Affective and Behavioral Assessment for Adaptive Intelligent Tutoring Systems

L. Marco-Gimenez, M. Arevalillo-Herraez, F.J. Ferri, S. Moreno-Picot
Dept. Informatica. Universitat de Valencia. Spain
margi4@alumni.uv.es,{miguel.arevalillo,francesc.ferri,salvador.moreno}@uv.es

J.G. Boticario, O.C. Santos, S. Salmeron-Majadas, M. Saneiro, R. Uria-Rivas
aDeNu, Dept. of AI, Comp. Sci. School. UNED. Madrid. Spain
{jgb,ocsantos,sergio.salmeron,marsanerio,raul.uria}@dia.uned.es

David Arnau
Dept. of Didactics of Math.
Universitat de Valencia. Spain
david.arnau@uv.es

José A. González-Calero
Dept. of Mathematics. Univ. of Castilla-La Mancha. Spain
jose.gonzalezcalero@uclm.es

Aladdin Ayesh
Faculty of Technology
De Montfort University, UK
aayesh@dmu.ac.uk

Raúl Cabestrero, Pilar Quirós
Dept. of Basic Psychology
Faculty of Psychology, UNED, Spain
{rcabestrero,pquiros}@psi.uned.es

P. Arnau-González, N. Ramzan
School of Eng. and Computing
Univ. of the West of Scotland, UK
pablo.arnau@uws.ac.uk

ABSTRACT

Adaptive Intelligent Tutoring Systems (ITS) aim at helping students going through the resolution of a given problem in a principled way according to the desired outcomes, the intrinsic capabilities of the student, and the particular context in which the exercise takes place. These systems should be capable of acting according to mistakes, boredness, distractions, etc. Several works propose different models to represent the problem being solved, the student solving it and the tutor guidance to the desired solution. The system complexity requires non trivial models whose corresponding parameters need to be estimated with differents kinds of data, usually requiring heavy and difficult sensing and recognition tasks. In this work, we present some of the work in progress in the BIG-AFF project. Between other issues, we deal with the use of low cost and low intrusive devices to gather contextual data to losely drive the actions of an ITS without constructing a fully structured student model with corresponding affective and behavioral states. The idea is to improve the students' learning outcome and satisfaction by progressively learning how to adapt the ITS in terms of the sensed data.

Keywords

Personalization; Adaptation; Intelligent Tutoring Systems; Word Problem Solving; Affective states; Emotion detection

1. INTRODUCTION

Previous research has provided solid evidence that emotions strongly affect motivation and hence play an important role in learning [1]. In the research project BIG-AFF, we build on the hypothesis that it is possible to provide learners with a personalised support that enriches their learning process and experience by using low intrusive (and low cost) devices to capture affective multimodal data that include cognitive, behavioural and physiological information.

Results from previous research [9, 40] led to the identification of open issues that need to be addressed to advance the scientific and technological knowledge regarding emotion recognition:

- Information sources. For the potential applicability of the methods in realistic learning environments, it is essential that devices are both low intrusive and inexpensive. This leads to considering quantified-self approaches and wearable technologies to gather diverse physical and physiological information about the user. In addition, computer vision methods [14, 8, 18, 38, 21] that can be run on inexpensive and ubiquitous hardware can complement this information.
- Context Modelling. Although commonly ignored or not given the importance that it deserves, the usage context has a relevant effect on the interaction (e.g. navigation patterns depend on the input devices, keyboard, mouse, touch screen, as well as the functional diversity of the user) [17].
- Individual influence. There is a personality and physiological influence of the individual when feeling and expressing emotions. This implies the need for designing appropriate experimental settings using both

inter-subject and intra-subject configurations to better understand the user impact on the detection process.

- Data processing. A major limitation for real time processing is the large amount of data simultaneously produced by independent devices. Deep learning methods for multimodal data fusion [23, 35, 7, 11] may help discovering more suitable abstract data representations. In addition, Big Data approaches provide a powerful framework for dealing with flows of large sets of unstructured data. They both may open up new possibilities in the way data is mined to find emotional and behavioural patterns.

Apart from evidencing the high complexity of the problem at hand, initial findings have revealed the need to further investigate experimental methodologies and related infrastructure support to help clarify the fragile and elusive nature of affect. According to the literature [27] and our background, which is grounded on affect-detection experiences at large (nearly 100 subjects) and small scale (intra-subject), there are no yet clear criteria for setting the appropriate affect detection and management experimental approaches. Within the context of the BIG-AFF project, we are deeply investigating into what types of experiences should be carried out as well as, how and when, depending on circumstances. We are also concerned with the application of user-centric evaluation to measure the impact of affect in the users' experience and learning gains, and test ecological validity. To this end, different contexts (individual, collaborative, ambient intelligence, enriched multimedia) are being considered. As a result, we have started to develop an extensive knowledge about the organization of the experiences, reliable (expert and user) data labelling and non-intrusive techniques to automatically recognize emotions.

2. PREVIOUS WORK BY THE CONSORTIUM

Project participants have extensive experience at learning technologies and computer vision. On the methodological side, TORMES methodology [28] allows to elicit educational oriented recommendations; and the Collaborative Logical Framework [29] allows to create effective scenarios that support learning through interaction, exploration, discussion, and collaborative knowledge construction. On the development side, extensive work has been performed in intelligent tutoring systems (ITS) and computer vision methods. In particular, an ITS for the arithmetic and algebra domains has been developed [5, 3, 4, 13]. Moreover, the potential applicability of traditional holistic techniques in the context of facial expression recognition using a method based on eigenfaces [21] has been analysed; and works have been developed on modelling the relation between facial action units and changes in the affective state, using kinect devices [6].

The consortium has an extensive research experience on some of the open issues identified above. This experience comes from the joint research carried out under the project MAMIPEC [30]. This project aimed to provide affective personalized support to learners on educational contexts, trying to identify learners' affective states from a multimodal approach based on mining data gathered from several input data sources. To this end, large-scale experiments were carried out in laboratory settings (nearly 100 participants, including visual impaired people). These aimed at building a database of emotional data, from where to analyse the viability of inferring learning emotions in an educational context [31]. Multimodal approaches were also used to combine these interactions with physiological signals [32, 24]. These research experiences aimed to detect changes in the user emotional states while solving mathematical problems. They served to identify an ad-hoc methodology to tag facial expressions and body movements that conform to changes in the affective states of learners while dealing with cognitive tasks in a real time learning process [26]. They also served to explore the viability of mouse and keyboard indicators for emotions detection [25]; and of other low cost devices and specific techniques related to the field of computer vision [21]. In a later experience, affective recommendations identified with TORMES methodology [33] were validated. In addition, AICARP open platform has been implemented at low cost with Arduino to detect changes in physiological signals that can be associated with stressful situations, and when this happens, it recommends the learner to relax by delivering modulated sensorial support in terms of light, sound, or vibration at a relaxation breath rate [34].

3. EMOTION DETECTION

3.1 Physiological sensing: EEG

In the task of identifying the emotion, numerous authors have used diverse kinds of physiological signals [15, 39, 36]. Electroencephalography (EEG) stands out as one of the most studied and yet less understood. In general, these EEG-based models are used to build a classification scheme that use EEG features as an input and yield a prediction related to the user's emotional state as an output [2, 20]. One major factor that affects the system's performance is related to the existing implicit relations between the selected features and the user's reaction to changes in the variables considered in the emotional model. In this sense, feature selection has a remarkable influence on the system's performance. As a first contribution aligned with the project objectives, we have developed a novel classification scheme that combines connectivity and energy EEG features. [36, 19, 17] The method is based on a feature reduction scheme that integrates a one-way ANOVA with a Principal Component Analysis (PCA) to yield two dimensional data which is fed into a non-linear Support Vector Machine (SVM). The method has been tested by using DEAP [16], a database commonly used as a benchmark in this type of applications. Results outperform the ones reported in [16], both in detecting valence and activation.

3.2 Video sensing

Processing images and videos using low-cost devices lead to effective ways of sensing affective states [23]. But instead of accurate, general-purpose recognisers we are interested in methods able to give appropriate hints about some affective variables in real time. Gabor Transform (GT) is a signal processing tool specially suitable when recognition is invariant to motion in certain dimensions. The use of GT as an appropriate feature extractor on large video streams is also motivated by new linear time algorithms able to be used in live scenarios [22, 37, 12].

The use of GT is specially useful in the context of sensing small local changes that are correlated with some affective variables as it is the case of slight changes in a moving face in front of a camera. Videos can therefore be segmented depending on local activity, even helping to segment the video in moving objects against a fixed background, or detect movement as opening of mouth, eyes, and parts of the face, helping to detect gestures directly.

Efficient GT has also been applied to recognize the 3D structure of moving objects, i.e. a face, by taking into account that there are fixed parts (forefront, ears, hair, etc) which behave together as rigid solid subject to perspective changes. Then, it is possible to recover the lost information in a way quite similar to a Markov filter, as only 6 variables are needed to locate the head in space. In this way it is possible to present the face in a normalized 3D way which can greatly help to decide which changes are due to perspective, and which ones are truthfully due to face expressions.

4. REACTING TO EMOTIONS

One major worry in ITS is related to the most adequate level of help that should be provided to the subject to optimize learning [10]. Arguments supporting or against intensive scaffolding methods, that always allows the user to reach the end of the problem, have been common in the literature. Researchers against this type of strategy claim negative effects because students could solve problems without engaging in their content. We reported the effect of an intensive scaffolding in the learning of algebraic word problem solving in our ITS [13]. In particular, we carried out an empirical study in which the effectiveness of two versions of the ITS (one with intensive scaffolding and another without it) were compared. The results show a significantly increase of the competence in algebraic word problem solving in the group that used the ITS with intensive scaffolding in comparison with the group that used the reduced version.

Nevertheless, we have recently extended this study to observe the impact of intensive scaffolding on variables other than learning, by using an adapted version of the ITS where the user inputs his/her state after solving each problem. While the effect on valence and activation does not show a significant correlation with the level of help supplied, the autonomy shows a correlation above 0.35, that we consider significant in statistical terms (as more help, as less gain in autonomy). This means that, although intensive scaffolding may not have a negative impact in learning in a direct way, it may have it in an indirect way by acting on affective variables.

Results derived from this research lead to the necessity to design appropriate methods that simultaneously improve learning and affective variables (valence, activation and autonomy). To this end, we are currently working on incorporating classification methods that decide on the most appropriate help level for each problem step, taking into account variables related to the current student knowledge and the difficulty of the problem at hand, such as the user's previous skill in solving problems, the number of help requests per step and the number of steps with help in the last solved problem, and information about changes in valence, activation and autonomy after finishing the current problem reported by the user through a SAM test. This information, in conjunction with precomputed average scores about users' skill, number of help requests per step and number of steps with help for the previous and current problems, constitutes the input for a set of SVM classifiers whose aim is to predict the most adequate help level for the next problem in terms of valence, activation and autonomy. The SVM classifiers with RBF kernel were implemented in Python using scikit-learn free software machine learning library[1], trained with a leave-one-out cross-validation technique, and optimized with an exhaustive Grid Search procedure to estimate the optimum C and G parameters. The parameters for each classifier and the accuracy in terms of recall and ROC area are shown in the table 1.

Classifier	C	G	Accuracy	ROC	Samples
Valence	100	0.01	70%	0.747	171
Activation	10	0.01	62%	0.643	179
Autonomy	10	0.01	64%	0.683	234

Table 1: Estimated parameters obtained by a Grid Search strategy, accuracy and ROC area for the implemented classifiers

At the time of writing this paper, experiments are being carried out at a secondary school to test the actual performance of these classification methods when used in a real environment. To perform the experiments, we have used a Ubuntu linux live distro with a Xfce desktop environment with data persistence, and with the appropriate and necessary tools to run the ITS in combination with the implemented classifiers.

5. EMOTION MODELING: COMBINING GLOBAL AND INDIVIDUAL MODELS

The performance of global affective models to detect emotions is limited by the subject's individualities [6], which are not taken into account in this type of settings. However, individual models generally suffer from the small sample problem, as there is a an intrinsic difficulty associated with gathering extensive data from the same user.

We are currently working on approaches that are able to use the whole set of data, while at the same time allow for user individualities to be taken into consideration. To this end, we have used two types of methods:

- Clustering methods. As a first approach, users were grouped by common and relevant characteristics that the models are compatible, such as age or gender. However, the results obtained did not significantly improve the state-of-the-art based on real modes. Our current line of work includes methods where the clustering is learned, rather than given according to user features.
- Refinement methods. In this approach, we construct a global model, which is then processed and adapted to each individual as data becomes available. It is in this approach where we get considerable improvement with respect to both, using individual models and global models.

[1] http://scikit-learn.org/stable/

6. CONCLUDING REMARKS AND ONGOING WORK

Teaching arithmetic word problem solving is a complex task. The ability of a teacher to provide feedback that is consistent with the current student reasoning and affective states is a major factor to arrive at the desired learning outcomes. In this paper, we have described strategies to incorporate this kind of skills into an adaptive ITS by considering low cost, low intrusive devices.

Acknowledgments.

This work has been partly supported by the Spanish Ministry of Economy and Competitiveness through project BIG-AFF with grants
TIN2014-59641-C{1,2}-1-P

7. REFERENCES

[1] M. Ainley. Connecting with learning: Motivation, affect and cognition in interest processes. *Educational Psychology Review 18*, pages 391âĂŞ–405, 2006.

[2] O. AlZoubi, R. A. Calvo, and R. H. Stevens. *AI 2009: Advances in Artificial Intelligence: 22nd Australasian Joint Conference, Melbourne, Australia, December 1-4, 2009. Proceedings*, chapter Classification of EEG for Affect Recognition: An Adaptive Approach, pages 52–61. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[3] M. Arevalillo-Herráez, D. Arnau, and L. Marco-Giménez. Domain-specific knowledge Representation and Inference Engine for an Intelligent Tutoring System. *Knowledge-Based Systems*, 49:97–105, 2013.

[4] D. Arnau, M. Arevalillo-Herráez, and J. Gonzalez-Calero. Emulating human supervision in an intelligent tutoring system for arithmetical problem solving. *IEEE Transactions on Learning Technologies*, To be published, April 2014.

[5] D. Arnau, M. Arevalillo-Herráez, L. Puig, and J. A. González-Calero. Fundamentals of the Design and the Operation of an Intelligent Tutoring System for the Learning of the Arithmetical and Algebraic Way of Solving Word Problems. *Computers & Education*, 63:119–130, 2013.

[6] A. Ayesh, M. Arevalillo-Herráez, and F. J. Ferri. Cognitive reasoning and inferences through psychologically based personalised modelling of emotions using associative classifiers. In *IEEE 13th International Conference on Cognitive Informatics and Cognitive Computing, ICCI*CC 2014, London, UK, August 18-20, 2014*, pages 67–72, 2014.

[7] J. N. Bailenson, E. D. Pontikakis, I. B. Mauss, J. J. Gross, M. E. Jabon, C. A. Hutcherson, C. Nass, and O. John. Real-time classification of evoked emotions using facial feature tracking and physiological responses. *International Journal of Human-Computer Studies*, 66(5):303 – 317, 2008.

[8] M. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, and J. Movellan. Fully automatic facial action recognition in spontaneous behavior. In *Automatic Face and Gesture Recognition, 2006. FGR 2006. 7th International Conference on*, pages 223–230, April 2006.

[9] R. Calvo and S. D'Mello. Affect detection: An interdisciplinary review of models, methods, and their applications. *Affective Computing, IEEE Transactions on*, 1(1):18–37, Jan 2010.

[10] S. K. D'Mello, S. D. Craig, and A. C. Graesser. Multimethod assessment of affective experience and expression during deep learning. *Int. J. Learn. Technol.*, 4(3/4):165–187, Oct. 2009.

[11] S. K. D'Mello and A. Graesser. Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-Adapted Interaction*, 20(2):147–187, 2010.

[12] H. G. Feichtinger and T. Strohmer, editors. *Gabor Analysis and Algorithms: Theory and Applications.* Birkhauser Boston, 1998.

[13] J. A. González-Calero, D. Arnau, L. Puig, and M. Arevalillo-Herráez. Intensive scaffolding in an intelligent tutoring system for the learning of algebraic word problem solving. *BJET*, 46(6):1189–1200, 2015.

[14] R. Kaliouby and P. Robinson. *Affective Computing and Intelligent Interaction: First International Conference, ACII 2005, Beijing, China, October 22-24, 2005. Proceedings*, chapter Generalization of a Vision-Based Computational Model of Mind-Reading, pages 582–589. Springer Berlin Heidelberg, Berlin, Heidelberg, 2005.

[15] K. H. Kim, S. W. Bang, and S. R. Kim. Emotion recognition system using short-term monitoring of physiological signals. *Medical and Biological Engineering and Computing*, 42(3):419–427, 2004.

[16] S. Koelstra, C. Mühl, M. Soleymani, J. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras. DEAP: A database for emotion analysis ;using physiological signals. *IEEE Trans. Affective Computing*, 3(1):18–31, 2012.

[17] M. B. Kostyunina and M. A. Kulikov. Frequency characteristics of eeg spectra in the emotions. *Neuroscience and Behavioral Physiology*, 26(4):340–343, 1996.

[18] I. Kotsia and I. Pitas. Facial expression recognition in image sequences using geometric deformation features and support vector machines. *Image Processing, IEEE Transactions on*, 16(1):172–187, Jan 2007.

[19] Y. P. Lin, C. H. Wang, T. P. Jung, T. L. Wu, S. K. Jeng, J. R. Duann, and J. H. Chen. Eeg-based emotion recognition in music listening. *IEEE Transactions on Biomedical Engineering*, 57(7):1798–1806, July 2010.

[20] F. Lotte, M. Congedo, A. LÃl'cuyer, F. Lamarche, and B. Arnaldi. A review of classification algorithms for eeg-based brainâĂŞcomputer interfaces. *Journal of Neural Engineering*, 4(2):R1, 2007.

[21] L. Marco-Giménez, M. Arevalillo-Herráez, and C. Cuhna-Pérez. Eigenexpressions: Emotion recognition using multiple eigenspaces. In *Pattern Recognition and Image Analysis - 6th Iberian Conference, IbPRIA 2013, Funchal, Madeira, Portugal, June 5-7, 2013. Proceedings*, pages 758–765, 2013.

[22] S. Moreno-Picot, M. Arevalillo-Herráez, and W. D. Villanueva. A linear cost algorithm to compute the discrete Gabor transform. *IEEE Transactions on*

Signal Processing, 58(5):2667–2674, 2010.

[23] M. Pantic and L. Rothkrantz. Toward an affect-sensitive multimodal human-computer interaction. *Proceedings of the IEEE*, 91(9):1370–1390, Sept 2003.

[24] S. Salmeron-Majadas, M. Arevalillo-Herráez, O. C. Santos, M. Saneiro, R. Cabestrero, P. Quirós, D. Arnau, and J. Boticario. Filtering of spontaneous and low intensity emotions in educational contexts. In *Artificial Intelligence in Education - 17th International Conference, AIED 2015, Madrid, Spain, June 22-26, 2015. Proceedings*, pages 429–438, 2015.

[25] S. Salmeron-Majadas, O. C. Santos, and J. Boticario. An evaluation of mouse and keyboard interaction indicators towards non-intrusive and low cost affective modeling in an educational context. In *18th International Conference in Knowledge Based and Intelligent Information and Engineering Systems, KES 2014, Gdynia, Poland, 15-17 September 2014*, pages 691–700, 2014.

[26] M. Saneiro, O. C. Santos, S. Salmeron-Majadas, and J. G. Boticario. Towards emotion detection in educational scenarios from facial expressions and body movements through multimodal approaches. *The Scientific World Journal*, 2014(Article ID 484873):14, 2014.

[27] O. C. Santos. Emotions and personality in adaptive e-learning systems: an affective computing perspective. In Tkalcic, DeCarolis, de Gemmis, Odic, and Kosir, editors, *Emotions and Personality in Personalized Services*, chapter 13, page (accepted). 2016.

[28] O. C. Santos and J. Boticario. Practical guidelines for designing and evaluating educationally oriented recommendations. *Computers & Education*, 81:354–374, 2015.

[29] O. C. Santos and J. G. Boticario. Involving users to improve the collaborative logical framework. *The Scientific World Journal*, 2014(Article ID 893525):15, 2014.

[30] O. C. Santos, J. G. Boticario, Arevalillo-Herraez, S. M., C. M., C. R., and S. Salmeron-Majadas. MAMIPEC - affective modeling in inclusive personalized educational scenarios. *Bulletin of the IEEE Technical Committee on Learning Technology*, 14(4):35âĂŞ–38, 2012.

[31] O. C. Santos, A. Rodríguez-Ascaso, J. Boticario, S. Salmeron-Majadas, P. Quirós, and R. Cabestrero. Challenges for inclusive affective detection in educational scenarios. In *Universal Access in Human-Computer Interaction. User and Context Diversity - 7th International Conference, UAHCI 2013, Held as Part of HCI International 2013, Las Vegas, NV, USA, July 21-26, 2013, Proceedings, Part I*, pages 566–575, 2013.

[32] O. C. Santos, S. Salmeron-Majadas, and J. Boticario. Emotions detection from math exercises by combining several data sources. In *Artificial Intelligence in Education - 16th International Conference, AIED 2013, Memphis, TN, USA, July 9-13, 2013. Proceedings*, pages 742–745, 2013.

[33] O. C. Santos, M. Saneiro, S. Salmeron-Majadas, and J. Boticario. A methodological approach to eliciting affective educational recommendations. In *IEEE 14th International Conference on Advanced Learning Technologies, ICALT 2014, Athens, Greece, July 7-10, 2014*, pages 529–533, 2014.

[34] O. C. Santos, R. Uria-Rivas, M. C. Rodriguez-Sanchez, and J. G. Boticario. An open sensing and acting platform for context-aware affective support in ambient intelligent educational settings. *IEEE Sensors*, 16(10):3865–3874, May 2016.

[35] N. Sebe, I. Cohen, T. Gevers, and T. S. Huang. Multimodal approaches for emotion recognition: a survey. In S. Santini, R. Schettini, and T. Gevers, editors, *Internet Imaging VI: Proceedings of the Society of Photo-Optical Instrumentation Engineers*, volume 5670, pages 56–67. SPIE, 2005.

[36] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic. A multimodal database for affect recognition and implicit tagging. *Affective Computing, IEEE Transactions on*, 3(1):42–55, Jan 2012.

[37] P. Søndergaard. Efficient algorithms for the discrete Gabor transform with a long FIR window. *Journal of Fourier Analysis and Applications*, 18(3):456–470, 2012.

[38] Y. Tong, J. Chen, and Q. Ji. A unified probabilistic framework for spontaneous facial action modeling and understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(2):258–273, 2010.

[39] J. Wagner, J. Kim, and E. Andre. From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification. In *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, pages 940–943, July 2005.

[40] Z. Zeng, M. Pantic, G. Roisman, and T. Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(1):39–58, Jan 2009.