

Determinação da Emoção Demonstrada pelo Interlocutor

Rafael A.M. Gonçalves¹, Diego R. Cueva¹, Marcos R. Pereira-Barretto¹, Fabio G. Cozman¹

¹Departamento de Engenharia Mecatrônica e Sistema Mecânicos
Escola Politécnica da Universidade de São Paulo (USP)
Av. Prof. Melo Moraes 2231 – São Paulo - SP – Brazil

marcos.barretto@poli.usp.br

Abstract. *This paper introduces a new model for the determination of conveyed emotion during a human-machine interaction, based on Kalman filtering of instantaneous facial expressions and the emotional trajectory over an emotional surface.*

Resumo. *Este artigo apresenta um modelo para a determinação da emoção demonstrada pelo interlocutor humano em uma interação homem-máquina de longa duração, a partir da expressão facial, baseado na trajetória emocional sobre uma determinada superfície, a partir de sinais instantâneos capturados e submetidos a um filtro de Kalman.*

1. Introdução

Na interação entre pessoas, a comunicação não verbal, como a realizada pelo gestual e pela demonstração de emoções, tem um papel de grande importância. É, portanto, de esperar-se que a comunicação homem-máquina possa também beneficiar-se deste tipo de comunicação. De fato, aplicações modernas incluem comandos gestuais ou por voz, tendo, talvez, iniciado uma revolução na chamada experiência centrada no usuário. O sucesso de tais aplicações é indício concreto de que o ser humano deseja e se beneficia de sistemas capazes de processar uma interação mais natural, o que é base para a Computação Afetiva (*Affective Computing*) e da Robótica Sociável (*Sociable Robots*).

Estudos sustentam que a emoção demonstrada na voz do interlocutor é corretamente reconhecida em aproximadamente 60% dos casos, enquanto que a demonstrada na face é reconhecida em 70% a 98% dos casos [Piccard, 1997], [Ekman, 2003].

No início da década de 1970, Ekman e colaboradores realizaram estudos extensivos sobre expressões faciais, desenvolvendo um método de descrição e identificação de unidades de atuação (*Action Units, AUs*) em expressões faciais, publicado sob a sigla FACS (*Facial Action Coding System*) [Ekman, 1978]. A universalidade dos *AUs*, apesar da forte contestação por quase duas décadas, encontrou suporte em estudos interculturais e com populações pré-literatas, levando à sua admissão [Ekman, 2003].

Os trabalhos acima inspiraram diversas abordagens computacionais de classificação a partir de fotografias ou sequências de vídeo como exposto por [Pantic et al., 2000]. Dentre os mais recentes, destacam-se o eMotion, desenvolvido pela Universiteit van Amsterdam [Azcarate, 2005] e o FaceDetect, desenvolvido no Fraunhofer Institute [Fraunhofer, 2011]. Todavia, a identificação de expressões emocionais em uma foto ou uma curta sequência de quadros em um vídeo não permite uma estimativa da emoção demonstrada pelo interlocutor durante uma interação; apenas

fornece indicações instantâneas e, como será demonstrado, ruidosas sobre ela. Assim, a detecção de um sorriso, em si, não é conclusiva para que o sistema classifique o usuário como feliz. Também é necessário salientar que as técnicas para avaliação de expressões faciais não contemplam a sobreposição com os efeitos de deformação da face induzidos durante a fala. Logo, aplicações como o eMotion ou o FaceDetect não são treinadas para estes casos e apresentam baixo desempenho quando utilizadas em sequências mais longas, de alguns segundos de interação. Um exemplo ilustrativo do problema pode ser observado pelas saídas do aplicativo eMotion quando aplicado ao vídeo, com quadros ilustrativos na Figura 1.



Figura 1. Imagens do vídeo ilustrativo. As imagens, da esquerda para a direita, foram classificadas pelo eMotion como felicidade (100% de certeza, aos 2,36s), tristeza (70,6%, 0,76s), medo (83,1%, 3,92s) e raiva (76%, 3,04s).

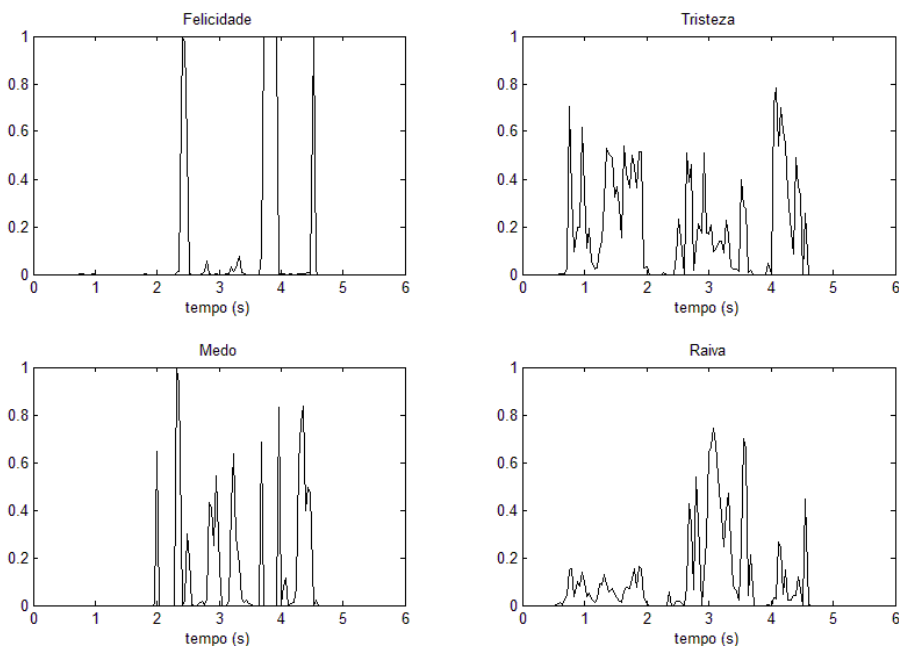


Figura 2. Saídas do eMotion para o vídeo ilustrativo.

Ao analisar as saídas do sensor (Figura 2), um observador humano possivelmente não chegaria a nenhuma conclusão ou até poderia concluir pela predominância da Tristeza. Entretanto, ao assistir ao vídeo, o observador humano facilmente concluiria que a emoção demonstrada é a Raiva.

Neste trabalho é apresentado um modelo para a determinação da emoção demonstrada pelo interlocutor, no qual as indicações instantâneas identificadas pelo aplicativo eMotion são, em certo sentido, integradas no tempo para a determinação da emoção predominante durante uma interação. Na sequência, apresenta-se uma revisão de trabalhos relacionados, para, a seguir, introduzir-se o modelo proposto e os resultados obtidos.

2. Referencial conceitual

As teorias comportamentalistas (*behavioural theories*) dominaram a cena acadêmica da Psicologia entre os anos 30 e 50. De acordo com estas teorias, as emoções eram não diferenciáveis, sendo apenas uma dimensão do comportamento correspondendo ao grau de energia ou atividade. O caráter determinístico e unidimensional da associação evento-emoção é um dos pilares destas teorias: determinado evento causará determinada emoção. As Teorias Cognitivas (*appraisal theories*) assumiram o centro das discussões em meados dos anos 80, embora as ideias iniciais tenham sido lançadas já nos anos 60. De forma simples, determinam que as emoções são elicitadas a partir de avaliações (*appraisals*) de eventos e situações [Roseman, 2001]. Em outras palavras, admitem a existência de processos de elicitação de emoções e que tais processos são comuns a todas as pessoas. Varia a forma com que cada processo se desenvolve individualmente, embora os processos sejam sempre os mesmos. Desta forma, oferecem um modelo que explica (ou descreve) as diferenças individuais, ao mesmo tempo em que determina os aspectos comuns a todas as pessoas.

Este trabalho baseia-se no conceito de *emoção* advindo das Teorias Cognitivas. Segundo Scherer (2001), emoções são “... *an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism*”.

A visão orientada a processos das Teorias Cognitivas levou, de forma natural, à modelagem usando conceitos de dinâmica de sistemas e à ideia de que a elicitação das emoções corresponde ao deslocamento sobre uma superfície, como a da fig.3, discutida originalmente em [Zeeman, 1976] e, depois, por exemplo, [Stewart, 1983], [Scherer, 2000], [van der Maas, 1992], [Sander, 2005].

A Teoria da Catástrofe, que leva à trajetória da Figura 3, foi central para a concepção do modelo apresentado neste artigo, embora não seja diretamente utilizada porque o foco deste artigo é na emoção *demonstrada* pelo interlocutor e não naquela *sentida* pelo receptor dos eventos. A emoção demonstrada permite modelos mais simplificados, como o aqui apresentado.

A utilização das expressões faciais para a inferência da emoção demonstrada fundamenta-se nos estudos de Ekman e seus colaboradores [Ekman, 2003]. As fortes evidências da universalidade das expressões faciais demonstradas permitem que se possa assumir a hipótese de que estas fornecem uma riqueza de indicações sobre as quais um estimador universal pode ser projetado.

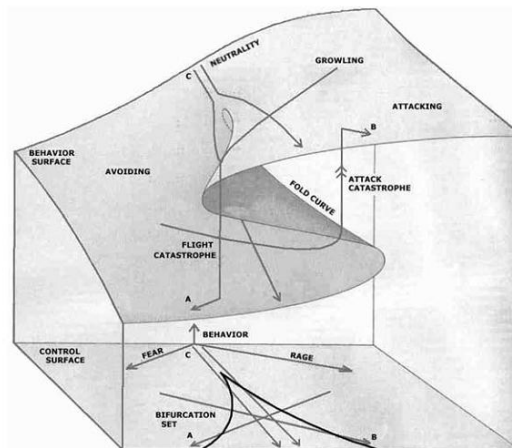


Figura 3. Superfície emocional [Zeeman, 1976].

3. Modelo proposto

O modelo proposto para a estimação da trajetória emocional do interlocutor a partir da análise das expressões faciais baseia-se na modelagem de uma partícula, inicialmente em repouso, sobre uma curva, sujeita a efeitos que alteram sua velocidade instantaneamente e proporcionalmente à probabilidade de detecção de uma expressão facial de conteúdo emocional.

De maneira geral, a curva sobre a qual a partícula se movimenta pode ser estendida a uma superfície N-dimensional contínua, a qual é, portanto, denominada superfície de dinâmica emocional (SDE), pois sua topologia representa os fenômenos envolvidos na evolução da estimação dos estados emocionais e também nas transições entre estes.

As principais vantagens da utilização da SDE no modelo são a possibilidade de abstração das dinâmicas envolvidas e a separação entre o problema de filtragem de sinais de entrada e o de representação dos fenômenos psicológicos das variações dos estados emocionais humanos. Assim, a modelagem de SDE poderia ser entendida como um estudo pertencente ao campo da psicologia onde ao menos duas questões afetam diretamente o desenvolvimento de robôs sociáveis e interfaces inteligentes: a parametrização de uma superfície comum, representativa das interações sociais universalmente aceitas, e a criação de experimentos capazes de reconhecer parâmetros e características das SDE individuais relativos a conceitos como temperamento e personalidade.

Sobre a SDE são posicionados atratores correspondendo a cada emoção a ser percebida.

A partícula, portanto, movimenta-se livremente sobre a SDE segundo efeitos de atração e escoamento, em direção ao estado neutro. Os efeitos de atração compreendem a intensidade com a qual a partícula é atraída para as projeções dos atratores na superfície. O efeito de escoamento é compreendido como a velocidade com a qual a partícula escoia sobre a superfície no sentido de assentá-la em mínimos locais. Assim, a velocidade da partícula é dada pela Eq.I.

$$V_p = V_e + \sum_{a=1}^N V_a \quad (\text{I})$$

onde V_p é a velocidade da partícula, V_e a sua velocidade inicial e V_a , a velocidade na direção de cada um dos atratores, computada de forma vetorial.

A intensidade instantânea de atração é dada por um detector instantâneo de emoções, como os de expressões faciais citados anteriormente, ou outros, como os baseados em EEG, condutância da pele, etc. Esta intensidade instantânea pode ser pré-filtrada, para eliminação de ruídos aleatórios, por exemplo.

4. Exemplo

Embora este trabalho apresente o conceito de superfície de dinâmica emocional como uma ferramenta para possibilitar a abstração, o estudo de uma SDE representativa do ser humano foge ao escopo dos trabalhos iniciais relacionados à obtenção de sinais e sua filtragem. Desta maneira, para fins de demonstrar o modelo proposto e realizar experimentos como prova de conceito, optou-se pela utilização de uma superfície simples. Respeitando as condições de continuidade e de existência de ao menos um ponto de mínimo, a superfície escolhida foi um parabolóide na forma:

$$z = f(x, y) = ax^2 + by^2 \quad (\text{II})$$

$$\gamma(x, y) = (x, y, ax^2 + by^2) \quad (\text{III})$$

$$a = b = 0,6 \quad (\text{IV})$$

Da mesma forma, o estudo detalhado do posicionamento dos atratores com relação aos aspectos psicológicos foge ao tema central. Assim, adotou-se a configuração mostrada na Tabela 1.

Tabela 1. Posicionamento dos Atratores

Emoção	Atrator
Felicidade	[60, 60, 0]
Raiva	[-60, 60, 0]
Tristeza	[-60, -60, 0]
Medo	[60, -60, 0]

Note-se que a localização dos atratores não corresponde ao seu posicionamento no plano Arousal X Valence (Plano AV) [Gunes, 2006]. De fato, a emoção Medo foi posicionada no quarto quadrante ao invés do terceiro, próxima à origem, seu local no Plano AV. Ressalta-se este aspecto para destacar que uma SDE, assim como a mostrada na Figura 3, não precisa ter correspondência direta com o Plano AV. Tal escolha teve

por base o modelo de [Zeeman, 1976] onde medo e raiva são comportamentos ortogonais e mostrou-se viável durante as fases de aprendizado e testes.

Um dos componentes centrais do modelo proposto é a decomposição das velocidades de atração em componentes X e Y, respeitando a topologia da superfície. As informações fornecidas pelos sensores, após pré-filtragem, são interpretadas no modelo como velocidades sobre a trajetória, S, que liga a posição atual da partícula, \vec{P} , à posição do atrator relativo à emoção analisada projetado na SDE, \vec{A} . Dessa forma, cabe observar que, a rigor, o módulo dos atratores não tem significado especial. Assim, para cada atrator:

$$\overline{AP} = \vec{A} - \vec{P} = [a_{px}, a_{py}, a_{pz}] \quad (\text{V})$$

$$S(x) = \gamma(x, rx) \quad (\text{VI})$$

$$r = \left| \frac{a_{py}}{a_{px}} \right|, a_{px} \neq 0 \quad (\text{VII})$$

$$escala = \left| \frac{dS(x)}{dx} \right| = \left| \frac{d\gamma(x, rx)}{dx} \right| = \sqrt{1 + r^2 + [2(a + br^2) * P_x]^2} \quad (\text{VIII})$$

$$V_{x,a} = \frac{dS}{dt} * \frac{\text{sin}(a_{px})}{escala} * \frac{dS}{dx} \cdot \vec{i} \quad (\text{IX})$$

$$V_{y,a} = \frac{dS}{dt} * \frac{\text{sin}(a_{py})}{escala} * \frac{dS}{dx} \cdot \vec{j} \quad (\text{X})$$

Após a decomposição dos efeitos de atração para cada uma das quatro emoções consideradas, o equacionamento para a modelagem do movimento da partícula sobre a SDE é bastante simplificado e foi omitido, por questões de espaço.

Enquanto a SDE pode ser discutida em nível conceitual, a pré-filtragem é fortemente relacionada com o sensor escolhido. Para validação do modelo, escolheu-se o aplicativo eMotion, para a análise quadro a quadro das expressões faciais. Observou-se que este fornecia dados com ruído e também tendência às emoções felicidade e, principalmente, tristeza. Após cautelosa revisão das publicações relativas à área de identificação de expressões, percebeu-se que o problema de eliminação dos ruídos ocasionados pela deformação da face durante a fala não é abordado. Foi possível identificar que a concavidade do arco definido pela boca como principal atributo na detecção de felicidade ou tristeza pelo eMotion. Assim sendo, adotou-se um filtro capaz de analisar as leituras levando em consideração diferentes confiabilidades da informação seria adequado ao tratamento dos dados, o que levou naturalmente ao filtro de Kalman.

A implementação do filtro de Kalman contempla a utilização de um modelo linear subjacente. Posto isso, optou-se por um modelo de primeira ordem da forma apresentada na Equação XI.

$$\frac{Y(s)}{R(s)} = \frac{K_k}{\tau s + 1} \quad (\text{XI})$$

Nesta forma, o filtro se apresenta como um sistema dinâmico de primeira ordem de modo a garantir que o sinal filtrado também seja contínuo. Mais do que isso, o sistema dinâmico constitui ainda um filtro passa-baixas que pode ser ajustado de forma a atenuar os ruídos. Esta escolha de projeto reduz a filtragem a sistemas do tipo *SISO* (*Single-Input/Single-Output*) e, portanto, não leva em consideração a influência entre os sinais.

Após a implementação dos filtros, valores para as variâncias de entrada (Q) e de saída (R) do filtro de Kalman foram ajustadas mediante testes com os vídeos de amostra. Para o sistema subjacente escolheu-se configurar frequências e ganhos idênticos para os quatro filtros. A Tabela 2 traz os parâmetros utilizados.

Tabela 2. Parâmetros dos Filtros de Kalman

	Q	R	K_k	τ
Felicidade	0,1	0,080	5	1,5
Raiva	0,1	0,100	5	1,5
Tristeza	0,1	0,035	5	1,5
Medo	0,1	0,010	5	1,5

É válido ressaltar que, uma vez que a implementação da rotina de MatLab tem como objetivo execução em tempo real- processando os dados à medida que chegam- os filtros de Kalman devem resolver a equação de Riccati iterativamente durante a execução. Nessa situação, uma estimativa inicial da variância e do ganho de Kalman é necessária para o início do processo. Seguindo uma prática comum, tais estimativas são configuradas como nulas.

5. Experimentos

5.1. Metodologia

O processo experimental computacional iniciou-se com a seleção de vídeos representativos das emoções. Após assistir todos os vídeos do *corpus eNTERFACE'05 Audio-Visual Emotion Database* [Martin, 2005], foram selecionados aqueles para os quais os autores deste trabalho julgaram que o conteúdo emocional transmitido pelas expressões faciais era condizente com a categorização do arquivo na base. Para tais vídeos, realizou-se o ajuste da malha e a análise. Durante o processo foi também observado se houve, em algum quadro, descolamento da malha, ocasionado por perda de algum dos pontos de referência da face. A ocorrência de tal fenômeno desqualificava os resultados, ocasionando exclusão do vídeo da base de testes. A alta taxa de descolamento, não reportada previamente pelos criadores do corpus, foi atribuída

principalmente às deformações faciais ocasionadas pela fala. Este método de seleção resultou na escolha de trinta vídeos.

Os dados quadro a quadro de todos os vídeos foram, em seguida, processados pela rotina desenvolvida em MatLab, a qual realizou dois tipos de tratamento nos sinais: filtro de média com janelamento igual à taxa de quadros do vídeo e aplicação do filtro de Kalman, apresentado anteriormente. Foram calculadas as médias dos dados conforme recebidos e obtidos após filtragem. Em todos os casos houve cuidado especial para análise da máxima variância do filtro de Kalman e também do máximo ganho de inovação (ganho de Kalman), garantindo a convergência do filtro.

Finalmente, utilizando o sinal pós-filtro de Kalman, determinou-se a trajetória sobre a SDE.

5.2. Aplicação ao vídeo ilustrativo

A aplicação dos filtros ao vídeo ilustrativo, cujas saídas em bruto foram mostradas na Figura 2, está demonstrada na Figura 4, onde a linha contínua representa o filtro de Kalman e a linha tracejada, o filtro de média móvel.

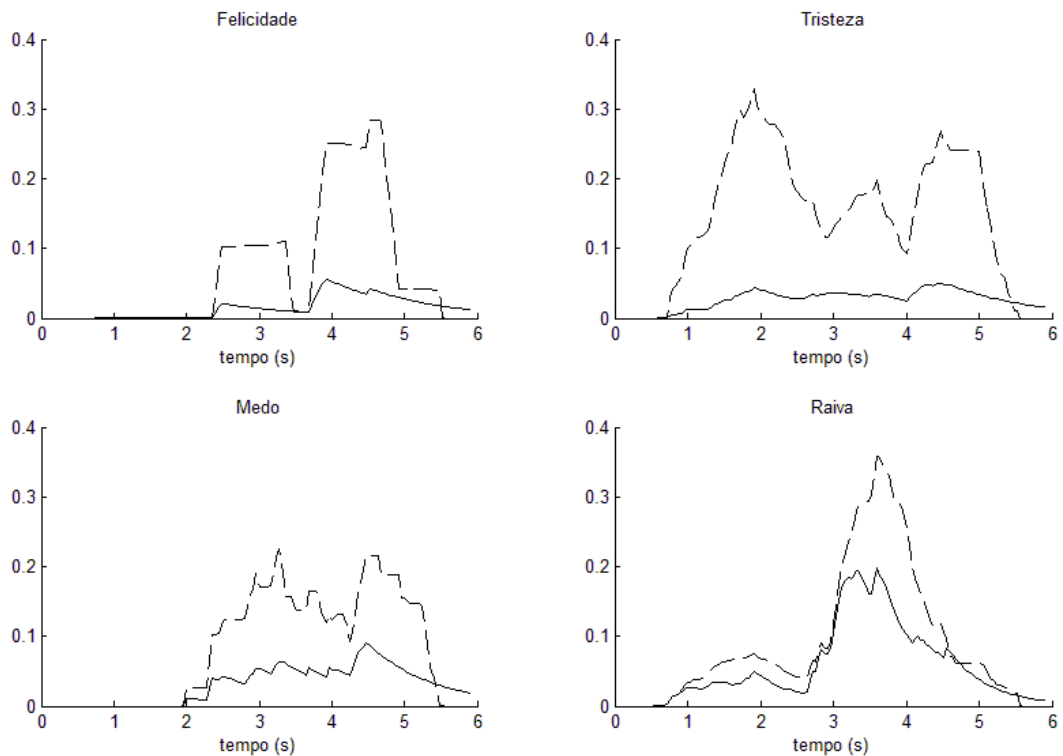


Figura 4. Saída do filtro de Kalman e do filtro de media móvel para o video ilustrativo.

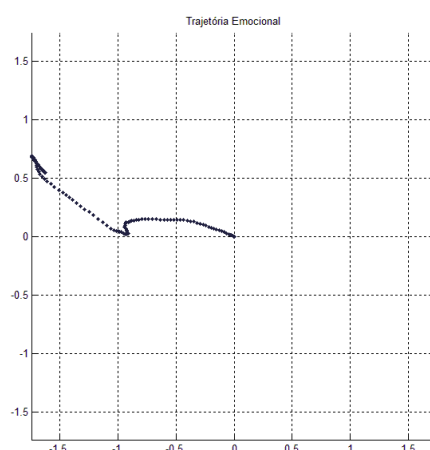
A Tabela 3 apresenta média e desvios padrão normalizados para os dados pré e pós-filtros.

Tabela 3. Análise estatística para o vídeo ilustrativo

Emoção	Original		Média Móvel		Kalman	
	μ	σ	μ	σ	μ	σ
Felicidade	0.175	0.634	0.175	0.237	0.114	0.127
Tristeza	0.377	0.532	0.377	0.254	0.207	0.108
Medo	0.211	0.544	0.211	0.206	0.234	0.203
Raiva	0.236	0.434	0.236	0.257	0.445	0.434

Como se pode perceber, a classificação correta, Raiva, é determinada pelo sistema com filtro de Kalman.

A Figura 5 mostra a trajetória sobre a superfície SDE e ilustra também como o sistema permite estimar a cada instante o estado emocional do interlocutor.

**Figura 5. Trajetória sobre a superfície SDE para o vídeo ilustrativo.**

A trajetória desenvolve-se no segundo quadrante, onde está posicionado o atrator Raiva. Há momentos mais fortes de tristeza, como se pode perceber.

5.3. Resultados

O vídeo ilustrativo, assim como os demais vídeos utilizados neste trabalho, foram obtidos a partir do *corpus eNTERFACE'05 Audio-Visual Emotion Database* [Martin, 2005], um banco de dados de cenas de curta duração, no qual indivíduos são convidados a expressar uma frase emocional da forma que mais desejarem.

Dos 30 vídeos selecionados, 16 foram utilizados para ajuste dos parâmetros do modelo. O experimento foi repetido para demais 14 vídeos selecionados. Com exceção de um caso (número 14), a trajetória desenvolveu-se corretamente nas amostras, como demonstrado na Tabela 4 onde classificação declarada é aquela registrada no corpus como correspondente ao vídeo e classificação avaliada é aquela obtida com o método descrito.

Tabela 4. Resultados do experimento

#	Arquivo	Classificação		#	Arquivo	Classificação	
		Declarada	Avaliada			Declarada	Avaliada
1	S1sa1	Tristeza	Tristeza	8	S43an4	Raiva	Raiva
2	S38an1	Raiva	Raiva	9	S43fe2	Medo	Medo
3	S38fe3	Medo	Medo	10	S42fe1	Medo	Medo
4	S42sa1	Tristeza	Tristeza	11	S43sa1	Tristeza	Tristeza
5	S43ha1	Felicidade	Felicidade	12	S43sa3	Tristeza	Tristeza
6	S43an2	Raiva	Raiva	13	S43sa4	Tristeza	Tristeza
7	S43an3	Raiva	Raiva	14	S43sa5	Tristeza	Raiva

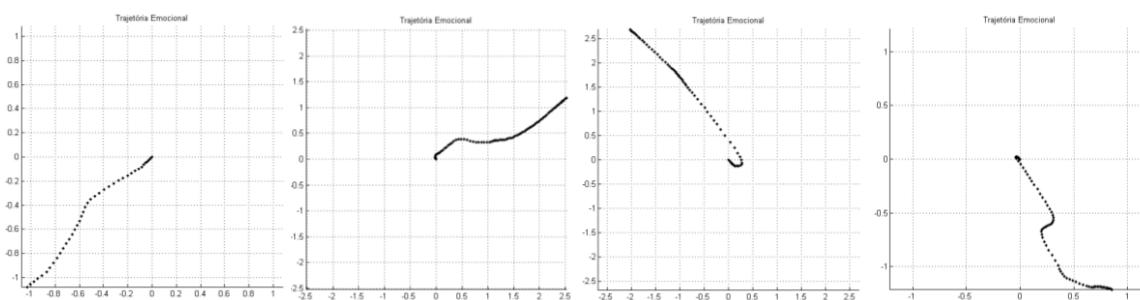


Figura 6. Trajetórias sobre a superfície SDE para vídeos 1, 5, 7 e 10.

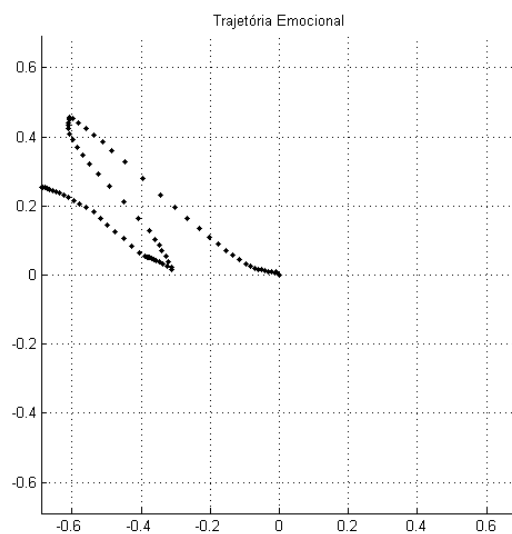


Figura 7. Trajetória sobre a superfície SDE para vídeo 14. Análise reportou oscilações em Raiva para um vídeo de Tristeza.

5.4. Discussão

A comparação entre a análise do sinal como provindo do eMotion e o sinal filtrado, seja pelo filtro de média móvel ou o filtro de Kalman, fornece subsídios para afirmar-se que os processos de filtragem possibilitaram uma classificação mais precisa dos vídeos. O uso de uma trajetória sobre uma superfície permite acompanhar a evolução da emoção no tempo, determinando-se não somente sua característica mais dominante como também permitindo que variações possam ser identificadas.

A experimentação também demonstrou que o “sensor” (no caso, o aplicativo eMotion) é um ponto crítico para a construção de sistemas robustos de determinação da emoção demonstrada pelo interlocutor. Em particular, acredita-se que sensores capazes de estimar mais adequadamente a face neutra do interlocutor e também de tratar as deformações na face induzidas pela fala, possibilitariam grandes ganhos no desempenho.

6. Conclusão

Este trabalho apresentou um método para modelagem e estimação da emoção demonstrada por um interlocutor, utilizando indicações fornecidas pela observação de expressões faciais, capturadas e analisadas preliminarmente por uma aplicação comercial.

O sistema proposto foi implementado em MatLab e testado sobre uma base de dados já existente e disponível para pesquisa, obtendo resultados significativamente melhores à interpretação dos sinais conforme obtidos do sensor. A implementação, mesmo com o Matlab, é executada em tempo real, o que fornece a certeza de que é aplicável em Robótica Sociável e Computação Afetiva.

Em trabalhos futuros, os autores acreditam que seja necessário implementar um sensor capaz de lidar com as principais fontes de ruído: a estimação da posição neutra da face e as deformações durante a fala. Além disso, alternativas de fusão sensorial inteligente e otimização das escolhas de parâmetros dos filtros oferecem possibilidade de melhor desempenho.

Por fim, os autores acreditam que a determinação de superfícies de dinâmica emocional possibilitariam modelar fenômenos mais complexos relacionados à trajetória estimada.

7. Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), ao Departamento de Engenharia Mecatrônica da Escola Politécnica da Universidade de São Paulo e à FAPESP, através do processo 2008/03995-5, pelo apoio financeiro à pesquisa.

8. Referências

- [Azcarate, 2005] Azcarate, A.; Hageloh, F.; Sande, K.; Valenti, R. “Automatic facial emotion recognition”. Universiteit van Amsterdam, 2005.
- [Ekman, 1978] Ekman, P.; Friesen, W.V. “Facial Action Coding System: a technique for the measurement of facial movement”. Consulting Psychologists Press, 1978.
- [Ekman, 2003] Ekman, P.; Friesen, W.V. “Unmasking the face”. Malor Books, 2003.

- [Fraunhofer,2011] Fraunhofer Facedetect. Disponível em:
<<http://www.iis.fraunhofer.de/en/bf/bv/ks/gpe/demo/>>. Acessado em: 10 de abril de 2011.
- [Gunes, 2006] Gunes, H; Piccardi, M. "Observer Annotation of Affective Display and Evaluation of Expressivity: Face vs. Face-and-Body". HCSNet Workshop on the Use of Vision in HCI (VisHCI 2006), Canberra, Australia, 2006.
- [Martin et al., 2005] Martin, O. Kotsia, I. Macq, B. Pitas, I. The eNTERFACE'05 Audio-Visual Emotion Database. Université Catholique de Louvain; Aristotle University of Thessaloniki, 2005.
- [Piccard,1997]Piccard, R.W. "Affective Computing". MIT Press, 1997.
- [Pantic, 2000]Pantic, M.; Rothkrantz, L.J.M. "Automatic analysis of facial expressions: state of art". IEEE Trans. On Pattern Analysis and Machine Intelligence, vol.22 no. 12, December, 2000.
- [Roseman, 2001] Roseman, I.J.; Smith, C.A. "Appraisal Theory - Overview, Assumptions, Varieties, Controversies". In "Appraisal Processes in Emotion – Theory, Methods, Research" editado por Scherer, K; Schorr, A.; Johnstone, T. Oxford University Press, 2001.
- [Sander, 2005] Sander, D.; Grandjean, D.; Scherer, K.R. "A systems approach to appraisal mechanisms in emotion". Neural Networks vol.18 pgs.317-352, 2005.
- [Scherer,2000] Scherer, K. "Emotions as episodes of subsystem synchronization driven by nonlinear appraisal processes". In "Dynamic Systems Approaches to Emotional Development", editado por Lewis, M.D.; Granic, I. Cambridge Press, 2000.
- [Scherer, 2001] Scherer, K. R. "Appraisal considered as a process of multilevel sequential checking". In "Appraisal Processes in Emotion – Theory, Methods, Research" editado por Scherer, K; Schorr, A.; Johnstone, T. Oxford University Press, 2001.
- [Stewart, 1983]. Stewart,I.N.; Peregoy,P.L. "Catastrophe theory modeling in Psychology". Psychological Bulletin vol. 94, no. 2, pags. 336-362, 1983.
- [van der Maas,1992] van der Maas, H.L.J.; Molenaar, P.C.M. "Stagewise cognitive development: an application of Catastrophe Theory". Psychological Review vol.99 no.2 pags. 395-417, 1992.
- [Zeeman, 1976]. Zeeman, E.C. "Catastrophe theory". Scientific American vol.4 no.254 pags. 65-83, 1976.