Video access and interaction based on emotions

Eva Oliveira

LaSIGE, University of Lisbon
FCUL, 1749-016 Lisbon, Portugal,
IPCA, 4750-117 Arcozelo BCL,
Portugal
+351 2175000533
eoliveira@ipca.pt

ABSTRACT

Films are by excellence the form of art that exploits our affective, perceptual and intellectual activity. Technological developments and the trends for media convergence are turning video into a dominant and pervasive medium and online video is becoming a growing entertainment activity on the web and iTV. Alongside, Human Computer Interaction research community has been using physiologic, brain and behavior measures to study possible ways to identify and use emotions in human-machine interactions. In our work we explore emotional recognition and classification mechanisms in order to provide video classification based on emotions, and to identify each user's emotional states so as to provide different access mechanisms. We also focus on emotional movie access and exploration mechanisms to explore ways to access and visualize videos based on their emotional properties and users' emotions and profiles.

Categories and Subject Descriptors

H.5.1 [Information Interfaces and Presentation (I.7)]: Multimedia Information Systems – video; H.5.2 [Information Interfaces and Presentation (I.7)]: User Interfaces – screen design;

General Terms

Design, Experimentation, Human Factors.

Keywords

Affective computing, Emotion-aware systems, Human-centred design, Psychophysiological measures, Pattern-recognition, Discriminant analysis, Support vector machine classifiers, Movies classification and recommendation.

1.INTRODUCTION

Video growth over the Internet changed the way users search, browse and view video content. Watching movies over the Internet is increasing and becoming a pastime. The possibility of streaming to TV Internet content, advances in video compression techniques and video streaming have turned this recent modality

of watching movies easy and doable. Films are by excellence the form of art that evolves affective, perceptual and intellectual activity. It is called as a way to transport us to new worlds, lives and fantasies by telling stories [11].

By combining diverse symbol systems, such as pictures, texts, music and narration, video is a very rich media type, often engaging the viewer cognitively and emotionally, and having a great potential in the promotion of emotional experiences. It has been used in different contexts: as a way to capture and show real events, to create and visualize scenarios not observable in reality, to inform, to learn, to tell stories and entertain, and to motivate; as movies, documentaries or short video clips. Isen et. al. [9] attested this potential, when she and her colleagues experimented the effect of positive affect in her patients, inducted by ten-minu-te comedy films. The study of films as an emotion induction method has some reports dated from 1996, as [12] reported, analyzing the mental operations of film viewers and discussing how emotions guide the motivation of perception and consequently the control of our attention by cinematographic narratives.

Emotion studies have been done over the last few years, since it became proved that they are fundamental in cognitive and creative processes. In fact, understanding emotions is crucial to understanding motivation, attention or aesthetic phenomena. There is an increasing awareness in the HCI community of the important role of emotion in human computer interactions and interface design, and new mechanisms for the development of interfaces that register and respond to emotions have been studied [3]. Gathering emotional information from users can contribute to create emotional context in applications interfaces. Rosalind Picard in [17] defends that systems that ignore the emotional component of human life are inevitably inferior and incomplete, and she states that systems that provide a proper and useful social and emotional interaction are not science fiction but a science fact.

Society's relation with technology is changing in such ways that it is predictable that, in the next years, Human Computer Interaction (HCI) will be dealing with users and computers that can be anywhere, and at anytime, and this changes interaction perspectives for the future. Human body changes, expressions or emotions would constitute factors that became naturally included in the design of human computer interactions [2]. There is a wide spectrum of areas that investigate emotions with different, but complementary, perspectives. For example, in the neurobiological area, [5] showed that emotions play a major role on cognitive and

decision making processes; HCI aims to understand the way users experience interactions and strives to stimulate the sense of pleasure and satisfaction by developing systems that focus on new intelligent ways to react to user's' emotions [10]. HCI is also concerned with evaluation and usability, which includes evaluating the extent and accessibility of the system's user interface, accessing a user's experience in interaction and identifying specific problems. The advent of rich multimedia interfaces has been providing new technological foundations to support these emotional trends in HCI. Currently, affective computing systems are being developed that are able to recognize and respond to emotions with the aim of improving the quality of human-computer interaction. Part of this research has concentrated on solving many technical issues involved with emotion recognition technologies. For example, [14] describe their work with sensors in the context of a study on emotional physiological response. According to [1] physiological measures such as galvanic skin response or pupil dilation constitute objective factors but are not easily correlated to particular emotions. Moreover, there are variations in rates which are due to normal individual differences among users, and intrusive wires or sensors may affect users' behaviors. To circumvent this, less intrusive devices were developed [17].

In our work we are developing a novel emotional recognition approach based on pattern recognition techniques, based on discriminant analysis and support vector machine classifiers, which are validated using movies' scenes selected to induce emotions ranging from the positive to the negative valence dimension, including happiness, anger, disgust, sadness, and fear. We present the system, iFelt, an interactive web video system designed to learn users' emotional patterns, create emotional profiles for both users and videos, and explore this information to create emotion based interactions. The iFelt system is composed of two components. The "Emotional Recognition and Classification" component performs emotional recognition and classification and semantic representation of emotions in order to provide video classification based on emotions, and to identify each user's emotional states so as to provide different search and access mechanisms. The "Emotional Movie Access and Exploration" component explores ways to access, to search, to represent and visualize videos based on their emotional properties and users' emotions and profiles.

2.PROBLEM

Emotions can be expressed in a variety of ways such as body expressions (facial, vocals, body posture), or neurophysiologic symptoms (respiration, heart-rate, galvanic skin response and blood pressure). Accordingly, the Human Computer Interaction research community has been using physiologic, brain and behavior measures to study possible ways to identify and use emotions in human-machine interactions [10]. However, there are still challenges in the recognition processes, regarding the effectiveness of the mechanisms used to induce emotions. The induction is the process through which people are guided to feel one or more specific emotions, which provokes body reactions. Some relevant works showed that films were one of the most effective ways to induce emotions. J. Gross et al. [6] tried to find as many films as possible to elicit discrete emotions and find the best films for each discrete emotion. In 1996, a research group [12] tested eleven induction methods and concluded that films are the best method to elicit emotions (both positive and negative).

The exploration of movies by their emotional dimensions can be used for entertainment, education or even medical purposes.

Considering the effect of the emotions in a person's attention, motivation and behavior, a scenario where it would be beneficial to have emotional impact videos to capture viewer's attention is in educational contexts, where video could capture students' attention in different ways, either to focus or to relax. The induction of emotions using movies has been largely used in psychology studies [8] and in health related studies. In fact, experimental studies confirmed that positive emotions can have a beneficial effect on physical health [15]. The development of new mechanisms to catalog, find and access movies based on emotions could help to assess videos' emotion impact, and to find movies or scenes that tend to induce a certain feeling in the users. It could also aid filmmakers to perceive the emotional impact of their movies and, in particular, the emotional impact of each scene and compare it to the intention they had for the scene impact, and relate it to the adoption of specific special effects, acting approaches and settings. Moreover, actors may also be able to perceive their impact in a specific act. Finally, movie consumers may be able to explore movies by the emotions stirred by the content in multiple ways, compare their emotional reactions with other users' reactions and see how they change overtime.

Other challenges in accessing video is the fact that it conveys a huge amount of audiovisual information that is not structured and that changes along time, and so, accessing all the data that a video can provide is often not an easy task. Semantic descriptors, like its emotional properties, either expressed on the movie or felt by the users, can be used to tag some information of the video. And once this information is collected, we can try to use it for a better and meaningful organization of the individual and collective video spaces, to search, and even to provide new forms of visualization and interaction [7,18]. Visualization techniques, emerged from research rooted primarily on visual perception and cognition [4], can actually help to handle the complexity and express the richness in these information spaces. Video visualization can be an intuitive and effective way to convey meaningful information in video [18]. These issues can be synthesized in the following problem statement addressed in this work:

Emotional classification based on physiologic information acquired from users when watching films, improves the relevance of movie search retrieval, contributes to enrich movie recommender systems and enables the design of emotional aware user interfaces for movie visualization by adapting their structure and their visualization elements and tools.

3. PHD OBJECTIVES

In order to address the issues identified in the problem statement presented above, and focusing on the research questions that emerged, four main goals were defined for this thesis, more specifically to:

- Improve movie search mechanisms, making information retrieval more relevant through the use of emotional profiles from users;

- Enrich recommender systems by adding emotional information allowing movie suggestions based upon emotional profiles of users and movie's emotional profiles;

- Access and visualize videos based on emotional characteristics of videos;

- Adapt user interface aspects with emotional awareness features that allow controlling the movie sequence

visualization, the complementary information available and even the way in which the collected information is displayed;

To address these objectives we are going to follow a methodology described in the next section.

4.METHODOLOGY AND PHD CONTRIBUTIONS

The methodology used to develop this work was the following. First, an extensive research literature review was performed so as to understand the role of emotions in the context of affective computing with the precise objective of identifying the importance of emotional approaches in Human Computer Interaction, on Multimedia Information Retrieval and in Video Processing, along with the clarification of associated problems and limitations. Background literature on emotional theories, emotion recognition, biosignal processing, classification techniques, video analysis and low-level feature extraction, emotional design and recommendation systems was also covered with the objective of providing a framework of key concepts and technologies on which to base the design of a system architecture that addresses the emotional classification of users and movies as well as its representation and access. We developed an interactive web video application - iFelt - developed to learn user's emotional patterns using movies' scenes selected to induce emotions. - The iFelt system has 2 main components aimed at: emotional movie Content Classification and emotional movie Access and Exploration. This last component aims to provide video access and visualization based on their emotional properties and users' emotions and profiles. We are designing different methods to access and watch the movies, at the levels of the whole movie collection, and the individual movies. The first prototype is focused towards the access based on the emotions felted by the user, to explore and evaluate emotional paradigm, on top of which we will later add the other perspectives. The design options are thoroughly addressed in [13]. Next we present our main contributions so far based on the problems stated before.

4.1.Emotional Recognition and Classification

Our emotional recognition and classification component is grounded in the induction of emotional states by having users watch movie scenes. The recognition process included two important phases, the training and the testing phases. Inspired by the works of [16], we based our testing phase in an induction of emotion using a set of emotional movie scenes. This component can be divided into two main modules, the Biosignal Recording module and the Pattern Recognition module. Biosignal recording uses biosensors for measuring Galvanic Skin Response (GSR), Respiration (Resp) and Electrocardiogram (ECG) and is responsible for users' biosignals recording and signal processing pipeline. These sensors were specifically chosen as they record the physiological responses of emotion, as controlled by the autonomous nervous system. The Pattern Recognition module uses discriminant analysis, support vector machine and K-nn classifiers to analyze the physiological data and it was validated by the usage of specific movie scenes selected to induce particular emotions. Our objective was to determine whether our classification engine is sufficiently accurate to automatically recognize emotional patterns from new data with a reasonable success rate. Another goal was to determine if the selected scenes had the same emotional impact in all the users in order to measure the importance of the scene for eliciting a specific emotion. Eight participants, averaged 34 year, were submitted to the experiment. In our study we are using the subjects' data obtained while watching movie scenes to create an engine to support user interaction, and to enhance automatic recognition of users' emotional states. We selected a set of movie scenes to induct subjects to feel five basic emotions (happiness, sadness, anger, fear and disgust) and the neutral one. Every subject watched 16 scenes (four of happiness, four of sadness, four of fear, two of disgust and two of anger) and one neutral scene. Based on their feedback, we associated the captured physiological signals with emotional labels, and trained our engine. Eight movies were randomly chosen from the total pool of 30. An average of two subjects watched these eight movies, and were classified by the system. With the SVM classifier, the overall average recognition rate is 69% (s.d. 5.0%), which represents a 49% improvement over random choice, whereas the k-NN classifier produced an overall average recognition rate of 47% (s.d. 9.3%). The SVM classification score shows promise that the iFelt system can be used to automatically evaluate human emotions.

4.2.Emotional Movie Access and Exploration

iFelt is an interactive web video application that allows to catalog, access, explore and visualize emotional information about movies. It is being designed to explore the affective dimensions of movies in terms of their properties and in accordance with users' emotional profiles, choices and states. Although iFelt supports any kind of video, we are focusing our analysis in movies.

The iFelt system has two main goals:

1) Emotional Movie content classification: to provide video classification based on emotions, either expressed in the movies, or felt by the users;

2) Emotional Movie access and exploration: to access and visualize videos based on their emotional properties and users' emotions and profiles.

In iFelt, we created different levels to access and explore movies: the 1) "movies space" where users get a view over the movies existing in the system, with information about their dominant emotions. We designed different representations, including movie lists and "emotional wheels", where the movies are represented by a colored circle, with their dominant emotion color, in ways that represent the level of emotion dominance in each movie.; the 2) "emotional scenes space", where users can obtain a view of the scenes of the movies based on the scenes dominant emotions, and allowing for e.g. to access the individual movies, but presenting only the scenes with the selected emotion, as emotional summaries of the movies; the 3) "individual movie level", where the movie can be watched and, in addition, information about its dominant emotions and emotional scenes can be viewed, for e.g. through an emotional timeline that represents the emotional scenes along the movies; and 4) users have an emotional profile, with emotional information about their movies, that is movies classified from their own perspective or view, and statistical information concerning their history, in terms of the emotional classification of movies they watched. The Emotional Movie Access and Exploration are thoroughly described in [13].

5.FUTURE WORK AND PHD JUSTIFICATION

Our ongoing research intends to support real-time classification of discrete emotional states from biosignals for multimedia content classification and user interaction mechanisms by developing emotional aware applications that react in accordance to user's' emotions. We are considering using emotion

recognition to automatically create emotional scenes, recommend movies based on the emotional state of the user and adjust interfaces according to user's emotions and based on emotional regulation theories. By creating emotional profiles for both movies and users, we are developing new ways of discovery interesting emotional information in unknown or unseen movies, compare reactions to the same movies among other users, compare directors intentions with users effective impact, analyze over time our reactions or directors tendencies.

Regarding visual exploration and access mechanisms of emotional information the next step would be to improve and extend the system in accordance with users feedback, our own evaluation of the current design and implementation, and some of the ideas we originally had and that were not yet included in the current version. Some of the future features include: extending the concept of video summaries to present movies in chosen emotional perspectives and preferences, with more criteria other then selecting scenes with one chosen emotion; summarizing or searching or recommending movies based on users current emotional states, or defined emotional criteria; to find movies by example, i.e. with emotional timelines similar to the timeline of a given movie; exploring the visual representation of huge amounts of movies and extend selecting and browsing methods based on more sophisticated and powerful filters and searches; and to include support for historical emotional information gathered along time, so we can witness the evolution of users' emotional reactions to movies over time, and compare it to other perspectives, including the actors and directors involved, in the several movies genres. We also intend to make all this information more available, or visible, on the web as a shared and recommender environment based on the emotional classification of movies, useful for the general public, as well as for more professional perspectives of directors and actors.

Finally, iFelt is currently focused in movies and the web environment, but this same approach can be useful and interesting to be explored with other types of videos, as is the case of advertisement videos that typically aim at specific emotional reactions from the viewers; and from interactive TV and video on demand services. The core functional and interface features could be the same, but some new requirements in these contexts might involve some adaptations or extensions.

6.REFERENCES

[1] Axelrod, L., Hone, K.S. Affectemes and all affects: a novel approach to coding user emotional expression during interactive experiences. Behaviour & Information Technology,25, 2, March-April, 159-173 (2006)

[2] Being Human: Human-Computer Interaction in the Year 2020. http://research.microsoft.com/hci2020/, 2007.

[3] Brava, S., Nass, C., & Hutchinson, K., 2005. Computers that care: investigating the effects of orientation of emotion exhibited by an embodied computer agent. International Journal of Human-Computer Studies, 62(2), 161-178.

[4] Card, S.K., Mackinlay J.D., and Shneiderman, B. 1999 Readings in Information Visualization: Using Vision to Think, San Francisco, California: Morgan-Kaufmann.

[5] Damasio, A. (1995). Descartes' Error. Harper Perennial;

[6] Gross, J. J. & Levenson, R. W. (1995). Emotion elicitation using films. Cognition & Emotion, 9(1), 87-108. 987-108

[7] Hauptmann, A. G. 2005. Lessons for the Future from a Decade of Informedia Video Analysis Research. Int. Conf. on Image and Video Retrieval, National Univ. of Singapore, Singapure, July 20-22. LNCS, vol 3568, pp.1-10, Aug.

[8] Huppert, F. 2006. Positive emotions and cognition: developmental, neuroscience and health perspectives. In Forgas J.P. (Ed.), Hearts and minds: Affective influences on social cognition and behavior., Psychology Press, New York.

[9] Isen A. M., Daubman K. A., and Nowicki G. P. (1987) Positive affect facilitates creative problem solving. Journal of personality and social psychology, 52:1122–31.

[10] Maaoui, C., Pruski, A., & Abdat, F. 2008. Emotion recognition for human-machine communication. 2008 IEEERSJ International Conference on Intelligent Robots and Systems, 1210-1215.

[11] Metz, C. & Taylor, M. 1991. Film language: A semiotics of the cinema. University of Chicago Press.

[12] Münsterberg, H. (1970) The film: A psychological study: The silent photoplay in 1916. Dover Public.

[13] Oliveira, E., Martins, P., Chambel, T. 2011. iFelt - Accessing Movies Through our Emotions. In Proceedings of EuroITV'2011, 9th. European Conference on Interactive TV and Video, ACM SIGWEB, SIGMM & SIGCHI, Lisbon, Portugal, Jun 29-1Jul, 2011.

[14] Peter, C., & Herbon, A. (2006). Emotion representation and physiology assignments in digital systems. Interacting With Computers: 18 (2), 139-170.

[15] Philippot, P., Baeyens, C., Douilliez, C., & Francart, B. (2004) Cognitive regulation of emotion: Application to clinical disorders. In P. Philippot & R.S. Feldman (Eds.). The regulation of emotion. New York: Laurence Erlbaum Associates.

[16] Picard, R. W., Vyzas, E., & Healey, J. 2001. Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1175-1191.1175-1191

[17] Picard, R.W., 1997. Affective Computing. MIT Press, Cambridge, MA.

[18] Teresa Chambel, Telmo Rocha and João Martinho, "Creative Visualization and Exploration of Video Spaces". In Proceedings of Artech'2010 : Envisioning Digital Spaces, 5th International Conference on Digital Arts, pp.1-10, Guimarães, Portugal, Apr 22-23, 2010.