Lecture Notes in Computer Science 6949

Pedro Campos   Nicholas Graham
Joaquim Jorge   Nuno Nunes
Philippe Palanque   Marco Winckler (Eds.)

Human-Computer Interaction – INTERACT 2011

13th IFIP TC 13 International Conference
Lisbon, Portugal, September 5-9, 2011
Proceedings, Part IV

Springer

Volume Editors

Pedro Campos
Nuno Nunes
University of Madeira
9020-105, Funchal, Portugal
E-mail: {pcampos, njn}@uma.pt

Nicholas Graham
Queen's University
Kingston, ON K7L 3N6, Canada
E-mail: graham@equis.cs.queensu.ca

Joaquim Jorge
Instituto Superior Técnico
1049-001 Lisbon, Portugal
E-mail: jaj@inesc.pt

Philippe Palanque
Marco Winckler
University Paul Sabatier
31062 Toulouse Cedex 9, France
E-mail: {palanque, winckler}@irit.fr

ISSN 0302-9743 e-ISSN 1611-3349
ISBN 978-3-642-23767-6 e-ISBN 978-3-642-23768-3
DOI 10.1007/978-3-642-23768-3
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2011935338

CR Subject Classification (1998): H.5.2, H.5.3, H.3-5, I.2.10, D.2, C.2, K.3-4

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Foreword

Advances in interactivity, computing power, mobile devices, large displays and ubiquitous computing offer an ever-increasing potential for empowering users. This can happen within their working environment, in their leisure time or even when extending their social skills. While such empowerment could be seen as a way of connecting people in their workspace, home or on the move, it could also generate gaps requiring larger effort and resources to fruitfully integrate disparate and heterogeneous computing systems.

The conference theme of INTERACT 2011 was "building bridges" as we believe human–computer interaction (HCI) is one the research domains more likely to significantly contribute to bridging such gaps. This theme thus recognizes the interdisciplinary and intercultural spirit that lies at the core of HCI research. The conference had the objective of attracting research that bridges disciplines, cultures and societies. Within the broad umbrella of HCI, we were in particular seeking high-quality contributions opening new and emerging HCI disciplines, bridging cultural differences, and tackling important social problems. Thus, INTERACT 2011 provided a forum for practitioners and researchers to discuss all aspects of HCI, including these challenges. The scientific contributions gathered in these proceedings clearly demonstrate the huge potential of that research domain to improving both user experience and performance of people interacting with computing devices. The conference also is as much about building bridges on the human side (between disciplines, cultures and society) as on the computing realm.

INTERACT 2011 was the 13th conference of the series, taking place 27 years after the first INTERACT held in early September 1984 in London, UK. Since INTERACT 1990 the conferences have taken place under the aegis of the UNESCO International Federation for Information Processing (IFIP) Technical Committee 13. This committee aims at developing the science and technology of the interaction between humans and computing devices through different Working Groups and Special Interests Groups, all of which, together with their officers, are listed within these proceedings.

INTERACT 2011 was the first conference of its series to be organized in cooperation with ACM SIGCHI, the Special Interest Group on Computer–Human Interaction of the Association for Computing Machinery. We believe that this cooperation was very useful in making the event both more attractive and visible to the worldwide scientific community developing research in the field of HCI.

We thank all the authors who chose INTERACT 2011 as the venue to publish their research This was a record year for the conference in terms of submissions in the main technical categories. For the main Technical Program there were a total of 680 submissions, including 402 long and 278 short papers, out of which we accepted 171 (111 long and 60 short submissions), for a combined acceptance rate of less than 25%. Overall, from a total of 741 submissions for all tracks, 290 were accepted, as follows:

- 111 Full Research Papers
- 60 Short Research Papers
- 54 Interactive Poster Papers
- 17 Doctoral Consortium Papers
- 16 Workshops
- 12 Tutorials
- 5 Demonstrations
- 6 Organizational Overviews
- 4 Industrial Papers
- 3 Special Interest Groups
- 2 Panels

Our sincere gratitude goes to the members of our Program Committee (PC), who devoted countless hours to ensure the high quality of the INTERACT Conference. This year, we improved the reviewing process by moving to an associate chair model. With almost 700 submitted papers, it is impossible for the PC Chairs to read every paper. We recruited 103 Associate Chairs (ACs), each of whom handled up to 12 papers. The ACs recruited almost 800 external reviewers, guaranteeing that each paper was reviewed by three to six referees. ACs also provided a meta-review. Internal discussion among all the reviewers preceded the final decision between the PC Chairs and the AC. This herculean effort was only possible due to the diligent work of many people. We would like to thank you all for the effort and apologize for all the bullying required to get the work done on time.

In addition, sincere thanks must be extended to those whose contributions were essential in making it possible for the conference to happen and for these proceedings to be produced. We owe a great debt to the Conference Committees, the members of the International Program Committee and the numerous reviewers who had to review submissions from the various categories. Similarly, the members of the conference Organizing Committee, the staff at INESC-ID, especially Manuela Sado, deserve much appreciation for their tireless help with all aspects of planning and managing the many administrative and organizational issues. We would like to especially thank Tiago Guerreiro for his dedication with the Student Volunteer program, and José Coelho who worked tirelessly to make the online program a reality. Thanks are also due to Alfredo Ferreira for keeping and single-handedly maintaining the website, and to Pedro Campos and Marco Winkler for the superb work done with the conference proceedings. Finally, our thanks go to all the authors who actually did the scientific work and especially to the presenters who took the additional burden of discussing the results with their peers at INTERACT 2011 in Lisbon.

July 2011

Nicholas Graham
Daniel Gonçalves
Joaquim Jorge
Nuno Nunes
Philippe Palanque

Table of Contents – Part IV

Long and Short Papers

Usable Privacy and Security

User Experience I

User Experience II

User Experience III

User Modelling

Visualization

Web Interaction

Demos

Doctoral Consortium

Industrial Papers

Interactive Posters

Organization Overviews

Panels

Special Interest Groups (SIGs)

Tutorials

Workshops

Recognizing Emotions from Video in a Continuous 2D Space

Sergio Ballano[1], Isabelle Hupont[1], Eva Cerezo[2], and Sandra Baldassarri[2]

[1] Aragon Institute of Technology, Department of R&D and Technology Services, Zaragoza. 5018, María de Luna 7-8, Spain
[2] Universidad de Zaragoza, Computer Science and Systems Engineering Department, Zaragoza. 50018, María de Luna 3, Spain
{sballano,ihupont}@ita.es, {ecerezo,sandra}@unizar.es

Abstract. 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. It makes use, for each video frame, of a classification method able to output the exact location (2D point coordinates) of a 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.

Keywords: Affective computing, facial expression analysis.

1 Introduction

Facial expressions are often evaluated by classifying still face images into one of the six universal "basic" emotions proposed by Ekman [1] which include "happiness", "sadness", "fear", "anger", "disgust" and "surprise". This categorical approach fails to describe the wide range of emotions that occur in daily communication settings and ignores the intensity of emotions.

Given that humans inherently display facial emotions following a continuous temporal pattern [2], 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.

This paper proposes a method for continuous facial affect recognition from videos. 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.

P. Campos et al. (Eds.): INTERACT 2011, Part IV, LNCS 6949, pp. 600–603, 2011.

2 Facial Images Classification in a Continuous 2D Affective Space

The starting point of the system is the method for facial emotional classification presented in authors' previous work [3]. 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. This initial method combines through a majority voting strategy [3] 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".

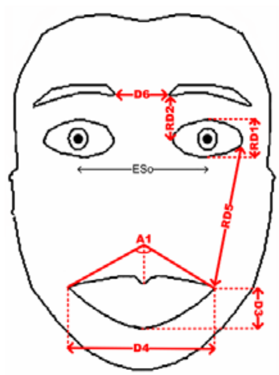

Fig. 1. System's facial inputs

To enrich the emotional output information from the system in terms of intermediate emotions, the evaluation-activation 2D model proposed by Whissell has been used. In her study, Whissell assigns a pair of values <evaluation, activation> to each of the approximately 9000 affective words that make up her "Dictionary of Affect in Language" [4]. 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.

The words corresponding to each of Ekman's six emotions have a specific location (x_i, y_i) in the Whissell space. 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 then calculated as the centre of mass of the seven weighted points.

3 From Still Images to Video Sequences

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.

3.1 Modeling Emotional Kinematics with a Simple Kalman Filter

For real-time "emotional kinematics" control, the Kalman filter is exploited. Analogously to classical mechanics, the "emotional kinematics" of the point in the Whissell space (x-position, y-position, x-velocity and y-velocity) is modeled as the system's state in the Kalman framework at time t_k. In this way, the Kalman iterative estimation process -that follows the well-known recursive equations detailed in Kalman's work [5]- 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). 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 so that, when a low level of confidence in the facial tracking is detected, 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 complex video sequences recorded in an unsupervised setting (VGA webcam quality, different emotions displayed contiguously, facial occlusions, etc.). A total of 15 videos from 3 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.

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 Yildrim [6]. 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, as shown in Table 1. As can be seen, the success rate is 61.90% in the most restrictive case, i.e. with ellipse criteria and rises to 84.92% when considering the activation axis criteria. Finally, Fig. 2 shows an example of emotional path obtained after applying the "emotional kinematics" model.

Table 1. Results obtained in an uncontrolled environment

	Ellipse criteria (success if inside the ellipse)	Quadrant criteria (success if in the same quadrant as the ellipse centre)	Evaluation axis criteria (success if in the same evaluation semi-axis as the ellipse centre)	Activation axis criteria (success if in the same activation semi-axis as the ellipse centre)
Success%	61.90%	74.60%	79.37%	84.92%

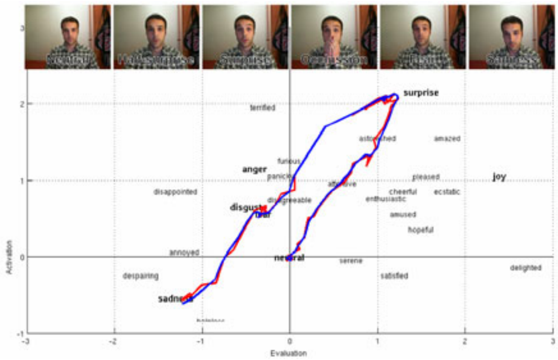

Fig. 2. "Emotional kinematics" model response during the different affective phases of the video and the occlusion period. In dashed red, emotional trajectory without Kalman filtering; In solid blue, reconstructed emotional trajectory using Kalman filter.

4 Conclusions

This paper describes an effective system for continuous facial affect recognition from videos. 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.

References

1. Keltner, D., Ekman, P.: Facial Expression Of Emotion. In: Handbook of Emotions, pp. 236–249. Guilford Publications, Inc., New York (2000)
2. Petridis, S., Gunes, H., Kaltwang, S., Pantic, M.: Static vs. dynamic modeling of human nonverbal behavior from multiple cues and modalities. In: Proceedings of the 2009 International Conference on Multimodal Interfaces, pp. 23–30 (2009)
3. Hupont, I., Cerezo, E., Baldassarri, S.: Sensing facial emotions in a continuous 2D affective space. Presented at the Systems Man and Cybernetics (SMC), Istanbul, Octubre 10 (2010)
4. Whissell, C.M.: The Dictionary of Affect in Language, Emotion: Theory, Research and Experience. Academic, New York (1989)
5. Kalman, R.: A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME – Journal of Basic Engineering, 35–45 (1960)
6. Kumar, P., Yildirim, E.A.: Minimum-Volume Enclosing Ellipsoids and Core Sets. Journal of Optimization Theory and Applications 126, 1–21 (2005)