

INTERACCIÓN
XII
2011

**XII Congreso Internacional de
Interacción Persona-Ordenador
2-5 de Septiembre de 2011
Lisboa (Portugal)**

<http://interaccion2011.aipo.es>

**ACTAS DEL XII CONGRESO INTERNACIONAL DE
INTERACCIÓN PERSONA-ORDENADOR (INTERACCIÓN 2011)**

EDITORES

**Nestor Garay Vitoria
Julio Abascal González**

**Actas del XII Congreso Internacional de Interacción Persona-Ordenador
Interacción 2011**

Editores: Néstor Garay Vitoria y Julio Abascal González

ISBN: 978-84-9281-234-9

IBERGARCETA PUBLICACIONES, S.L., Madrid, 2010

Edición: 1ª

Impresión: 1ª

Nº de páginas: 392

Formato: 17 x 24

Materia CDU: 004 Ciencia y tecnología de los ordenadores. Informática

Reservados los derechos para todos los países de lengua española. De conformidad con lo dispuesto en el artículo 270 y siguientes del código penal vigente, podrán ser castigados con penas de multa y privación de libertad quienes reprodujeren o plagieren, en todo o en parte, una obra literaria, artística o científica fijada en cualquier tipo de soporte sin la preceptiva autorización. Ninguna parte de esta publicación, incluido el diseño de la cubierta, puede ser reproducida, almacenada o transmitida de ninguna forma, ni por ningún medio, sea éste electrónico, químico, mecánico, electro-óptico, grabación, fotocopia o cualquier otro, sin la previa autorización escrita por parte de la editorial.

Diríjase a CEDRO (Centro Español de Derechos Reprográficos), www.cedro.org, si necesita fotocopiar o escanear algún fragmento de esta obra.

COPYRIGHT © 2011 IBERGARCETA PUBLICACIONES, S.L.

info@garceta.es

Actas del XII Congreso Internacional de Interacción Persona-Ordenador Interacción 2011

Derechos reservados ©2011 respecto a la primera edición en español, por LOS AUTORES

Derechos reservados ©2011 respecto a la primera edición en español, por IBERGARCETA PUBLICACIONES, S.L.

1ª Edición, 1ª Impresión

ISBN: 978-84-9281-234-9

Depósito legal: M-33924-2011

Maquetación: Los Editores

Coordinación del proyecto: @LIBROTEX

Impresión y encuadernación: PRINT HOUSE, S.A.

IMPRESO EN ESPAÑA -PRINTED IN SPAIN

Nota sobre enlaces a páginas web ajenas: Este libro puede incluir referencias a sitios web gestionados por terceros y ajenos a IBERGARCETA PUBLICACIONES, S.L., que se incluyen sólo con finalidad informativa. IBERGARCETA PUBLICACIONES, S.L., no asume ningún tipo de responsabilidad por los daños y perjuicios derivados del uso de los datos personales que pueda hacer un tercero encargado del mantenimiento de las páginas web ajenas a IBERGARCETA PUBLICACIONES, S.L., y del funcionamiento, accesibilidad y mantenimiento de los sitios web no gestionados por IBERGARCETA PUBLICACIONES, S.L., directamente. Las referencias se proporcionan en el estado en que se encuentran en el momento de publicación sin garantías expresas o implícitas, sobre la información que se proporciona en ellas.

PRESENTACIÓN

Este libro recoge los trabajos aceptados en el XII Congreso Internacional de Interacción Persona-Ordenador “Interacción 2011” celebrado del 2 al 5 de septiembre de 2011 en Lisboa (Portugal).

Interacción 2011 es un congreso internacional, organizado por la Asociación para la Interacción Persona-Ordenador (AIPO), que tiene como principal objetivo promover y difundir los avances recientes en el área de la Interacción Persona-Ordenador (IPO), tanto a nivel académico como profesional.

La serie de congresos Interacción pretende difundir el conocimiento de la Interacción Persona-Ordenador a todos los ámbitos y disciplinas universitarias afectados, Informática, Sociología, Comunicación Audiovisual, Ciencias Empresariales, Bellas Artes, Psicología Cognitiva, etc., y dar a conocer a la Asociación para la Interacción Persona-Ordenador (AIPO). AIPO viene organizando este congreso desde el año 2000 (Granada), y se ha mantenido en sucesivas ediciones en Salamanca (2001), Leganés (2002), Vigo (2003), Lleida (2004), Granada (en el CEDI 2005), Puertollano (2006), Zaragoza (dentro del CEDI 2007), Albacete (2008), de forma conjunta Barcelona (España) y Armenia (Colombia) (2009) y Valencia (integrado en el CEDI 2010). En este tiempo, Interacción ha puesto en contacto a diferentes grupos de trabajo de todo el mundo, principalmente hispano y luso parlantes, para establecer vínculos de colaboración, y ha potenciado la relación entre la universidad y la empresa.

El objetivo de la XII edición de Interacción es mejorar el conocimiento mutuo de las comunidades científicas portuguesa y española que trabajan en el área de la IPO y facilitar colaboración entre ambas. La organización local ha estado a cargo de INESC-ID de la Universidade Técnica de Lisboa. Además, en 2011 Interacción se celebra junto con el congreso INTERACT 2011, organizado por el IFIP TC13, lo que permitirá aumentar la proyección internacional de los participantes.

Este libro de actas refleja la evolución de la investigación y los desarrollos en IPO, a través de trabajos de cooperación entre universidades y aportaciones de empresas del sector. Se ha recibido un total de 66 contribuciones. El panel de revisores ha seleccionado 24 artículos largos de entre los 55 recibidos (con un 43,6% de aceptación), 18 artículos cortos y 3 posters, que forman parte de este volumen. A ellos se les añaden las dos conferencias invitadas al congreso a cargo de los profesores Gerrit Van deer Veer (Universidad Abierta de los Países Bajos) y Simone DJ Barbosa (Pontificia Universidade Católica de Rio de Janeiro, Brasil), a los que debemos agradecer su entusiasta colaboración.

Los temas tratados en los artículos que se presentan aquí muestran una disciplina cambiante y muy dependiente de los avances tecnológicos: realidad virtual y aumentada, sistemas para la colaboración (CSCW/CSCL), desarrollo de interfaces basadas en modelos, interacción para personas con discapacidad, evaluación de sistemas interactivos, accesibilidad de la información, herramientas y entornos para el desarrollo de interfaces de usuario, diseño gráfico, bellas artes e interacción, experiencias en empresas, o interacción, aprendizaje y enseñanza, etc.

Confiamos en que los lectores encuentren este volumen una referencia útil dentro del área de la Interacción Persona-Ordenador.

Sandra Baldassarri, Javier Marco, Cecilia Sanz, Andrea Guisen, Armando de Giusti, Eva Cerezo

Realidad virtual y aumentada II

Espejo aumentado: sistema interactivo de Realidad Aumentada basado en Kinect.....	347
Lucia Vera, Jesus Gimeno, Inmaculada Coma, Marcos Fernández	
Continuous Facial Affect Recognition from Videos.....	357
Sergio Ballano Pablo, Isabelle Hupont, Eva Cerezo, Sandra Baldassarri	

Pósteres

MOVE-ON: otro sistema de interacción.....	369
Agustín Linares, Juan Galán Páez, Jose María Alonso Calero, Jesus Marín Clavijo	
Propuesta de estandarización de la configuración en los magnificadores de pantalla.....	371
Elena Ballesteros, Mireia Ribera, Elena Sánchez	
Guía de Especificación para el EUD.....	373
José Antonio Macías Iglesias, Silvia Teresita Acuña Castillo, Miguel Ángel Mora Rincón, Juan de Lara Jaramillo, Roberto Moriyón Salomón	
Índice de autores.....	375

Continuous Facial Affect Recognition from Videos

Sergio Ballano¹, Isabelle Hupont¹, Eva Cerezo² and Sandra Baldassarri²

¹ Aragon Institute of Technology, Department of R&D and Technology Services,
Zaragoza. 5018, María de Luna 7-8, Spain

² University of Zaragoza, Computer Science and Systems Engineering Department,
Zaragoza. 50018, María de Luna 3, Spain
{sballano, ihupont}@ita.es, {ecerezo, sandra}@unizar.es

Abstract. The interpretation of user facial expressions is a very useful method for emotional sensing and constitutes an indispensable part of affective HCI designs. This paper proposes an effective system for continuous facial affect recognition from videos. The system operates in a continuous 2D emotional space, characterized by evaluation and activation factors, enabling a wide range of intermediary affective states to be worked with. It makes use, for each video frame, of a classification method able to output the exact location (2D point coordinates) of the still facial image in that space. It also exploits the Kalman filtering technique to control the 2D point movement along the affective space over time and to improve the robustness of the method by predicting its future locations in cases of temporal facial occlusions or inaccurate tracking. The system has been tuned with an extensive universal database and preliminary evaluation results are very encouraging.

Keywords: Affective computing, facial expression analysis.

1 Introduction

Facial expressions are the most powerful, natural and direct way used by humans to communicate affective states. Thus, the interpretation of facial expressions is the most common method used for emotional detection and forms an indispensable part of affective Human Computer Interface designs.

Facial expressions are often evaluated by classifying still face images into one of the six universal “basic” emotions or categories proposed by Ekman [1] which include “happiness”, “sadness”, “fear”, “anger”, “disgust” and “surprise” [2-4]. There are a few tentative efforts to detect non-basic affective states, such as “fatigue”, “interested”, “thinking”, “confused” or “frustrated” [5, 6]. In any case, this categorical approach, where emotions are a mere list of labels, fails to describe the wide range of emotions that occur in daily communication settings and ignores the intensity of emotions.

To overcome the problems cited above, some researchers such as Whissell [7] and Plutchik [8], prefer to view affective states not independent but rather related to one another in a systematic manner. They consider emotions as a continuous 2D space whose dimensions are evaluation and activation. The evaluation dimension measures

how a human feels, from positive to negative. The activation dimension measures whether humans are more or less likely to take some action under the emotional state, from active to passive.

For many years, a lot of effort was dedicated to recognize facial expressions in still images. Given that humans inherently display facial emotions following a continuous temporal pattern [9], more recently attention has been shifted towards sensing facial affect from video sequences. The study of facial expressions' dynamics reinforces the limitations of categorical approach, since it represents a discrete list of emotions with no real link between them and has no algebra: every emotion must be studied and recognized independently. Dimensional approach is much more able to deal with variations in emotional states over time, since in such cases changing from one universal emotion label to another would not make much sense in real life scenarios.

Continuous dimensional annotation is best suited if the annotated sample has more than one emotional apex or blended emotions [10] but this can be very time consuming, with a very poor inter-annotator agreement and can make difficult the posterior evaluation when working with long videos. An intermediate approach is to annotate only certain moments in time (key-frames). The key-frames, selected by the user which expressed the emotions, will usually correspond to the onset, apex and offset of an emotion and any other moment that the user may find interesting especially in case of blended emotions which don't pass through the neutral state.

This paper proposes a method for continuous facial affect recognition from video. The system operates in a 2D emotional space, characterized by evaluation and activation factors. It combines a classification method able to output, frame per frame, the exact location (2D point coordinates) of the shown facial image and a Kalman filtering technique that controls the 2D point movement over time through an "emotional kinematics" model. In that way, the system works with a wide range of intermediary affective states and is able to define a continuous emotional path that characterizes the affective video sequence. The system is capable of analyzing any subject, male or female of any age and ethnicity, and has been validated considering human assessment.

The structure of the paper is the following: Section 2 describes the method for facial images classification in a continuous 2D affective space. In Section 3 the step from still images to video sequences through the "emotional kinematics" model is explained in detail and Section 4 comprises the conclusions and future work.

2 A Novel Method for Facial Images Classification in a Continuous 2D Affective Space

This section describes a novel method for sensing emotions from still facial images in a continuous 2D affective space. The facial images classification method starts with a classification mechanism in discrete emotional categories that intelligently combines different classifiers simultaneously to obtain a confidence value to each Ekman universal emotional category (Section 2.1). Then, this output is subsequently expanded in order to be able to work in a continuous emotional space and thus to consider intermediate emotional states (Section 2.2).

2.1 Classifiers Combination for Discrete Emotional Classification

The starting point of the system is the method for facial emotional classification presented in authors' previous work [4]. The inputs to this method are the variations with respect to the "neutral" face of the set of facial distances and angles shown in Fig. 1. In that way, the face is modeled in a computationally simple way without losing relevant information about the facial expression. The facial points that allow to calculate the facial distances and angles are obtained thanks to faceAPI [11], a commercial real-time facial feature tracking program.

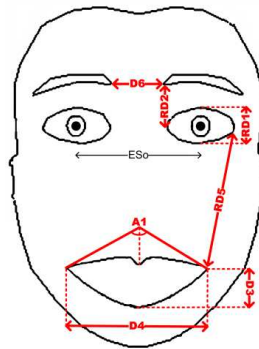


Fig. 1. System's facial inputs.

This initial method combines through a majority voting strategy [4] the five most commonly used classifiers in the literature (Multilayer Perceptron, RIPPER, SVM, Naïve Bayes and C4.5) to finally assign at its output a confidence value $CV(E_i)$ of the facial expression to each of Ekman's six emotions plus "neutral". It has been well-tuned and tested with a total of 1500 static frames selected from the apex of the video sequences from the well-known FG-NET [12] and MMI [13] facial expression databases. Therefore, it has been validated with a large database of individuals of all races, ages and genders.

Table 1 shows the confusion matrix obtained when applying the initial discrete facial emotional classification method to the 1500 selected static frames. As can be observed, the success rates for the "neutral", "joy", "disgust", "surprise", "disgust" and "fear" are very high (81.48%-97.62%). The lowest result is for "sadness", which is confused with the "neutral" emotion on 20% of occasions, due to the similarity of their facial expressions.

Table 1. Confusion matrix obtained after applying the discrete emotional classification method to the 1500 selected static frames.

Emotion --> is classified as	Disgust	Joy	Anger	Fear	Sadness	Neutral	Surprise
Disgust	94,12%	0,00%	2,94%	2,94%	0,00%	0,00%	0,00%
Joy	2,38%	97,62%	0,00%	0,00%	0,00%	0,00%	0,00%
Anger	7,41%	0,00%	81,48%	0,00%	7,41%	3,70%	0,00%
Fear	3,70%	0,00%	0,00%	85,19%	3,70%	0,00%	7,41%
Sadness	6,67%	0,00%	6,67%	0,00%	66,67%	20,00%	0,00%
Neutral	0,00%	0,00%	2,00%	2,00%	2,00%	94,00%	0,00%
Surprise	0,00%	0,00%	0,00%	2,22%	0,00%	2,22%	95,56%

In his work, Plutchik [8] assigned “emotional orientation” values to a series of affect words. For example, two similar terms (like “joyful” and “cheerful”) have very close emotional orientation values while two antonymous words (like “joyful” and “sad”) have very distant values, in which case Plutchik speaks of “emotional incompatibility”. According to Plutchik’s findings, the obtained results can be considered positive as emotions with distant “emotional orientation” values (such as “disgust” and “joy” or “neutral” and “surprise”) are confused on less than 2.5% of occasions and incompatible emotions (such as “sadness” and “joy” or “fear” and “anger”) are never confused.

2.2 Emotional Mapping to a 2D Continuous Affective Space

To enrich the emotional output information from the system in terms of intermediate emotions, one of the most influential evaluation-activation 2D models has been used: that proposed by Whissell. In her study, Whissell assigns a pair of values <evaluation, activation> to each of the approximately 9000 selected affective words that make up her “Dictionary of Affect in Language” [7]. Figure 2 shows the position of some of these words in the evaluation-activation space. The next step is to build an emotional mapping so that an expressional face image can be represented as a point on this plane whose coordinates (x,y) characterize the emotion property of that face.

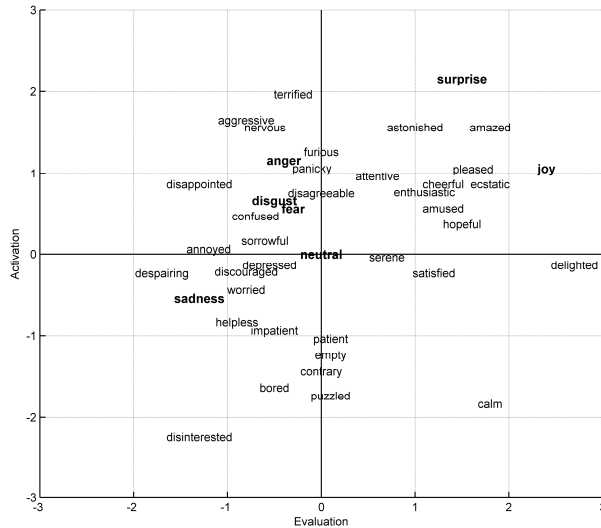


Fig. 2. Simplified Whissell's evaluation-activation space.

It can be seen that the words corresponding to each of Ekman’s six emotions have a specific location (x_i, y_i) in the Whissell space (in bold in Fig. 2). Thanks to this, the output of the classifiers (confidence value of the facial expression to each emotional category) can be mapped onto the space. This emotional mapping is carried out considering each of Ekman’s six basic emotions plus “neutral” as weighted points in the evaluation-activation space. The weights are assigned depending on the confidence value $CV(E_i)$ obtained for each emotion. The final coordinates (x,y) of a given image are calculated as the centre of mass of the seven weighted points following equation (1) (see Fig. 3). In this way, the output of the system is enriched with a larger number of intermediate emotional states.

$$x = \frac{\sum_{i=1}^7 x_i CV(E_i)}{\sum_{i=1}^7 CV(E_i)} \quad \text{and} \quad y = \frac{\sum_{i=1}^7 y_i CV(E_i)}{\sum_{i=1}^7 CV(E_i)} \tag{1}$$

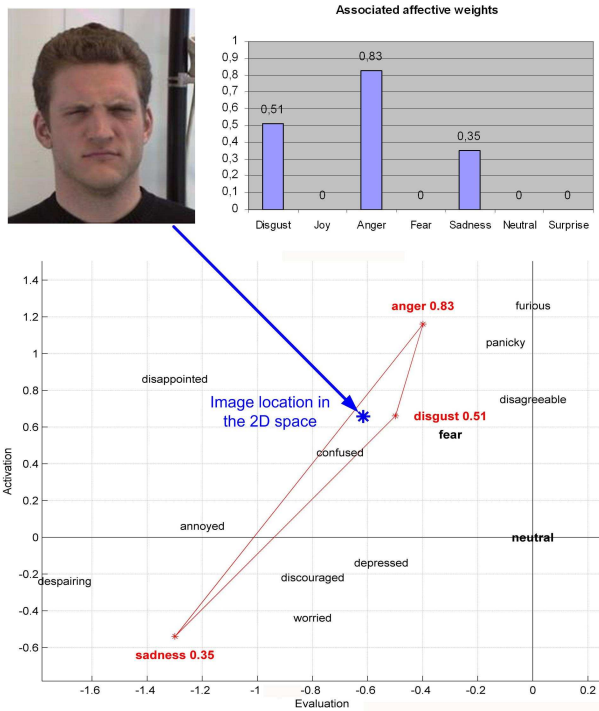


Fig. 3. Diagram for obtaining the location of a facial image in the 2D emotional space. A graphic illustration of the 2D emotional mapping process is included as an example.

Fig. 4 shows several images of the database with their nearest label in the Whissell space after applying the proposed emotional mapping.



Fig. 4. Example of images from the database with their nearest label in the Whissell space after applying the 2D emotional mapping.

3 From Still Images to Video Sequences through 2D Emotional Kinematics Modeling

As pointed out in the introduction, humans inherently display facial emotions following a continuous temporal pattern. With this starting postulate and thanks to the use of the 2-dimensional description of affect, which supports continuous emotional input, an emotional facial video sequence can be viewed as a point (corresponding to the location of a particular affective state in time t) moving through this space over time. In that way, the different positions taken by the point (one per frame) and its velocity over time can be related mathematically and modeled, finally obtaining an “emotional path” in the 2D space that reflects intuitively the emotional progress of the user throughout the video. In Section 3.1, a Kalman filtering technique is proposed to model the “emotional kinematics” of that point when moving along the Whissell space and thus enable to both smooth its trajectory and improve the robustness of the method by predicting its future locations (e.g. in cases of temporal facial occlusions or inaccurate tracking). Section 3.2 presents the results obtained when applying the emotional kinematics model to different complex video sequences.

3.1 Modeling Emotional Kinematics with a Simple Kalman Filter

For real-time “emotional kinematics” control, the well-known Kalman filter is exploited [14]. Kalman filters are widely used in the literature for estimation problems ranging from target tracking to function approximation. Their purpose is to estimate a system’s state by combining an inexact (noisy) forecast with an inexact measurement of that state, so that the most weight is given to the value with the least uncertainty at each time t .

Analogously to classical mechanics, the “emotional kinematics” of the point in the Whissell space (x-position, y-position, x-velocity and y-velocity) are modeled as the system’s state in the Kalman framework at time t_k . The output of the 2D classification system described in Section 2 is modeled as the measurement of the system’s state. In this way, the Kalman iterative estimation process -that follows the well-known recursive equations detailed in Kalman’s work [14]- can be applied to the recorded user’s emotional video sequence, so that each iteration corresponds to a new video frame (i.e. to a new sample of the computed emotional path). For the algorithm initialization at t_0 , the predicted initial condition is set equal to the measured initial state and the 2D point is assumed to have null velocity.

One of the main advantages of using Kalman filter for the 2D point emotional trajectory modeling is that it can be used to tolerate small occlusions or inaccurate tracking. As pointed out in Section 2.1, the input facial feature points of the classification method are obtained thanks to the commercial facial tracker faceAPI [11]. In general, existing facial trackers do not perform the detection with high accuracy: most of them are limited in terms of occlusions, fast movements, large head rotations, lighting, beards, glasses, etc. Although faceAPI deals with these problems quiet robustly, on some occasions its performance is poor, especially when working in real-time. For that reason, its measurements include a confidence weighting, from 0 to 1, allowing the acceptability of the tracking quality to be determined. Thanks to it,

when a low level of confidence is detected (lower than 0.5), the measurement will not be used and only the filter prediction will be taken as the 2D point position.

3.2 Experimental Results

In order to demonstrate the potential of the proposed “emotional kinematics” model, it has been tested with a set of emotionally complex video sequences, recorded in a natural (unsupervised) setting. These videos are complex owing to three main factors:

- An average user’s home setup was used. A VGA resolution webcam placed above the screen is used, with no special illumination, causing shadows to appear in some cases. In addition, the user placement, not covering the entire scene, reduces the actual resolution of the facial image.
- Different emotions are displayed contiguously, instead of the usual neutral→emotional-apex→neutral pattern exhibited in the databases, so emotions such as surprise and joy can be expressed without neutral periods between them.
- Some facial occlusions occur due to the user covering his/her face or looking away during a short period of time, causing the tracking program to lose the facial features. In these cases, only the prediction from the Kalman filter is used, demonstrating the potential of the “emotional kinematics” filtering technique.

15 videos from three different users were tested, ranging from 20 to 70 seconds from which a total of 127 key-frames were extracted to evaluate different key-points of the emotional path. The key-frames were selected by the user who recorded the video, looking for each of the emotional apex and neutral points.

These key-points were annotated in the Whissell space thanks to 18 volunteers. The collected evaluation data have been used to define a region where each image is considered to be correctly located. The algorithm used to compute the shape of the region is based on Minimum Volume Ellipsoids (MVE) and follows the algorithm described by Kumar and Yildirim [15]. MVE looks for the ellipsoid with the smallest volume that covers a set of data points. The obtained MVEs are used for evaluating results at four different levels:

1. *Ellipse criteria.* If the point detected by the system is inside the ellipse, it is considered a success; otherwise it is a failure.
2. *Quadrant criteria.* The output is considered to be correctly located if it is in the same quadrant of the Whissell space as the ellipse centre.
3. *Evaluation axis criteria.* The system output is a success if situated in the same semi-axis (positive or negative) of the evaluation axis as the ellipse centre. This information is especially useful for extracting the positive or negative polarity of the shown facial expression.
4. *Activation axis criteria.* The same criteria projected to the activation axis. This information is relevant for measuring whether the user is more or less likely to take an action under the emotional state.

The results obtained following the different evaluation strategies are presented in Table 2. As can be seen, the success rate is 61.90% in the most restrictive case, i.e. with ellipse criteria. It rises to 84.92% when considering the activation axis criteria.

Table 2. Results obtained in an uncontrolled environment.

	Ellipse criteria	Quadrant criteria	Evaluation axis criteria	Activation axis criteria
Success Rate	61.90%	74.60%	79.37%	84.92%

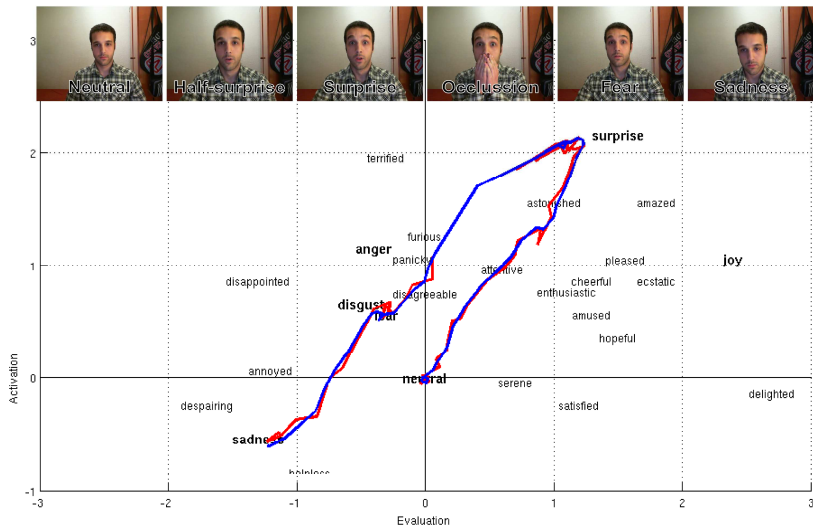


Fig. 5. “Emotional kinematics” model response during the different affective phases of the video and the occlusion period. In dashed red, emotional trajectory without Kalman filtering; In solid blue, reconstructed emotional trajectory using Kalman filter.

4 Conclusions and Future Work

This paper describes an effective system for continuous facial affect recognition from videos. The inputs are a set of facial parameters (angles and distances between facial points) that enable the face to be modeled in a computationally simple way without losing relevant information about the facial expression. The system makes use, frame per frame, of a classification method able to output the exact location (2D point coordinates) of a still facial image in the Whissell evaluation-activation space. The temporal consistency and robustness (to occlusions or inaccurate tracking) of the recognized affective sequence is ensured by a Kalman filtering technique that, through an “emotional kinematics” model, controls the 2D point trajectory when moving along the Whissell space.

The main distinguishing feature of our work compared to others is that the output does not simply provide a classification in terms of a set of emotionally discrete

labels, but goes further by extending the emotional information over an infinite range of intermediate emotions and by allowing a continuous dynamic emotional trajectory to be detected from complex affective video sequences. Another noteworthy feature of the work is that it has been tuned with an extensive database of 1500 images showing individuals of different races and gender, giving universal results with very promising levels of correctness.

References

1. Keltner, D., Ekman, P.: Facial Expression Of Emotion. Handbook of emotions. pp. 236-249. New York: Guilford Publications, Inc. (2000).
2. Hammal, Z., Couvreur, L., Caplier, A., Rombaut, M.: Facial expression classification: An approach based on the fusion of facial deformations using the transferable belief model. *International Journal of Approximate Reasoning*. 46, 542-567 (2007).
3. Soyel, H., Demirel, H.: Facial Expression Recognition Using 3D Facial Feature Distances. *Image Analysis and Recognition*. pp. 831-838. Springer Berlin / Heidelberg (2007).
4. Hupont, I., Cerezo, E., Baldassarri, S.: Sensing facial emotions in a continuous 2D affective space. Presented at the Systems Man and Cybernetics (SMC) , Istanbul Octobre 10 (2010).
5. Kapoor, A., Burleson, W., Picard, R.W.: Automatic prediction of frustration. *Int. J. Hum.-Comput. Stud.* 65, 724-736 (2007).
6. Yeasin, M., Bullo, B., Sharma, R.: Recognition of facial expressions and measurement of levels of interest from video, (2006).
7. Whissell, C.M.: *The Dictionary of Affect in Language, Emotion: Theory, Research and Experience*. New York Academic (1989).
8. Plutchik, R.: *Emotion: a Psychoevolutionary Synthesis*. Harper & Row (1980).
9. Petridis, S., Gunes, H., Kaltwang, S., Pantic, M.: Static vs. dynamic modeling of human nonverbal behavior from multiple cues and modalities. *Proceedings of the 2009 international conference on Multimodal interfaces*. 23-30 (2009).
10. Ellen Douglas-Cowie, Laurence Devillers, Jean-Claude Martin, Roddy Cowie, Suzie Savvidou, Sarkis Abrilian, Cate Cox: *Multimodal Databases of Everyday Emotion: Facing up to Complexity*. Ninth European Conference on Speech Communication and Technology. pp. 813-816. , Lisbon, Portugal (2005).
11. Face API technical specifications brochure, <http://www.seeingmachines.com/pdfs/brochures/faceAPI-Brochure.pdf>.
12. Facial Expressions and Emotion Database. Technische Universität München, <http://cotesys.mmk.e-technik.tu-muenchen.de/isg/content/feed-database>.
13. Pantic, M., Valstar, M., Rademaker, R., Maat, L.: Web-based database for facial expression analysis. *Proc. Int'l Conf. Muntimedia and Expo*. 317-321 (2005).
14. Kalman, R.: A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME – Journal of Basic Engineering*. 35-45 (1960).
15. Kumar, P., Yildirim, E.A.: Minimum-Volume Enclosing Ellipsoids and Core Sets. *Journal of Optimization Theory and applications*. 126, 1-21 (2005).