refine the interaction between user and system. In academia, this classification can be used in virtual learning environments in order to motivate the student to from his emotion. This paper presents a real-time system for the recognition of seven basic emotions (happiness, surprise, anger, fear, disgust, sadness, and neutral). For this, the detection of the facial expressions was made by inserting points on the face based on the motion units. After extraction of features found in the face, it used the technique Decision Tree for the precise classification of emotions. The developed tool has been integrated into Moodle and the recognition of emotions has been tested and validated with static and videos generated from the Cohn-Kanade database images. The results were satisfactory for all emotions, achieved an overall percentage of* 86.4% *accuracy for recognizing them.*

Resumo. *A classificação de emoções através de expressões faciais em ambientes computacionais tem sido largamente investigada, pois permite ampliar e refinar as interações entre usuário e sistema. No âmbito acadêmico, esta classificação pode ser utilizada em ambientes virtuais de aprendizagem a fim de motivar o estudante à partir de sua emoção. Este trabalho apresenta um sistema em tempo real para o reconhecimento das sete emoções básicas (alegria, surpresa, raiva, medo, desgosto, tristeza e neutra). Para isto, a detecção das expressões faciais foi feita através da inserção de pontos na face tendo como base as unidades de movimento. Após a extração de características na face encontrada, foi utilizada a técnica Árvore de Decisão para a classificação precisa das emoções. A ferramenta desenvolvida foi integrada ao Moodle e, o reconhecimento das emoções foi testado e validado com imagens estáticas e vídeos gerados da base de dados Cohn-Kanade. Os resultados obtidos foram satisfatórios para todas as emoções, atingido um percentual geral de* $86,4\%$ *de acurácia.*

1. Introdução

A computação afetiva investiga como extrair características e ações sociais de indivíduos, tais como, reconhecer automaticamente as emoções, e por conseguinte, responder a cada

DOI: 10.5753/cbie.sbie.2015.1102

uma delas. Com base neste contexto, pesquisas sobre a detecção e classificação das emoções dos estudantes em sistemas tutores inteligentes (STI) vêm sendo desenvolvidas com o objetivo de observar o estudante motivando-o com base nas emoções. Os STIs e a Computação Afetiva podem auxiliar em propostas de solução dos grandes desafios em educação através de métodos e técnicas que permitam desenvolver programas que forneçam assistência individualizada e inteligente ao aluno, proporcionando uma maior interação social [Jaques et al. 2012].

As emoções são a chave para proporcionar melhores formas de efetuar uma comunicação humano-computador. Nessa perspectiva, a computação afetiva pode aprimorar este processo em várias dimensões. Um exemplo de uma aplicabilidade é num curso voltado a computador, onde estudantes, com o auxílio de uma webcam, a qual captura expressões faciais deles, mapeando suas emoções. Dessa maneira é possível inferir como será um melhor método de ensino aplicado à ele, adaptando o conteúdo para o estudante baseado em seu perfil emocional num determinado momento.

Darwin foi um dos primeiros naturalistas a estudar reconhecimento facial com seres humanos como fator social. Essas ideias de Darwin fez com que fossem desenvolvidas teorias sobre o reconhecimento de padrões com finalidade de auxiliar o ser humano em suas atividades [Garcia 2009]. A partir das ideias Darwinistas, surgiu a teoria Facial Action Coding System - FACS [Ekman and Friesen 1978]. Nesta teoria comprova-se que existem emoções universais entre grupos sociais e raças distintas, expressas pelo ser humano.

No trabalho de Ekman e Friesen [Ekman and Friesen 1978] foi descrito todos os possíveis movimentos faciais visíveis ao inferir determinada emoção definidos em 44 Unidades de Ação (UA's). Cada UA tem seu código numérico associado a cada movimento da face, essas UA's foram estudadas e observadas cuidadosamente ao inferir uma emoção. Além disso, foi desenvolvido o modelo Emotion FACS (EMFACS), esse sistema mapeia e seleciona as UA's observadas na teoria FACS com determinada emoção.

Em [Pantic and Rothkrantz 2000] efetuou-se um estado da arte do reconhecimento automático de expressões faciais e técnicas computacionais utilizadas para classificação das emoções. Neste trabalho afirma-se que a detecção dinâmica das emoções através das UA's é um problema desafiador devido a complexidade de detectar todas às 44 UA's expressas no sistema FACS. Assim, [Azcarate et al. 2005], desenvolveu um sistema computacional para reconhecimento da sete emoções básicas baseado teoria FACS. A detecção dos pontos da face foi feita através de Unidades de Movimento (UM's) que estão associadas com algumas UA's reconhecidas por [Ekman and Friesen 1978]. Porém, às UM's não são equivalentes às UA's.

Assim, este trabalho descreve um sistema em tempo real capaz de detectar e classificar as sete emoções básicas(alegria, tristeza, medo, surpresa, desgosto, raiva e neutra) através das expressões faciais. Para isto, após detectar a face humana pelo fluxo do vídeo, os pontos inseridos na face estão relacionados com às UM's propostas por [Azcarate et al. 2005]. Em seguida, efetua-se a classificação das emoções através das UM's detectadas com o classificador Árvore de Decisão para o reconhecimento dinâmico de todas as emoções.

O artigo está organizado da seguinte maneira: na Seção 2 mostra alguns trabalhos

relacionados; na Seção 3 é descrito o modelo dinâmico para reconhecimento das emoções, juntamente com as técnicas computacionais e ferramentas utilizadas para o desenvolvimento do classificador; na Seção 4 são mostrados os testes e validação do sistema, na Seção 5 são mostrados os resultados e, na Seção 6 são apresentadas as considerações finais.

2. Trabalhos Relacionados

No trabalho de [Melo et al. 2014] foi realizado um estudo comparativo referente a três técnicas computacionais - Redes Neurais tipo MLP, Rede Neural Artificial tipo RBF e Redes Bayesianas - com a finalidade de verificar a melhor técnica para a classificação das emoções baseada na teoria *Facial Action Coding System* - FACS [Ekman and Friesen 1978]. Assim, foi feita uma análise ao Banco de Dados (BD) proposto por [Kanade and Cohn 2005]. Este BD é composto por imagens relacionadas a inferência das emoções com diversidade de sexo, etnias e idade. Após a análise feita ao BD escolhido desenvolveu-se uma base de dados composta pelo agrupamento das UA's relacionada a determinada emoção. E, por fim foram executados os experimentos para a classificação das emoções. Os resultados das técnicas foram comparados e obtiveram uma precisão geral de 86%, 75,72% e 71,76% para as Redes Bayesianas, Rede Neural Artificial tipo MLP e Rede Neural Artificial tipo RBF, respectivamente.

No trabalho de [Kohonen et al. 2001] é descrita uma técnica onde as emoções constantemente orientam e modulam a racionalidade, que desempenham um papel essencial na forma como as pessoas comportam de forma inteligente ao interagir com outros seres humanos, bem como máquinas. A técnica descrita por [Kohonen et al. 2001] fornece uma interface entre humanos e máquinas que usam expressões faciais. Esta técnica pode ser utilizada para permitir que as máquinas incorporem uma interpretação das emoções humanas em seus princípios de racionalidade, o que poderia resultar em uma interação mais inteligente com os seres humanos. Nesta técnica, são mostrados 15 valores calculados a partir dos 18 pontos conjunto de imagens faciais. Ele usa abordagem baseada agrupamento e a rede de mapas auto-organizáveis supervisionados para a classificação emoção. A novidade desta técnica é que ela usa uma forma modificada de FACS [Ekman and Friesen 1978] para obter 15 vetores de características faciais de uma imagem. Cinco emoções que foram consideradas: neutra, raiva, feliz, triste e surpresa.

O trabalho de [Azcarate et al. 2005] mostra um sistema de reconhecimento emoções através de expressões faciais exibidas em ao vivo em vídeo e sequências de vídeo. o sistema baseia-se no *Piecewise Bézier Volume Deformation* (PBVD) e foi extendido para um detector rosto, usado para localizar, inicialmente o rosto humano automaticamente. Os experimentos de [Azcarate et al. 2005] usando *Naive Bayes* e classificadores *Tree-Augmented-Naive*. Além disso, o autor associou Unidades de Movimentos com as Unidades de Ação proposta pela teoria FACS. Os testes foram feitos com as pessoas dependente do banco de dados Cohn-Kanade [Kanade and Cohn 2005] e com pessoas independente da base de dados utilizada para treinamento. Os resultados obtidos foram satisfatórios para o reconhecimento das emoções através da expressão facial com as UM's..

O artigo de [Rázuri et al. 2013] centra-se em um sistema de reconhecimento de emoção do ser humano a partir da detecção da face. A informação de saída alimenta a

entrada para uma "máquina" capaz de interagir com as habilidades sociais, no contexto da construção social sistemas inteligentes. A metodologia faz uso de uma técnica de classificação de informações para uma nova imagem fundida que é composta de dois blocos integrados pela área dos olhos e da boca, que são consideradas áreas muito sensíveis às mudanças de expressão do ser humano e que são particularmente relevantes para a decodificação de expressões emocionais. [Rázuri et al. 2013] usam a imagem mesclada como uma entrada que foi propagada usando *feed-forward* de uma rede neural treinada por *backpropagation*. Tal análise de imagens fundidas tornou possível, obter informações relevantes através da combinação dos dados apropriados da mesma imagem e reduzir o tempo de treinamento, enquanto fora preservada a taxa de classificação. Foi demonstrado por [Rázuri et al. 2013] que o algoritmo proposto pode detectar emoção com boa precisão.

No trabalho proposto por [Rao and Koolagudi 2013] foi desenvolvido um sistema de reconhecimento de emoções utilizando características acústicas e faciais. Foi observado que tal combinação evidencia a melhora do desempenho do reconhecimento num nível de 93,62%. O desempenho do sistema de reconhecimento de emoção desenvolvido utiliza modelos de redes neurais com camada oculta, comparado com Modelos de Markov. Os recursos e os modelos propostos são avaliadas em banco de dados interativos de emoções.

A partir desses trabalhos correlatos, surge a ideia de desenvolver um modelo dinâmico para reconhecimento de emoções, denominado *Facial E-motion*, com a finalidade de classificar as emoções básicas do estudante em um ambiente virtual de aprendizagem (AVA) e, ao classificar o estudante com uma emoção prejudicial ao aprendizado o sistema seja capaz de estimular cada estudante com intuito de promover à aprendizagem. Além disso, pode-se relacionar o modelo com os estilos de aprendizagem, conforme proposto no modelo de [Moura et al. 2013].

3. Modelo dinâmico para reconhecimento de emoções

Neste trabalho foi desenvolvido um sistema em tempo real para classificação dinâmica das sete emoções básicas utilizando UM's [Azcarate et al. 2005] para ser integrado em um AVA. As UM's foram utilizadas para insertir os pontos nas regiões de interesse da face encontrada e, em seguida utilizou-se a técnica Árvore de Decisão como classificador para o reconhecimento das emoções. Os principais módulos deste sistema estão detalhados na Figura 1.

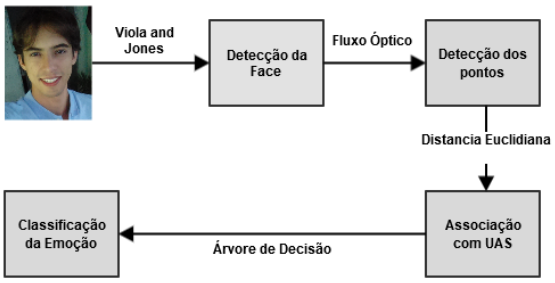


Figura 1. Representação dos módulos do sistema e técnicas computacionais utilizadas.

Numa primeira instância localiza-se a face do estudante, o primeiro módulo denominado "Detecção da Face" foi desenvolvido com auxílio do classificador em cascata proposto por [Viola and Jones 2001] para localizar as regiões de interesse (olhos, boca, nariz e contorno da face). E, além disso, neste módulo para reduzir o tempo computacional utilizou-se o algoritmo Adaboost [Schapire 1990]. Para cada área de interesse foram desenvolvidos classificadores específicos para o tratamento de oclusões na face, onde as situações de oclusão podem ser definidas como a presença de óculos e barba. Os classificadores foram desenvolvidos a partir do *HaarCascade*.

No próximo passo, o módulo "Detecção dos pontos" é feito através da detecção das regiões de interesse na face. O sistema desenvolvido foi implementando para ser integrado *on-line* , assim para que fosse possível a detecção desses pontos em tempo real foi utilizado o algoritmo de Fluxo Óptico para mapear as regiões de interesse cinematicamente. Em seguida, foram identificados e inseridos na face um total de 12 pontos nas regiões de interesse da face adquirida. Os pontos foram inseridos na face conforme as 12 UM's [Azcarate et al. 2005] e são exibidos na Figura 2.

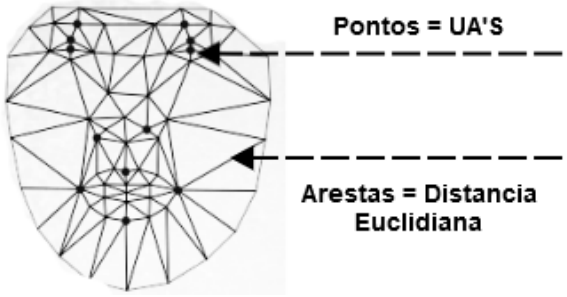


Figura 2. Máscara facial com os 12 pontos inseridos nas regiões de interesse.

Conforme exposto na Figura 2, o sistema ao detectar a face desejada insere 12 pontos na face obtida. Deste modo, foram inseridos 4 pontos na região da boca, 2 pontos na região do nariz, 4 pontos na região dos olhos e, 2 pontos na região da sobrancelha. Através destes pontos calculou-se a Distância Euclidiana entre os pontos referentes às UM's. Este cálculo foi utilizado para determinar se uma determinada UM que está relacionada à uma determinada UA foi inferida em um dado instante. A associação das UA's através das UM's e seus pontos estão descritos na Tabela 1.

Tabela 1. Associação das Unidades de Movimento com Unidades de Ação. [Azcarate et al. 2005]

UM	UM Contraída	Ponto Associado
1	O movimento vertical superior do centro do lábio	P51
2	O movimento vertical inferior do centro do lábio	P57
3	Movimento horizontal do canto esquerdo da boca	P48
4	Movimento vertical do canto esquerdo da boca	P48
5	Movimento horizontal do canto direito da boca	P54
6	Movimento vertical do canto direito da boca	P54
7	Movimento vertical da sobrancelha direita	P19
8	Movimento vertical da sobrancelha esquerdo	P25
9	Levantamento da bochecha esquerda	P02
10	Levantamento da bochecha direita	P12
11	Piscar olho direito	P33
12	Piscar olho esquerdo	P28

Ao observar a Tabela 1 observe que a Unidade de Movimento 1 está associada com o ponto 51 na face, este ponto representa o movimento vertical superior do centro do lábio. Sendo assim, os 12 principais movimentos faciais que representam as sete emoções báscias foram associados com os 12 pontos inseridos na face.

Após a detecção e inferência das UM's, e de realizar o cálculo das Distâncias Euclidianas referente a cada movimento, o próximo passo a ser feito é a classificação das sete emoções através dessas caracteríticas obtidas. Assim, foi utilizado a técnica Árvore de Decisão como um classificador para que dessa maneira fosse possível efetuar o reconhecimento automático das emoções.

Uma Árvore de Decisão é uma estrutura composta por três elementos básicos, sendo eles: nós de decisão, que corresponde aos atributos; bordas ou ramos, que correspondem aos diferentes valores de atributos possíveis e, as folhas que incluem objetos que pertencem à mesma classe ou que são muito semelhantes. Assim, é possivel induzir regras de decisão que serão usadas para classificar novas instâncias. Cada um dos caminhos a partir da raiz para uma determinada folha corresponde a um conjunto de atributos de teste. A maioria das árvores de decisão é composta por dois procedimentos principais: a construção (de indução) e os procedimentos de classificação (inferência)[Jenhani et al. 2008].

Para a classificação das emoções através desta técnica computacional cada nó corresponde a determinada emoção que deseja classificar, os ramos são compostos pelo cálculo das distâncias euclidianas referente às seis emoções básicas e, por fim as folhas representam a emoção induzida à partir dos ramos da árvore. A Figura 3 mostra a Árvore de Decisão para a classificação das emoções *desgosto* e *raiva*.

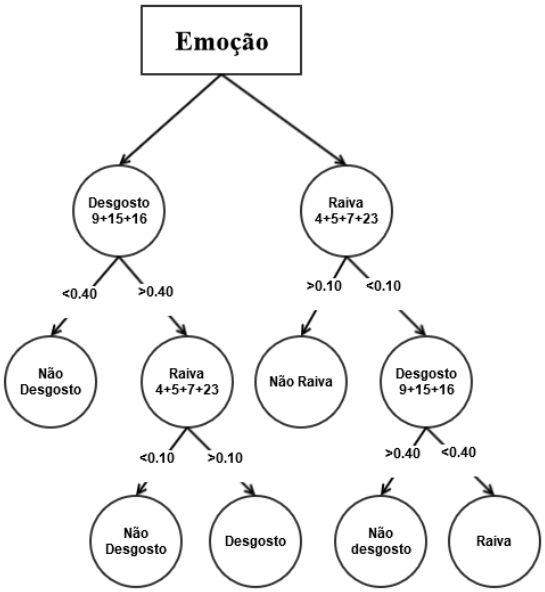


Figura 3. Árvore de Decisão para classificação das emoções.

Nesta ramificação da árvore o nó principal é a emoção a qual deseja-se classificar, os ramos são as Distâncias Euclidianas obtidas através das UM's assim, essas distâncias

são conduzidas até as folhas através dos arcos que contém os pesos que satisfazem determinada condição de classificação. Algumas emoções possuem mais ramificações do que as outras pelo fato dessas emoções possuírem unidades de movimento semelhantes em diferentes emoções .

4. Testes e Validação do Sistema

Para o treinamento da árvore foi desenvolvido um conjunto de testes composto por 20 imagens extraídas da base de dados [Kanade and Cohn 2005] relacionadas a cada uma das seis emoções. Essa base possui imagens de diferentes etnias e sexo, além disso, essas imagens estão em sequências que representam cada emoção no seu estado inicial até inferir o máximo de cada emoção.

A seleção das imagens para o treinamento foi feita utilizando um agrupamento das imagens com início, meio e fim da sequência que representa a emoção. Esta escolha foi utilizada para evitar que árvore de decisão construa pesos somente para a emoção em seu estado máximo ou mínimo, podendo assim obter uma classificação errônea.

Para efetuar testes com imagens foram utilizadas 1452 imagens. Para os testes com vídeos foi necessário um trabalho minucioso na base de dados [Kanade and Cohn 2005]. Primeiramente, separou imagens por emoção e por pessoa e, em seguida foi construído uma sequência de frames que compõe cada vídeo. Assim, foram gerados 194 vídeos, a Tabela 2 mostra a quantidade de imagens e vídeos referente a cada emoção utilizadas para testes de classificação das emoções no modelo proposto.

Tabela 2. Total de imagens e vídeos gerados para cada emoção

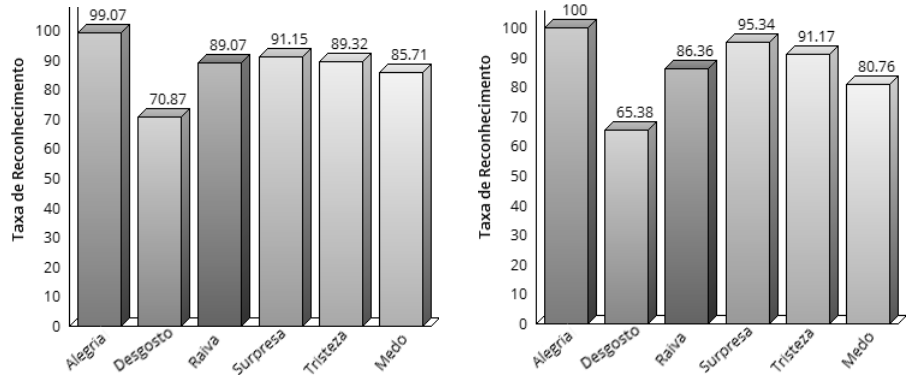
Emoção	Imagens	Vídeos
Alegria	324	43
Tristeza	253	34
Raiva	183	22
Surpresa	328	43
Desgosto	182	26
Medo	182	26
Total	1452	194

Através das imagens e vídeos relacionadas a cada emoção o sistema foi implementado e, a validação foi feita através de duas parte de testes feitos com imagens e vídeos. Os resultados obtidos para classificação das emoções, exceto para a emoção neutra, para os dois testes estão mostrados a seguir.

5. Resultados

A Figura 4 mostra dois gráficos comparativos com resultados obtidos através dos testes efetuados com imagens e vídeos. Note que, os resultados para classificação das emoções foram satisfatórios tanto para emoções expressas por imagens quanto por vídeos.

Uma exceção é para a emoção “desgosto” que obteve menor acurácia comparada as demais nos dois diferentes testes. Isso se deve ao fato desta emoção possuir muitas unidades de ação correlacionadas com outras emoções, gerando assim, falsos positivos durante o reconhecimento das emoções.



a) Taxa de Reconhecimento por Imagem. b)Taxa de Reconhecimento por Vídeo.

Figura 4. Gráfico comparativo de Taxa de Reconhecimento por Imagem e Vídeo.

Os resultados parciais para os testes com imagens obtiveram, no geral, uma média de acertos de 87,53%. Para a emoção "feliz" o total de acertos foi de 99,07%, para a "tristeza" 70,87%, para a "raiva" o total de acertos foi de 89,07%, o "desgosto" obteve um total de 91,15% acertos, "surpresa" 89,32% acertos e a emoção "medo" obteve um total de acertos na margem de 85,71%. Para a emoção "desgosto" num total de 182 imagens, 4 imagens foram confundidas com a emoção "raiva", 2 com a emoção "medo" e 47 para "tristeza", conforme mostrado na Tabela 3, que expõe a matriz de confusão obtida para os testes com imagens, podendo-se observar os demais resultados.

Tabela 3. Matriz de confusão obtida com testes para imagens.

Total de imagem	Alegria	Tristeza	Desgosto	Surpresa	Raiva	Medo	Neutro
324	**321**	0	0	0	0	0	3
253	0	**233**	7	0	8	12	0
182	0	47	**129**	0	4	2	0
328	0	2	10	**299**	0	17	0
183	0	12	8	0	**163**	0	0
182	17	2	6	0	1	**156**	0

Os testes com vídeos obtiveram, no geral, uma média de acertos de 86.50%. Para a emoção "feliz" o total de acertos foi de 100,00%, "tristeza" 91,17% de acertos, "raiva" 86,36% de acerto, o "desgosto" obteve um total de 65,38% acertos, "surpresa" 95,34% acertos e, "medo", o total de acertos foi de 80,76%. Para a emoção "tristeza" – exemplo, linha 2 – foram usados um total de 34 vídeos, onde foram confundidas 3 vídeos para a emoção "desgosto", conforme mostrado na Tabela 4, que mostra a matriz de confusão obtida para os testes com vídeos, podendo observar os demais resultados.

Tabela 4. Matriz de confusão obtida com testes para vídeos.

Total de vídeos	Alegria	Tristeza	Desgosto	Surpresa	Raiva	Medo
43	**43**	0	0	0	0	0
34	0	**31**	3	0	0	0
26	0	4	**17**	0	4	1
43	0	0	0	**41**	2	0
22	0	0	0	3	**19**	0
26	0	0	0	2	3	**21**

Ao analisar os testes, é possível observar que o reconhecimento das emoções

através das Unidades de Movimentos também é uma fator complexo, pois além de detectar os pontos na face é necessário tratar a presença de pontos iguais em emoções diferentes. Mas, para as outras emoções os resultados foram satisfatórios e, a partir deles é possível reconhecer com precisão as emoções básicas. Este sistema de reconhecimento dinâmico de emoções foi integrado ao [Moodle 2015] com a finalidade de detectar as emoções do estudante.

6. Considerações Finais

Este trabalho apresentou um sistema em tempo real para detecção e classificação das emoções básicas. Para isto, utilizou as unidades de movimento para a detecção dos pontos na face e, a classificação das expressões faciais em emoções foi feita através do cálculo da Distância Euclidiana e Árvore de Decisão.

O sistema foi implementado, validado e integrado ao ambiente virtual de aprendizagem com intuito de reconhecer as emoções de cada estudante durante o processo de aprendizagem e, à partir disso observar o comportamento do estudante através de suas emoções. Os resultados obtidos através dos testes feitos por imagens e vídeos foram satisfatórios. Como trabalho futuro pretende-se efetuar testes com estudantes e observá-lo durante o processo de ensino e de aprendizagem.

Referências

Azcarate, A., Hageloh, F., Sande, K. V., and Valenti, R. (2005). Automatic facial emotion recognition. *Universiteit van Amsterdam*.

Ekman, P. and Friesen, W. (1978). Facial action coding system. *Alto: Consulting Psychologist Press*.

Garcia, L. S. L. (2009). A expressão das emoções no homem e nos animais. In *Traduzido por: Leon de Souza Lobo Garcia – A Origem das Espécies, publicado por Charles Darwin – 1859*, pages 33–55, São Paulo, Brasil. Companhia de Bolso. http://www.letras.ufmg.br/arquivos/matte/bib/darwin.pdf.

Jaques, P., Nunes, M. A. S., Isotani, S., and Bittencourt, I. (2012). Computação afetiva aplicada a educação: Dotando sistemas tutores inteligentes de habilidades sociais. In *Anais do Workshop de Desafios da Computação Aplicada a Educação*, pages 50–59.

Jenhani, I., Amor, N. B., and Elouedi, Z. (2008). Decision trees as possibilistic classifiers. *International Journal of Approximate Reasoning*, 48(3):784–807.

Kanade, T. C. and Cohn, J. (2005). Au-coded facial expression database.

Kohonen, T., Schroeder, M. R., and Huang, T. S., editors (2001). *Self-Organizing Maps*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 3rd edition.

Melo, S. L., Moura, F. F., Macedo, K., Alves, F. S. R., and Fernandes, M. A. (2014). Estudo comparativo de técnicas computacionais para classificação de emoções. *XXV Simpósio Brasileiro de Informática na Educação (SBIE 2014)*, 25:456–465.

Moodle (2015). Moodle. http://www.moodle.org/.

Moura, F. F., Franco, L. M., and Melo, S. L. Fernandes, M. A. (2013). Evolução das inteligências múltiplas pelo método da espiral de aprendizagem de kolb utili-

zando pso. *XXIV Simpósio Brasileiro de Informática na Educação*. http://www.br-ie.org/pub/index.php/sbie/article/view/1794/1555.

Pantic, M. and Rothkrantz, L. J. M. (2000). Automatic analysis of facial expressions: The state of the art. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 22(12):1424–1445.

Rao, K. S. and Koolagudi, S. G. (2013). Recognition of emotions from video using acoustic and facial features. *Signal, Image and Video Processing*, 6:1863–1703.

Rázuri, J. G., Sundgren, D., Rahmani, R., and Cardenas, A. M. (2013). Automatic emotion recognition through facial expression analysis in merged images based on an artificial neural network. In *MICAI (Special Sessions)*, pages 85–96. IEEE.

Schapire, R. E. (1990). The strength of weak learnability. *Mach. Learn.*, 5(2):197–227.

Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. pages 511–518.