HAL
open science

Toward a generic architecture for UI adaptation to emotions

Julián Andrés Galindo, Éric Céret, Sophie Dupuy-Chessa

To cite this version:

Julián Andrés Galindo, Éric Céret, Sophie Dupuy-Chessa. Toward a generic architecture for UI adaptation to emotions. 29ème conférence francophone sur l'Interaction Homme-Machine, AFIHM, Aug 2017, Poitiers, France. 10 p., 10.1145/3132129.3132156 . hal-01578221

HAL Id: hal-01578221

https://hal.science/hal-01578221v1

Submitted on 28 Aug 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Toward a generic architecture for UI adaptation to emotions

Julián Andrés Galindo
Univ. Grenoble Alpes, CNRS, LIG
F-38000, Grenoble, France
Escuela Politécnica Nacional, Facultad de Ingeniería de Sistemas, Quito, Ecuador
julian.galindo@univ-grenoble-alpes.fr

Éric Céret
Univ. Grenoble Alpes, CNRS,
Grenoble INP, LIG
F-38000, Grenoble, France
eric.ceret@univ-grenoble-alpes.fr

Sophie Dupuy-Chessa
Univ. Grenoble Alpes, CNRS,
Grenoble INP, LIG
F-38000, Grenoble, France
sophie.dupuy@univ-grenoble-alpes.fr

RÉSUMÉ

L'adaptation des interfaces homme-machine à l'exécution est une problématique bien connue en interaction homme-machine qui devient encore plus prégnante lors qu'il s'agit de prendre en compte des caractéristiques dynamiques de l'utilisateur telles que les émotions. En guise de

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.
IHM'17, August 29-September 1, 2017, Poitiers, France
© 2017 Copyright is held by the owner/author(s).
ACM ISBN 978-1-4503-5109-6/17/08. https://doi.org/10.1145/3132129.3132156

contribution à l'adaptation des interfaces aux émotions, nous proposons Perso2U, une architecture de personnalisation à l'exécution. Cette approche s'appuie sur des outils de reconnaissance des émotions, ce qui conduit à se questionner sur la précision et la similarité, entre outils, des émotions détectées pour conduire l'adaptation. Cet article a pour but de montrer que des résultats similaires peuvent être obtenues à partir de reconnaissance faciale d'émotions afin de montrer l'indépendence aux outils du moteur d'inférence des émotions et plus généralement de l'architecture. En ce sens, l'article présente une expérimentation qui compare trois outils de détection d'émotions.

MOTS CLÉS

Adaptation des interfaces homme-machine, architecture, détection d'émotions

ABSTRACT

Adapting at runtime user interfaces is a well-known requirement in human computer interaction which becomes a very challenging task when taking into account dynamic user properties such as emotions. To address the question of adapting user interfaces to emotions, we propose Perso2u is an architecture to personalize user interfaces with user emotions at runtime. This approach relies on emotion recognition tools which raises the question of accuracy This paper aims at showing that it is possible to obtain similar emotion results from several tools based on face recognition to emphasize the independence of the emotion inferring engine and more globally of the architecture. To achieve this goal, this paper reports on the results of an experiment to compare three emotion detection tools.

CCS CONCEPTS

Human-centered computing → **Interaction design** → **Interaction design process and methods** → User centered design; User interface design • **Software and its engineering** → *Software architectures*

KEYWORDS

User interface adaptation, architecture, emotion detection.

ACM Reference format:

Julián Andrés Galindo, Éric Céret and Sophie Dupuy-Chessa. 2017. Toward a generic architecture for UI adaptation to emotions. In *Proceedings of 29eme Conference Francophone sur l'Interaction Homme-Machine , Poitiers- Futuroscope, FR, Août 2017 (IHM 2017),* 9 pages.
DOI: 10.1145/3132129.3132156

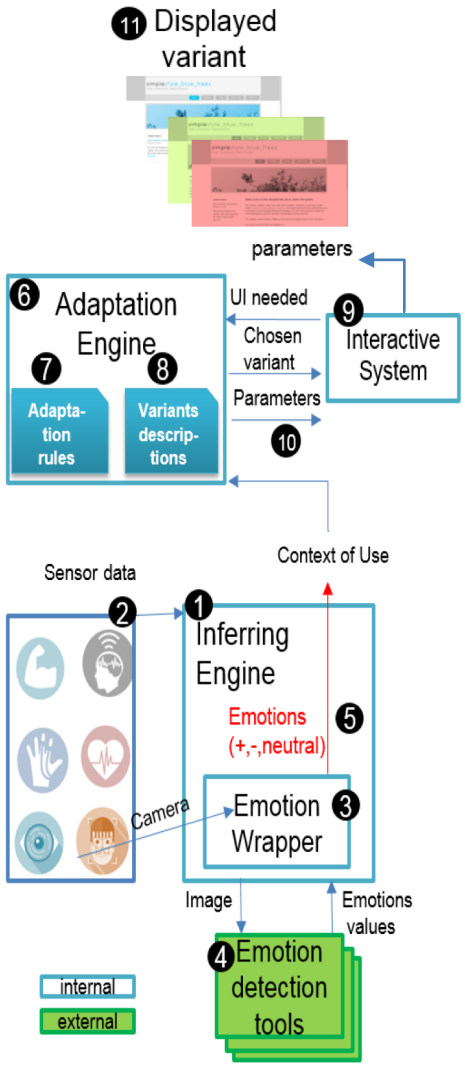


Figure 1: Global schema of the Perso2U architecture.

1 INTRODUCTION

Human Computer Interfaces must now be able to dynamically adapt to changes which may occur in their context of use (user, platform and environment) while preserving usability [1]. Such adaptation can deal with dynamic user properties such as emotions felt by users during interaction. As a matter of fact, emotions are recognized as important factors for considering user experience [2, 3]. So an interesting approach is to use emotions for user interfaces (UI) adaptation [4]. As previous works in UI [5, 6], such approaches need to be generic. One element of genericity is the architecture. In [7], we propose an architecture, Perso2U, for UI adaptation to emotions. This architecture seems to be generic as adaptation is based on rules that do not depend on emotion detection tools. However, the success of UI adaptation depends on the finding of relevant emotions. Hence a crucial part of the architecture is emotion detection and particularly its independence from emotion recognition tools.

Nowadays, there are many emotion recognition tools such as Microsoft API[1] ,FaceReader[2] and Affdex[3] .These tools usually provide seven basic emotions with some uncertainty. There is clearly a risk of divergence between them because of the accuracy differences in face detection algorithms [8]. These divergences could lead to erratic emotion identification and thus to unreliable and unstable adaptations. So our question is related to the possibility to use any of these tools in our architecture. That is the point studied in this paper to show the possibility of a generic interpretation of emotions. In this paper, we propose to compare three face emotion recognition tools used in Perso2u. We hypothesize that from these tools, the inferring engine can provide similar emotion results (positive, negative, surprise and neutral) during the interaction. To identify this similarity, we conduct an experiment with 40 users across 4 websites. This paper describes the results of this experiment and its impact on the PersoU2 architecture.

The following section of the paper presents the related work about UI adaptation to emotions. The third section describes the Perso2u architecture. Next we present an experiment that compares three emotion recognition tools. The last section summarizes our results and presents future work.

2 RELATED WORK

Users' emotions have been modeled in many works [12, 13]. In such works, emotions are studied in relation to other users' features such as learnability or performance. However, none of them deals with reusing this knowledge about the user for exploring deeply UI adaptation. Few other proposals use emotion recognition to adapt the UI [4, 11, 12]. In this paper, we are particularly interested in the genericity of their architecture, i.e. in the independence of their adaptation process from the emotion recognition tool used. First, the Affect and Belief Adaptive Interface

[1] www.microsoft.com/cognitive-services/en-us/emotion-api
[2] www.noldus.com/human-behavior-research/products/facereader
[3] www.affectiva.com/solutions/affdex/

Figure 2: Four variants of a web site user interface.

System (ABAIS) [10, 11] is based on four modules: User state assessment, impact Prediction, strategy Selection, and GUI Adaptation. The first one is in charge of the emotion inference, which combines user information and a rule-based approach to infer the affective user state. Even if it is not explicit in the papers, the process appears to be independent from recognition tools thanks to these rules. But only one emotion, anxiety, is considered and UI adaptation only affects icons, displays, notifications and custom configuration. Adaptation does not consider the structure (layout) of the UI.

Second, Nasoz proposes an adaptive intelligent system [4] relying on the recognition of affective states from physiological signals (sensors) based on a Multimodal Affective User Interface architecture [13]. The emotion inference analyses data given by emotion recognition tools (e.g. facial analysis) while the adaptation relies on a machine learning technique that considers some user's attributes such as recognized emotion, personality trait, age, gender and emotion frequency. Independence from recognition tools could be ensured thanks to the independence of the inference of emotions from these tools. However, adaptation rules should be made able to deal with variations in user parameters like emotion frequency (e.g. from rarely to often). To summarize, only two main architectures exist. If they seem to be generic, they have (1) limited emotions and UI adaptations and (2) fixed adaptation rules. To avoid this drawback, we propose an architecture called Perso2U.

2.1 Perso2U ARCHITECTURE

The Perso2U architecture (Fig. 1) is composed of three components: the Inferring Engine ❶, the Adaptation Engine ❻ and the Interactive System ❾. The Inferring engine is in charge of dynamically deducing the value of the context of Use (users with their emotions, platform, environment). Emotions are detected thanks to external Tools such as FaceReader. Then the emotion Wrapper ❸ filters and aggregates the acquired emotions values to find whether the current user emotion is positive (happiness), negative (anger, contempt, disgust, fear, and sadness), surprise or neutral. Finally inferred emotions are sent to the Adaptation Engine with other contextual information. This engine elicits the right UI adaptation: from the contextual information [14] ❺, it (a) selects a UI variant (i.e. a predefined UI structure) among all available variants for the needed UI and (b) computes parameters (e.g. color, font) for even better adapting the chosen variant to the context of use. Then it displays the variant and gives it the computed parameters to the interactive module, which applies the parameters, thus displaying the appropriate adapted UI. To extend, a more complete analysis of Perso2u which includes descriptions of functions, components, examples, current prototype and feasibility results is evidenced in [7].

Table 1: User interaction data collected by clicks and images.

	Clicks	*Images*
w1	515	226
w2	549	280
w3	944	354
w4	746	380
Total	2754	1240

Table 2: User interaction data collected by avg. interaction time and page views. w1=web site 1, w2=web site 2, w3=web site 3 and w4=web site 4.

	avg. Interaction time (minutes)	*page views*
w1	32	289
w2	29	172
w3	27	216
w4	34	228
Total	122	905

3 EXPERIMENTAL STUDY

The experimental study aims at exploring one initial and promising criterion for the genericity of the architecture: the independence from the way of detecting emotions. So the inferring engine must be able to send appropriate emotions to the adaptation engine whatever the detection tool is. We will show that from three different detection tools, the inferring engine can infer similar results.

3.1 Goal and hypothesis

The objective of this experiment is to examine how similar are emotions, i.e. contempt, happiness, anger, disgust, fear, sadness, surprise and neutral [4] ,detected by different face emotion recognition tools. For getting reliable comparison, we needed that these tools start from the same input, the users' pictures taken during interaction. We also wanted the tools to provide similar sets of emotions, so that their aggregation is meaningful. Three tools could be identified that answer these constraints: Affectiva, FaceReader and Microsoft emotion API. So the experiment hypothesis is:

H1: the emotions detected by Affectiva, FaceReader and Microsoft emotion API are similar enough to be able to infer from them similar aggregated emotions: positive (happiness), negative (anger, contempt, disgust, fear and sadness), surprise and neutral.

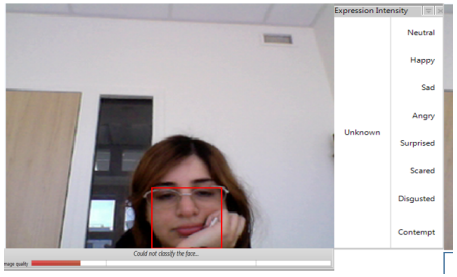
3.2 Experimental method

Four UI versions of a same travel website were designed and developed [15]. The websites differ by two factors that influence users' emotions while interacting with a UI: usability and aesthetics. So by combining two values (low vs. high) for these two factors (aesthetics and usability), four versions of the website were created (Fig. 2).

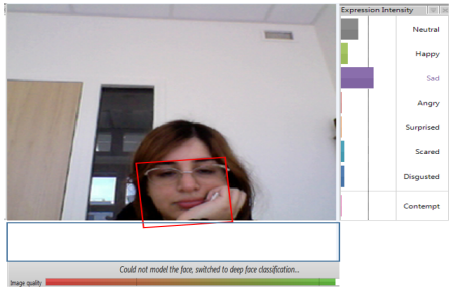
3.3 Procedure, tasks and participants

Each participant was assigned to one version of the travel website. While he/she was interacting with the website, his/her emotions were detected based on user taken photos every 10 seconds. Second, to ensure that participants are involved enough in the interaction, they had to answer an online questionnaire corresponding to three tasks to achieve: (1) look for a 5-stars hotel offering a visit in the desert, (2) look for a specific hotel with a breakfast plan and (3) look for a hotel circuit in a specific location. Lastly, 40 persons participated to the experiment, 10 participants per website. The participants were 50% male from 21 to 67 years old and 50% female from 23 to 63 years old (mean male age 34, σ=13.95 and mean female age 27 with σ=12.34).

[4] Not available in Affectiva (JS SDK).



A. Before (300x260)



B. After (300x320).

Figure 3: FaceReader image correction (A and B).

3.4 Interaction Data and Images Correction

During the experiment, the following triple data is captured per user: User, UI and the Survey which results in a set of 8 emotions, 67 UI elements and 41 Survey attributes across 40 users. To clarify, user features such as identifier (1 to 40), age, gender and a set of emotions: happiness, contempt, fear, anger, sadness, disgust, surprise and neutral (when available) with values from 0 to 1. Furthermore, table 1 and 2 illustrates interaction data for each web site: the number of clicks, the number of images detected, avg. time and page views per website. These measures reached in total 2754,1240,122 and 905 respectively.

3.5 Measures and method

User pictures were sent to the 3 emotions detection tools for getting the emotions detected in real-time for FaceReader and Microsoft in contrast to Affectiva where post analysis was done. Before being able to compare emotions detected by tools, we had to make some corrections. As a matter of fact, to get realistic values from tools during the interaction, participants were not required to have a predefined position (rigid head or hand positions) as it is usually done. But some images were not good enough for an interpretation with FaceReader. So a face image correction was done for FaceReader for every image. The correction was done by adding a white rectangle at the bottom of the image without affecting the user face as illustrated in Fig.3. As a benefit, it was possible to improve the image detection, quality and gather emotion values (Fig3. B).

Finally, the 1226 captured images were sent to Microsoft API and FaceReader and Affdex, which could recognize respectively 880, 703 and 706 of them. In other words, each tool produced respectively an error rate of 28.22%, 42.65% and 42,41%. Each tool produced for each picture a vector of emotion measures with values from 0 to 1: happiness, contempt, fear, sadness, anger, disgust, surprise and, for Microsoft API and FaceReader, neutral. Then distances for each emotion measure across all tools were computed by using Euclidean distances to compare how similar they are [16]. This process relies on a distance function applied to the three vectors: emotionDistance=distance {VfaceReader,i, VMSAPI,i, VAffDex,i}/n where Vx,i is the value of emotion I in vector of tool X, and n refers to the number of emotions measures across all vectors. This division was necessary to obtain a central value per data set (vector). To illustrate, happiness distance is equal to the distance among VfaceReader,happiness, VMSAPI,happiness, VAffDex,happiness which is then divided by the number of vector items

Table 3: Distances among emotion recognition tools in surprise and neutral emotions.

tools	*surprise*	*neutral*
a-f	0,0096	n.a.
a-m	0,0075	n.a.
f-m	0,0069	1,62%
avg	0,80%	

a: affectiva, f: faceReader and m: microsoft

Table 4: Distances among emotion recognition tools in positive emotion.

tools	*happiness*
a-f	0,0094
a-m	0,0079
f-m	0,0081
avg	0,84%

a: affectiva, f: faceReader and m: microsoft

4 RESULTS

4.1 Feasibility of a generic architecture

The results confirm that there are similar recognized emotions among the three tools: low distances were found with values under 2.0% for all emotions (happiness, contempt, disgust, sadness, fear, anger and surprise) as shown in Fig.4 and tables 3, 4 and 5. In fact, more similarity is shown in negative than positive emotions with a global average of 0.96 and 0.84% for each group (Table 5 & 4). While neutral was almost a double than surprise (0,8%, Table 3).

Table 5: Distances among emotion recognition tools in negative emotions.

tools	*disgust*	*sadness*	*fear*	*anger*	*contempt*	*avg.*
a-f	0,0092	0,0151	0,0070	0,0085	0,0204	1,20%
a-m	0,0102	0,0119	0,0036	0,0045	0,0190	0,98%
f-m	0,0035	0,0115	0,0061	0,0067	0,0070	0,70%
avg.	0,77%	1,28%	0,55%	0,66%	1,55%	0,96%

a: affectiva, f: faceReader and m: microsoft.

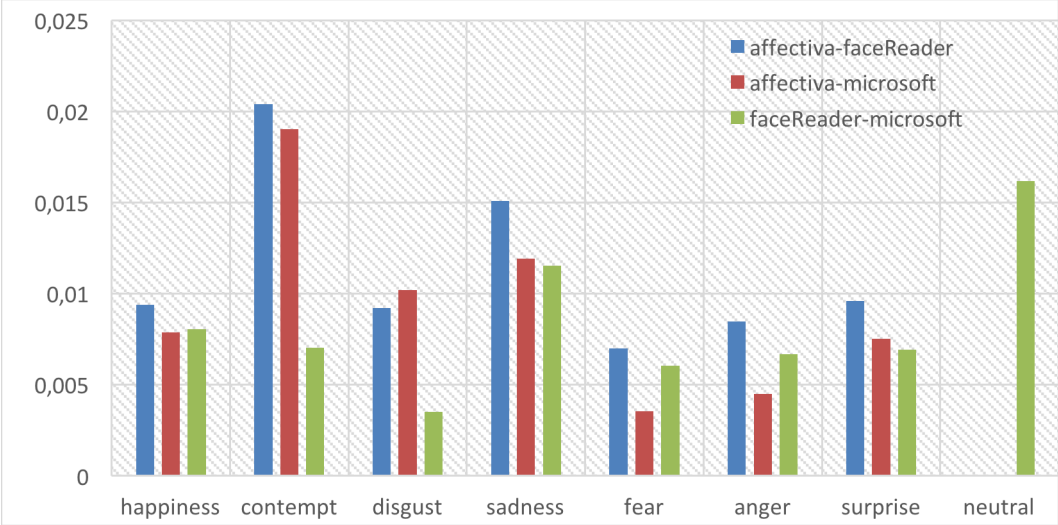
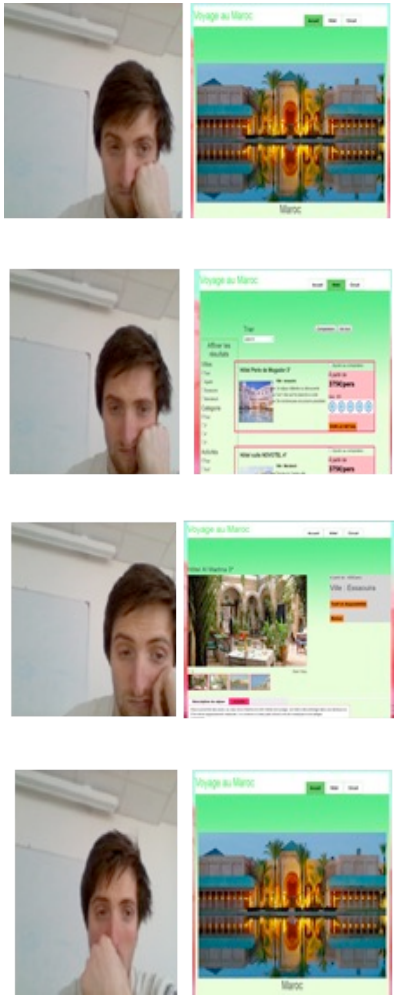


Figure 4: Distances among emotion recognition tools.



4.2 Failed emotion recognition

Fig. 5 evidences that there are cases where the emotion recognition fails for a specific time slot. In fact, during this user interaction, there were 2, 4, and 2 failed emotion detections for Affectiva, FaceReader and Microsoft in different time slots and UI variances. In such case, the adaptation may need to consider a time range rather than a time slot to define the user response at interacting with the UI.

Figure 5. A User Happiness detection for each tool.

4.3 Experiment limitations

Although these results are encouraging, some restrictions should be noted. First, all presented tools rely on the same emotion detection method (user face image) without considering others such as eye tracking, heart rate or video trace analysis. Second, other elements in the context of use have not been considered such as noisy and light environment or different platforms (e.g. mobile). Lastly, correlations between users' declared and recognized emotions will extend the inferring engine validation.

5 CONCLUSIONS AND PERSPECTIVES

This paper presents an architecture named Perso2u for UI adaptation emotions and an experiment to show its independence from emotions detection tools. It validates that it is feasible to have similar emotion values with different tools as first step (inferring) in a generic architecture to drive UI adaptation based on emotions. More experiments will be necessary to prove the correlation between the declared emotions and the inferred ones in order to validate the inferring engine. From the inferred emotions, some adaptation rules must be implemented and experimented to validate the global architecture.

REFERENCES

[1] G. Calvary, J. Coutaz, D. Thevenin, Q. Limbourg, L. Bouillon, and J. Vanderdonckt, "A unifying reference framework for multi-target user interfaces," *Interact. Comput.*, vol. 15, no. 3, pp. 289–308, 2003.

[2] B. Reeves and C. Nass, *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge university press Cambridge, UK, 1996.

[3] S. Carberry and F. de Rosis, "Introduction to special Issue on 'Affective modeling and adaptation,'" *User Model. User-Adapt. Interact.*, vol. 18, no. 1–2, pp. 1–9, Jan. 2008.

[4] F. Nasoz, "Adaptive intelligent user interfaces with emotion recognition," University of Central Florida Orlando, Florida, 2004.

[5] A. García Frey, E. Céret, S. Dupuy-Chessa, G. Calvary, and Y. Gabillon, "UsiComp: an extensible model-driven composer," in *Proceedings of the 4th ACM SIGCHI symposium on Engineering interactive computing systems*, 2012, pp. 263–268.

[6] F. Paterno, C. Santoro, and L. D. Spano, "MARIA: A universal, declarative, multiple abstraction-level language for service-oriented applications in ubiquitous environments," *ACM Trans. Comput.-Hum. Interact. TOCHI*, vol. 16, no. 4, p. 19, 2009.

[7] J. Galindo, S. Dupuy-Chessa, and É. Céret, "Toward a UI Adaptation Approach Driven by User Emotions," presented at the ACHI07, Nice, France, 2017.

[8] G. Brodny, A. Ko\lakowska, A. Landowska, M. Szwoch, W. Szwoch, and M. R. Wróbel, "Comparison of selected off-the-shelf solutions for emotion recognition based on facial expressions," in *Human System Interactions (HSI)*, 2016, pp. 397–404.

[9] J. Rowe, B. Mott, S. McQuiggan, J. Robison, S. Lee, and J. Lester, "Crystal island: A narrative-centered learning environment for eighth grade microbiology," in *workshop on intelligent educational games at the 14th international conference on artificial intelligence in education, Brighton, UK*, 2009, pp. 11–20.

[10] K. Forbes-Riley and D. Litman, "Designing and evaluating a wizarded uncertainty-adaptive spoken dialogue tutoring system," *Comput. Speech Lang.*, vol. 25, no. 1, pp. 105–126, Jan. 2011.

[11] E. Hudlicka and M. D. McNeese, "Assessment of User Affective and Belief States for Interface Adaptation: Application to an Air Force Pilot Task," *User Model. User-Adapt. Interact.*, vol. 12, no. 1, pp. 1–47, Feb. 2002.

[12] E. Hudlicka, "Increaing Sia Architecture Realism by Modeling and Adapting to Affect and Personality," in *Socially Intelligent Agents*, Springer, 2002, pp. 53–60.

[13] C. Lisetti and G. Bastard, "MAUI: a Multimodal Affective User Interface Sensing User's Emotions based on Appraisal Theory-Questions about Facial Expressions…," in *1st Workshop of the EC HUMAINE Network of Excellence on Research on Emotions and Human-Machine Interaction, Geneva, Switzerland*, 2004.

[14] E. Céret, S. Dupuy-Chessa, G. Calvary, and M. Bittar, "System and method for magnetic adaptation of a user interface," TPI2015053 déposé via France INPI le 7 juillet 2015, 2ème dépôt le 7 juillet 2016.

[15] S. Dupuy-Chessa, Y. Laurillau, and E. Céret, "Considering aesthetics and usability temporalities in a model based development process," in *Actes de la 28ième conférence francophone sur l'Interaction Homme-Machine*, 2016, pp. 25–35.

[16] J. Kalita and K. Das, "Recognition of Facial Expression Using Eigenvector Based Distributed Features and Euclidean Distance Based Decision Making Technique," *ArXiv13030635 Cs*, Mar. 2013.