UNIVERSIDADE DE LISBOA
Faculdade de Ciências
Departamento de Informática

Multimedia Interaction and Access Based on Emotions: Automating Video Elicited Emotions Recognition And Visualization

Eva Ferreira de Oliveira

DOUTORAMENTO EM INFORMÁTICA

ESPECIALIDADE ENGENHARIA INFORMÁTICA

2013

UNIVERSIDADE DE LISBOA
Faculdade de Ciências
Departamento de Informática

Multimedia Interaction and Access Based on Emotions: Automating Video Elicited Emotions Recognition And Visualization

Eva Ferreira de Oliveira

Tese orientada pela Prof. Maria Teresa Caeiro Chambel e co-orientada pelo Prof. Nuno Jorge Gonçalves de Magalhães Ribeiro, especialmente elaborada para a obtenção do grau de doutor no ramo de Informática, especialidade de Engenharia Informática.

2013

This dissertation is lovingly dedicated to my husband, Miguel, my love.
His support, encouragement, and constant love have sustained me throughout this work.

Acknowledgements

To Professor Teresa Chambel for her time, encouragement, expertise, exquisite attention to detail and demand for excellence. Without her, the pursuit of this advanced degree would be enormously more difficult.

To Professor Nuno Magalhães. Without him, his insight, his knowledge on so many topics, the lessons he taught me without giving me the answers, his tolerance, his belief in my abilities (even in the absence of my own), none of this would be possible.

There are people in my live whose existence make my life rewarding. My husband Miguel, my mother, my father, my sister Carla, my brother-in-law Jorge, my stepdaughter Matilde, my nieces Francisca and Margarida. You were my happiness during this time.

To IPCA, for supporting and encouraging me throughout all these years.

In this hard and long process there were people who helped me whenever I needed to. To Mitchel Benovoy and Pedro Martins for their contribution in this work. To the members of the VIRUS Research Project for their collaboration in the project, and to Dr. Pedro Gonçalves all his support at the Lasige lab. My thanks to Prof. Pedro Henriques and Prof. Vitor Santos for every text review. To Konstantinos Chorianopoulos for our discussions.

The last but not the least, my friends. Andrea and Antonio Vieira, Sandra and Bruno Dias, Raquel Costa, Ana Claudia Fernandes, Isabel Dourado and Carlos Pires, thank you for all you friendship and fellowship during this period. You are the best!

This work has been developed in the context of the VIRUS (PTDC/EIA-EIA/101012/2008) research project, the Human Computer Interaction and Multimedia Group (HCIM) – of the Large-Scale Informatics Systems Laboratory (LaSIGE), a research unit of the Informatics Department (DI) of the Faculty of Sciences (FC), University of Lisbon (UL), and the Multimedia Studies and Resources Research Centre (CEREM) of University Fernando Pessoa (UFP). This work was partially supported by FCT through PROTEC 2009 SFRH/BD/49475/2009 and LASIGE Multiannual Funding.

Abstract

Films are an excellent form of art that exploit our affective, perceptual and intellectual abilities. 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. Alongside, physiological measures are making it possible to study additional ways to identify and use emotions in human-machine interactions, multimedia retrieval and information visualization.

The work described in this thesis has two main objectives: to develop an Emotions Recognition and Classification mechanism for video induced emotions; and to enable Emotional Movie Access and Exploration. Regarding the first objective, we explore recognition and classification mechanisms, in order to allow video classification based on emotions, and to identify each user's emotional states providing different access mechanisms. We aim to provide video classification and indexing based on emotions, felt by the users while watching movies. In what concerns the second objective, we focus on emotional movie access and exploration mechanisms to find ways to access and visualize videos based on their emotional properties and users' emotions and profiles. In this context, we designed a set of methods to access and watch the movies, both at the level of the whole movie collection, and at the individual movies level.

The automatic recognition mechanism developed in this work allows for the detection of physiologic patterns, indeed providing valid individual information about users emotion while they were watching a specific movie; in addition, the user interface representations and exploration mechanisms proposed and evaluated in this thesis, show that more perceptive, satisfactory and useful visual representations influenced positively the exploration of emotional information in movies.

Keywords: Movie Affective Classification, Movie Emotional Perspectives, Emotional Impact on Users, Emotional Movie Access and Exploration, Emotional Recognition and Classification, Psychophysiological Measures, Emotions, Emotion-Aware Systems, Affective Computing, Information Visualization.

Resumo

Os filmes constituem um meio de comunicação e uma forma de arte excelentes para explorar as nossas capacidades perceptivas, cognitivas e emocionais. A evolução tecnológica e as tendências para a convergência dos media estão a transformar o vídeo num meio cada vez mais dominante, e o vídeo on-line está a tornar-se uma atividade de entretenimento cada vez maior na Web. Paralelamente, as medidas fisiológicas tornam possível novas formas de identificação das emoções nas áreas de interação pessoa-máquina, na recuperação de informação multimédia e na visualização da informação.

O trabalho descrito nesta tese tem dois objetivos principais: desenvolver um mecanismo de reconhecimento e de classificação de emoções; e desenvolver diferentes formas de acesso e de exploração de filmes baseados em emoções. Em relação ao primeiro objectivo exploram-se mecanismos de reconhecimento e de classificação, a fim de permitir a classificação de vídeo a partir de emoções, identificando os estados emocionais de cada utilizador. Temos como objectivo a classificação e indexação de vídeo baseadas em emoções sentidas pelos utilizadores enquanto visualizam filmes. No que diz respeito ao segundo objectivo, pretende-se explorar o acesso a aspectos emocionais de filmes e mecanismos de exploração para encontrar diferentes formas de visualizar e explorar vídeos com base nas suas propriedades emocionais.

O mecanismo de reconhecimento automático desenvolvido neste trabalho permite a detecção de padrões fisiológicos, fornecendo informações individuais sobre as emoções dos utilizadores enquanto assistem a filmes específicos. Além disso, os elementos da interface e os mecanismos de exploração propostos e avaliados nesta tese, sugerem que as representações visuais influenciam positivamente a exploração de informação emocional em filmes.

Esta tese analisa os desafios que se colocam ao reconhecimento automático das emoções e de que forma o vídeo, mais concretamente os filmes, podem ser explorados no sentido de tornar mais útil a sua pesquisa, acesso ou organização, baseadas em emoções. Tendo como base estes objectivos, esta investigação passou por se concentrar nos dois tópicos centrais: Reconhecimento automático de emoções; e Exploração das

representações visuais do impacto emocional de filmes; complementados por um terceiro: Estrutura semântica de filmes.

(1) Reconhecimento automático de emoções

Procuramos desenvolver uma nova forma de detectar emoções baseada em reconhecimento de padrões, na análise discriminante e classificadores SVM (support vector machine) que são validados usando cenas de filmes selecionadas especialmente para induzir emoções, que vão desde emoções positivas a negativas. Estas incluem alegria, raiva, nojo, tristeza e medo. Esta abordagem ao reconhecimento automático permite anotar automaticamente vídeos e utilizadores de acordo com o impacto emocional dos vídeos. A indução de emoções através de filmes também tem sido amplamente utilizada em estudos de psicologia e em estudos relacionados com a saúde. De facto, estudos experimentais confirmam que as emoções positivas podem ter um efeito benéfico sobre a saúde física.

As emoções expressam-se de várias formas, como expressões faciais ou corporais (vocais, postura corporal), ou sinais neurofisiológicos (respiração, frequência cardíaca, resposta galvânica da pele e pressão arterial). Assim, a comunidade de IPM (Interacção Pessoa-Máquina) tem vindo a utilizar medidas fisiológicas, cerebrais e comportamentais para estudar as possíveis formas de identificar e usar as emoções em interações pessoa-máquina. No entanto, há ainda dificuldades no processo de reconhecimento, nomeadamente em relação à eficácia dos mecanismos usados para induzir emoções. A indução de emoções é o processo através do qual as pessoas são orientadas a sentir uma ou mais emoções específicas. Alguns trabalhos relevantes mostraram que os filmes são uma das formas mais eficazes de indução de emoções. Por exemplo, Gross e Levenson tentaram encontrar o maior número de filmes possível para provocar emoções discretas no sentido de identificar os melhores filmes para cada emoção. Em 1996, um grupo de investigação (Westermann, Spies, Stahl & Hesse) testou onze métodos de indução e concluiu que os filmes são o melhor método para obter emoções positivas e negativas.

(2) Exploração das representações visuais do impacto emocional de filmes

A exploração de filmes através das suas dimensões emocionais pode ser utilizado como forma de entretenimento, educação e até mesmo fins médicos. Considerando-se o efeito das emoções na atenção, na motivação e no comportamento humano, um cenário em que seria benéfico conhecer o impacto emocional de vídeos é em contextos educativos. Aqui o vídeo poderia capturar a atenção dos alunos de diferentes maneiras, seja para os motivar, ajudar a concentrar ou relaxar, facilitando o processo de aprendizagem. Assim, o segundo objetivo desta pesquisa foi o desenvolvimento de novos mecanismos para catalogar, encontrar e aceder a filmes, através do seu impacto emocional, e também encontrar filmes ou cenas que tendencialmente ajudem a induzir um certo sentimento, ou emoção. Esta informação emocional pode estar disponível em ambientes de recuperação de informação, mais especificamente no contexto dos motores de busca. Na verdade, há mais de quatro milhões de filmes na base de dados do IMDB, incluindo séries de TV que poderiam ter esta informação associadas aos seus filmes. Tendo filmes emocionalmente classificados permitiria a procura e a descoberta de filmes por emoções. No entanto, até hoje não existe um sistema que represente esta informação, e a procura numa enorme quantidade de informação pode ser muito difícil, se pensarmos nos filmes que não estão em lugares destacados de qualquer ranking, mas que poderiam ser emocionalmente interessantes para uma pessoa. Representar emoções em filmes e representar enormes quantidades de filmes de forma a serem explorados pelo seu impacto sobre as emoções humanas também requer pesquisa na área de visualização da informação. Técnicas de visualização, fundadas na percepção e cognição visual, foram analisadas neste trabalho para nos ajudar a lidar com a complexidade e a expressar a riqueza em tais espaços de informação.

(3) Estrutura semântica de filmes

O terceiro tópico abordado neste trabalho é motivado pelo facto de o vídeo possuir uma enorme quantidade de informação audiovisual que não está estruturada e, assim, aceder a todos os dados que um vídeo pode conter não ser uma tarefa fácil. Descritores semânticos, com as suas propriedades emocionais, expressas no filme ou sentida pelos utilizadores, podem também constituir informação relevante para a catalogação dos vídeos. Na verdade, a recuperação de informação multimédia (RIM) ainda procura soluções para extrair informações do conteúdo de meios como o vídeo, por se sentir uma enorme necessidade de novos mecanismos para a produção de descrições (meta-dados) quer do conteúdo, quer das estruturas dos objetos a analisar, e assim preencher a lacuna semântica entre utilizadores e máquinas. Em essência, a investigação na RIM pretende, a partir de objetos de media (vídeo, imagens, áudio, texto) e dos utilizadores (preferências, gostos, comentários, descrições, anotações), descobrir formas de apreender informações disponíveis, tornando-os pesquisáveis e acessíveis. Embora os métodos baseados em análise de conteúdo possam melhorar grandemente a anotação e a procura de resultados, é depois essencial anotar a informação com metadados, e a RIM investiu em paradigmas centradas no utilizador, como a capacidade dos utilizadores fazerem pesquisas pela sua própria terminologia ou o desenvolvimento de estudos que exploram os níveis de satisfação do utilizador no processo de recuperação.

Esta tese também contribuiu nesta dimensão, através do desenho e desenvolvimento de uma estrutura baseada em Emotion ML, uma norma do W3C, para a informação adquirida e armazenada no sistema experimental. Desta forma, é adoptada uma abordagem que facilita que o sistema seja aberto a outros sistemas, para acesso e partilha da informação emocional subjacente.

A análise e as contribuições nestes três tópicos ajudaram-nos a validar a hipótese da tese:

A classificação emocional de vídeos, adquirida a partir dos próprios utilizadores, automaticamente através de informação psicofisiológica enquanto vêem filmes, ou manualmente através da etiquetagem do utilizador, tem o potencial de enriquecer a exploração de informação emocional associada a vídeos, quer através do enriquecimento de critérios de pesquisa, quer através de representações visuais mais perceptíveis e úteis que permitam a exploração de vídeos por emoção.

O mecanismo de reconhecimento automático desenvolvido neste trabalho permite a detecção de padrões fisiológicos, fornecendo informações individuais sobre as emoções dos utilizadores enquanto assistem a filmes específicos. Além disso, os elementos da interface e os mecanismos de exploração propostos e avaliados nesta tese, sugerem que representações visuais perceptivas, satisfatórias e úteis, influenciam positivamente a exploração de informação emocional em filmes.

Palavras-chave: Classificação Afectiva de Filmes, Perspectivas Emocionais de Filmes, Impacto Emocional nos Utilizadores, Acesso e Exploração Emocional de Filmes, Reconhecimento e Classificação Emocional, Medidas Psicofisiológicas, Emoções, Sistemas Baseados em Emoções, Computação Afectiva, Visualização de Informação.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

1.1.1 Video Usage as an Entertainment Source

Video growth over the Internet changed the way users search, browse and view digital video content. Watching movies over the Internet is an activity that has been steadily increasing in the past years and is becoming an authentic pastime for people around the world. The possibility of streaming Internet content to a television set, and recent advances in digital video compression techniques and digital video streaming, have turned this recent modality of watching movies into an easy and doable activity. Thus, this rapid growth demands for new techniques to assign semantic value to video, both computationally and humanly understandable, in order to enable video contents to become more searchable and accessible to users according to their characteristics.

Video usage proliferated in the last years due to a set of circumstances such as the creation of websites like Netflix[1] where video plays a central role. The popularity of such websites can be verified by the massive uploads made per day. Recent reports (Madden 2009; Purcell 2010) show that video usage is growing over the Internet as an entertainment source and as an education resource. Specifically, the report from Pew Internet & American Life Project (Purcell 2010) states that seven in ten adult Internet users have used it to watch or download videos, movies or TV shows. The viewership by adult Internet users increased since 2007 to 2010 from 16% to 32%, including educational videos. Results from this study show that the most significant increase lies in comedy and humorous videos that have risen from 31% to 50% of adult Internet users. Moreover, video compression and streaming techniques, larger bandwidths and larger digital archives are transforming video into one of the most dominant media, and online movie visualization is becoming one of the key trends for the next years (Thomas, Mulligan & Wiramihardja, 2010). These facts justify the necessity of having new and more powerful video access techniques that help users in finding, among others, entertainment and educational videos on the Internet.

By combining diverse symbol systems, such as pictures, texts, music and narration,

[1] Available online at: http://www.netflix.com

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. Moreover, video 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; taking such forms as movies, documentaries or short video clips. Psychologists Isen, Daubman and Nowicki (1998) attested this potential, when they experimented the effect of positive affect in their patients, induced by ten-minute comedy films. The study of films as an emotion induction method has some reports dated from 1996, as presented in Westermann, Spies, Stahl and Hesse (1996), who analyze the mental operations of film viewers and discuss how emotions guide the motivation of perception and consequently the control of a person's attention by cinematographic narratives.

1.1.2 The Role of Emotions in User Experience with Media

Emotion has been the subject of a multitude of studies, since it has been shown that emotions are fundamental in cognitive and creative processes. In fact, understanding emotions is crucial to understand motivation, attention or aesthetic phenomena. There is an increasing awareness in the HCI (Human-Computer Interaction) 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 (Maaoui, Pruski & Abdat, 2008). Thus, gathering emotional information from users can contribute to create an emotional context in applications interfaces. Picard and Wexelblat (2002) defend that systems that ignore the emotional component of human life are inevitably inferior and incomplete, and they state that systems that provide a proper and useful social and emotional interaction "are not science fiction but a science fact" (p. 1).

Society's relation with technology is changing in such ways that it is predictable that, in the near future, HCI will be dealing with users and computers that can be anywhere and anytime, and this fact changes interaction perspectives for the future. Human body changes, expressions or emotions constitute factors that will become naturally included in the design of human-computer interactions (Human, 2007). There is a wide spectrum of areas that investigate emotions with different, but complementary, perspectives. For example, in the neurobiological area, Damásio (1995) showed that emotions play a major role in cognitive and decision making processes. On its own side, 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 users' emotions (Dix, Finlay & Abowd, 2004). HCI is also concerned with evaluation and usability, which includes evaluating the extent and accessibility of the system's user interface, assessing 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 that are being developed 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, Axelrod and Hone (2006) suggest that physiological measures, such as galvanic skin response or pupil dilation constitute objective factors, although not easily correlated with 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.

1.2 Research Objectives

1.2.1 Automatically Recognize Emotions from Users

One of the objectives of this research is to develop a novel emotion recognition approach based on pattern recognition techniques, founded 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. These include happiness, anger, disgust, sadness, and fear. This recognition approach allows to automatically annotate videos and users according to the emotional impact of videos. The induction of emotions using movies has also been largely used in psychology studies (Huppert 2006) and in health related studies. In fact, experimental studies confirmed that positive emotions can have a beneficial effect on physical health (Philippot, Baeyens, Douilliez & Francart, 2004).

Emotions are expressed in a variety of ways, such as body expressions (facial, vocals, body posture), or neurophysiological signals (respiration, heart-rate, galvanic skin response and blood pressure). Accordingly, the HCI research community has been using physiological, brain and behavior measures to study possible ways to identify and use emotions in human-machine interactions (Maaoui, Pruski & Abdat, 2008). However, there are still challenges in the recognition processes, regarding the effectiveness of the mechanisms used to induce emotions. The induction of emotions is the process through which people are guided to feel one or more specific emotion. Some relevant works showed that films were one of the most effective ways to induce emotions. For example, Gross and Levenson (1995) tried to find as many films as possible to elicit discrete emotions and to find the best films for each discrete emotion. In 1996, a research group (Westermann, Spies, Stahl & Hesse, 1996) tested eleven induction methods and concluded that films are the best method to elicit both positive and negative emotions. The research work described in this thesis draws on contributions from affective computing, and consists of a thorough analysis of emotional theories and emotion recognition techniques in order to propose an elicitation mechanism and an automatic recognition method for emotions.

1.2.2 Explore Visual Representations of Movies' Emotional Impact

The exploration of movies by their emotional dimensions can be used for entertainment, education and 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 videos with emotional impact is in educational contexts. Here video could capture students' attention in different ways, either to focus or to relax, facilitating the learning process. Thus, the second aim of this research is the development of new mechanisms to catalog, find and access movies based on emotions that could help to assess videos' emotional impact, and also to find movies or scenes that tend to induce a certain feeling, or emotion, in 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 analyze how they change overtime. Afterwards, this emotional information could become available to information retrieval environments, more specifically to search engines contexts. In fact, there are more that one million four thousand movies in IMDB database[2], including TV series. Having movies emotionally classified allows for searching and discovering of movies by emotions. However, until the present there is no system that represents such information, and searching in such a huge amount of information can also be very difficult if we think about those movies that are not on higher places of any ranking but could be fond for an individual person. Representing emotions of movies and representing a huge amount of movies to be explored by their impact on human emotions also requires research in the information visualization area. Visualization techniques, that emerged from research rooted primarily on visual perception and cognition (Card, Mackinlay & Shneiderman, 1999), can actually help to handle the complexity and express the richness in such information spaces as the ones addressed in this work.

1.2.3 Semantic Structure of Movies Based on Standards

The third objective addressed in this research lies in the fact that video 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, such as its emotional properties, either expressed on the movie or felt by the users, can be used to tag some information of the video. In fact, Multimedia Information Retrieval (MIR) research is still trying to find solutions to extract content information from media as a result of a huge need of new mechanisms for the production of descriptions (metadata) of contents and object structures that are computationally and humanly understandable, given that systems should understand the semantics of a query, and bridge the semantic gap between users and machines (Lew, Sebe, Djeraba & Jain, 2006). In essence, MIR research seeks for gathering information both from media objects (video, images, audio, text) and from users (preferences, likings, comments, descriptions, annotations) in order to make a large amount of information available, making it more searchable and accessible (Lew, Sebe, Djeraba & Jain, 2006). Although content analysis based methods can highly improve annotation and search results, MIR has invested in human-centered paradigms, such as the ability

[2] Available online at: http://www.imdb.com/stats visited october 2012

of making queries by users' own terminology and the development of studies that explore user satisfaction levels, namely in the retrieval process (Lew, Sebe, Djeraba & Jain, 2006).

The research objectives just presented can be summarized in the following problem statement addressed in this thesis.

1.3 Thesis Statement

The emotional classification of videos, acquired from users, automatically based on physiological information when watching films and manually through user labeling, has the potential to enrich the ***exploration*** *of emotional information of videos, both by enriching the set of video search criteria and by allowing a more perceptive and useful visual representation for browsing video collections.*

1.4 Thesis Objectives

In order to address the issues identified in the problem statement, three main goals were defined:

- To contribute to the automatic emotional classification of movies based on the recognition of physiological patterns.
- To semantically represent the gathered emotional information and potentiate its expressiveness and its use by any external application.
- To enrich the access to, and search of, emotional information contained in videos by adding new criteria and by improving its visual representation, resulting in a more perceptive and useful way to enable the access, search, browsing, recommendation and visualization of emotional information related to both videos and its users.

1.5 Research Methodology

In order to fulfill these objectives, the following research methodology was applied.

- To understand the role of emotions in the context of Affective Computing, specifically in what concerns videos and movies.
- To understand the role and the use of emotions in movie classification, as well as its problems and limitations referred in literature from the areas of Human-Computer Interaction, Multimedia Information Retrieval and Video Processing.
- To identify suitable emotion recognition techniques based on biosignal processing and classification algorithms.
- To determine technologies on which to base the design of a system architecture that addresses emotional recognition and classification.
- To identify and determine how to explore key concepts of the emotional classification of users and movies.
- To investigate a set of alternatives in order to produce useful, satisfactory and easy to use ways to represent and access emotions in movies.
- To develop an interactive web video application in order to conduct experiments enabling us to evaluate
    - A mechanism for emotional movie content classification by users derived from physiological measures.
    - A set of different information visualization strategies designed to access and explore movies by emotions.

The following section presents the main contributions of the research work developed towards the objectives proposed above.

1.6 Contributions

This research work has already produced the following results and contributions:

- The development of an emotional recognition and classification component grounded in the induction of emotional states by having users watch movie scenes. 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.

- On the video access and exploration part we designed iFelt, an interactive web video application that allows to catalog, access, explore and visualize emotional information about movies.
- We also defined an XML schema for the semantic description of emotions oriented towards the user experience when watching videos, considering the user implicit assessments (by collecting physiological signals), the user explicit assessment, and video content analysis.

All the contributions presented above have been peer-reviewed, accepted for publication and, except for the journal papers, also presented in the following conferences:

- Eva Oliveira, Teresa Chambel and Nuno Ribeiro, “Sharing video emotional information in the Web”. International Journal of Web Portals (IJMP), IGI Publishing (to appear), 2013.
- Eva Oliveira, Pedro Martins and Teresa Chambel, “Accessing Movies Based on Emotional Impact”. Special Issue on Social Recommendation and Delivery Systems for Video and TV Content, ACM/Springer Multimedia Systems Journal (to appear), 2013.
- Teresa Chambel, Eva Oliveira, and Pedro Martins, "Being Happy, Healthy and Whole Watching Movies that Affect our Emotions". In Proceedings of ACII 2011, 4th International Conference on Affective Computing and Intelligent Interaction, Springer: Berlin Heidelberg. pp. 35-45, Memphis, TN, USA, Oct 9-12, 2011.
- Eva Oliveira, Mitchel Benovoy, Nuno Ribeiro and Teresa Chambel, "Towards Emotional Interaction: Using Movies to Automatically Learn Users' Emotional States". In Proceedings of Interact'2011: "13th IFIP TC13 International Conference on Human-Computer Interaction", pp.152-161, Lisbon, Portugal, September 5-9, 2011.
- Eva Oliveira, Pedro Martins, and Teresa Chambel, "iFelt: Accessing Movies Through Our Emotions". In Proceedings of EuroITV'2011: "9th International

Conference on Interactive TV and Video: Ubiquitous TV", in cooperation with ACM SIGWEB, SIGMM & SIGCHI, pp.105-114, Lisbon, Portugal, June 29-July 1, 2011. (Best Paper Award)

- Eva Oliveira, "Video Access and Interaction Based on Emotions". Doctoral Consortium. EuroITV'2011: "9th International Conference on Interactive TV and Video: Ubiquitous TV", in cooperation with ACM SIGWEB, SIGMM & SIGCHI, Lisbon, Portugal, June 29-July 1, 2011. (Nominee for Best PhD Award)
- Thibault Langlois, Teresa Chambel, Eva Oliveira, Paula Carvalho, Gonçalo Marques, and André Falcão, "VIRUS: Video Information Retrieval Using Subtitles". In Proceedings of Academic MindTrek'2010 : Envisioning Future Media Environments, in cooperation with ACM SIGCHI & SIGMM, poster session, pp.197-200, Tampere, Finland, Oct 6th-8th, 2010.
- Eva Oliveira, Teresa Chambel, and Nuno Ribeiro, "A Semantic Representation of Users Emotions when Watching Videos". In Proceedings of XATA '2010, XML Aplicações e Tecnologias Associadas, pp.149-158, ESEIG, Vila do Conde, Portugal, May 19-20, 2010.
- Eva Oliveira, Nuno Ribeiro and Teresa Chambel, "Towards Enhanced Video Access and Recommendation through Emotions". In "Brain, Body and Bytes: Psychophysiological User Interaction" Workshop, at ACM CHI'2010, 6 pgs, Atlanta, GA, USA, Apr 10-15, 2010.
- Eva Oliveira and Teresa Chambel, "Emotional Access and Interaction with Videos". In Proceedings of Academic MindTrek'2009 : Everyday Life in the Ubiquitous Era, in cooperation with ACM SIGCHI & SIGMM, poster session, pp.218-219, Tampere, Finland, Sep 30th-Oct 2nd, 2009.
- Eva Oliveira, and Teresa Chambel, "Emotional Video Album: getting emotions into the picture", emotion-in-hci'2008, The 4th Workshop on Emotion in Human-Computer Interaction: Designing for People, at HCI'2008, the 22nd BCS HCI Group conference on HCI: Culture, Creativity, Interaction, Liverpool, UK, September 1-5, 2008.

1.7 Thesis Structure

The remaining of this dissertation is organized as follows. Chapter 2 presents a set of background concepts on the emotion theories and recognition methods in general and then focused specifically on video. Chapter 3 details the state-of-the art of the visual exploration of videos, as well as an overview of the semantic structure of emotions on videos. Afterwards, Chapter 4 details experimental studies on eliciting and recognizing emotions, conducted with the purpose of assessing the recognition rate of the approach suggested in this thesis. Chapter 5 presents another set of experimental studies on the visual exploration of emotional information, carried out with the goal of identifying new ways for exploring such information spaces. Finally, Chapter 6 revisits the hypothesis raised earlier in this chapter, offering a set of conclusions in the light of the results obtained with this work, and discusses future directions for this research.

Chapter 2

Emotion Recognition in Video

2.1 Introduction

This chapter reviews the state of the art of the main topics involved in this work, specifically the importance of emotions in HCI, as well as theories of emotions underlying affective systems and especially those that involve digital video. Capture and measurement of emotions will be reviewed considering current emotion recognition techniques and processes as well as emotion induction methods.

2.2 HCI and Emotions

In the last decade, emotion research has been steadily growing and became, nowadays, a firmly established field within HCI, as researchers became more and more aware of the importance of emotions in the human-computer interactions. For example, a number of researchers have pointed out the need to take into account the emotional characteristics of users in the context of user interfaces usability (Reeves & Nass, 1996). Moreover, product designers have also been exploring emotional considerations in their contexts, as is evident in the body of work developed by Norman (2004), Desmet (2002) and Picard (1999).

The reason for this growing interest in emotion research lies precisely in the importance of the emotions for human beings. According to Axelrod and Hone (2006), emotions are currently regarded as important for human development, in social relationships and in the context of thinking processes such as reasoning, problem solving, motivation, consciousness, memory, learning, and creativity. Notably, it has been pointed out that emotions enrich persons' living experiences with either a pleasant or unpleasant quality (Desmet, 2003), being an important aspect that guides behavior and confers meaning to everyday existence (Winkielman & Cacioppo, 2001). Thus, the relationship between people and the world they live in is clearly emotional. Nowadays, with technological development, the world in which people live also implies computers and their applications. When considering this particular relation it becomes evident that we must support user experiences that are engaging and enjoyable.

Because movies are loaded with a great amount of emotional information, they can be explored to improve the development of emotion aware systems. One example of this improvement is that movies can be useful to boost people's emotions and, having the necessary means, it is possible to measure this information, which turns out to be valuable information to be treated and applied for emotion aware applications.

Sometimes we search for movies that make us feel happy or fearful or romantic. We can say that we search for emotional states or particular moods. With emotional information associated to movies, we could then search not only for particular scenes that made us sad, happy or afraid, but also for movies that can have a great probability of making us feel in the way we desire, or even make us better understand our own emotions resulting from events that occur in our own lives. As emotional states can now be assessed, movie access and exploration can be paradigmatically changed, and movies can now be searched by emotions, in particular the specific emotions they induce in viewers. By knowing a user's emotional state, emotion aware systems and applications can adapt their functionality in accordance with the user's moods. For example, in a web movie social network, recommendation and other users' emotional profiles can be used to search and explore movies by their affective dimensions.

Clearly, emotions are always present in the mind of users because they are a natural part of the way people react to their surroundings. And as importantly, frustration, confusion, anger, anxiety and other similarly negative emotional states may adversely affect the way users explore entertainment applications, much in the same way as it negatively affects productivity, learning, social relationships and, ultimately, people's well-being (Chambel, Oliveira & Martins, 2011). These are some of the reasons why HCI is concerned in integrating emotions from various perspectives in their studies on interaction (Sharp, Rogers & Preece, 2007). Following the emotional design approach advocated by Norman (2004), a number of researchers and practitioners started to consider the impact of emotions on users' experiences when interacting with computers. Also, the way we explore and visualize information, namely in terms of colors, shapes, information presentation and representation also influence our affect (Sharp, Rogers & Preece, 2007). These conclusions have highlighted the importance of such considerations given that emotional reactions caused by computer-based systems may indeed influence the acceptance and usage of such systems (Hassenzahl, 2005). Nowadays, we use computers in everyday activities, not only for work, but also for leisure, learning, entertainment and communication. Thus, usability researchers figured out that users' emotions are of critical importance in various domains, from online entertainment, to e-learning or online shopping, to name just a few.

But emotions are not straightforward to understand and are very complex to model because an emotional state may arise as the result of a combination of a set of emotional responses to a succession of events. For example, when watching a movie, a viewer may experience different kinds of emotions, such as happiness, amusement, anger, fear and disappointment. Thus, instead of just a single emotion, it is the combination of these emotions that determines the emotional timeline of each person, defining the emotional classification attributed by that particular person to the experience of watching a particular movie. However, as Blythe et al. (2003) point out, the problem is that very little is still known about how people respond emotionally to computers and about the specific aspects of design or interaction that trigger emotional responses.

As stated in this section, human emotions are increasingly understood to be fundamental in HCI and in movie access and exploration. In particular, being emotions studied within different areas, from psychology to neurophysiology, in the following section we review a number of emotional theories in order to understand how emotions are defined and how they can be assessed and processed in the scope of computer science research.

2.3 Emotional Theories

We organize our lives to maximize the experience of positive emotions and minimize the experience of negative emotions.

Paul Ekman (2004)

After a long academic debate on the primacy of cognition (Lazarus, 1984), versus the primacy of emotion (Zajonc, 1980, 1984), there is nowadays the perspective that cognitive and emotional processes are essentially linked. For example, Frijda (1987) defends that an emotion experience is the result of an appraisal process about the objects or events that triggered the emotion, and the intensity of the experience will influence memory, in the way that we easily remember emotional intense moments. Many authors (e.g. Lang, Bradley & Cuthbert, 1998; LeDoux, 1996; Whybrow, 1999) state that emotions are survival artifacts, as they regulate biologic patterns to deal with specific environments. For example, anger or fear can be triggered by internal changes like low glucose in blood or by the perception of danger (mental changes). Therefore, our preferences and criteria are signed in our body and are also a product of the experiences of our lives.

These internal mechanisms are then responsible for the classification of things as being good or bad. In essence, we can say that emotions classify everything as being positive, neutral or negative. Damásio also corroborates this idea – the somatic markers hypothesis – which attributes to emotions the role of marking every cognitive activities as good or bad, positive or negative. In fact, even William James (1950/1890) defended that emotions give color and heat to life, giving the sense of pleasantness or unpleasantness to every event or interaction with the world, and labeling it “to avoid” or “to search”. Moreover, without emotions life would be dull and empty (Oatley, Keltner & Jenkins, 2006). Positive emotions enhance cognitive capacities and creative problem solving tasks, as well as physical and mental health; while negative emotions, which narrow the individual’s repertoire of thought and action, have a valuable survival strategy.

From philosophy, Plato, and Aristotle wrote on the topic of emotions. More recently, Descartes established the current basis of neurophysiology and presented six fundamental emotions, correlating them with neurophysiologic basis in his book “Les passions de l’âme” (Descartes, 1989/1649). The comprehension of emotions has been explored scientifically since the 19th century, being the work from Charles Darwin (1998 /1872) “Expression of the Emotions in Man and Animals” one of the most notorious book on emotions. William James (1950/1890) dedicated his life into dissecting how emotions affect every aspect of a human being leading to the famous James-Lange theory of emotions. But still lively issues, such as what is an emotion, or how emotions are triggered and even how we can regulate emotions, continuously inspire research. Curiously, when trying to define emotion, we encountered a multitude of definitions and models in the literature. As stated by Fehr and Russell (1984) "everyone knows what an emotion is, until asked to give a definition." (p. 464). Despite this fact, we have some main theories defining emotions and how they can be understood. Throughout our review, we examine findings in two different perspectives:

the categorical and the dimensional one.

The categorical perspective, also known as the Darwinist or discrete model of emotions (e.g. Ekman, 1992), defines emotions as discrete states that identify a certain behavior and experience. The dimensional perspective is based on a small number of scales or dimensions to define an emotion, where the most common are valence (positive/negative), arousal (low/high) and potency (strong/weak) (Fontaine, Scherer, Roesch & Ellsworth, 2007). In the following sections, we provide an introduction for each perspective, focusing in the different ways of identifying and categorizing emotions.

2.3.1 Categorical Models

The categorical or discrete model of emotions defines emotions as discrete states. These states are well-defined units that identify a certain behavior and experience. One of the major research programs on emotional expressions was developed by Paul Ekman (1992) (Figure 1). His multicultural research project tested thousands of people from all over the globe. It turns out as commonly accepted that six basic emotions (anger, fear, surprise, sadness, disgust, happiness) represent innate and universal emotions recognized in face expressions across cultures. However, surprise has been withdrawn from the basic emotions since his later article (Ekman, 1999).

Figure 1 - Ekman six basic facial expressions.
(From up to down left to right) anger, fear, disgust, surprise, happiness and sadness

Also, Damásio (1995) agrees in the same five of the six Ekman's basic emotions:

Happiness, Sadness, Fear, Anger, Disgust, which appear to be a ground truth for discrete emotional models. However, Damásio distinguishes primary and secondary emotions: primary emotions are innate, pre-programmed emotions, whilst secondary emotions are a blend of primary emotions that result from the reflection about an event or experience. For example, contempt is regarded as anger plus disgust.

Besides Ekman experiment with facial expressions, there are several scientific experiments that measured these categorical emotional states through physiological changes. The first studies that related emotions to body changes were made by two persons separately, William James and Carl Lange (1922), which leads to the famous James – Lange theory. Both experiments suggest that emotion is the corporal sensation of our perceptions. This theory has however been refuted because of its sequential process which claims that the emotion occurs after the body reacts to an external stimulus (such as a bear), while modern theories have proved that mental and body reactions occur at the same time. For example, Damásio (1995) defends that an emotion is the combination of a mental evaluation process and body dispositional responses to that process, which produce physical changes in our body and mind

To sum up, almost every categorical theorist considers the existence of a limited number of basic emotions. Secondary emotions, such as guilt, pride, jealousy and embarrassment, are considered junctions (Damásio, 1995), or a blend of the primary ones and are the result of a cognitive process.

2.3.2 Dimensional Model

One of the most notorious dimensional theorists is Russel (1980), who proposed a circumplex model of emotions, in which one central dimension defines the polarity (positive/negative or pleasant/unpleasant) of the emotion and the other defines the level of activation. Figure 2 shows the variety of human emotions defined by the x-axis of valence and the y-axis of arousal arranged in a circle around a two-dimensional space. Russel bases his theory on a layman's conceptualization of affect and on multivariate analysis of self-reported affective states, by showing that there are two predominant dimensions that participants use to communicate and label their own emotional experience.

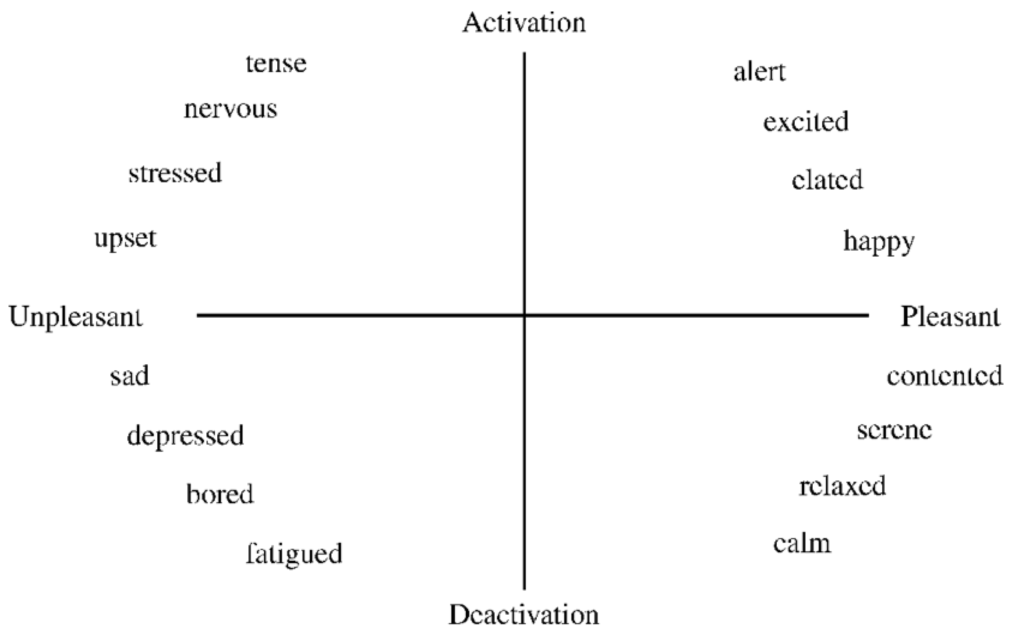

Figure 2 - Russell's circumplex model of affect

Russel (1980) has also argued that there are nonverbal emotions, which are interrelated other than independent. Like the color spectrum, emotions have no well-defined borders, and they are not completely independent (Fehr & Russell, 1984).

In the following section we discuss other theories of emotions, which are not clearly categorical or dimensional, having influences from both models, namely the appraisal model as proposed by Scherer (2000) and the model of Plutchik (1980).

2.3.3 Appraisal Model

The appraisal theory of emotions (Figure 3) is essentially a categorical emotional model that suggests that emotion elicitation and differentiation depend on people's own evaluation of events and its circumstances. In fact, it was Magda Arnold 1960 who first called "appraisal" for categorizing emotions based on people's own judgment and according to their experiences (Scherer, 1999). Klaus Scherer (see also Lazarus, 1984) defends that emotions are triggered by people's own interpretation of events, i.e. by their subjective interpretations and considers that appraisals emerges from the following sequence: novelty, pleasantness, own goals, causes and compatibility with standards.

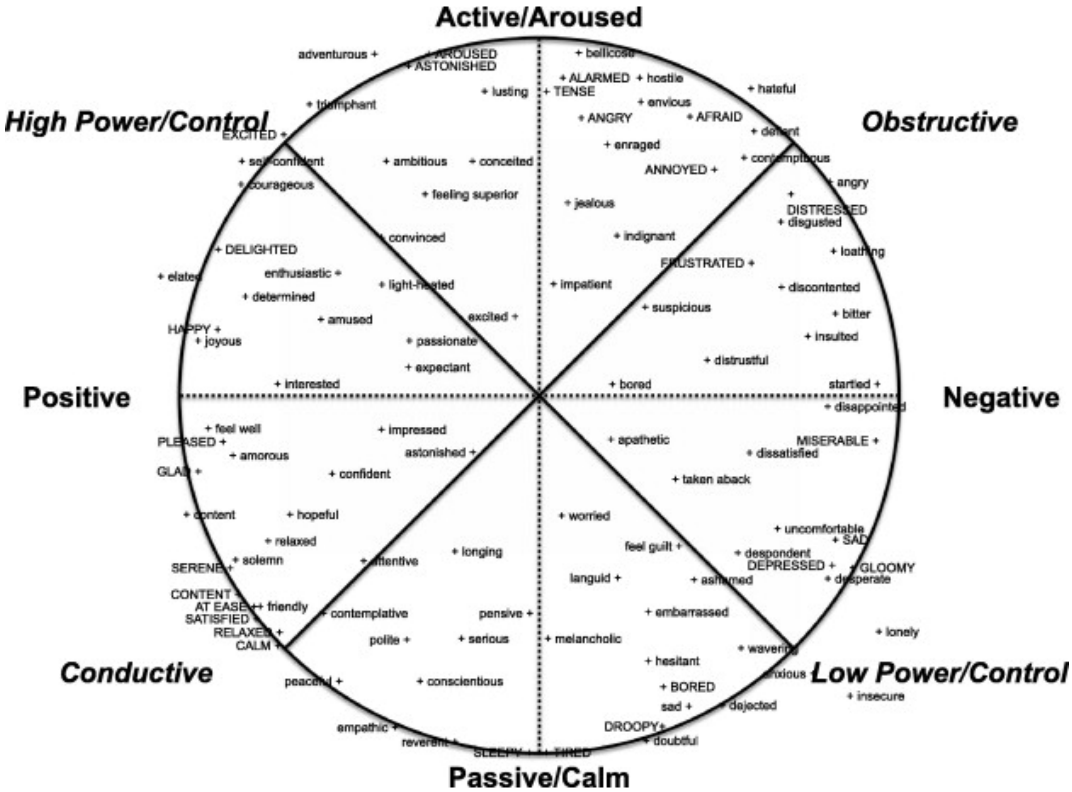

Figure 3 - The Geneva emotional wheel

The causes of the event are next evaluated and, finally, there is an analysis if the event is standard and if any social rule has been broken. In Scherer's theory, it is the answer to all these questions that defines every emotion. In his later work (Scherer, 2005) he proposes a 2D instrument of sixteen emotion families labeled by natural language categories to measure forced choice self-reports of emotional experience – The Geneva Emotional Wheel (Figure 3). With his instrument, appraisal categories are related with some dimensions such as goal conduciveness translated into positive/negative and coping potential, identified in the figure by strong/weak control distributed emotions all over the wheel, where intensity is represented by the distance to the center of the wheel.

2.3.4 Other Models

Another theorist – Plutchik (1980) (Figure 4) - mixes both categorical and dimensional models by proposing a multidimensional model of emotions. The author considers two levels of emotions, the primary ones, which are biological and primitive, and every other emotion as variants of those primary with different similarities, intensities and polarities. This perspective leads to a three dimensional model (polarity, similarity, intensity) with eight primary emotions - anger, fear, sadness, disgust, surprise, anticipation, trust, and joy.

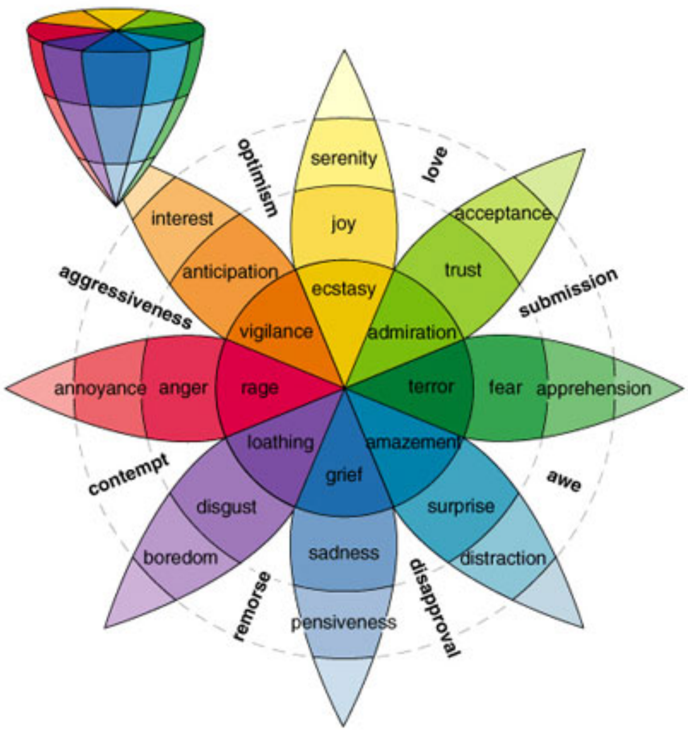

Figure 4 - Plutchik three-dimensional circumplex emotional model

For example, in his model, he describes rage as a stronger variant of anger, which in turn is a stronger variant of annoyance. Another example is ecstasy, joy or serenity, which form the same family of emotions with different intensity degrees. Regarding the polarity, Plutchik established a set of basic opposites for the basic emotions, for example, joy as opposite to sadness, trust to disgust, or anticipation to surprise, among others.

In Plutchik's perspective, an emotion is a complex reaction that includes cognitive evaluation. One of the Plutchik most known diagrams (Figure 4) is found in his book "Emotion - A Psychoevolutionary Synthesis" (Plutchik 1980). Figure 4 presents a radial diagram analogous to a color circle to represent different emotions with different degrees of intensity, with the eight more intense in the center (primary). Opposite emotions are represented in opposite sides of the wheel. The circle assumes the degree of similarities among emotions and the vertical dimension represents the intensity, the eight colors represent the eight primary emotions suggested by Plutchik.

2.4 Emotion Expressions

In this section, we present a review on the state of the art techniques for emotion detection and recognition, which is how emotions can be assessed. Emotions can be expressed in a variety of ways such as expressions (facial, vocal, body posture), or neurophysiological signs (respiration, heart-rate, galvanic skin response, blood pressure).

2.4.1 Facial Expressions and Speech

The study of emotions recognition through facial expressions was one of the first studies of emotion recognition, when Charles Darwin published his notorious work – *The expression of the emotions in man and animals* (Darwin, 1998 /1872) – from which emerged a taxonomy of emotions from expressions. Paul Ekman conducted in 1992 a study inspired in Darwin's proposal, in which a Facial Action Coding (FAC) system (see Figure 5) was developed to describe facial expressions by their muscles and their movements. This system allows the development of several other studies on emotion recognition (e.g. Pantic & Rothkrantz, 1999).

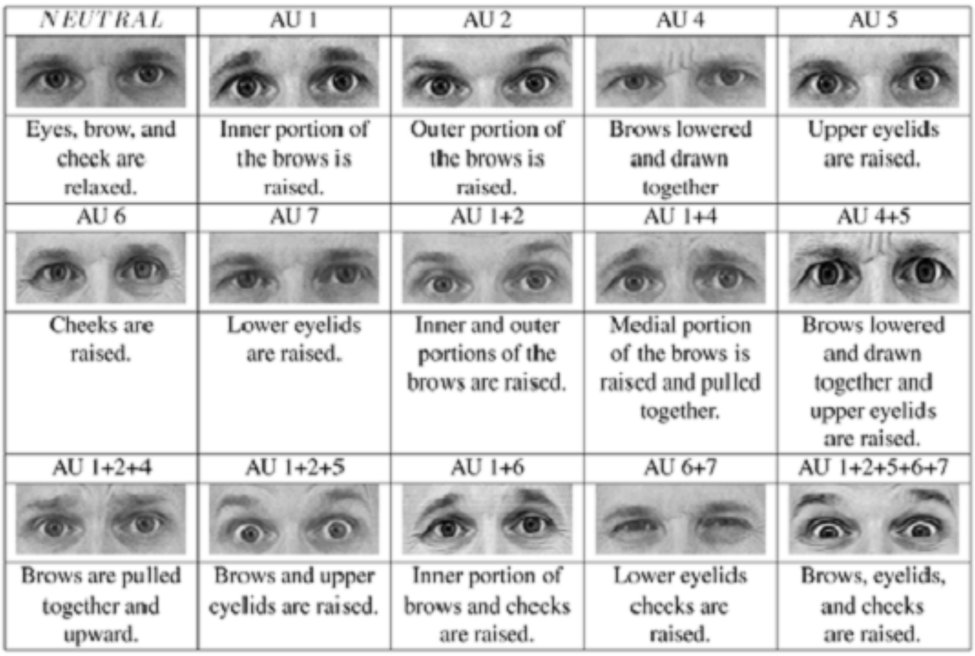

Figure 5 - Facial action coding

Other studies focus on the recognition of emotion through speech analysis. Speech can express emotions by its pitch, duration and the intensity of utterance (Murray & Arnott, 1993) and also by the specific words that are spoken (Douglas-Cowie, Campbell, Cowie & Roach, 2003).

2.4.2 Physiological Responses

Since William James (1984), it is known that corporal changes are correlated with emotions, which turned possible the modulation of emotions through the evaluation of physiological patterns (Rainville, Bechara, Naqvi & Damásio, 2006) which, in turn, unveiled the possibility of the automatic recognition of emotions from such patterns.

The Autonomic Nervous System (ANS) is part of the peripheral nervous system that has the function of conducting and regulating corporal processes such as sensory impulses from the blood, heart, respiration, salivation, perspiration or digestion. This system is divided into two parts: the parasympathetic nervous system (PSNS) responsible to calm our system down, and the sympathetic nervous system (SNS) that has the function to prepare our body to stress conditions. For example, meditation is characterized by the activation of the PSNS, while in running situations the SNS is activated using body energy (blood pressure increases, heart beats faster, digestion slows down) to react accordingly.

Some authors claim that there are distinctive patterns of ANS for anger, fear, disgust, sadness or happiness (Ekman, Levenson & Friesen, 1983b). These evidences led to many advances in the study of emotion, and the most common ways to measuring it include the following: self-reports, autonomic measures (physiological measures), startle responses magnitudes (verbal reaction of sudden stimuli), brain states and behavior (vocal, facial, body postures)

2.5 Recognizing Emotions in Humans by Physiological Sensors

Each emotion expression provokes a reaction in the ANS, one of the human body subsystems responsible for the regulation of physiological processes. We now examine how we may evaluate such reactions.

2.5.1 Physiological Sensors

Biosensors or physiological sensors are used to gather these body changes. In fact, it was precisely through the analysis of the variations of the ANS that Damásio (1995) tested the somatic markers hypothesis which states that, before we rationalize any decision, our body reduces our options by corporal and somatic changes based in our past experiences. Thus, from this study, we conclude that emotions can be characterized by somatic patterns and that every somatic pattern could be evaluated by the analysis of ANS components variation (Mauss, 2009).

From Ekman's work, it was also evident that an emotion-specific autonomic pattern could be distinguishable besides its valence and arousal, because emotions of the same valence appear very distinguishable when analyzing hear-rate, skin temperature, skin resistance and forearm flexor muscle. Table 1 shows an overview of some conclusions from research regarding the most common physiological signals and their correlation with emotions in a dimensional perspective. As noted in Table 1 there are different types of sensors, measuring different activities, such as heart-rate, skin conductance, temperature, blood pressure, respiration, just to mention a few. From Table 1, we can also attest that there are some physiological features that characterize the valence and arousal of emotions.

Several other studies show a direct correlation between emotion categories and physiological pattern as noted on Table 2, which represents physiological indicators of primary emotions. For example heart rate has been used to differentiate positive from negative emotions, suggesting that several emotions can be predicted and defined by the analysis of patterns of cardio (heart-rate) and respiratory activity.

Besides these evidences, the characterization of emotions by physiological patterns faces some problems, despite its advantages when compared to other recognition methods. The main limitation is the difficulty in discriminating emotions from physiological signals. Other problem lies in finding the adequate elicitation technique to target a specific emotion (Rainville, Bechara, Naqvi & Damásio, 2006).

Emotions are time, space, context and individual based, so trying to find a general pattern for emotions can be difficult. Moreover, it is also difficult to obtain a "ground truth". These are some of the major problems, when compared to facial and vocal recognition.

Table 1 - Physiological sensors description and dimensional emotional indicators

Sensor	Description	Emotion Indicators
Skin Conductance Level (SCL)	Measures the electrical conductivity of the skin, given by the sweat glands distributed on the skin. It is also called galvanic skin response or skin conductance level.	Good indicator of arousal, the higher the value the more intense is the emotion (Gomez & Danuser, 2004; Gomez, Stahel & Danuser, 2004; Greenwald, Cook & Lang, 1989; Schupp et al., 2000; Steinbeis, Koelsch & Sloboda, 2006).
Respiration (Resp)	Measures the respiration rate and depth through a large velcro belt around the chest and an elastic which stretches when the chest expands.	It gives indications about arousal and valence (Kreibig, Wilhelm, Roth & Gross, 2007).
Electrocardiogram (ECG)	Measures heart activity by recording the electric waves generated during heart activity. It gives information about heart rate (HR), inter-beat intervals (IBI) and heart rate variability (HRV)	It differentiates positive from negative states (Kreibig, Wilhelm, Roth & Gross, 2007).
Blood Volume Pulse (BVP)	(Or Plethysmograph) Measures blood pressure, heart rate.	It can be an indicator of valence as in negative states, the BVP is lower while in positive it is increased (Money & Agius, 2009).
Skin Temperature (SKT)	Measures skin temperature.	Variations in the temperature reflect autonomic nervous system activity and are an indicator of emotional status (Kreibig, Wilhelm, Roth & Gross, 2007; Rimm-Kaufman & Kagan, 1996).
Electromyography (EMG)	Register the electric potential generated by muscle cells. EMG Frontalis and the Zygomaticus measures face muscles activities.	Helps to recognize the movements of facial muscles states of happy (smiling, laughing) (Duchenne & Cuthbertson, 1990), anger or fear (Kim & André, 2008; Rainville, Bechara, Naqvi & Damasio, 2006b).
Posture Sensors	Foam-based pressure sensor which measures the Compression of movements.	Can indicate bored, agitated, stressed or surprised states. Surprised states (Winters, 2005).

The major advantage is that the characterization of emotions from physiological patterns cannot be intentional, as people cannot trigger the autonomic nervous system (ANS) contrarily to the so called "poker face" where people disguise expressions as well as vocal utterances.

Another common argument against physiological measurements for emotion analysis is the fact that sensors might be invasive. However, nowadays, sensors are less intrusive (e.g. wearable, made of rubber) and they allow protecting peoples' identity or appearance, which could be an advantage when compared with techniques like, for example, facial recognition.

Table 2 - Physiological characteristics of five categorical emotions

Categorical Emotion	Physiological Indicators
Fear	High heart rate levels coupled with robust decreases in heart-rate variability (HRV) (Friedman & Thayer, 1998; Rao & Vikram, 2001). High decrease in respiratory period and HRV within respiratory cycles (Rainville, Bechara, Naqvi & Damásio, 2006). Heart-rate acceleration and larger skin conductance in fear and disgust than in happiness (Levenson, Ekman & Friesen, 1990).
Anger	Increase in heart rate without noticeable changes in high frequency HRV (Rainville, Bechara, Naqvi & Damásio, 2006).
Happiness	Less variability in respiration (period and amplitude) (Rainville, Bechara, Naqvi & Damásio, 2006a). Larger skin conductance in fear and disgust than in happiness (Levenson, Ekman & Friesen, 1990).
Sadness	Less HRV than anger, a longer respiratory period, larger respiratory sinus arrhythmia, and a smaller RR decrease than fear and more respiratory variability than happiness (Rainville, Bechara, Naqvi & Damásio, 2006). Heart rate acceleration (Levenson, Ekman & Friesen, 1990).
Disgust	Larger skin conductance in disgust than in happiness (Levenson, Ekman & Friesen, 1990).

Physiological data allows for the development of systems that facilitate the understanding of cognitive, emotional and motivational processes, by giving access to emotional information that can inform about the type of engagement of a user when interacting with a computer, for example, when performing a task or watching a video.

Although emotions can be characterized by some physiological variations, machine-learning techniques may be applied in automatic emotion recognition methods, based on pattern recognition, to deal with huge amount of data needed to be analyzed. The following section discusses how recent studies are processing physiological data to recognize emotions.

2.5.2 Physiological Signal Classification Techniques

The collection of physiological data produced while users are watching movies was recently developed in psychology to test whether films can be efficient emotional inductors, which could help psychologists in specific treatments (Kreibig, Wilhelm, Roth & Gross, 2007), or in the computer science area to automatically summarize videos according to the emotional impact on viewers (Soleymani, Chanel, Kierkels & Pun, 2009).

In light of the evidence of distinct physiological responses of emotion, the machine learning and HCI communities have each investigated the automatic recognition of emotions. Picard, Vyzas and Healey (2001) pioneered this area by showing that some emotional states can be recognized automatically using physiological signals and pattern recognition methods. In their report, Zwaag and Broek (2010) made an overview, summarized in Table 3, about the most relevant studies regarding emotion recognition, using different physiological signals and different classification methods.

Table 3 - Summarization of the most important studies on emotion recognition

Ref	Signals	Participants	Features	Select./Red.	Classifiers	Target	Results	Elicitation
1	C,E,R,M	1	40	SFS, Fisher	LDA	8 emotions	81%	images
2	C,E,R,M	3	110	SBS	LDA	4 emotions	70%	music
3	C,E,R,M	40	5	-	SVM	5 emotions	47%	movies
4	C,E,R,FE	10	18	-	LDA, SVM, RVM	3 emotions	51%	movies

Ref: 1 - Picard, Vyzas & Healey (2001); 2 - Kim & André (2008); 3 - Lichtenstein, Oehme, Kupschick & Jürgensohn (2008); 4 - Soleymani, et al. (2009). Signals: C: cardiovascular activity; E: electrodermal activity; R: respiration; M: electromyogram and; S: skin temperature; FE: facial expressions. Selection: SFS: Sequential Forward Selection; SBS: Sequential Backward Selection; Fisher: Fisher projection; Classifiers: SVM: Support Vector Machine; RVM: Relevance Vector Machines; LDA: Linear Discriminant Analysis;

We selected the works more related to our own work and also inserted in Table 4 an elicitation column. Curiously, in these automatic emotion recognition tests the methods that have used movies had the worst results, which lead us to speculate that there is a demand for new ones that can more effectively explore the power of movies to induce emotions. This is even more relevant given that, as suggested by Westermann, Spies, Stahl and Hesse (1996), movie scenes are considered to be one of the best methods to induce emotions. Picard et al. (2001) experiment involved five physiological measures (heart rate, galvanic skin response, muscle activity, temperature and respiration). These measures were processed through the Sequential Floating Feature Selection (SFFS) algorithm used to choose the best features of each physiological measure and Linear Discriminant Analysis (LDA) to build a statistical model of each emotional class. In Picard's work these tools were used to build a learning model capable of recognizing eight different emotions with 81% accuracy, in a single participant who used personal imagery and event recollection to induce the emotions. In their work, four physiological sensors were used: blood volume pulse (BVP), galvanic skin response (GSR), electromyography, and respiration.

2.6 Recognizing Emotions in Video Content

2.6.1 Automatic Extraction of Emotions from Video

Recent research (Kang, 2003) exploits new forms of video affective segmentation based on low-level video analysis using pattern recognition algorithms. Video classification based on affective content has also been studied in their results into a 2D dimensional emotional space (valence and arousal). They focused only in fear, sadness and joy due to the difficulty of discriminating fear from anger by using just color or motion and shot cut rate.

In another work, Hanjalic and Xu (2005) developed a framework to represent and model the affective content in videos. They based their content analysis in cinematographic techniques, such as motion analysis, vocal effects, shot length, sound and rhythm analysis, being perfectly aware that the emotional result stems from the directors' intention about the expected emotion, and not necessarily corresponding to the viewer's actual feelings. One of the relevant results of this work is the creation of links between the 2D dimensional emotion space and the low-level features of the video.

More recently, Soleymani et al. (2008) proposed an emotional classification of movies, which they have called complementary approach of emotional models, to analyze affective content. The authors used a cinematographic perspective to analyze low-level features of the film from both audio and visual perspectives and map them into basic emotions categories. They then used the dimensional perspective to correlate each dimensional output with other emotional descriptors to give meaning to the selected category.

2.6.2 Video Emotional Classification by Content Analysis in the VIRUS Project

The emotional recognition and classification of video content has been addressed in VIRUS Project - Video Information Retrieval Using Subtitles (Langlois et al., 2010). In this project, a set of algorithms has been built that, given video, audio and subtitles streams of information, will allow a user to extract movie scenes (from a video library) that feature some specific characteristic (e.g. love scenes, violent scenes, funny situations, scenes with great tension). Video material classification is based mostly on its audio and subtitles contents, where most of the semantics is expressed and relies on more traditional video processing techniques. An example of these are shot detection, motion and color histograms to help the indexing of the most relevant video units, to enrich the classifications based on visual properties (e.g. to detect night or day outdoors scenes).

Regarding the analysis of the subtitles, two main approaches were followed. First, an approach based on statistics assigns semantic keywords to terms. The statistical analysis of keyword frequencies, n-grams and the use of ontologies allows to identify the main topics of the scene. Complementarily, a linguistic-based approach helps the detection of emotions in text. In particular, a set of fine-grained emotional resources, such as lexicons, grammars and annotated corpora, is being semi-automatically created for recognizing and classifying the sentiment predicates related to the topics of the scenes previously identified (Godbole, Srinivasaiah & Skiena, 2007). Simultaneously, the audio stream is analyzed in order to search for specific events: yelling, screaming, gun shots, car noises, and so on, in order to give additional hints about the nature of the scene. The information taken from the audio stream will help disambiguate some situations that are particularly hard to deal with when using a text-only source. For example, if a gunshot is detected, then if the "murder" topic is detected, we will have a good indication that the scene is actually a murder scene and not some people talking about a murder that happened years ago. It is believed that the dual approach to semantic analysis (statistical/linguistic) and the simultaneous analysis of the audio stream is an innovative approach to video indexing that maximizes the quality of search results. We also think that the fact that the semantic analysis relies on several sources of information reduces the risks, because a potential weakness of the analysis of one source will be compensated by the results obtained through the other sources.

It is also focused on the automatic classification of the subjective dimension of emotion, which is the impact on viewers, as explained in the next subsection, that can be combined to set integrated views of expressed and felt emotions (e.g. when there is a gunshot the user feels fear).

2.6.3 Video Classification Based on Users Emotions

The automatic recognition of users' emotions elicited by video content is another perspective of the analysis of videos concerning affective aspects. For instance, Soleymani et al. (2009) use the automatic classification of movie scenes to improve video indexing and retrieval from the analysis of physiological features such as Galvanic Skin Response (GSR), Blood Volume Pressure (BVP), Electro cardiogram (ECG), Respiration, Skin Temperature and Electromyograms (EMG). A central result of this study was that, from the correlation between physiological measures and participants' self-reports, the authors were able to identify the central physiological features for affective ranking of movies. Money and Agius (2009) are also clear that there is a strong relationship between some movie genres and specific biometric artifacts, being horror/thriller films and comedies those which have most impact on users. In this study, horror movies are associated with increases in respiration rates and decreases in respiration amplitudes. Several other genres had a clear relationship with electro dermal response (EDR). Thus, horror movies had higher EDR, while comedy and drama had a reduced EDR. Moreover these movies had parallel EDR patterns.

In another experiment, Smeaton and Rothwell (2009) measured physiological signals of viewers in order to classify the emotional impact of films from the viewers' perspective. With the objective of testing if video emotional highlighting can be detected from physiological signals, they tested whether the film emotional experience is different in a group context or in an individual context. Using the categorical model of emotions, they manually classified the supposed evoked emotions of the prepared films with 21 emotional categories. They also analyzed low-level audio features of each film in order to distinguish speech from music and from silent frames, so as to improve the accuracy of emotional classifications. The output classification was obtained by comparing the detected biometric peaks with the manually introduced emotional categories for each movie segment. This study had three interesting findings: (1) music is associated with users' emotional highlights; (2) participants have similar physiological responses, when watching films in small groups; and (3) the emotions easier to detect were fear/anger/distress and hope/pride/joy, and these were the most influenced by music.

Another study proposes a multimedia retrieval system that allows to identify content relevance based on user feedback that derives from real-time facial expression recognition and physiological signals (Arapakis, Konstas & Jose, 2009). Physiological signals were used to assess the intensity of the categorical emotion observed by the facial recognition. Intense emotions were associated with relevant contents, while not so intense emotions were associated with irrelevant contents, and these two measures (facial and physiological) were the feedback information used to classify videos.

Although these works feature advanced research in the affective classification of videos, they still do not provide the differentiation of emotions along a full movie. Therefore, to improve the search, access and create new ways of watching and interact with videos we believe we should explore information visualization methods and techniques. Thus, in the following section, we present a set of different perspectives for video presentation and exploration.

2.7 Video Impact on Human Emotions

The impact of video on human emotions is being studied for some decades now (Münsterberg, 1970; Philippot, 1993; Rottenberg, Ray & Gross, 2007) but the analysis of such an emotional impact can only be measured from an automatic perspective if emotions can be elicited. Any emotion recognition process needs to address the problem of how to induce emotions, in what conditions, and then it is necessary to establish how to evaluate and measure the emotional responses. As Picard claimed, emotional inducement is everywhere around us (Picard, 1997). When we see our loved ones happy, we feel happy; when at work, we accomplished something, we feel good. In the scientific field, this induction is being study by analyzing different sources, such as actors acting or our imagination. In media studies of affect, movies have already been studied as emotions inductors. In fact, to automatically recognize emotions we need to have an inductor method. Regarding media studies, the most common methods for eliciting emotions are images (Bradley & Lang, 1994; Junghöfer, Bradley, Elbert & Lang, 2001; Lang, Bradley & Cuthbert, 1997; Pollatos, Herbert, Matthias & Schandry, 2007), music (Kim & André, 2008; Scherer & Zentner, 2001), or films (Christie & Friedman, 2004; Gross & Levenson, 1995). Often, it is difficult to induce intense emotions in a laboratory context (Marston, Hart, Hileman & Faunce 1984; Martin 1985) and, according to Rottenberg, Ray and Gross (2007), it is extremely difficult to induce anger. In 1996, a research group (Westermann, Spies, Stahl & Hesse) tested eleven induction methods and concluded that films are the best method to elicit emotions (positive and negative), in particular when participants (not studying psychology) are treated individually and are introduced to the purpose of the study. This study also concluded that it is easier to induce negative emotions than positive, but the authors argued that this could be a result of the positive mood of the participants when entering the experiment, which becomes harder to enhance than to reduce. However, this is not hard to understand given that there are more negative than positive basic emotions. In fact, from the five consensual basic emotions (Damásio, 1995; Ekman, 1994; Oatley & Johnson-Laird, 1987) four are negative (anger, fear, sadness, disgust) and only one is positive (happiness).

The study of films as an emotion induction method had its first appearance in 1916 when Hugo Münsterberg analyzed the mental operations of film viewers (Münsterberg, 1970) and discussed how emotions guide the motivation of perception and consequently the control of our attention by cinematographic narratives. More recently, other researchers also used films to induce emotions with different goals. One of the first works is the study of Philippot (1993), who tried to find films that induce different emotional states (dimensional space). Gross and Levenson (1995) tried to find as many films as possible to elicit discrete emotions and to identify the best films for each discrete emotion. In their study, sixteen movie scenes were the most productive in the induction of seven emotions (amusement, anger, contentment, disgust, fear, sadness, surprise) plus a neutral state.

Although these studies differ in the number of participants and in the emotional model used to evaluate the stimulus, both of them used the self-assessment method to recognize the emotional results from each scene. Table 4 lists Gross and Levenson's sixteen movie scenes

Table 4 - List of the selected movie scenes to induce emotional states

Emotion	Movie	Scene
Sadness	The Champion	Boy cries at father's death
	Bambi	Mother deer dies
Anger	My bodyguard	Bullying scene
	Cry Freedom	Police charging against protesters
Amusement	When harry met Sally	Discussion of orgasm in cafe
	Robin Williams Live	Comedy routine
Disgust	Pink Flamings	Person eats dig faces
	Amputation	Amputation of arm
Fear	The shinning	Boy playing in hallway
	Silence of the Lambs	Basement chase
Surprise	Capricorn one	Agents burst through door
	Sea of Love	Person startled by pigeons
Contentment	Unknown movie	Waves
		Beach Scene
Neutral	Unknown movie	Abstract shapes
		Color Bars

For instance, from Gross's study we have Amusement, Contentment, Surprise, Sadness, Anger, Disgust, Fear and Neutral.

2.8 Summary

In this chapter we analyzed the role of emotions in the HCI context and presented the main models of emotion to be considered in this work. We have also analyzed different types of emotions recognition techniques, deepening on physiological responses. We also explained different perspectives of emotion recognition in video. After having revisited emotional theories and recognition processes and issues, we present in the following chapter a number of possibilities by which emotions and videos are being visually represented in the literature.

Chapter 3

Visual Exploration of Emotions in Video

3.1 Introduction

Nowadays, huge amounts of data are being uploaded to the Internet, ranging from texts to audio or video. Such data, when properly organized, accessible and reachable, represents information and meaning (Zhang, 2008), but without this organization it could mean just chaos and noise. Also, relations between data represent more information that, in huge, complex, dynamic and heterogenic information systems, demands for suitable visualization and interaction techniques (Zhang, 2008).

In this work, we are exploring the emotional dimensions contained in videos, more specifically in movies. In what emotions are concerned, videos can be seen from two perspectives, a) by the *Objective Emotion* that is conveyed, e.g. a video or a scene showing happy people; or b) by the *Subjective Emotion* that it induces on the viewer, for e.g. sadness, because the user relates those specific kind of situations with a sad event in her life.

Objective emotions can be classified either manually or with the aid of some automatic process: through video processing techniques (Snoek, 2007; Yoo & Cho, 2007), and emotion recognition methods (Hanjalic & Li-Qun, 2005; Xu, 2008) based in video low-level feature analysis, combined with audio and subtitles processing (Langlois, 2010), as noted in the previous chapter. On the other hand, subjective emotions can also be classified manually, by the user, in accordance with the emotions the user felt along time, or automatically recognized with biometric methods based on brain signals and physiological signals, such as respiration, heart rate, galvanic skin response, detected on users, or through recognition of facial emotional expressions (Mauss & Robinson 2009).

In this chapter we analyze how information visualization techniques help human cognition and then we investigate how emotions and videos are both being represented and explored by current systems.

3.2 Cognition and Information Visualization

As suggested by Neisser (1976), "cognition is the activity of knowing: the acquisition, organization, and the use of knowledge". Perception is the first step of the cognitive activity and it consists in the apprehension of stimulus. Visual perception enables to successfully navigate and explore through an environment, which is why it is so important the way information and interactive mechanisms are represented. Information visualization is the discipline that presents and explores information for better perception and understanding. In our times, computers play a major role as a cognitive tool, and the representation of information in their displays is an important mean for computers to communicate with humans (Ware, 2004). In his book, *Information Visualization – Perception for Design*, Colin Ware (2004) enumerates five advantages of computed information visualization, including the following:

- It provides an ability to comprehend huge amounts of data.
- It allows the perception of emergent properties of the data that were not anticipated.
- It facilitates understanding of both large-scale and small-scale features of data.
- It facilitates hypothesis formation, i.e. it facilitates the creation of knowledge beyond the one explicitly represented.

One of the 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. Because of this, 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 thus be used to tag some information of the video. And once this information is collected, we may 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 (Hauptmann, 2005b; Martinho & Chambel, 2009). Visualization techniques that emerged from research rooted primarily on visual perception and cognition (Card, Mackinlay & Shneiderman, 1999) 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 (Martinho & Chambel, 2009). In the context of this work, we are focused on the emotional properties of movies, thus, in the following section, we describe the actual emotional models and their visual representation.

3.3 Emotion Information

Human beings often use metaphors to understand the meaning of abstract concepts, or to process new information (Lakoff & Johnson 1980). These metaphors are used both in daily life (e.g. “life is heavy”) and in scientific knowledge (e.g. “processes of repairing tissues”), and they help dealing with a new reality, using a pre-existing one. Thus, representing mechanisms that we already know help human perception to easily detect the meaning or functionality of new information. For example, desktop metaphors interfaces in operating systems draw a trash basket resembling its exact appearance in our offices and, thus, people automatically recognize its functionality. Metaphors are also reflected, for example, in the way applications such as iTunes choose to organize the music around the music album metaphor. This section analyses how emotions are being worked in actual systems and how they are being represented.

3.3.1 Emotion Information Visualization

It is our intention to comprehend how emotional models represent their information. In any used representation the aim is to further analyze how visual design can enhance the perception of the information they represent. As we stated above, in section 2.2, there are two main models of emotion, the categorical and the dimensional. Although there is no standard visual proposal to represent categorical emotions, there is a correspondence of these emotions in Russel’s circumplex model (see Figure 6). Happiness is in the top right side of the circle while Surprise, and Fear, Anger and Disgust in the top left. In the down left side is where we find Sadness, while there is no emotion in the down right side of the circle.

The Appraisal Model is also categorical and consists of a set of words describing emotional descriptors. Each emotional descriptor is defined as the evaluation of the interaction between someone and their goals, beliefs and the environment: Scherer (2005) defined emotional descriptors based on these assumptions to create the Geneva Affect Label Code (Scherer 2005). There is also a visual representation correspondence of these emotions in Russel’s circumplex (see Figure 6).

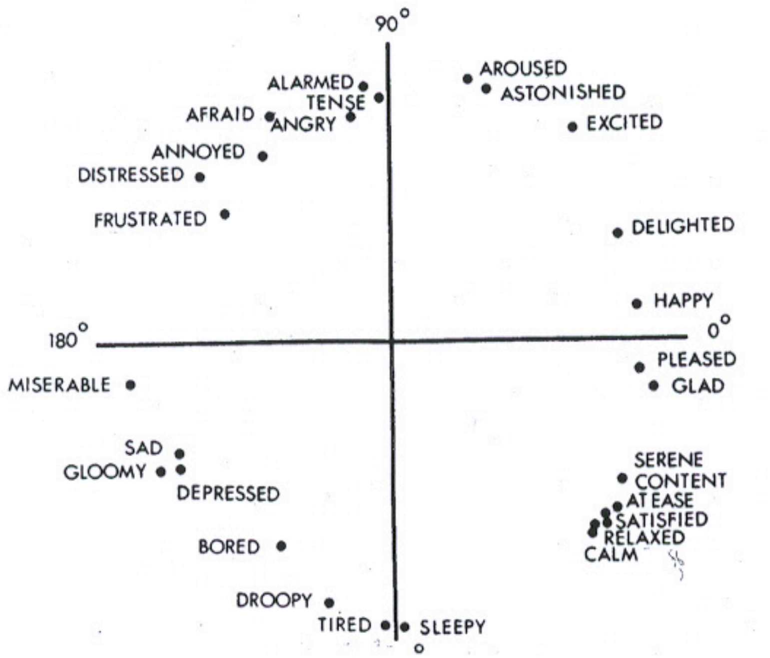

Figure 6 - Russel's circumplex model of affect

Another emotion theorist, which uses both categorical and dimensional models, is Plutchik (1980) who defends a three dimensional model (polarity, similarity, intensity) with eight primary emotions: anger, fear, sadness, disgust, surprise, anticipation, trust, and joy.

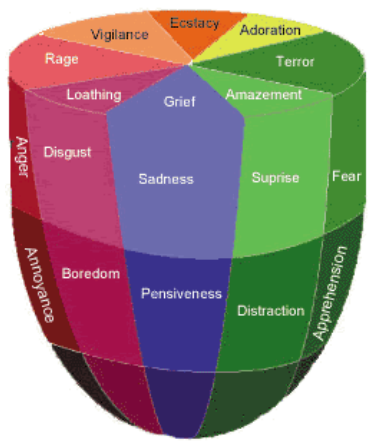

Figure 7 - Plutchik's diagram of blend emotions

In Plutchik's Three-Dimensional Circumplex Model (see Figure 7), the relations between emotional concepts are comparable to the colors of a color wheel. The circle assumes the degree of similarities among emotions and the vertical dimension represents the intensity.

The design for emotion explores color and its properties such as saturation as well as the combination between colored shapes and neutral graphics. In fact, the emphasis in emotional characteristics in a graphical visual representation directs a person attention. In the context of this work, we developed automatic means to detect emotions on users, to visually represent them, and explored new ways to access and interact with this emotional information.

However, emotions automatic recognition methods can only use the dimensional or the simpler categorical models, due to limitations in differencing emotions from physiological, facial or brain signals (Maaoui, Pruski & Abdat, 2008). However, the Appraisal Model gives a wider range of words for movie emotional impact and content descriptions. For example, 'amused' is a word often used to describe an emotional state after watching a movie (Rottenberg, Ray & Gross, 2007), which gives to manual classification methods a wide range of visual representation possibilities to identify emotions on movies.

The visual representation of emotional models should be explored in the context of video emotional classification systems to improve the perception and understanding of this type of information, along with interactive mechanisms for improving its usability, which is precisely the main motivation of this thesis.

3.3.2 Visualization Systems Based on Emotional Information

In the context of emotion elicitation and visualization there are some recent works but, to our knowledge, only few of these were developed on video. The "We Feel Fine" (Harris & Kamvar, 2009) system, depicted in Figure 8, harvests human feelings from a large number of weblogs, searching in newly posted entries the occurrences of the phrases "I feel" and "I am feeling", and identifies the "feeling" expressed in that sentence (e.g. sad, happy, depressed, better) as well as available information about the author (e.g. age, gender and geographical location) and local weather conditions at the time it was written, allowing to answer such questions as the following: "do Europeans feel better more often than Americans?"

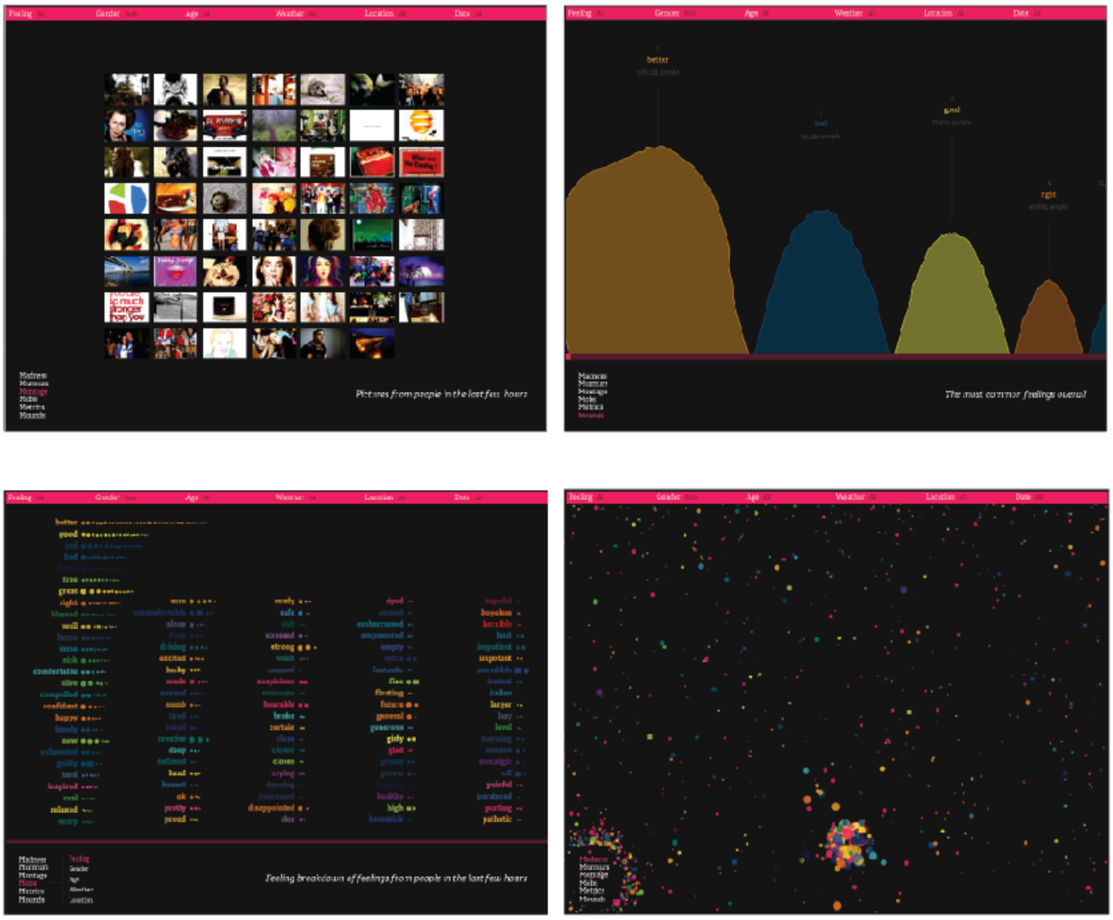

Figure 8 – "We Feel Fine" emotion representations (by image, mobs, word listing and bubbles)

The user interface of this system is based on a physical particles system, where each particle represents one feeling expressed by an individual (showing the whole sentence, when clicked), with properties such as color, shape, size and opacity representing the nature of the feeling.

Synesketch[3] is a textual emotion recognition and visualization software based on the concept of synesthesia, dynamically converting text into animated visual patterns. Hooloovoo (see Figure 9), one of its demos, provides simple and minimalist visualizations

[3] Available online at: http://www.synesketch.krcadinac.com

of emotions found in texts through colored squares. Emotions influence colors, saturation, size and frame rate.

Figure 9 - Hooloovoo scheme of a happy phrase

For example, 'no emotion' is represented by dim grey in small squares, weak disgust in more saturated grey scale in small squares, while strong happiness is represented with bigger squares in brighter colors with dominance of yellow and red components.

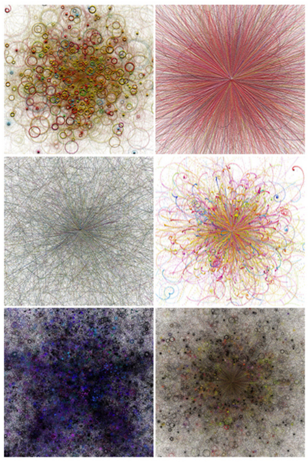

Figure 10 - Six images corresponding to six basic emotions
(From left to right up to down: happiness, anger, fear, surprise, sadness and disgust)

Synemania (Figure 10), another demo of Synesketch, adopts a more complex visualization scheme based on a particles system that creates abstract visual patterns with their movement. Patterns depend on the corresponding emotion and intensity, reflected on the particle color and type. The EmotionallyVague[4] project addresses the relationship between body and emotion and how people feel emotions such as anger, joy, fear, sadness and love.

[4] Available online at: http://www.emotionallyvague.com

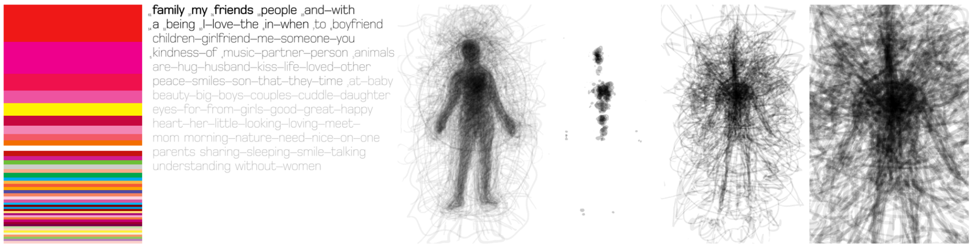

Figure 11 - EmotionallyVague representation of love

The results obtained from a survey revealed patterns of feeling that were represented by colors and body locations, showing, for example, that anger and love (see Figure 11) manifest mostly in the head and the chest, respectively, while joy is mostly associated with bright and saturated colors, and in opposite emotions, like anger and love, the dominant color tends to be the same.

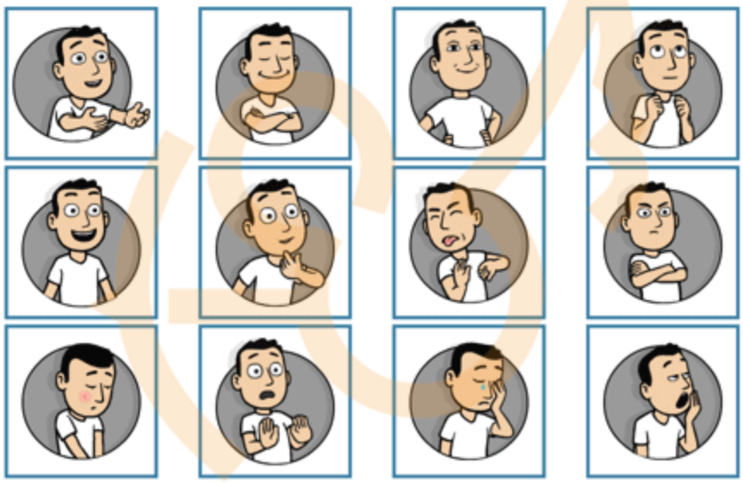

Figure 12 - Premo classification mechanism

Premo (Figure 12) is an emotional classification mechanism developed to be a non-verbal, multicultural, way of inferring about emotional experiences in interaction with products (Desmet, Porcelijn, & van Dijk, 2007; Desmet, 2007). It differs from the other systems in that it represents emotions by using cartoons in motion, and even with sound, expressing seven positive and seven negative emotions, that represent low intensity complex emotions. There are other research works, which we review in the following section, that explore information visualization in terms of the contents of videos.

3.4 Video Information

Video is a medium that communicates information using visual, audio, speech and text cues. The message a video communicates is intrinsically linked with its type of content. In what regards movies, there is a great amount of information that would be useful to be readily accessible, like its general information (e.g. genre, director, cast, country), as well as the emotional content, or the emotional impact it produces on viewers. While the former is already accessible in mainstream movie sites (such as IMDB), the later is still unavailable.

3.4.1 Video Information Visualization

In the context of video visualization, Film Finder (Ahlberg & Truvé, 1995), depicted in Figure 13, allows users to search for certain films with different types of visualization support based on the movie duration (minimum and maximum), genres, titles, actors and directors. The selected movies are presented in a star field graphic based on their date and popularity. In this approach, users may find movies based on their tastes or preferences, and explore the movies space by zooming in and out of more or less detailed information. Most visualization tools and applications found in recent surveys, such as the one in Perez (2008), do not however address video. Exceptions include: 1) "Call and Response[5]" experimental project, that visualizes a communication network made through short videos among art students, representing videos by one keyframe, and focusing on the communication; 2) "YouTube", that was available in a previous version until about three years ago; 3) "Video Sphere" (Bestiario, 2008); 4) Yasiv, a visualization tool for YouTube related videos. These last three are the ones most related to the objectives of the work described in this thesis. From each video on YouTube, the user could access a 2D view that represented videos as circular scattered still images, providing access to the traditional page where users can watch the video. It allowed for visual neighborhood navigation, but provided limited functionality and information about the videos or the video space. On the other hand, VideoSphere (see Figure 14) represents TED's videos as a video space around a 3D sphere, with links among the videos, reflecting semantic compatibility, and allowing navigation around, inside and

[5] Available online at: http://www.risd.tv/callresponse/leptonRun.html

outside the sphere.

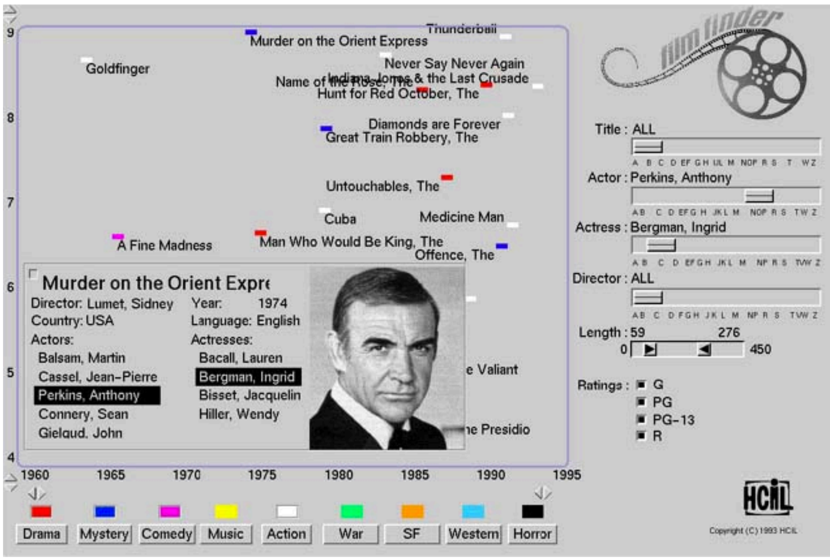

Figure 13 - Film Finder interface

Figure 14 - VideoSphere

In this case, the visualization is restricted to the videos represented on that sphere, with the focus on exploring semantic relations, without any special support for the visualization of the videos other than still keyframes and traditional video play. A more recent work, YASIV [6], is a search mechanism that displays YouTube related videos, sizing each video window according to its number of views. The relations between videos reflect YouTube's relation policy; what Yasic does is to graphically arrange the YouTube related video list, by displaying them in clouds as we can see in Figure 15.

Figure 15 - Yasiv YouTube video related

In our research team, previous work (Rocha & Chambel, 2008) provides interactive 3D visualization and navigation of video, as an art installation (Figure 16), to explore cultural properties and links among different videos, in semantic categories such as countries, themes and authors, at the level of the videos space and the individual videos.

Another system - Color in Motion - is a 2D interactive system based on a physical particles system (Martinho & Chambel, 2009) that could visualize and explore the same videos but stressing features such as their color dominance, rhythm and movement (see Figure 17). These systems allow for capturing, experiencing, and expressing videos' properties and relations, providing the means to gain new insights into our culture and to influence the expression of its intrinsic aesthetics in creative ways at the crossroads of information access, culture and digital art.

[6] Available online at: http://blog.yasiv.com/2012/02/introducing-YouTube-visualization.html

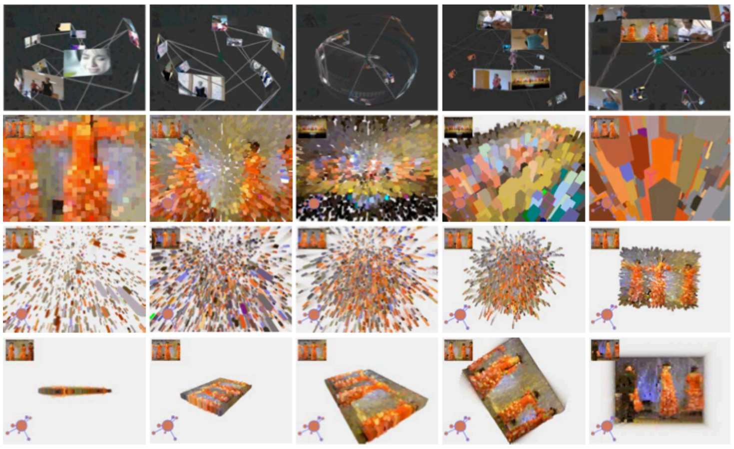

Figure 16 - Video Space visualizations

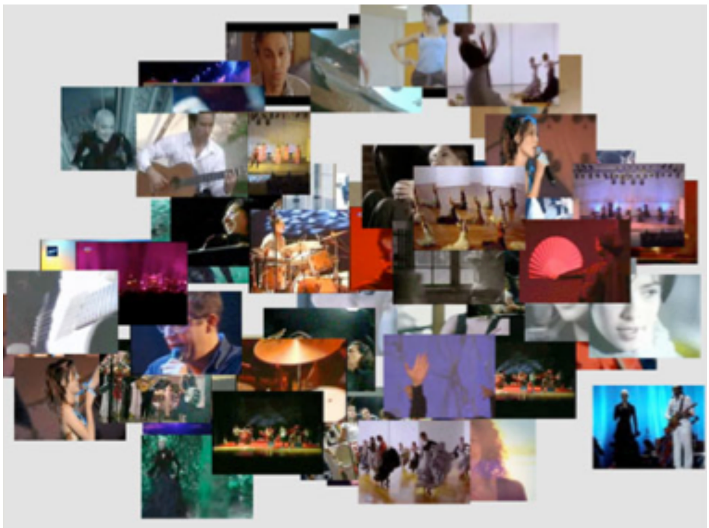

Figure 17 - Colors in Motion – video loops view

All these representations confer an identity to video that can only be reached if these characteristics are well organized and semantically well defined. However, these works address video visualization but they do not consider emotions. The semantic representations of emotions in videos will constitute the basis for the search of emotional dimensions of videos, both from the content perspective and from the impact on users. In this context, some advances have already been made in the semantic organization of emotional information in different contexts of computational applications. Accordingly, the following section reviews the semantic representation languages of emotions in video.

3.4.2 Video Annotation and Classification

Some research developments in video processing and indexing are already taking emotions into account to classify videos. Video content analysis and indexing for browsing, exploration, or summarization is an area of research that deals with techniques and mechanisms to extract information from video and to find meaningful segments, making videos easily discoverable. Some techniques are based in the extraction of low-level features such as color, texture, lighting, motions, sounds and lexicon analysis, inspired by cinema theorists' tools like narrative descriptions or sound and image properties. The retrieval process is done through an iterative process that creates new populations of videos through crossover, by searching most similar solutions in the video database, based on human evaluations of scene types: action, excitement, suspense, quietness, relaxation and happiness. Another example is the detail-on-demand video mechanism, explored by Doherty, Girgensohn, Helfman, Shipman and Wilcox (2003).

Other techniques reported in the literature are essentially manual, also using video descriptors such as the ones proposed in MPEG-7, or typical tags like those used in YouTube or Vimeo to describe the videos. Therefore, in order to be able to search for a video, users need that video was previously annotated. Video annotation is a technique that uses the information extracted from content analysis and links it to the content itself by an external file or by encapsulation. It allows accessing, browsing and discovering videos by the information stored in its annotated files or structures. Media annotation is based on metadata, which can be defined by ontologies. One example of a standard ontology is the Ontology for Media Resources that is defined by the World Web Consortium (W3C). The main goal is to define a core vocabulary for media to enable a bridge between different descriptions of media resources on the web, providing a core set of metadata properties along with a mapping with existing metadata formats such as the one used by YouTube, MPEG-7, or Dublin Core.

IMDb, the Internet Movie Database, is probably the most well-known online database of information about movies and TV shows. It provides information about actors, directors, production crew, video games and fictional characters. This information (metadata) is stored in a database based in XML and is collected by staff that processes the general information about movies, but also by people in movie industry and users in general. Metadata related with ratings and preferences or comments represents the major part of their information (70%) and is fed by online visitors. Films are characterized by film genres, and the system allows users to classify movies based on their appreciation. Other online services, such as Netflix[7], allow accessing and watching movies through on-demand video streaming over the Internet and online rental of DVD and Blu-ray videos, also providing some common information about the movies. Netflix uses, like IMDB, a large database to collect all the information (metadata) about movies, which is also manually introduced by their staff and stored in a database.

In this context, YouTube is probably the most famous and used online service to publish and watch videos. Users publish and annotate videos providing a title, description, tags and categories that can be used in video search and to select videos most similar to the video the user selects to watch. Users may evaluate, comment and share videos in social networks, and the system also provides recommendations, tendencies, highlights, and most recently watched movies. Other systems in this category include Vimeo. More specifically, YouTube uses the Google Data API protocol to represent videos information (metadata) and to provide developers the access to general and user introduced information. The annotation of videos in these systems is based on their manual classification. However, none of the previously reviewed systems is automatic by itself or does it provide emotional information about movies.

Another innovative movie based site – Jinni[8]– based on manual tagging, uses a specific set of moods according to the genre of the movie, and encourages users to vote in a mood. One disadvantage is that viewers are restricted to a narrow set of moods, and for instance one cannot classify with a negative mood if the movie genre is comedy. Also, in Jinni, where movies are categorized by moods, when choosing bittersweet titles viewers will find 2168 movies categorized with that mood. The way they present those results is in pages of about 23 movies per page, where the size of the thumbnails varies according to the number of classifications for that title. Viewers clearly see the most bittersweet movies, by their size. The problem is that there are many thumbnails of the same size and viewers don't actually know the difference between them. Also, when

[7] Available online at: https://www.netflix.com
[8] Available online at: http://www.jinni.com/

viewers move from page to page, they see again a movie as big as the most bittersweet of the first page.

In the following sections, we analyze recent research works on what regards the automatic classification of videos in an emotional perspective and by both content and emotional impact on its viewers.

3.5 Semantic Representation of Emotions and Videos

As stated above, the advent of rich interactive multimedia content over the Internet in educational or entertainment environments, among others, and the way users use online multimedia content, such as film viewing, video and image sharing, asks for new ways to access, explore and interact with this information (Purcell, 2010). Video on the Web has been in explosive growth, which improves the fullness of the user experience but leads to new challenges in content discovery, searching and accessing. Information needs to be labeled or annotated to be accessible, shared and searchable.

The Multimedia Information Retrieval (MIR) research area is still trying to find solutions to automated content and users analysis techniques, and annotation techniques for media, as a result of a huge need of descriptors (metadata) of contents that can be understood by computers and accessible to humans. We need automatic methods for gathering information both from multimedia objects (video, images, audio, text) and from users (preferences, emotions, likings, comments, descriptions, annotations), and subsequently making this information available, searchable and accessible (Lew, 2006).

In the literature, there are several studies that have attempted to define standards to establish structures of descriptors and concepts for affective applications in their categorization issues (Devillers, Vidrascu & Lamel, 2005; Douglas-Cowie et al., 2007; Luneski & Bamidis, 2007; Schroder et al., 2007). One of the first works developed towards this goal was the HUMAINE[9] database, despite the fact that it did not propose a formal definition to structure emotions, but simply identified the main concepts.

The attempt to define a standardized description of emotions faces several challenges

[9] Available online at emotion-research.net

and some questions naturally arise. As we described in section 2.1, there is more than one theory of emotions and there is also no common agreement on the quatity and the actual names of emotions. Despite this fact, there is some agreement that some recognition techniques might be used, such as physiological patterns, brain and speech changes, or facial expressions. In the following sections, we review several approaches that allow structuring these different types of emotional information in a standard and comprehensive way.

In the context of this thesis, we focus on the semantic representation of videos and movies, regarding them both from users and from the content points of view, although the standards for video have semantic approaches for many other application cases. Representing emotional information about video content in both content and users perspectives requires a structure and a set of descriptions and concepts. We now introduce the two main semantic approaches to categorize emotional information in computer application environments.

3.5.1 EARL - Emotion Annotation and Representation Language

One of the first works (Schroder, Pirker & Lamolle, 2006) towards the standardization of the representation of emotions was the Emotion Annotation and Representation Language (EARL), which aims to represent emotions suitable for the most common use cases: 1) manual annotation of emotional content, 2) affect recognition systems, and 3) affective generation such as that performed by speech synthesizers. Because there is no agreed upon model of emotions, and there is more than one way to represent an emotion (e.g. dimensional, categorical), EARL leaves freedom for users to manage their preferred emotion representation, by creating XML schemas for domain–specific coding schemas embeded in EARL, that serves application cases which demand for specific data categorization (Schroder et al., 2006). This constitutes one of the advantages of EARL, and the other is the fact that being an XML-based language, it standardizes the representation of data, allowing for re-use and data-exchange with other technological components (Schroder et al., 2007).

In order to provide suitable descriptions among the different use-cases, EARL assumes descriptions for the following information presented in Table 5. Schroder et al. (2007) argue that these are the informative topics needed to describe simple use-cases that use only emotional labels to process their information.

Table 5 - EARL emotional description minimal requirements

Emotional Data	Description
Emotion descriptor	Any emotional representation set
Intensity	Intensity of an emotion expressed in numeric, discrete values
Regulation types	Which encode a person’s attempt to regulate the expression of her emotions
Scope of an emotion label	Link to an external media object, text, or other modality
Combination	Co-occurrence of emotions and their dominance
Probability	Labeler’s degree of confidence

An EARL example of a description that needs to use a simple emotion to represent a text that aims to sound pleasurable, could be written in the following simple structure:

```
<emotion category=”pleasure”>Nice to meet you!</emotion>
```

Complex emotions, as considered in the context of EARL, are composed of several elements, such as more than one emotion and the probability of each one occurring. In the following example, we have a description of an image with two emotional categories and their intensities:

```
<complex-emotion xlink:href="face12.jpg">
  <emotion category="pleasure" probability="0.5"/>
  <emotion category="friendliness" probability="0.5"/>
</complex-emotion>
```

An example of video annotation with a simple emotion could be as follows:

```
<emotion category="pleasure" probability="0.4" start="0.5" end="1.02"
xlink:href=”video.avi”/>
```

While a video annotation with a complex emotion could be represented as follows:

```
<complex-emotion start="0.5" end="1.02" xlink:href="video.avi">
  <emotion category="pleasure" intensity="0.7"/>
  <emotion category="worry" intensity="0.5"/>
</complex-emotion>
```

Other example, that is important to the context of this work, is the representation of emotions recognized from sensors while users watch a video clip, such as the following:

```
<complex-emotion xmlns="http://emotion-research.net/earl/040/posneuneg"
xlink:href="clip123.avi">
   <complex-emotion modality="biosignal" start="0.387" end="0.416">
      <emotion category="positive" probability="0.1"/>

   <complex-emotion modality="biosignal" start="0.417" end="0.597">
      <emotion category="neutral" probability="0.4"/>

   <complex-emotion modality="biosignal" start="0.598" end="0.897">
      <emotion category="negative" probability="0.05"/>
    </complex-emotion>

  </complex-emotion>
```

EARL is constructed with modules, which allow for the development of new dialects by combining a base XML schema with other XML schema “plugins”. Three examples of plugins are the emotional definitions, such as the categorical, the dimensional and the appraisal, so any application can use these EARL dialects according to their needs, but can also develop different EARL documents with new specifications. In fact, EARL can also be used integrated with other languages like the Extensible Multimodal Annotation markup language (EMMA, 2010), used to represent information automatically extracted from a variety of inputs such as speech, natural language text and ink input, providing a set of elements and attributes focused on enabling annotations on user inputs and interpretations of these inputs.

The integration of EARL with EMMA can be useful to add an emotion component on EMMA's user input interpretations and can be made via EMMA's extensible content representation mechanism. For example, an interpretation of emotion on an EMMA element using EARL complex emotion representation could be represented as follows:

```
<emma:emma version="1.0" xmlns:emma:="http://www.w3.org/2003/04/emma"
   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
   xmlns:earl="http://emotion-research.net/earl/040/posneuneg">
  <emma:interpretation id="int1">
   <earl:complex-emotion>
    <earl:complex-emotion modality="biosignal">
     <earl:emotion category="positive" probability="0.1"/>
     <earl:emotion category="neutral" probability="0.4"/>
     <earl:emotion category="negative" probability="0.05"/>
    </earl:complex-emotion>

</earl:complex-emotion>
  </emma:interpretation>
</emma:emma>
```

EARL was the first movement towards the creation of a standard semantic representation of emotion. The second step resulted in the Emotion Markup Language, which we present in the following section (Schröder et al., 2011).

3.5.2 Emotion ML - Emotion Markup Language

The Emotion Markup Language (Emotion ML) was born to define a general-purpose emotion annotation and representation language. After the definition of EARL, the respective working group (Schröder et al., 2011) moved to the World Wide Web Consortium (W3C) in the form of two incubator groups. First, an Emotion Incubator group was set to analyze use cases and requirements of a markup language, and then the Emotion Markup Language Incubator group to dissect the requirements of such a language. Like EARL, the first W3C incubator group defined a standard for representing and processing emotions related information in technological environments.

The work group opted to use XML as the semantic representation language because its standardization allows the easy integration and communication with external applications by having a standard emotional representation structure that enables other technological artifacts and applications to understand and process data. In fact, the most recent markup language is defined in XML, which appears to be a good choice. Emotion ML is presently in the Second Public Working Draft (*W3C working draft*) after the First Public Working Draft (FPWD), which was published in 2009. The FPWD proposed elements of a syntax to address use-cases in what regards emotion annotation, the automatic recognition of user-related emotions and the generation of emotion-related system behavior.

As EARL, Emotion ML does not enclose annotated emotions, but uses XML attributes to represent the type of data to be presented. For example, there is an attribute – *expressed-through* – that is used when considering annotation or recognition use-cases to specify where the emotion is expressed, whether by face, voice, posture or physiological data.

As an example of the application of this attribute, we might use it in a video annotation in Emotion ML such as the following:

```
<emotionml xmlns=http://www.w3.org/2009/10/emotionml"
xmlns:meta="http://www.example.com/metadata"
category-set=http://www.example.com/custom/emotv-labels.xml expressed-
through="biosignal">

<emotion>
    <category name="irritation" value="0.46"/>
    <reference uri="file:ext03.avi?t=3.24,15.4">
</emotion>
<emotion>
    <category name="despair" value="0.48"/>
    <reference uri="file:ext03.avi?t=5.15,17.9"/>
</emotion>
</emotionml>
```

An example of the representation of the automatic recognition of emotions using physiological sensors could be written as follows:

```
<emotionml xmlns=http://www.w3.org/2009/10/emotionml
category-set="http://www.w3.org/TR/emotion-voc/xml#everyday-categories">
…
<emotion start="1006526160" end="1268647330"
 expressed-through="physiology">
      <category name="excited"/>
      <reference uri="http://www.example.com/physiodb#t=19,101"/>
</emotion>


<emotion start="2346526520" end="5438647330" expressed-through="physiology">
        <category name="angry"/>
      <reference uri="http://www.example.com/physiodb2#t=2,6"/>
</emotion>
</emotionml>
```


As can be seen in the semantic representations above, the development of technological support by using XML structures for human activities that are based on emotional aspects requires the representation of different components. The Emotion Markup Language notably considered a proper body of use cases to set a general structure that supports and enables the communication between technological components regarding the emotional properties of, among others, images, videos, text, automatic recognition of emotions, emotion synthesis and speech. For example, SMIL (Bulterman, Dick & Rutledge, Lloyd, 2008) is a presentation-level XML language that combines animation and media objects into a single presentation in an interactive way. It controls the temporal and spatial characteristics of media and animations elements in an interactive interface. Emotion ML can be used as a specialized plug-in language for SMIL as an emotion engine to thrill facial and vocal expressions.

3.5.3 Other Semantic Representations

De Carolis et al. (2001) present an Affective Presentation Markup Language (APML) which is also an XML based language used to represent face expressions of their agents dialogs, in their work on Embodied Animated Agents (ECA). Other research work related with emotion concepts representation is ALMA (Gebhard, 2005), a layered model of affect, also called AffectML, whose main purpose is to represent affect and moods in their virtual humans' project. The interesting aspect of this project is that, in a first phase, they use emotions as a medium-term-affect representation, and later, as a long-term representation of the personality of their characters. Both APLM and AffectML were developed within affective applications to endure the affective needs.

3.6 Summary

In this chapter we reviewed the most relevant research works that focus in the representation of emotions, analyzing how emotional models present emotions and how actual systems that explore emotional information use such representations in their work. Video information was particularly covered highlighting visual explorations of videos and the interface mechanisms encountered to represent a structure of video information. From the review presented in this chapter, we may conclude that there is no report in current literature concerning the representation of movies either by the emotional impact on viewers or by its content. This review also made clear that there is an effort to develop standards concerning the representation of emotions, but also that there is not yet a closed solution for this representation. The following chapter describes an empirical work carried out on the automatic recognition of emotions felt by users while watching movies using physiological sensors.

Chapter 4

An Empirical Study on Automatic Emotion Recognition and Annotation

4.1 Introduction

In this chapter, we present a set of proposals designed to explore the benefits and potentialities of automatic emotion recognition and annotation, facing the challenges and overcoming the problems and difficulties identified in the previous chapter. We propose a system that allows for emotion recognition and for the assessment and exploration of the different dimensions of emotions on video (Oliveira, Benovoy, Ribeiro & Chambel, 2011; Oliveira, Martins & Chambel, 2011). For this purpose, we developed an emotion recognition engine as our experimental environment, along with an emotion induction procedure, and present preliminary results of the recognition of emotions. We also present the information requirements for a semantic representation for users and movies regarding the emotional impact of movies on users based on web standards, allowing the usage of collected emotional data by other systems.

4.2 Experimental Environment Description: iFelt

The experimental environment developed for testing the hypothesis suggested in this thesis was designated by iFelt. First, we present an experiment that automatically learns emotional patterns felt by viewers while watching movies. We then present two experimental interfaces developed for enabling the visual exploration and representation of the learnt emotional information, and present a semantic structure to organize all such information. The first prototype was developed in Flash, while the second prototype is being presently developed in PHP/MSQL and JavaScript.

Thus, the iFelt system was designed and developed to support both experiments. Although the two components were tested separately, in this section we present the architecture that supported both experiments.

As may be noted in Figure 18, we developed two main components for iFelt: 1) the Recognition and annotation component, which enabled us to test a recognition engine and to structure the gathered emotional information in an XML file; and 2) the Emotion representation and exploration component, a visual user interface component which enabled us to test a set of different ways to represent emotions and to create mechanisms to explore them in the context of movies collections.

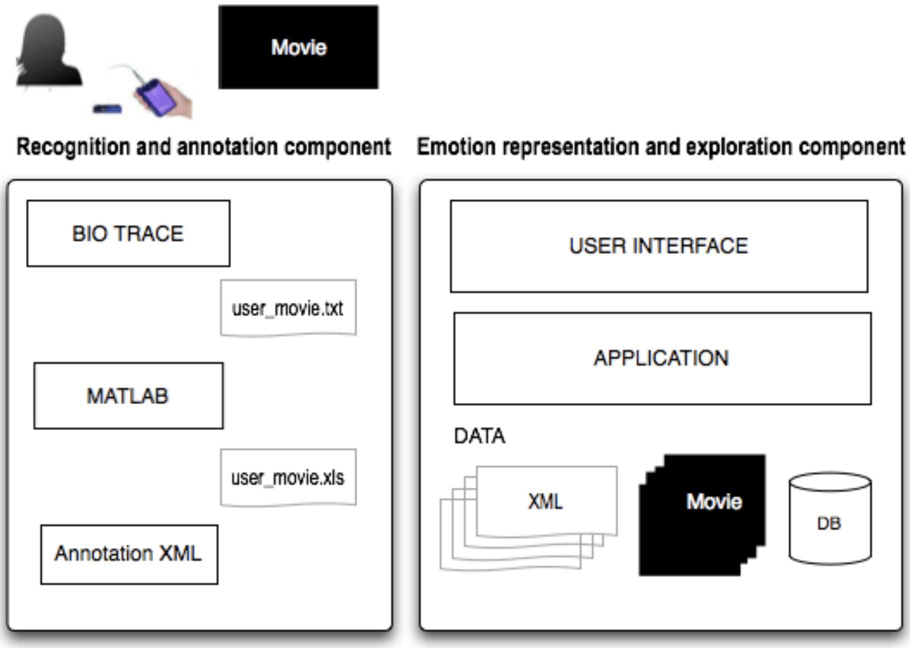

Figure 18 - iFelt system architecture

The technological artifact chosen to measure physiological measures was Nexus-4[10], a portable system with a bluetooth interface that supports the three physiological sensors we have used (Galvanic Skin Response, EEG and Respiration), together with BioTrace[11], a software application that transforms each sensor's signal into bio measures and saves each session into a text file, which is then fed to the recognition algorithm implemented in Matlab. The Matlab module transforms the text file information into emotionally meaningful information, as detailed later in this chapter. For this experiment, we exported the resulting information to an MS Excel file and analyzed the data to create an XML schema that organizes and structures all the data. As stated before, we created the XML schema to enable the emotional data to be communicated to other systems and applications as explained in section 4.8.

The Emotion representation and exploration component is composed by the user interface, fully described in Chapter 5 of this thesis, and by the application layer, which is responsible for the operationalization of the system and connection to the database.

[10] http://www.mindmedia.nl/CMS/de/products/nexus-systems/item/170-nexus4.html

[11] http://www.mindmedia.nl/CMS/en/products/software/item/305-biotrace%20.html

The application layer was not developed for the purposes of this thesis because the focus of this work was to test if it was possible to perform an automatic recognition of users' emotions while watching movies, i.e. whether emotions could be measured within our experiment environment.

Even so, this experimental system allowed us to test and evaluate our proposals in what two key aspects are concerned: a) the automatic recognition of emotions; and b) the exploration of the emotional information about videos, regarding the ways in which it can be presented, accessed and used in a useful way.

4.3 Experimental Objectives

As suggested by previous research works, reviewed above, given that movies can be one of the best methods to elicit emotions, we selected a collection of sixteen movie scenes for eliciting five of the emotions of the Ekman's categorical model and seven movies to test the emotion recognition engine. In this context, we performed video classification based on biometric methods, employing digital signal processing and pattern recognition algorithms to learn from users, inspired by statistical techniques used by Picard et al. (2001). For the recognition engine, we implemented support vector machines (SVM) and k-NN classifiers. After the classification procedure, movies were automatically annotated with their emotional information. This process and its results are thoroughly described below.

Our objective was to determine whether our classification engine is sufficiently accurate to automatically recognize emotional patterns from new data with a reasonably successful rate. Another goal was to evaluate whether the selected scenes had the same emotional impact in all users, in order to measure the importance of the scene for eliciting a specific emotion. Two classification algorithms, described below on section 4.6.3 were used to analyze user's biosignals.

4.4 Movie Scene Selection for Training

We selected scenes that were intense and representative of the set of emotions needed to test our engine, and also based on the experiments done by Gross' team (Gross and Levenson 1995; Rottenberg, Ray, & Gross, 2007) in finding the most intense movie scenes for a range of eight emotional categories. Table 6 lists our scene selection. Although Rottenberg et al. suggest that there should only be one or two movie scenes per emotion, so that users do not became fatigued, we have used a maximum of four scenes because, for automatic recognition, the training phase is crucial for obtaining good recognition rates, and having more movie scenes for emotion detection is crucial for this process of training. Thus, we have selected 3 scenes from Gross' work, and our own selection of 13 additional scenes with an average duration of 2 minutes and 22 seconds per scene. In order to elicit the neutral state, a video with some circles moving around was shown to each participant for 1.5 minutes, as suggested by Gross (Rottenberg et al. 2007).

Table 6- Movie scenes list used to elicit five emotions

(a) means that the scene is from Gross and Levenson study

Movie	**Expected Emotion**
Noncommercial Drool Film	Disgust
Noncommercial Surgery Film (a)	Disgust
Last Days (2 scenes)	Sadness
Champ (a)	Sadness
Noncommercial Child Cancer	Sadness
Le Fabuleux Destin d'Amélie Poulain	Happiness
Two Days in Paris	Happiness
When Harry Met Sally	Happiness
And Now Something Completely Different	Happiness
The Ring	Fear
I know what you did Last Summer	Fear
Shinning (a)	Fear
Shinning	Fear
Irreversible	Anger
Irreversible	Anger
Circles moving around	Neutral

We have selected each movie scene to target only one emotion. For example, it was expected that the scene of "The Ring" only elicited fear. This watching procedure is explained in section 4.6, below. It is important to note that Anger is a difficult emotion to induce; that is why we only have two movie scenes that address this particular emotion (Rottenberg et al. 2007). Also, for Disgust, we only introduced two scenes, since we did not want to disturb the participants, creating an unnecessary burden with performing the experiment, given that disgust usually elicits very strong reactions.

4.5 Movie Selection for Testing

To test our pattern recognition engine, we chose a collection of seven full movies, different from the movies from which we extracted the scenes used in the training phase. As we are in a movie context, we have opted to experiment our classification upon the data of 8 movies chosen on a multicultural selection basis, ranging from noncommercial to commercial blockbusters, and including different genres. The list of films used in our experiment is listed in Table 7.

Table 7 - List of movies used to test our recognition engine

Movie	Director	Country
Cashback	Sean Ellis	UK
De battre mon coeur s'est arrête	Jacques Audiard	FR
Cinema Paradiso	Giuseppe Tornatore	IT
Requiem of a Dream	Darren Aronofsky	USA
28 Days Later	Danny Boyle	UK
Mary and Max	Adam Elliot	AU
Crash	Paul Haggis	USA
Psycho	Alfred Hitchcock	USA

4.6 Experimental Protocol

Eight participants, averaged 34 years old computer literate, between 26 and 56 (6 female, 2 male), volunteered to participate in this study (convenience sample). They were connected to three sensors (see below on section 4.6.1) to test the recognition engine. Because we presented full-length feature films to our participants, and this was a very time-consuming activity, we were able to recruit only eight participants. However, given the exploratory nature of this study, this seems acceptable and sufficient for a first attempt to classify users' emotional reactions. In this experiment, we used a 20 inches monitor, 5 feet distant from the participant in a regular room. To capture user's responses, a portable system - Nexus 4 (*Nexus-4*) - was used with 3 inputs channels for ECG, Respiration and GSR.

After the participants arrived, the electrodes were attached and the recording system was checked. A web interface (depicted in Figure 19) was developed to easily perform the learning procedure, which began by asking the user to rest for 3 minutes, in a quiet mode. This interface presented the selected movie scenes, where the experimenter chose the scene to be watched. After each visualization, the interface prompted for the emotion the user has felt.

At the beginning of the sessions, a 3-minute silent baseline was recorded, while the participants engaged in focused relaxation by limiting their concentration to their respiration.

Figure 19 - iFelt emotion learning web interface

This pre-stimulus relaxation time was critical to stabilize the physiology to a homeostatic state, as some participants initially exhibited anxiety at being the focus of attention and being wired to the sensors. Then, the neutral scene was shown to collect the neutral state right after the baseline. The sequence of 16 movie scenes began with the happiest scenes first, alternating with fear related ones. Then, we showed the disgust related scenes, alternating with fear related ones. The last set included the sad scenes, alternating with anger, but chosen in a way that the most intense sad emotions were presented towards the end of the experimental session.

At the end of each scene, participants answered a digital questionnaire, identifying the emotion, from a set of five basic emotions, which they considered was the dominant one, reported its intensity (using a Likert scale from 0-9) and, finally, indicated whether they enjoyed watching it. This questionnaire was used to compare each user's answers about each scene. An algorithm was developed to link the captured biodata with the emotion a user felt when watching a specific scene. Based on their feedback, we associated all measured physiological signals with emotional labels, and trained our engine.

At the end of the experiment, the experimenter rated, on a minute per minute basis, the dominant emotion expected to be felt by the user in every movie that was presented. This rating was then used to compare with the results of the recognition.

4.6.1 Biosignal Capture

In order to analyze participants' physiological data, biosignal recording biosensors were used for measuring Galvanic Skin Response (GSR), Respiration (Resp) and Electrocardiogram (ECG) which were responsible for users' biosignals recording and signal processing pipeline, each sampled at 256 Hz, since it is the required sampling rate which prevents signal aliasing or distortion (Oliveira, Benovoy, Ribeiro & Chambel, 2011). These sensors were specifically chosen as they record the physiological responses of emotion, as controlled by the autonomous nervous system. Moreover, previous studies demonstrated that these sensors allow for measuring our five basic emotions (happiness, sadness, anger, fear and disgust) (Maaoui, Pruski & Abdat, 2008). As discussed in section 2.5.1, the evaluation of specific measures of these three sensors provides categorical information about emotions. For example, ECG is related with anger while respiration plus ECG gives feedback about fear and sadness. Skin conductance plus respiration informs about happiness and disgust, while skin conductance on its own can vary with fear and anger. From the heart, we measured heart rate, heart rate acceleration, and heart rate variability. We also measured the first and second derivative of the heart rate. For respiration, we measured the rate, the amplitude and the first and second derivative of the rate. For GSR, we measured the stats plus the number of Skin Conductance Responses (SCRs).

4.6.2 Trainning Phase - Classification of Video Segments

A Hanning Window low pass filter was used to eliminate high-frequency components of the signals that are considered to be noise (Oppenheim, 1989). For the GSR signal a cutoff frequency of 2.0 Hz was used to filter out the recording noise as the GSR response waveform is approximately 1 Hz, whereas for the ECG and respiration signals, cutoffs of 128 Hz and 10 Hz were used in the filters, since frequencies outside that spectral ranges are considered noise (Rangayyan, 2002). To account for inherent physiological differences between participants, the mean of the 3-minute silent baseline data preceding each stimulus onset was subtracted from the active data and the signal range was adjusted to a [0-1] interval.

We extract six common statistical features from each type of the filtered biosignals of size N (Xn, n $\in$ [1...N]) where N is the interval of time of the signal in seconds and Xn is each one of the signals. We measured (1) the filtered signal mean, (2) the standard deviation of the filtered signals, (3) the mean of the absolute values of the first differences of the filtered signals, (4) the mean of the absolute values of the first differences of the normalized signals, (5) the mean of the absolute values of the second differences of the filtered signals, and (6) the mean of the absolute values of the second differences of the normalized signals. First derivatives give the measure of velocity (speed) of the signal, while second derivatives give a measure of the acceleration of the process we are measuring. These derivative measures approximate signal gradients, while the z-score normalized signals (4) and (6) may remove day-to-day differences between a participant's physiological response that can be affected by hand-washing or sensor placement.

1. Filtered signal mean:

$$\mu_X = \frac{1}{N}\sum_{n=1}^{N} X_n \qquad (1)$$

2. Filtered signal standard deviation:

$$\sigma_x = \left(\frac{1}{N-1}\sum_{n=1}^{N}(X_n - \mu_X)\right)^{1/2} \qquad (2)$$

3. Filtered signal mean of absolute value of the first difference:

$$\delta_x = \frac{1}{N-1}\sum_{n=1}^{N-1}\left|X_{n+1} - X_n\right| \qquad (3)$$

4. Z-scored normalized signal mean of absolute value of the first difference:

$$\tilde{\delta}_x = \frac{1}{N-1}\sum_{n=1}^{N-1}\left|\tilde{X}_{n+1} - \tilde{X}_n\right| = \frac{\delta_x}{\sigma_x} \qquad (4)$$

5. Filtered signal mean of absolute value of the second difference:

$$\gamma_x = \frac{1}{N-1}\sum_{n=1}^{N-1}\left|X_{n+1} - X_n\right| \qquad (5)$$

6. Z-scored normalized signal mean of absolute value of the second difference:

$$\tilde{\gamma}_x = \frac{1}{N-1}\sum_{n=1}^{N-1}\left|\tilde{X}_{n+1} - \tilde{X}_n\right| = \frac{\gamma_x}{\sigma_x} \qquad (6)$$

Where $\tilde{X}_n$ represents the z-score normalized signal (zero-mean, unit variance):

$$\tilde{X}_n = \frac{X_n - \mu_x}{\sigma_x} \qquad (7)$$

A total of 6 x 3 = 18 features are computed from the three types of biosignals. These features were chosen to cover the typically measured statistics in physiological recordings.

For the classification engine, we implemented linear discriminant analysis (LDA) since previous experiments with physiological data have shown that LDA produces high classification rates. LDA was selected to dimensionally reduced data, by building a statistical model for each emotional class and then cataloguing novel data to the model that best fits. We were thus concerned with finding which classification rule (discriminant function) best separates the emotion classes. LDA finds a linear transformation Φ of the x and y axes that yields a new set of values, providing an accurate discrimination between the classes. The transformation thus seeks to rotate the axes with parameter *v*, so that when the data is projected on the new axes, the difference between classes is maximized. We specifically chose statistical features, as these are computationally easy to produce, which opens the way to future real-time systems, choosing only classifier-optimal features, followed by Fisher dimensionality reduction.

4.6.3 Testing Phase – Classification of Full Movies

To analyze unclassified physiological data and recognize an emotional pattern support vector machine (SVM) and k-Nearest Neighbor (K-NN) classifiers (Bishop, 2006) were used and validated by the usage of specific movie scenes selected to induce particular emotions. These two techniques were selected in order to compare results. We employed digital signal processing and pattern recognition, inspired by statistical techniques used by Picard et al. (2001) due to its good results in testing with one person, in particular in our use of sequential forward selection (a variant of sequential floating forward selection). In fact, the greedy sequential forward floating selection (SFFS) algorithm was used to form automatically a subset of the best n features from the original large set of m (n < m). SFFS starts with an empty feature subset and, on each iteration, exactly one feature is added. To determine which feature to insert, the algorithm tentatively adds to the candidate feature subset one that is not already selected and tests the accuracy of a k-NN classifier built on this provisional subset. A feature that results in the highest classification accuracy is permanently included in the subset, while a poor feature is deleted. The process stops after an iteration where no feature additions or deletions cause an improvement in accuracy. The resulting feature set is now considered optimal.

The SVM classifier generates parallel separating hyperplanes that maximize the margins between the participants' data, which has the effect of minimizing generalization error. Because SVMs are binary classifiers by nature, we used a one-versus-all decision strategy to perform multiclass classification. This technique decomposes C-class classification problem, i.e. the single multiclass problem into *c* binary classifiers where each classifier separates the rest of the remaining classes, and the one with the highest recognition confidence value assigns the final label (Duda, Hart & Stork, 2001). We trained the SVMs with a Radial Basis Functions (RBF) kernel for which the parameters were determined concurrently using an iterative grid selection technique that finds the best combination using the training error of the classifier as a performance metric. The SVM module outputs the emotion of the recognized pattern, along with a classification confidence value based on the distance between the feature vector of the probe and the hyper-margin of the closest participant. The k-NN classifier used here classifies a novel object by a majority of "votes" of its neighbors, assigning to this object the most common class amongst its *k* nearest neighbors, using the Euclidean distance as metric. It was found through experimentation that a value of k = 5 resulted in the best possible selected feature subset.

4.7 Evaluation and Discussion of the iFelt Recognition System

The results of the learning phase exemplified in Figure 20 illustrates the class clustering of five emotional states: happiness, anger, sadness, fear and disgust projected on the 2D Fisher space along with the SVM class-boundaries of one person. The Fisher Projection is a technique that allows the representation of multi-dimensional spaces into two dimensions by depicting all bonds into vertical and horizontal lines of a space. To evaluate the pattern recognition process, we compared the manual-labeled movies with the output of the automatically classified data from our system. This labeling was done in a consecutive blocks of one-minute length with the emotion that is supposed to evoke on a user (Oliveira, Benovoy, Ribeiro & Chambel, 2011). To compare with the output of the recognition system, which produced classification scores on consecutive five-second windows, the median score was computed over minute long segments of the classified data. At least two participants watched each of the eight movies, and eight users were classified by the system, in a total of sixteen movies.

With the SVM classifier, the overall average recognition rate was 69% (SD= 5.0), which represents a 49% improvement over random choice. Since we are classifying 5 classes, the random probability is 1/5 (20%), which is what the simplest classifier could do. Our k-NN classifier produced an overall average recognition rate of 47% (SD= 9.3).

The SVM classification score shows promise that the iFelt recognition system can be used to automatically evaluate human emotions. Two key positive aspects of the system emerged: (1) the use of easily computed statistical features, which can be used to develop real-time classification systems (Oliveira, Benovoy, Ribeiro & Chambel, 2011); and (2) a quite reasonable recognition rate, with only three sensors when compared with the works listed in section 2.3.2 In fact, we achieved an optimization of 18% when compared with the best recognition rate of emotions elicited by movies (Soleymani et al., 2009).

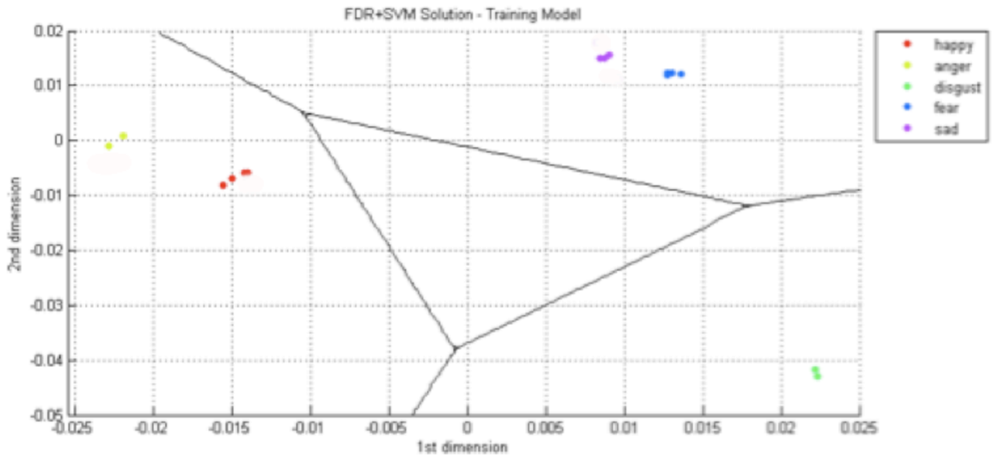

Figure 20 - Class clustering of five emotional states

Also, a larger range of categories was achieved since we obtained five and Soleymani et al. (2009) recognized only three. In the following section we focus on the work developed so far in what concerns the emotion annotation.

4.8 Emotion Annotation

There is a difficulty in what regards the selection of an appropriate annotation scheme, together with a proper set of semantic labels, to structure and identify emotions (Devillers, Vidrascu, & Lamel, 2005), because different application contexts demand for specific labels for emotional data annotation. As stated before, in section 3.4, some works already focused this issue along with other issues such as the development of semantic representations and descriptions of emotions in different contexts (Devillers, Vidrascu & Lamel, 2005; Douglas-Cowie et al., 2007).

In this section, we first propose a set of annotation requirements for emotion-oriented video applications in order to define the specific labels needs in this context and then we proceed by presenting a first semantic representation proposal based on XML, a standard language that can be used in multiple technological environments.

We started this task by carrying out the requirements analysis of our specific video applications and found some guidelines that we believe are important and should therefore be followed. In this section, we present such a set of classification considerations for emotion annotation related with video affective properties in order to help creating a mechanism for obtaining meaningful affective information from the perspective of viewers affective impact. Subsequently, we present an XML schema consisting of a set of semantic descriptors based on both user physiological signals, captured while watching videos (to properly characterize our classification method) and also based on video low-level features extraction.

4.8.1 Classification Requirements

We now discuss a set of classification considerations related with the representation of video affective properties in order to help creating a mechanism for obtaining meaningful affective information from the perspective of viewers and movies affective impact. Addressing the problem of the different annotation requirements, we established the most relevant aspects that must be included in a classification procedure, in order to enable the creation of emotion-oriented applications that include video, users and their affective relationship.

One of the first aspects to consider is the choice of a standard semantic representation for emotional information: in fact we believe that such a representation must possess the following characteristics (Oliveira, Ribeiro, & Chambel, 2010):

- It must be simple enough so that emotions can be captured from the diversity of existing emotional data gathering methods;
- It should be readable by any module of an emotion processing system;

- It should be structured in accordance with the W3C Emotion Markup Language guidelines (Schroder, Wilson, Jarrold & Evans, 2008) so as to enable the communication among future web services.

The second aspect to take into consideration is the importance of the movie collection, that should cover each and every basic high intensity emotion in order to ensure that we have, at least, one recognized emotional category. Additionally, a neutral state must also be included in the list of basic emotions, which corresponds to the occurrence of an emotionless moment in the video.

The third aspect concerns the information requirements of such a classification system. In fact, the classification of a user's subjective emotions, acquired both from physiological signal analysis and from user's manual classifications implies having classified at least the following information:

- The user's affective perspective for every video from emotional impact perspectives.
- The basic emotion for each scene in the video, in order to have a complete affective description of the whole video along its complete duration.
- An emotional profile constituted by all user's subjective emotional (user's emotional impact) classification based on physiological data and manual input of every movie.
- A video emotional profile constituted by all users' classifications.
- User emotional profiles are constructed over time by collecting and analyzing all emotional user data detected for each video scene.
- The two major emotional models (section 2.3) should be represented, if possible, in every classification of every scene, improving the access and interaction with emotional data.

As described above, we have a considerable amount of emotional data to process, which demands for suitable structures that allow the annotation of this information in videos and users' profiles. In the following section we describe the information requirements for such a classification system.

4.8.2 Information Requirements

The exploration and access of movies by their emotional dimensions involves a great amount of information. Here, we first present the information that needs to be structured in order to be accessible in an emotional perspective by users' point of view, videos and their relationships. And then we propose a semantic description of emotion oriented towards the user experience while watching videos, considering the user implicit assessment (by acquiring physiological signals) and the user explicit assessment (manual classification). In this way, we gather information about users, videos and their relationships (user-video): Table 8, Table 9 and Table 10 detail such information.

Table 8- Emotional information: User

Subjective Emotions (User Felt Emotion)	Physiologic Sensors	Most felt category /valence
		Less felt category /valence
		Most recent category/valence

Table 9 - Emotional information: Video

Subjective Emotions (User Felt Emotion)	Physiologic Sensors	All users dominant emotions (Mean | SD)
		All users most felt emotions per scene (Mean | SD)
		All users dominant emotions along a movie (Mean | SD)
		All users valences per movies and scenes (Mean | SD)
		All users intensities per movies and scenes (Mean | SD)
	User self-Assessment	All users self assessment about felt dominant emotion (Mean | SD)
		All users self assessment about felt intensity of dominant emotion (Mean | SD)
		All users self assessment about Felt valence (Mean | SD)
		All users preferences about the movie (0-9) (Mean | SD)

In each table, we represent the subjective emotional information and we explicitly indicate the way used to obtain such information – either through the classification of the physiological signals or through the manual input from users. Some studies use movies' emotional information to automatically summarize videos according to the emotional impact (Soleymani, Chanel, Kierkels & Pun, 2009), or to detect video emotional highlightings from physiological signals (Smeaton and Rothwell, 2009). This novel compilation of information presented in tables 8, 9 and 10 is focused on classification and annotation to explore and share movie impact on users. Thus, we structured it, in a XML file, by defining a schema and using the Emotion ML specification in order to make this information organized and sharable. Such a structure is presented in the following section.

Table 10 - Emotional information: User-video relationship

Subjective Emotions (User Felt Emotion)	Physiologic Sensors	Most felt category /valence
		Less felt category /valence
		Category felt each 5 seconds
		Global percentage of categorical/valence
		Main variances between emotional states (in every movie and all over the time)
	User self-assessment	Dominant felt emotion all over the movie
		Intensity of the felt dominant emotion
		Valence felt
		Intensity of the valence felt
		Preference about the movie (0-9)

It is important to note that there is also non-emotional information that we suggest should be included as meta-information about videos. The other types of information that we believe are also needed can be gathered through a standard API from online services like IMDb. More specifically, we are considering using the title of the video, a link to the IMDb description, the year of release, a URL to the cover, users' ratings registered in IMDb, the first of the directors, the genres, plot, and information on all the actors (or the ones selected) listed in the main title page (name, link to actor page, link to picture and character). Also interesting is to include the movie reviews from the New

York Times[12] critics that can be gathered to complete the movie profile general information.

After determining the information requirements, we then defined a data structure such as the one discussed in the following section.

4.8.3 Emotion ML Structure for iFelt

The Emotion Markup Incubator Group proposed, in May of 2012, the first W3C Candidate Recommendation of Emotion ML. In this research work, we followed this recommendation for the classification of iFelt movies, their users and the relationships users/movies. Emotional ML, being an XML based language, standardizes the representation of emotional data, allowing the re-use and data-exchange with other technological components, following the guidelines we suggested above for emotional semantic representations. This information is linked as metadata to every video in our system (see Figure 18) but it was still not used or even tested for the purposes of this thesis.

In order to have our emotional labels well specified, we created a vocabulary, as Emotion ML guidelines suggest. For that reason, inside the main vocabulary file of Emotional ML we inserted our set of emotions, like the ones in Figure 21. The vocabulary element has two main attributes: a) type, to specify the type of emotions between the traditional emotional models (see section 2.3) and b) id, to identify the list of emotions to be used in external files.

We have created three Emotion ML files and an XML schema to ensure that our data structure can be divided into a user emotional profile, a movie emotional profile and a user-movie classification. This division further enables the reuse of the information of just the user, just the movie emotional information, or just the relationship among both of them, or even the three simultaneously.

The XML schema we created has, as its main objective, to structure the information

[12] http://developer.nytimes.com/docs/movie_reviews_api

about users, movies and user-movies relationships. It also includes additional information that our system requires, such as last emotions, lists of users, lists of movies and users' preferences.

```
<vocabulary type="category" id="iFeltEmotions">
        <item name="happy" />
        <item name="amused"/>
        <item name="involved"/>
        <item name="inspired"/>
        <item name="tender"/>
        <item name="surprised"/>
        <item name="astonished" />
        <item name="curious"/>
        <item name="melancholic"/>
        <item name="bored"/>
        <item name="compassioned"/>
        <item name="fear"/>
        <item name="disturbed" />
        <item name="amused"/>
        <item name="scared"/>
        <item name="disgust"/>
        <item name="embarrased"/>
        <item name="anger"/>
        <item name="irritated" />
</vocabulary>
```

Figure 21 - IFelt emotional vocabulary

As an example, the following code snippet (Figure 22) illustrates the definition of the movie information. The whole specification is in the appendix B of this thesis.

```
<xsd:element name="MovieInformation" type="xsd:string"/>
        <xsd:complexType>
            <xsd:sequence>
                <xsd:element name="MovieId" type="xsd:id" use="required"/>
                <xsd:element name="Title" type="xsd:string" use="required"/>
                <xsd:element name="Director" type="xsd:string" use="required"/>
                <xsd:element name="Country" type="xsd:string" use="required"/>
                <xsd:element name="Year" type="xsd:string" use="required"/>
                <xs:element name="Actors" maxOccurs="unbounded">
                  <xs:complexType>
                    <xs:sequence>
                      <xs:element name="name" type="xs:string"/>
            </xs:sequence>
        </xs:complexType>
        </xs:element>
    </xsd:sequence>
        </xsd:complexType>
```

Figure 22 - iFelt schema: Movie information definition

In this example, we have, as an XML element, the tag *MovieInformation* that includes a *movieId*, a title, a director, a country, a year and a list of actors. This tag is going to be used in the movie Emotion ML file inside the *info* tag (Figure 23), which is the way

Emotion ML allows us to introduce metadata into their files. We also have a list of users related to this specific movie, which can easily inform us about each user classification for this movie. The second part of the movie.xml file is concerned with the emotion classification (Figure 24).


```
<info>
    <iFelt:MovieInformation>
        <iFelt:MovieID>MMMMMM</iFelt:MovieID>
        <iFelt:Title>Cashback</iFelt:Title>
        <iFelt:Director>Sean Ellis</iFelt:Director>
        <iFelt:Year>2009</iFelt:Year>
        <iFelt:Actors>
            <iFelt:Actor>Tom B</iFelt:Actor>
            <iFelt:Actor>Ann A</iFelt:Actor>
        </iFelt:Actors>
    </iFelt:MovieInformation>
</info>

<info>
    <iFelt:ListUsers>
        <iFelt:UserID>XXXXXX</iFelt:UserID>
        <iFelt:UserID>XXXXX0</iFelt:UserID>
        <iFelt:UserID>XXXXX1</iFelt:UserID>
        <iFelt:UserID>XXXXX2</iFelt:UserID>
        <iFelt:UserID>XXXXX3</iFelt:UserID>
        <iFelt:UserID>XXXXX4</iFelt:UserID>
    </iFelt:ListUsers>
</info>
```


Figure 23 - iFelt Emotion ML specification part 1: movie.xml


```
<emotion expressed-through="text">
        <category name="curious" value="0.5"/>
</emotion>

<emotion expressed-through="biosignal">
        <category name="happy"/>
</emotion>

<emotion expressed-through="text">
        <dimension name="arousal" value="0.3"/>
</emotion>


<emotion expressed-through="text">
        <category name="involved" value="0.4"/>
</emotion>

<emotion expressed-through="text">
        <category name="happy" value="0.3"/>
</emotion>

<emotion expressed-through="text">
        <iFelt name="inspired" value="0.2"/>
</emotion>

<emotion expressed-through="text">
        <category name="sad" value="0.18"/>
</emotion>

<emotion expressed-through="text">
        <category name="melancholic" value="0.1"/>
</emotion>

<emotion expressed-through="text">
<dimension name="arousal" value="0.3"/>
</emotion>

</emotionml>
```


Figure 24 - iFelt Emotion ML specification part 2: movie.xml

In Figure 24, the *emotion* tag has only the global classifications of a movie including all forms of classification: manual for both iFelt emotional labels and the dimensional model measures (arousal, valence), and automatic by biosensors. The attribute *expressed-through* allows specifying how information was gathered. We only presented an example for each case and also for an iFelt vocabulary.

Every file header includes the Emotion ML namespace, the reference to the vocabularies to be used, as well as the schema created (see Figure 25). The user's information has the same structure of the movie's information (see Figure 26), being the first part about the iFelt information requirements: 1) User general information, as name, number of movies watched, number of neighbors and country; 2) Last Emotions, presents a list of movies and the associated emotion; and 3) List of Movies that this user watched.

```
<?xml version="1.0" encoding="UTF-8"?>
<emotionml xmlns="http://www.w3.org/2009/10/emotionml"
           category-set="http://www.w3.org/TR/emotion-voc/xml#big6"
           appraisal-set="http://www.w3.org/TR/emotion-voc/xml#scherer-appraisals"
           dimension-set="http://www.w3.org/TR/emotion-voc/xml#scherer-dimension"
           iFelt-set="/Users/oliveira/Dropbox/_Tese/XMLImplementation/vocabulary.xml"
           xmlns:iFelt="/Users/oliveira/Dropbox/_Tese/XMLImplementation/iFelt.xsd">
```

Figure 25 - iFelt XML headers

The second part, related to user information, also presents global classifications of all forms of classification but from a user's perspective: manual, for both iFelt emotional labels and the dimensional model measures (arousal, valence), and automatic, through the usage of biosensors as illustrated in Figure 26. The third file, *user-movie.xml*, specifies every emotional classification (categorical or dimensional) along the movie, acquired through biosensors, and the final classification manual, which was manually inserted by the user. Thus, besides the general information about users and movies, we have the emotional classification as illustrated in Figure 27.

```xml
<info>
  <iFelt:UserInformation>
  <iFelt:UserID>XXXXXX</iFelt:UserID>
  <iFelt:name>MaryFond</iFelt:name>
  <iFelt:NumberMovies>123</iFelt:NumberMovies>
  <iFelt:NumberNeighbors>34</iFelt:NumberNeighbors>
  <iFelt:Country>Portugal</iFelt:Country>
  </iFelt:UserInformation>
</info>

<info>
  <iFelt:LastEmotions>
    <iFelt:Position>1</iFelt:Position>
    <iFelt>Curious</iFelt:Emotion>
  <iFelt:MovieTitle>24 Hour Party People</iFelt:MovieTitle>
  </iFelt:LastEmotions>
    <iFelt:LastEmotions>
    <iFelt:Position>2</iFelt:Position>
    <iFelt:Emotion>Involved</iFelt:Emotion>
  <iFelt:MovieTitle>Black Swan</iFelt:MovieTitle>
  </iFelt:LastEmotions>
    <iFelt:LastEmotions>
    <iFelt:Position>3</iFelt:Position>
    <iFelt:Emotion>Tender</iFelt:Emotion>
  <iFelt:MovieTitle>Breakfast in Pluto</iFelt:MovieTitle>
  </iFelt:LastEmotions>
</info>
<info>
  <iFelt:ListMovies>
    <iFelt:MovieID>MMMMMM</iFelt:MovieID>
    <iFelt:MovieID>XXXXX1</iFelt:MovieID>
    <iFelt:MovieID>XXXXX2</iFelt:MovieID>
    <iFelt:MovieID>XXXXX3</iFelt:MovieID>
    <iFelt:MovieID>XXXXX4</iFelt:MovieID>
  </iFelt:ListMovies>
</info>
```

Figure 26 - iFelt Emotion ML specification part 1: user.xml

```xml
<info>
    <iFelt:MovieInformation>
    <iFelt:MovieID>MMMMMM</iFelt:MovieID>
        <iFelt:name>Cashback</iFelt:name>
        <iFelt:Director>Sean Ellis</iFelt:Director>
        <iFelt:Year>2009</iFelt:Year>
        <iFelt:Actors>
            <iFelt:Actor>Tom B</iFelt:Actor>
            <iFelt:Actor>Ann A</iFelt:Actor>
        </iFelt:Actors>
    </iFelt:MovieInformation>
</info>


<emotion  expressed-through="biosignal">
<reference uri="movieId.avi#t=105,110"/>
<category name="happy"/>
</emotion>

<emotion  expressed-through="biosignal">
<reference uri="movieId.avi#t=115,120"/>
<category name="curious"/>
</emotion>

<emotion expressed-through="text">
<category name="happy" value="0.7"/>
</emotion>

<emotion expressed-through="text">
<category name="curious" value="0.3"/>
</emotion>

<emotion expressed-through="text">
<dimension name="arousal" value="0.3"/>
</emotion>

<emotion expressed-through="biosignal">
<dimension name="valence" value="0.9"/>
</emotion>
```

Figure 27 – iFelt Emotion ML specification part 2: user-movie.xml

Thus, Figure 27 shows three types of emotional information about a user and a specific movie. The first two *emotion tags* specify, emotionally, a movie in 5 second range intervals by using a relative time, following Emotion ML specification. Clipping the movie by time, is denoted by the name *t*, and specified as an interval with a begin time and an end time. In this case, we have an example from the second 110 until the second 115, with an emotional classification of "happy", gathered through biosensors. Manual input can be specified using the *text* value for the *expressed-through* attribute. In fact, this value can be used for both categorical and dimensional (valence, arousal) classifications. In this example, the dimensional classification value represents the intensity of the dimension, being 0.3 a low-than-average arousal, and a 0.9 value a very positive valence. Having this information annotated and available, we believe it is now easier to develop an interface with other systems that may wish to use this information. Affective information of users acquired while watching videos has specific data requirements to be represented. Having video emotionally classified will further allow for finding interesting emotional information in unknown or unseen movies, and it will also allow for the comparison among other users' reactions to the same movies, as well as comparing directors' intentions with the effective emotional impact on users, and also analyze users' reactions over time. This classification can subsequently contribute to the specification of a software tool, which enables the personalization of TV guides according to individual emotional preferences and states, and recommend programs or movies based on those preferences and states.

4.9 Summary

In this chapter we presented the iFelt system, an experimental environment with a special emphasis on its emotional recognition component, capable of automatically recognizing a set of human emotional states using psychophysiological measures and pattern recognition techniques based on discriminant analysis and support vector machine classifiers.

We have also analyzed the information requirements for creating a semantic representation of a system based on the exploration of the emotional components of movies, regarding their implicit impact on users and in the affective categorization of content by users perspective (manual input). In the following chapter, we detail our proposal for visual representation and interaction mechanisms and report on preliminary usability evaluations.

Chapter 5

An Empirical Study on Visual Exploration of Emotions on Movies

5.1 Introduction

The representation of emotional aspects in everyday applications is quite uncommon. Movies are by definition one of the most emotional media, because of their "impression of reality" that turn users more empathetic and vulnerable to their content; Movies can indeed affect viewers' emotions and perceptions. (Metz, 1974; Plantinga, 2009). Their impact is obviously related to a diversity of individual differences such as historical periods, viewers' pleasures, desires, affects and moods (Plantinga, 2009). Whenever people think of motion pictures, they often remember the higher emotional impact movies and then they remember those scenes that turned that particular movie so unforgettable. But sometimes, even those scenes become impossible to recall, among hundreds or even thousands of movies people watch in a few years. Thereof, in this work we tried to collect and present to users a set of emotional characteristics about movies, describing the feelings of the person who watched a particular movie, along with their movie profile, and we also created some visual interface mechanisms to help in exploring such emotional information. Specifically, we developed a system –iFelt – that consists of an interactive web video system designed to learn users' emotional patterns and make use of this information to create emotion based interactions.

5.2 iFelt User Interface Description

As stated above, our experimental system – iFelt – is an interactive web video application that allows for visualizing, searching, accessing, exploring, recommending and cataloguing users' emotional information about movies. It was 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 focused our analysis on movies. In this chapter, we first describe the system main goals and then describe the assumptions that served as the basis for the design of the user interface for our experimental system.

5.2.1 User Interface Goals

The iFelt user interface has the following five main goals:

1) To visually represent movies by emotions in such a way as to facilitate user understanding of the inner concept, i.e., users should easily understand the graphical symbolisms used in the system's interface and its interactive mechanisms.
2) To enable accessing, searching or finding movies based on emotions felt by users, i.e., the system should provide novel, perceptive and simple mechanisms in order to enable users to access movies by emotional characteristics.
3) To facilitate browsing millions of movies from an affective perspective, i.e., the system should be designed in such a way that users can browse lists containing large numbers of movies and visually detect those who are more relevant given a particular emotion, always having the perception of the emotions that categorize each movie and their intensity.
4) To enable recommending movies by analyzing users with similar movies playlists and emotional classifications.
5) To be easy, useful and fun to use and designed to follow established usability guidelines.

We focused on these five goals to create novel representations and mechanisms for our experimental system in order to allow us to analyze their usefulness and even the engagement caused on users by providing them with information about the emotional impact of movies in their viewers. As we are considering users' emotions when watching movies, obtained both automatically and manually, we also intended to explore the relationships among users and new ways to emotionally explore movies.

To evaluate our proposals, we performed user interface usability tests in order to assess our choices, functionalities, the ease-of-use, usefulness and satisfaction with the system. Moreover, we also intended to evaluate the perception of users about our design choices. Before beginning the description of the user interface, we first introduce the assumptions we took into account for designing the interface in what regards the selection of emotions used to classify movies in this system.

5.2.2 Design Assumptions

5.2.2.1 Types of Emotions

Our first assumption is that a movie, or a movie scene, can be searched by the *subjective emotion* that it induced in a user, e.g. sadness, because the user might relate that specific kind of situation with a sad event in her life. A movie can also be searched by the *objective emotion* that is conveyed, e.g. a video or a scene showing happy people. For example, if in a sweet romantic scene, the user felt sad and classified it as *sad*, we then tag that scene as a *sad* one in the user profile, while the objective emotion it conveyed could be tenderness. Also, for the categorization of movies by subjective emotions, we considered automatic and manual classifications. The automatic process is based on the biometric methods described above in Chapter 4 and the manual mechanism is based on the input of users as described below in section 5.7.1.

5.2.2.2 Selection of Emotions

Another assumption we took into account is related with the emotional labels we used. The emotional labels used in the automatic process correspond to the categorical labels: happiness, sadness, fear, anger and disgust. This choice is based on the facts that (1) the characterization of emotions by physiological patterns still has some limitations regarding its differentiation (Rainville et al., 2006); and (2) the Ekman's basic categorical emotions are the most agreed upon. So, despite the importance of automatic recognition of valence and arousal dimensions, we are not addressing this automatic recognition in the context of this thesis. However, we are considering the support of the dimensional and appraisal models in the manual classification mechanism of the iFelt system. Respecting the manual emotional labels, we introduced "surprise" to complete the basic emotions and added 13 additional labels besides the basic ones, due to the fact that those basic six emotions are too narrow to properly characterize the emotional complexity of movies. In our study regarding viewers' attitudes, awareness and preferences about the emotional impact of movies (Chambel, Oliveira & Martins, 2011), it was interesting to note some tendency in preferences of emotions like surprise, fun, feeling good, happiness, and mostly, imagining, dreaming, inspiration and motivation, towards the search of engagement and meaning beyond positive emotions, and suggesting the need to address wider models of emotional impact. Having followed these results and the literature both from psychology, based on the work of Scherer, (2005), and from movie eliciting studies, based on the work of Gross and Levenson (1995) and Rottenberg et al. (2005), we chose, for the second case study, the additional 13 labels that in our opinion best describe movies. Based on the 36 affect categories from all those studies we reviewed, we then created 6 main categories which correspond to the 6 Ekman's basic emotions (happiness, surprise, sadness, fear, disgust, anger) and a set of labels to specify each one that fits in each emotion. We now present the full list of our proposed emotional labels to categorize movies:

1. Happiness: involved, amused, inspired, tender.
2. Surprise: curious, astonished.
3. Sadness: melancholic, compassioned, bored.
4. Fear: scared, disturbed.
5. Disgust: embarrassed.
6. Anger: irritated.

We decided to provide more emotional labels for the "Happiness" category because there are more negative than positive emotions in our list of basic emotions.

5.2.2.3 Types of Classification

Another important note about the classification, and specifically about the kind of classification we applied, automatic and manual, must be explained. As we stated in the description of this work, and following this thesis goals, we are exploring ways to present emotions gathered through both automatic and manual methods. In the second case study, we opted to have more emotions than the automatically recognized, which left us with the problem of determining what or who classified a movie: a sensor or a user. In fact this raised the problem that movies can have a sad narrative all along the movie, but in the end viewers may have the sensation of inspiration. So, in the second interface we assumed that timelines should be constructed through automatic methods, with basic emotions, and the user may subsequently change them if not in agreement with the automatic result, changing it to any other emotion available in the system. This procedure is however not included in this work. But when we ask users to classify a movie as a whole, it is their final and global appreciation, which can be different from the one obtained from the automatically generated timeline.

5.3 Design Rationale

We based our design choices in usability heuristics, especially those concerning an aesthetic and minimalist design, familiarity and low cognitive workload, with the aim to design an interactive system that users would find useful, satisfying, easy to use and perceptive. As described in more detail in section 5.2.1, we also wanted users' affective assessment of the interface in order to evaluate its fun, engagement and motivation factors. Thus, we performed two user interface studies, fully detailed in sections Case Study I and Case Study II.

The first study was conducted with the goal of testing some of the main features of our work and to assess the main issues regarding the proposed content. The second user interface case study addressed some of the problems identified in the results of our first evaluation, exploring the visual representation of huge amounts of movies, extended selecting and browsing methods based on more sophisticated filters and searches, thus, incorporating all functionalities proposed in this thesis.

On both cases, we created many user interface elements for representing emotions in different functions and meanings. Thus, they can represent, among others, the dominant emotion in a movie, the movie emotional timeline, a set of emotional scenes in a movie and users' emotional profiles. These representations were inspired in the different representations of emotions used in the emotional models we reviewed that were described above in Chapter 3.

Also, on both cases, we have adopted round shapes and circular organizations inspired by the Geneve Emotion Wheel (GEW), organized by valence and arousal. In fact, we were also inspired by the work of Norman (2004), and used rounded shapes and smooth or symmetrical objects, as some of the interface characteristics that may induce positive states.

Regarding the "funology" of our system, we intended to provide some additional features hoping that users would find them enjoyable, inspired by the statement of Schneiderman (2004) that users should be more engaged and pushing affect and emotions with fun features, which he considered to be: alluring metaphors, compelling content and attractive graphics. Based on these assumptions, we decided to use the album metaphor as the preferred way of organizing the videos, given that the album is the traditional means where people collect personal and favorite photos, and people, thus, tend to develop an affective connection with it.

5.4 System Functionalities

This section describes all the system's functionalities organized by their emotional goals.

5.4.1 Visual Representation of Movies by Emotions

At the individual movie level, emotions may represent movies through the following visual features:

1) The movie space, where movies are visually distributed in a surface, ordered or filtered, and identified by their dominant emotion.

2) The most dominant emotion. Each movie has a signal representing the main emotion a specific user felt, or the weighted mean of emotions felt by all the users registered in the system, as well as their neighbors.

3) The dominant emotions, represented by the percentage of dominance of each emotion in the movie. This information can be presented in each user's profile and in the movie's profile.

4) The emotional timeline, which represents the emotional scenes along time, in a sequential form, in both users' and movies' profiles. Each user can also have the possibility of comparing his or her timeline with any users' or neighbors' timelines.

5) The current emotion, which represents the emotion felt in a particular time of the movie play.

6) A set of movie emotional scenes, which is a non-sequential representation of the emotional scenes contained within each movie, and can be represented in every movie profile, according to a specific user.

5.4.2 Movie Search and Access by Emotions

The iFelt user interface presents a set of mechanisms to search, access and find movies at an individual level and at a movie album level.

Regarding the **search**, users have the possibility of typing search terms on a typical text input box existing in every page of the user interface. Users may search for emotions, or for movie specific information, such as title, director and cast. Whenever they search for movie genres or movie specific information, they can then select an additional emotional criterion to complement the initial search terms. Another way to search is by using our novel search wheel of emotions based on the Geneve Emotion Wheel (GEW) (Scherer, 2005), where users can select a set of emotions and their intensity to find movies according to their preferences.

Regarding the **access** to movies, search results can be presented in two different ways: a) a wheel organized either by time or intensity, depending on the user's preference, where a user can access a collection of hundreds of movies per page, represented by colored circles with different sizes, depending on emotion and the chronology or the intensity; b) a typical list of movies with the information about the dominant emotion and its corresponding intensity. Users can also access movies by visiting their albums or their neighbors' albums. For this purpose, we provided a menu item called "Movies" in every page of the user interface that leads users to a searching page which is another way to access movies. Finally, there is also a recommendation page where users can find movies similar to another one, based on the emotions, which is yet another way to access movies based on emotions.

We distinguish the term "search" from the term "find" by the explicitly of the act of searching for movies, and the unexpectedness of finding movies that users never suspected they existed. Regarding the finding mechanism we propose recommendation features based on: a) the movies a user didn't watch from the list of the most similar neighbors, and, whenever reasonable, b) when a user is searching for emotions, we present primarily those movies in our results.

5.4.3 Browsing Large Quantities of Movies from an Affective Perspective

Considering the IMDB database, there must be more than two millions movies according to IMDB statistics[13]. Therefore, in order to browse, or navigate, in a perceptive and comprehensive way, we also needed to provide a browsing mechanism based on emotions. For example, when a person looks for drama movies in IMDB from the year of 1992, the system returns 6.039 titles. If a person wants to have a global perception of user ratings for those thousands of movies the system provides a list of 50 movies per page. Thus, ordered lists by any kinds of filters is a possible solution that we also adopted for our system, but we created a novel navigation strategy to show hundreds of movies in a single view, based on emotions ordered by time or intensity. By default, our wheel can represent more than 250 movies that, depending on the context of the search, can reflect the last 250 movies watched or categorized by emotion, or the last 250 movies watched from the strongest intensity and categorized by emotion. If, for example, the user is looking for "fear" related movies, this wheel shows a set of 250 movies sorted by the intensity of fear, or by the chronology of the movie collection, from the most recent ones to the older ones. This logic appears in every type of search such as the one described below in section 5.6.2.

5.4.4 Recommending Movies by Analyzing Emotional Neighbors

Even though we do not explore recommendation techniques in this thesis, we present user interface elements based on the knowledge that big efforts are being made to turn recommendation as accurate as possible. Some common methods applied to movies, such as collaborative filtering, are based only on the activity of users and content-based methods based on metadata information (Pilászy & Tikk, 2009). In 2006, Netflix launched a competition trying to improve its movie recommendation features to its consumers. The winner's algorithm used matrix factorization methods, which basically characterizes both items and users using vectors of factors inferred from item rating patterns (Koren, Bell & Volinsky, 2009). As stated above, we do not intend to explore those algorithms but, instead, to present novel interface elements that, based on those methods, allow us to recommend movies based on emotions. Although the Netflix

[13] Available online at: http://www.imdb.com/stats

algorithm also uses implicit information such as demographic information from users and the time in which the feedback was given, in our system we consider **emotional neighbors:** users that have a great percentage of movies similar to a given user, and whose emotional classifications are also similar in the type of emotion and in the same order of intensity. Given that the chronologic information is always present, users can use that information to choose a film to watch. Users that have 100% of similarity in the list of movies, but show completely different emotional classifications, are not considered as neighbors, but are still useful in the way that a user can see what others have felt, by contrast.

As such, we propose a recommendation system based on the following information:

1. By emotional neighbor:
   a) A list of movies the user didn't watch – this list might recommend movies that would fulfill wanted emotional states.
   b) A list of the most similar movies – can provide the user with an idea about the kinds of movies watched by people with similar interests.
   c) A list of the most different classification movies – this list provides information about the kinds of movies that are in disagreement with neighbors.
2. By movie: we suggest movies that have the same categorized emotions as the proposed movie, giving a heavy weight to those user neighbors' movies that match the criteria, and then to movies that, in the system, match the emotional profile of the first proposed movie.
3. By emotions: we recommend movies by giving a heavy weight to those user neighbors' movies that match the criteria and that the user didn't watch, and then to movies that match the user profile.

Users emotional profiles are constructed over time by collecting and analyzing all emotional user data detected for each video allowing the system to compare profiles.

5.5 User Interfaces Studies and Usability Evaluation

We developed the iFelt system in two phases. The first phase consisted in the development of a prototype in order to test, through a usability evaluation, how useful and perceptive it would be to have a system that explores movies by affective perspectives. In the second phase we developed a prototype using the feedback we collected from the first usability evaluation, adding new methods to explore, access and visualize movies based on emotional aspects.

5.5.1 Usability Goals

We now proceed to present two usability studies of user interfaces regarding the exploration of emotions on movies, the users that watched those movies, and the mechanisms to search and find as well as the recommendation strategies. These studies addressed the evaluation of the usefulness, satisfaction and ease-of-use, but also the joy when using the iFelt system. In this section, we first present our experimental objectives and then we describe in detail the evaluation method followed to assess the system's functionalities, goals and assumptions.

For this purpose, the design of our system's interface took into account usability guidelines and user experience (UX) concepts. We followed the definition of usability given in ISO 9241-11 standard, which states that usability is about "effectiveness, efficiency and satisfaction in a specified context of use". In this definition, effectiveness refers to the accomplishment of the task, efficiency to the way users perform each task. We also tried to assess the user experience (UX) in the perspective of how fun it is to use our system. We apply the term UX because we agree with Nielsen Normal Group's definition of user experience when they argue that UX are services or products "that are a joy to own, a joy to use" besides some other characteristics. Nevertheless, we want to make clear that we did not intend to make the assessment of the user experience our first priority, we just wanted to have some feedback about the joy and fun on using and exploring the functionalities we designed for our experimental system.

Regarding the usability evaluation, we have defined a set of questions based on the USE model (Lund, 2001) that measures the utility, satisfaction and the ease-of-use of a user interface. The usability dimensions underlying the USE questionnaire meet the requirements and design evaluation we have defined for the iFelt user interface which include evaluating the following properties:

- Usefulness - How useful can the exploration of emotional information be? Which are the best ways to access this information? Were the results provided by the system perceived as useful?
- Satisfaction – Do users find iFelt fun to use? Do users have a good experience using it?
- Ease of use: Do users find iFelt easy to use?

In what the perceptiveness of the information provided by the system is concerned, as a supplement to the satisfaction and ease of use dimensions of this USE usability analysis, we included the following additional questions:

- Is the information representation easy to understand?
- Which was the best way to understand the information?

We also wanted to assess the entertainment aspects related to the user experience of: (a) having movies represented by the emotion they induce, (b) exploring and accessing this emotional information about movies and movie scenes, and (c) recommending movies based on other users profiles with emotional similarities to a given user's profile. For this purpose we included the following questions:

- Did you think it was fun and cool to use this application?
- Do you consider recommending it to other people?

Our observations helped on having feedback on how fun and engaging were users' interactions with the user interfaces developed for our system. A parallel objective was to determine if users had specific and/or global comments and suggestions in what concerns functionalities, access mechanisms or information representation alternatives to suggest.

5.5.2 Usability Evaluation Methodology

We performed the evaluation based mainly on two experimental methods: observation – using paper and a pen to register observations, and semi-structured interviews.

Participants were informed about the characteristics of the iFelt experimental system, the nature of the questions and the purpose of the test. Then, the interview session began with a set of questions to gain insights about user's movie watching background, to help us in the later analysis of the results. Next, users watched two scenes of the same movie outside the iFelt system, and were asked to evaluate their dominant emotion as a reaction (T0), to increase their awareness about the concept of emotional evaluation of movies and movie scenes.

The next step was a task-oriented activity with iFelt, where errors, hesitations and timings were observed. At the end of each task, users were confronted with USE-based questions, to be answered in a five point Likert scale, and they were given the opportunity to provide qualitative feedback through comments and suggestions for the iFelt features involved in that task. In this step, interviewers also wrote down their own judgment in terms of the USE aspects, by observing users performance in completing the task, to help detecting less thoughtful answers. In fact, satisfaction parameters can be more reliable when observed, because interviewers easily perceive users attitudes and postures in the interaction. In the task, whenever there was more than one version of the same feature, we randomly chose the order to present them in each test, to change the routine of the tests towards more equitable evaluation conditions of the different versions. Finally, users were asked to verbally answer to another USE-based questionnaire, this time about the overall application overview.

Regarding the recommended number of participants, it depends on the objectives and methods adopted. Following Nielsen (1994) or Virzi (1992) five to eight participants are sufficient to understand 80% of usability problems and design options. In fact, this was our main goal in these two qualitative evaluation, along with gaining insights about users acceptance, opinions and preferences as well as their whole experience when interacting with our system. Also regarding participants and again based on Nielsen, we focused on having the right type of participants. In our case, movies engage people from 8 to 80 years old, turning our main concerns to have computer literate and movie enthusiasts.

5.6 Case Study I

In our first case study we wanted to assess if users were sensible to this way of exploring movies, and if users would find it useful (Oliveira, Martins & Chambel, 2011). Our concerns in this phase were to decide which information to provide, how to provide it, and to create four main pages: the video space, the scene space, the user profile and the movie profile, where we implemented our proposed mechanisms to visually represent videos by affective dimensions and the different ways to access emotional properties of a movie and of users.

The emotional representation was inspired in the different representations of emotions used by the models described in section 2.3. We adopted a representation of emotions based on colors, a strategy similar to the one used in Plutchik's model (1980). Thus, emotions such as Anger, Disgust, Fear, Sadness, Surprise and Happiness, were represented, respectively, by the pink, purple, green, blue, light blue, and yellow colors, as in Plutchik's. These colors were also used in the different representations of the emotional aspects of the movies, in the current prototype design, where we also intended to consistently adopt round shapes and circular organizations inspired by the 2D circumplex representation of emotions from Russel's dimensional model (Russell, 1980).

In the following sections, we present the user interface of the first prototype organized by our main goals, which consisted of assessing: a) the visual representation of movies emotions in our system; and b) the movie search, access and finding mechanisms. The remaining two goals we defined for our system were only evaluated in the second case study.

5.6.1 Visual Representation of Movies by Emotions

Regarding the visual representation of movies by emotions we propose a movie space, scene spaces, most dominant emotions, the most dominant movie emotions, emotional timelines and the current emotion. We now describe in detail the objective for, and how we designed, each one of these representations.

5.6.1.1 Movie Space

The movie space is where the user obtains an overview for a set of somehow filtered movies existing in the system, along with information about their dominant emotions, in addition to the most traditional information such as the title, as well as other relevant information (see Figure 28). In this representation, the movies contained in the movie space selection are represented by a colored circle, in which the color corresponds to their dominant emotion, and placed on the wheel accordingly, close to the movies with the same dominant color, creating six sections corresponding to the six emotions included in our emotional model. In addition, the distance to the center represents the level of emotion dominance: the farther to the center, the more dominant that emotion is in the movie.

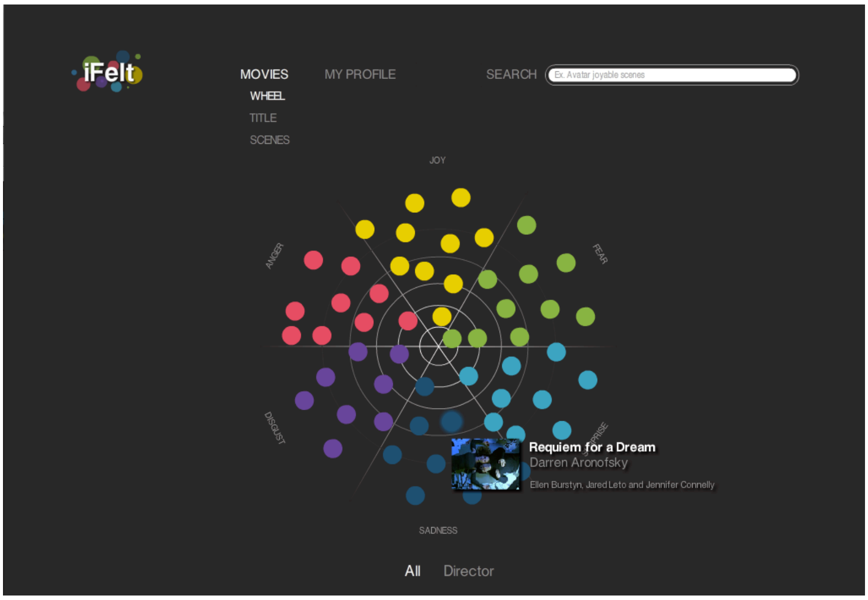

Figure 28 – Movie space - Emotional wheel

Furthermore, on mouse over, each circle shows the corresponding movie information (see Figure 28). This mechanism was designed to allow the visualization of the movies space organized by dominant emotions, in order to get a better perception of the dominance of emotions in the selected movies and also to help users finding and browsing movies based on these criteria.

In order to increase the perceptiveness and effectiveness of this representation, four versions of the wheel were designed and further tested, as depicted in (see Figure 29):

V1) Same brightness & size in emotion dominance + circular grid:

Each circle in the wheel has the same size and the same brightness for every emotion, independently of the level of emotion dominance. There is a grid in the background stressing the separation of emotional regions (in addition to the different colors used in each wheel section) and suggesting some gradient along the circle radius through concentric circles (V1 in Figure 29).

V2) Different size in emotion dominance + circular grid:

In this version, the size of the circles reflects the percentage of emotion dominance, in addition to the distance to the center. In this way, the farther from the center, the larger the circles will get. It adopts the same grid as the one used in V1 (See V2 Figure 29).

V3) Different brightness in emotion dominance + smoother grid:

In V3, circles have the same size, but their color brightness reflects the emotion dominance: the brighter being the more dominant, in addition to the distance to the circle center. The grid here is smoother, without the concentric circles (See V3 in Figure 29).

V4) Different size and brightness in emotion dominance + no grid:

This version combines both brightness and size to convey emotion dominance information and it does not use a grid (See V4 in Figure 29).

In all the suggested versions, when the cursor is over any of the circles, the information that is presented for each movie in the title list is shown here for the selected movie, close to the cursor. Also, the movie space can be displayed in one of three possible views: “My”, “All users” and “Director’s view”, being “All users” the default one.

Figure 29 - Movies emotional wheel - the 4 versions

5.6.1.2 Individual Movies – Movie Emotional Profile

At the individual movie level (Figure 30), the movie can be represented with information about its dominant emotions and its emotional scenes:

1) The most dominant emotion, in the current view, is represented on top of the video by a large circle filled with the same color used to represent it in the movies space.

2) The dominant emotions, in the current view, represented by the percentage of dominance of each emotion in the movie, are presented to the right of the video. We designed several representations for this purpose, but here we present and tested only one based on bars, as an alternative to the circles used in the other representations.

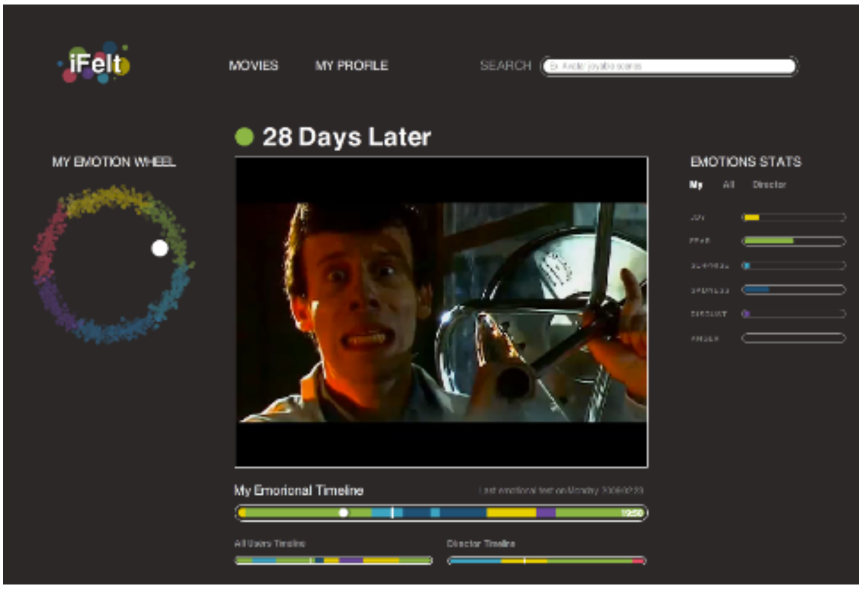

Figure 30 - Movie profile

3) The emotional timelines represent the emotional scenes along time, below the video. The top timeline (in the current view depicted in Figures 30 and 31) represents the emotions in accordance to “My view”, i.e. the current user’s view, the bottom left represents the “All users” view, and the bottom right represents the “Director’s’ View”. A person can use this information to gain more awareness of the emotions involved in the movie, and of how different the three views (mine, all users, and director’s) are. To allow for displaying the current view as the dominant at each moment, and to allow the coexistence of the three views, the bottom two were made smaller. To ease the comparison, when selected, the smaller timelines expand to the same size as the top one (see Figure 31). These timelines also allow to access scenes based on their dominant emotions: by clicking the colored timeline, the video starts to play in the corresponding time, in a scene classified with the corresponding emotion.

4) The current emotion, in accordance to the current highlighted view (the top timeline), is represented: a) in this timeline as a pointer which travels along the timeline while the movie plays; and b) in a Circle of Emotions, to the left of the movie, dynamically presenting a white circle moving to the current emotion represented in the circle (see Figure 30).

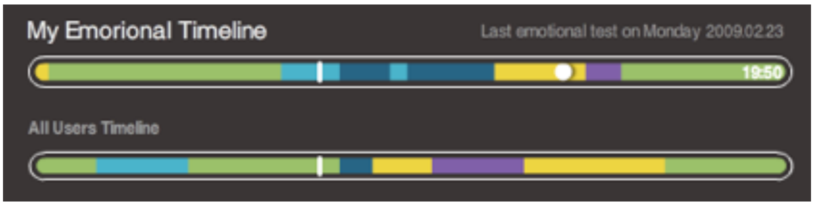

Figure 31 - Emotional timelines after mouse over All Users Timeline

5.6.1.3 Scene Space

In the movies emotional Scenes space, users can get a view of the scenes of the movies based on the dominant emotions existing in those scenes. As in the Movie space, the Scenes spaces can be viewed from the three views: My, All users (default) and Director's view. Scenes with a certain dominant emotion are represented by a circle, with the correspondent color, and with its size matching the percentage of dominance of that emotion in that movie. Scenes with the same dominant color are grouped together. As in the movie wheel, when the cursor is over any of the circles, the information that is presented for each movie in the title list is shown here for the selected movie, close to the cursor, and, in addition, the circles corresponding to the same movie in the different emotions are highlighted while the others become dim (see Figure 32).

Figure 32 - Scene space

5.6.1.4 User's Emotional Profile

From the user's emotional profile, the current users registered in the iFelt system can obtain information about the movies they have watched, i.e. about movies classified from their own perspective. The user's profile contains the following information:

1) My personal information: photography and name, most dominant emotion felt in the movies already classified from "My perspective", and the date the last classification was made (see top left in Figure 33).

2) My dominant felt emotions: represented by colored circles, where the size reflects the percentage of felt emotion dominance in all the movies already classified by the current user (see bottom left in Figure 33).

3) My last classified movies: each movie is represented by an image of the movie (similar to the movies list), tainted with a color filter corresponding to the dominant emotion felt from his view or perspective (see top right in Figure 33). With this design we intended to explore still another representation for emotions in movies.

4) My classified movies space: similar to the Movies space, but presenting only the movies that the current user classified through felt emotions. Users may also choose movies from the wheel or title list to view (see bottom right in Figure 33).

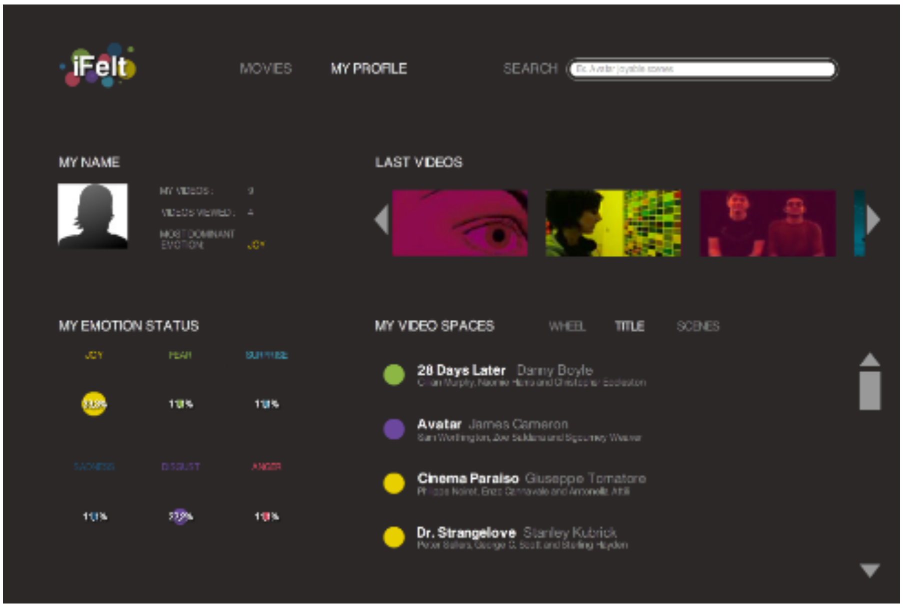

Figure 33 – User's emotional profile

5) My classified emotional scenes: similar to the emotional scenes, but presenting only the movies scenes that the current user classified through felt emotions. This is

presented in the same region as the Movies space (see bottom right in Figure 33), in accordance with the user selection.

5.6.2 Movie Search and Access by Emotions

Regarding the search and access mechanisms to find movies we also created the set of mechanisms described below.

5.6.2.1 Movie Space

From the movie space, the user may navigate to any of the represented movies to access and watch it individually. In the current prototype, there are two representations of the movie space the user can choose from: an emotional wheel (see Figure 28) and a list of movie titles (see Figure 34).

5.6.2.2 Movie Title List

The movie title list allows users to access movies based on their title and image, in addition to their dominant emotion. For this purpose, we provide a list with an image and the title of the movie preceded by a colored circle representing the movie's dominant emotion (see Figure 34).

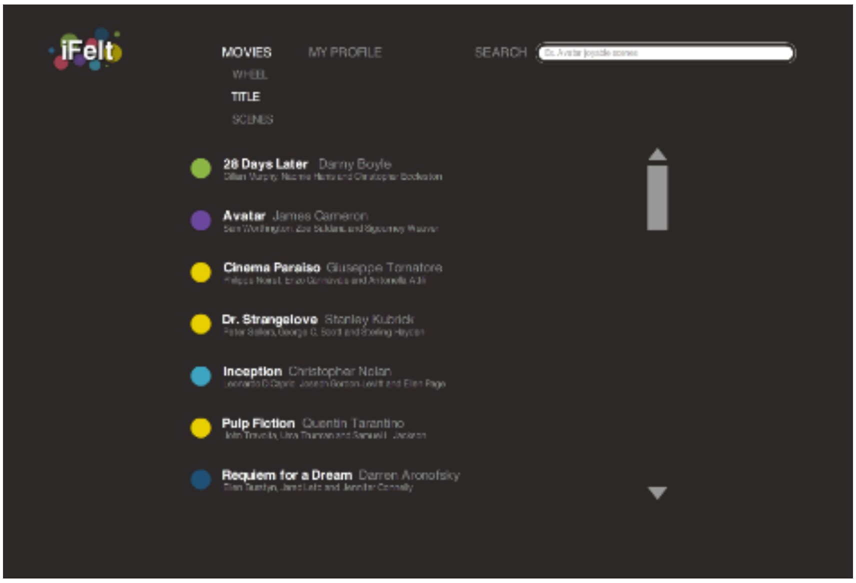

Figure 34 - Movie title list

5.6.2.3 Scene Space

To access a movie by the movie scene, when the user clicks on any of the circles, this state is kept independently of the cursor position, allowing users to click in any of the selected movie scenes, and, as a consequence, to be directed to that individual movie, but presenting only the scenes with the selected emotion.

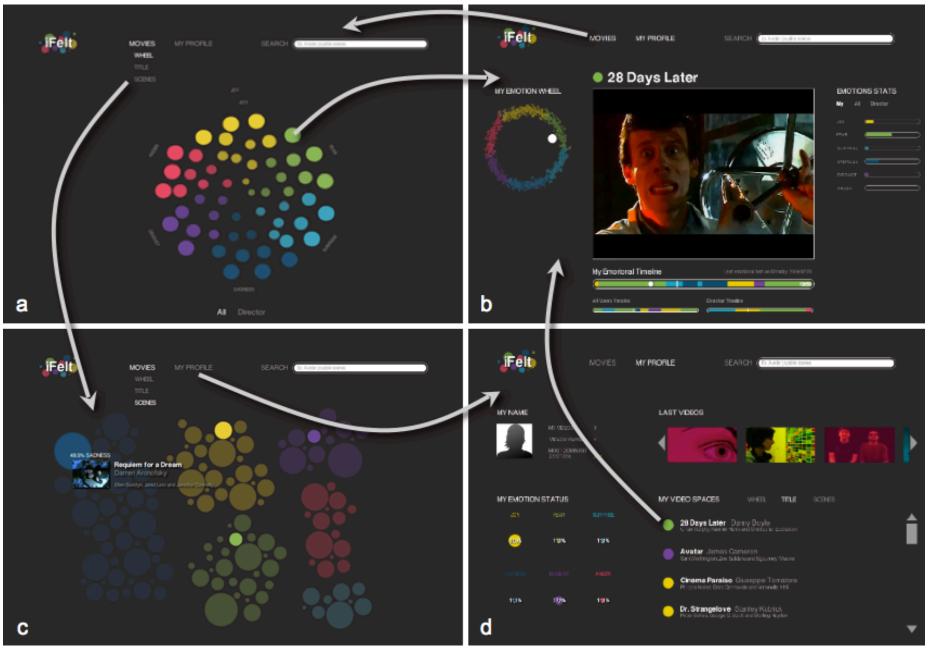

Figure 35 – iFelt interface:

a) Movie space (in movie wheel); b) Movie profile; c) Emotional scenes space; d) User profile (In title list)

The scene space intends to help users to visualize the dominance of colors across the scenes in the selected movies, and to help them to perceive, find and browse movies based on the dominance of emotions, other than the most dominant emotions. It also allows users to access movies in such a way that the scenes are filtered based on their emotions – as emotional movie summaries (Doherty et al., 2003) (e.g. scenes where most users felt sad) (see Figure 35 c). In the example presented in Figure 35, sadness was the emotion chosen (Figure 35 c) and the movie only presents the sad scenes in the perspective of "All users", because that was the view in the scenes space, reflected in the "All users" emotional timeline (Figure 35 b) in the movie profile.

5.6.3 Usability Evaluation – Case Study I

The usability evaluation for the first case study was carried out through a set of tasks prepared in order to assess our research questions, with the goals and methodology described in section 5.4 (Oliveira, Martins & Chambel, 2011). For each assessment we itemize the questions made and present the results for each of them. The mean (M) and the standard deviation (SD) achieved from the analysis of the USE results, based on 5 five point Likert scales, on the proposed tasks (Tn) and underlying questions are presented in tables 11 to 14, complemented with the presentation and discussion of the most relevant results and comments provided by participants, in addition to our own observations.

For this study, we recruited 10 computer literate participants (6 female, 4 male) with ages between 21 and 56 years old, to perform the tests of the experimental system. The results of the evaluation are presented and discussed below.

5.6.3.1 Results

5.6.3.1.1 Movie space evaluation

In order to analyze the usability of the navigation mechanisms, as well as the information representation perceptiveness in the movie space wheel, and to evaluate the search, access and navigation of individual movies from the movies space, we prepared three tasks:

T1) Find the five happiest movies in the collection.
T2) Find the directors' perspectives of the same movies.
T3) Find and watch the "28 Days Later" movie (T3).

Table 11 - USE evaluation in the movie space (scale: 1-5)

USE	Usefulness		Satisfaction		Ease of Use	
Features in task	M	SD	M	SD	M	SD
T1	4.6	0.9	4.9	0.7	5	0.9
T2	4.4	0.9	4.9	0.7	4	0.9
T3	4.3	0.8	3.9	0.7	5	0

We had prepared four versions for the movie wheel representation, as described in section 5.6.1.1. We concluded that version 4 presented above, with the size and brightness conveying emotional intensity and no grid, was the preferred for 90% of the participants, and the participants tended to associate the wheel, or grid, center to less intense emotions.

We also observed that participants felt amused with the possibility of having their collection of movies emotionally classified with viewers' emotional feedback, including their own (T1), and also with the director's perspective of the movies they liked (T2). Concerning T3, the task in which participants were asked to find and watch the "28 Days Later" movie, we observed that 80% of the participants used the Movie Title List, while the remaining 20% of the participants used the Movies Emotional Wheel. In fact, to search for a specific movie, most of the participants opted for the Title List facility, while to search or explore the emotional properties of movies, 90% of the participants resorted to the Wheel.

The movie dominant emotion representation, through a colored circle preceeding the movie titles in the Title List, was not very satisfactory because users were not very familiar with the color of emotions and there were no labels in this view. Moreover, users also found the features in the Movies Space very useful, satisfying and easy to use (shown in Figure 28).

5.6.3.1.2 Movie emotional profile

After accessing the movie emotional profile in task T3, three other tasks were performed regarding that particular movie, in order to enable us to get further information about the ease of use and perceptiveness of the displayed information.

The tasks performed were:

T4) Which were the three first emotions of the movie?

T5) Which are the dominant emotions of the movie?

T6) What can you tell about your classification of the movie compared to all users?

Table 12 - USE evaluation in the movie emotional profile (scale: 1-5)

USE	Usefulness		Satisfaction		Ease of Use	
Features in task	M	SD	M	SD	M	SD
T4	4.4	0.7	3.6	1.5	5	0
T5	4.9	3.1	4.6	0.6	5	0
T6	4.9	0.3	4.9	0.3	5	0

When asked about the emotional scenes sequence of the movie (task T4) (shown in the timeline), we observed an initial difficulty in fully understanding the concept of emotional scenes in this context. As a consequence, for this feature, we got the lowest, but still positive, usability result regarding satisfaction about any of the features we designed for our experimental system. Some participants suggested having also a "traditional scene" representation, in order to facilitate the identification of the emotional scenes by the differentiation between the two concepts. Still, this feature was considered very useful and easy to use, and the highest score was actually given to the possibility of viewing different emotional timelines in task T6. In fact, we observed that 90% of participants were visibly enjoying the exploration of timelines, in order to compare their own emotional classifications with directors' and all users' perspectives.

A substantial number of participants (80%) referred that this was a "cool" functionality to share in a social network environment. Also, for this item, users gave high scores to the usability and usefulness of having the bar based representations of dominant emotions, when asked to find the lowest felt emotion in this movie (task T5).

5.6.3.1.3 Emotional scenes space

The emotional scenes space was evaluated on the usability of the mechanism we devised to access and explore movies by the scenes emotional information, and the participants' perception of the ease of use of the visual representation of the scenes emotional information. We asked participants to perform the following tasks:

T7) Find the movie with the highest percentage of sad scenes.
T8) What do you think is represented on these circles?
T9) Which are the other emotions of that film?

T10) Access to all the sad scenes of that film.

Table 13 - USE evaluation in the emotional scenes space (scale: 1-5)

USE	Usefulness		Satisfaction		Ease of Use	
Features in task:	M	SD	M	SD	M	SD
T7	3.7	1.3	2.8	1.3	3.3	1.2
T8	4.8	0.4	4.4	1	3.7	1
T9	4.4	0.7	4.8	0.7	4.6	0.8
T10	4.6	0.4	4.6	1	4.9	0.9

We observed that only 60% of participants could finish the task T7 and that participants had some difficulties in understanding the visual representation of the scenes space. The access to this scene-based emotional space turned out to be strange for users, which was the reason why almost half of the participants could not finish the task. Based on participants' feedback, we can conclude that this fact occurred because participants were used to the wheel representation of the movie space and did not understand right away the concept of this novel, although somehow similar, representation for a different purpose.

On the other hand, after discovering the concept of the emotional scenes, the possibility of exploring scenes by their emotional impact (task T8) turned out to be found very useful by the participants, and the utility of having, through mouse over or a click in any circle, the view of all other emotional scenes of that same movie (T9) was considered visibly pleasurable and appreciated. In this respect, we observed that participants were especially motivated with the possibility of viewing a representation of the full movie by its emotional scenes. Regarding the access of the movie profile with a summary of the emotional scenes, according to the emotion selected from this scenes view (task T10), it was classified as useful (mean=4.6, SD=0.4), satisfying (mean=4.6 SD=1), and very easy to use (mean=4.9, SD=0.9).

5.6.3.1.4 User's emotional profile

For the evaluation of the user's emotional profile, we asked users to perform the following tasks to assess the perception of emotional information regarding their own profile:

T11) Identify your most and less felt emotion while watching movies.

T12) How many movies did you classify?

T13) Which are your dominant emotions?

T14) Which were the last movies you saw?

Table 14 - USE evaluation in the user's emotional profile (scale: 1-5)

USE	Usefulness		Satisfaction		Ease of Use	
Features in task	M	SD	M	SD	M	SD.
T11	4.8	0.5	4.7	0.5	4.8	0.3
T12	4.8	0.4	3.5	1	4.8	0.9
T13	4.8	0.4	3.5	1	3.6	0.9
T14	4.4	0.5	4.5	1.2	4	1.6

The user's emotional profile was globally satisfactory; users understood the profile concept and found the possibility of having films classified by their own emotions a very useful functionality. In summary, from tasks T11 and T13, we conclude that the "My dominant Felt Emotions" is a perceptive way to represent the emotional information about movies, although users were not sure if this view only included the movies they had already classified emotionally (task T12). The visual representation of the movies in the "My Last Classified Movies" (task T14) also fulfilled usability requirements but, when asked, users said they preferred the circle-based representation for emotional information, instead of the colored tainted movie image.

The following section presents the results of the second study we carried out, where the system was modified to take into account the results just presented for the first case study, also including other functionalities and improvements that we describe below.

5.7 Case Study II

We now describe a user interface that meets more closely the objectives presented in Chapter 1 for this thesis. This second user interface addressed the usability issues identified in the first usability study previously performed, and was completed with the remaining features that were not included in the experimental system's interface developed for the first usability study.

Indeed, our first study described above was very important to provide user feedback on the affective exploration of movies, and also to enable us to understand whether users would enjoy using and exploring emotional properties in movies by using such a system. As we concluded through the first evaluation, the system's usefulness, ease-of-use and satisfaction had overall good results, and provided us with further insights on how to continue to develop the system's interface and its remainder functionalities.

In this section, we first itemize the features that, in our opinion, had to be changed in order to improve the weaknesses identified in the first prototype, and then we describe the design, development and evaluation of a set of new functionalities based on the main goals we set for such a system.

The representations and functionalities we have changed due to the first usability evaluation results, our own observations and participants' suggestions include the following:

1. The first study's participants opted to use the movie title list in 80% of the cases when they wanted to find movies by its general information and 90% resorted to the wheel when they wanted to find an emotion. As we are exploring movies by emotions, we tried to create a more effective search mechanism that could provide more engaging results information.
2. We also observed that users felt that there was no easy way to browse into the wheel; while looking at the list they could sense there was access to a much larger number of movies. This happened because in this first prototype we did not design a filtered searching mechanism.

3. Given that the movie dominant emotion representation (accomplished through a colored circle preceding the movie titles in the Title List) was not very satisfactory, we decided to include a label in every page on this second user interface.
4. In the first study, we also noted that there was a considerable difficulty for users in understanding the scene space. Therefore, we decided to remove this feature as it was initially implemented, and to create a different scene exploration functionality within the context of each movie, providing the possibility of having albums of emotional scenes.
5. The first study's participants also had enormous difficulties in understanding the concept of the emotional scenes by the timeline feature, but found it nevertheless very interesting to enable them to compare emotions. Thus, we changed the way we present information (time and labels) in the timeline design, and created a mechanism for its exploration that turns emotional scenes more perceptive and clear and, at the same time, may create a more engaging functionality.
6. The first study's users could not decide whether the movie they were exploring was already part of their personal albums. We thus created and opted to use a specific symbol to provide such visual feedback.
7. A main suggestion provided by all users of the first study was to be able to share the emotional information with other users, and the preferred functionality for them was to compare their emotions with other users. Therefore, such a mechanism has been created for this second user interface.

Taking into account these recommendations and the results of the previous usability evaluation we then developed a new interface, with a new design and new features that, in our opinion, resulted in a more comprehensive system for the purpose of exploring movies by emotions. Regarding the graphical design of our two proposals we decided to change from a dark background to a lighter one for aesthetics reasons (as shown in Figure 36).

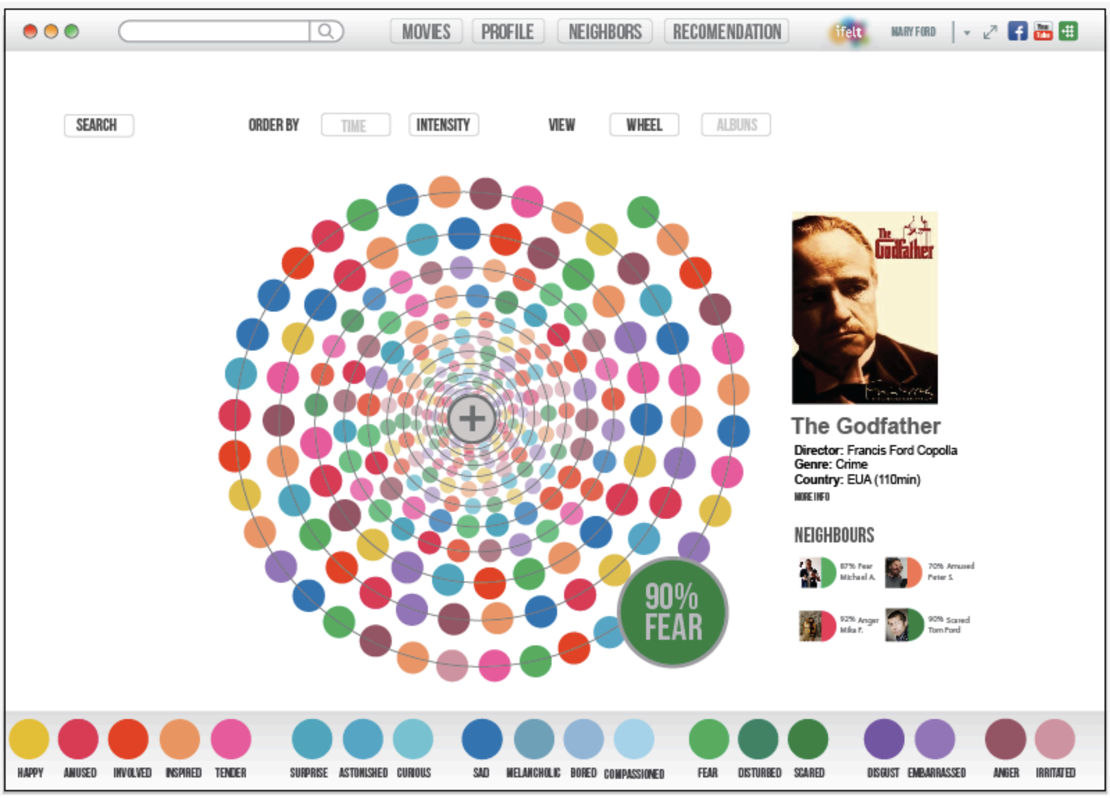

Figure 36 - iFelt main interface

For this user interface we developed two types of menus: global and context. The **global** menu, in the top of the page, allows to navigate to: 1) Movies, which is also the homepage for users; 2) Profile, which is the link to the user's profile page; 3) Neighbors, which is a link to the page where users can explore their most similar iFelt users, i.e., the users that watched the same movies and felt similar emotions; and 4) Recommendation, which is a link to a page that recommends movies based on keywords provided by the user. The context menu appears every time the user mouse over an element of the interface with the specific commands for that element.

We have also tried to explore an additional set of features regarding the fun aspect of the system, and created some exploratory movie options that we describe in the following sections. The following section presents a full description of the system according to our initial goals (see section 5.4).

5.7.1 Visual Representation of Movies by Emotions

As in the first study, we adopted a representation of emotions based on colors, like the model used by Plutchik (1980). However, in this case, besides the basic emotions, we have also used, for each main emotion, a subset of other emotions inspired from the work of Scherer (2005), which in our opinion represents better an emotion triggered by a movie, as described in section 5.2.2.2. We have also included these emotions after a user study conducted to learn about emotional impact of movie watching (Chambel, Oliveira & Martins, 2011). To represent the subset of emotions we used variations of the color, which corresponds to the main associated emotion (see Figure 37). This information is now systematically presented in every page of the new interface.

Figure 37 - List of color and its associated emotions

5.7.1.1 Movie Space

In the new prototype, we designated our representation of the movie space as "Movie Space Wheel" (MSW). We preserved the rounded shapes and color to identify the movie dominant emotions, but we decided to provide the wheel with an order, to enable users to sense the emotional continuity from a browsing perspective. In fact, the movie space has presently more than 250 movies, and can be sorted by the intensity of overall emotions or by a chronological perspective. The following figures illustrate the MSW organized by intensity (Figure 38) and by time (Figure 39). In these illustrations, each circle represents a movie and its color represents the dominant emotion. Also, in the MWS organized by intensity (Figure 38), circles size decrease as we move towards the center of the wheel, i.e., meaning that the emotional intensity of each movie decreases towards the center. In the second wheel, organized by time (Figure 39), the first circle corresponds to the most recent movie added to the system (in this case) and continuing browsing the wheel shows older movies. We have also a representation of the wheel organized by time in the user's profile, where the first movie represents the last movie watched by this user.

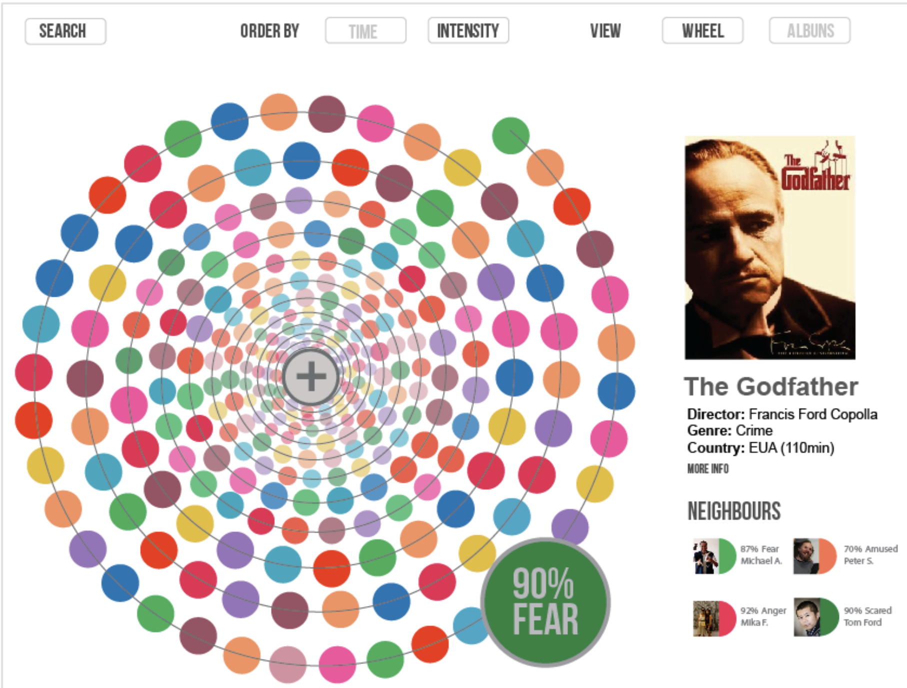

Figure 38 - Movie space wheel organized by intensity

As we move the cursor upon each circle, the systems displays the main information of the movie (title, director, country, duration), a link to more information about the movie (which links to the movie's profile page) and the neighbors that also watched the movie (with the information about the main emotion they felt).

The movie space wheel allows users to search more effectively among hundreds of movies. Although we are focusing on emotions, we also decided to provide a common search mechanism by typing keywords or by specifying other general information such as title, director, cast and year.

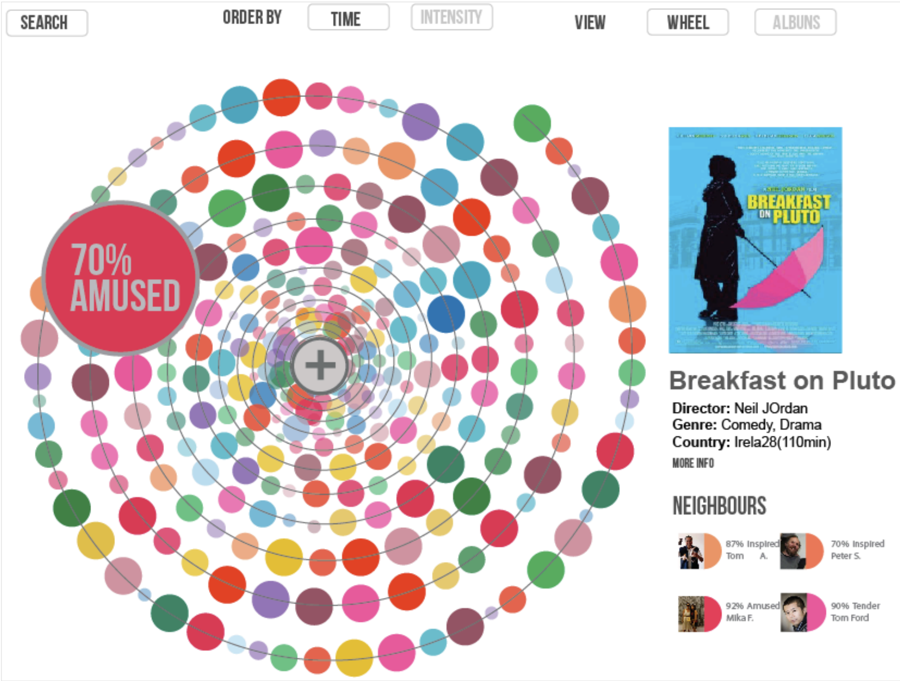

Figure 39 – Movie space wheel organized by time

5.7.1.2 Album List

As a complement, we also included a list to represent a search instead of the MSW, providing a view option where users can opt to explore by the wheel or by a list of albums (illustrated in Figure 40). As we are focusing on emotions, we listed movie posters with the emotional information and not the movie title. Nevertheless, we have also included a list of movies with titles, for the cases where people decide to perform searches through keywords (shown in Figure 40). For example, the album list presented in Figure 40 shows a list of albums filtered by the emotion "happiness", but when the user moves the cursor over each movie's thumbnail, the system shows, just as the wheel, a summary of the information related with that particular movie (e.g. title and director).

Despite the different design, for the second prototype we adopted the same emotional representation for the dominant emotion in a movie or user's profile, as well as for the main emotions and the design of the timeline.

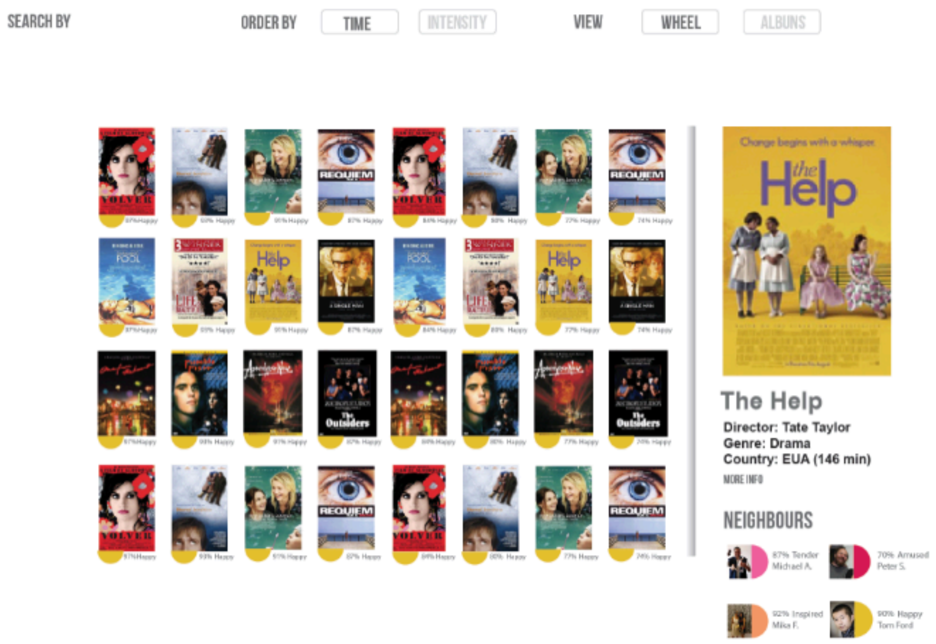

Figure 40 - Album list

5.7.1.3 Individual Movies – Movie's Emotional Profile

At the individual movie level, in this prototype we eliminated the possibility of watching the movie in the movie's profile page, and created a different page for this single purpose, where the movie appears in full screen mode. In order to explore the movie's current emotions we designed two additional functionalities: a) Explore timeline, and b) Explore movie scenes, that also provide new exploration mechanisms of emotions for each movie, as we describe below. However, as in the first prototype, a movie can still be represented with the most dominant emotion in the current view, being represented on the right of the movie poster by the largest circle (using the same color that was used to represent it in the movies space). The dominant emotions, in the current prototype, are represented by the percentage of dominance of each emotion in the movie through the use of circles (as depicted in Figure 41).

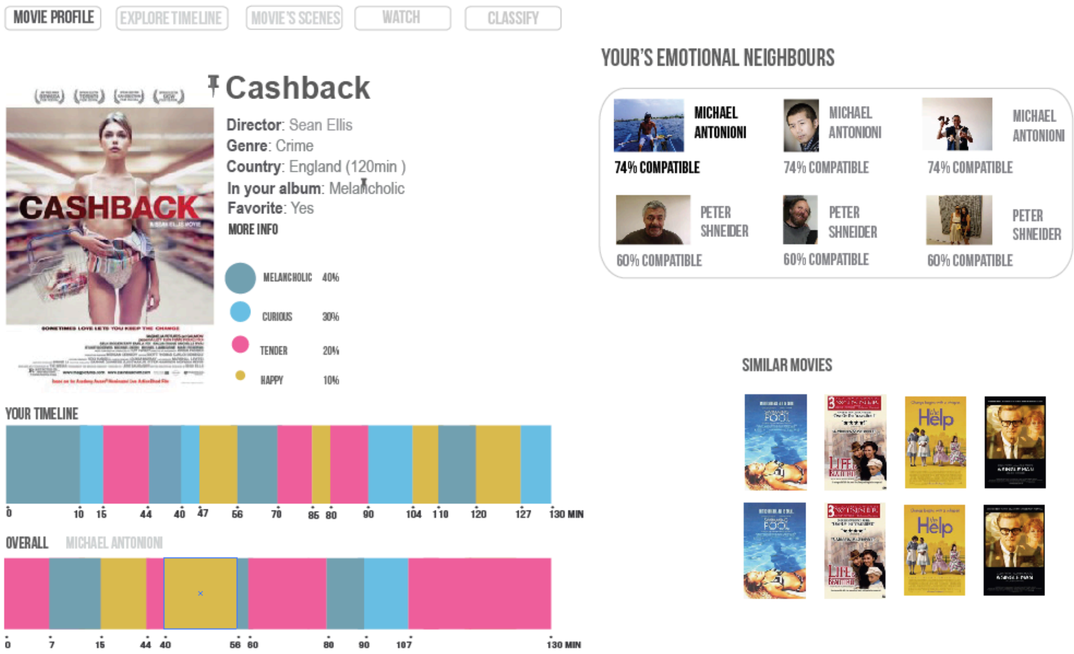

Figure 41 - Movie individual profile

In this new prototype, the emotional timeline is represented in a different way (see Figure 42). As explained in the beginning of this section, for the second prototype we decided to introduce a mechanism where on mouse over the emotion label appears, besides the emotional labels that are always present. We also designed a timer along the timeline that visually indicates the time of the various scenes that compose a movie.

In this prototype, we also decided to have a second timeline below the user's timeline to enable us to represent the most similar neighbor who has watched the current movie being displayed. If for example there are no neighbors that have watched the movie, then the overall classification will appear instead. In fact, there is always the possibility of changing the view to the overall timeline. In complement, some new features were also added to the second prototype to contemplate a broader set of emotional information and functionalities, such as the following:

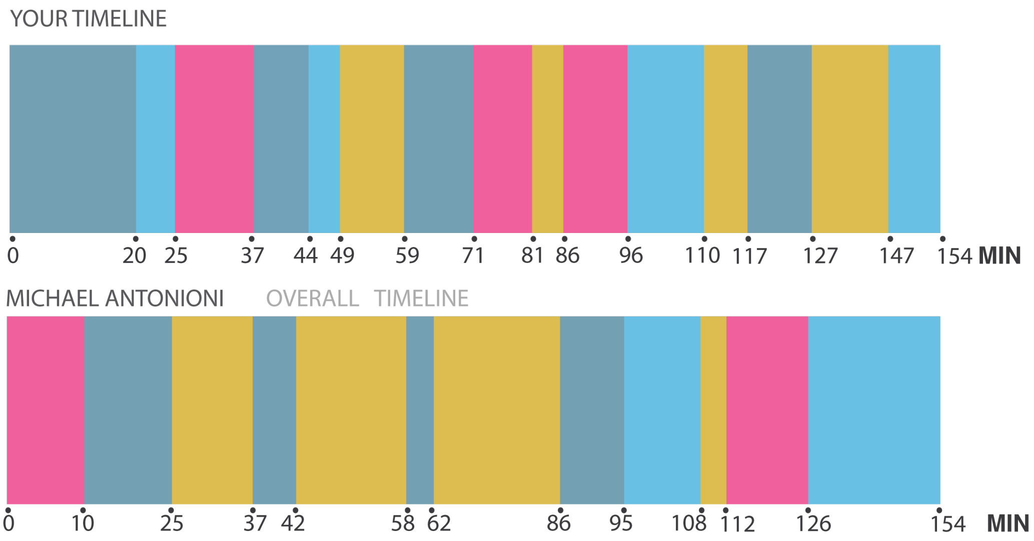

Figure 42 - iFelt emotional timeline

1) A **list of user's neighbors** that also watched the current movie. Whenever the cursor is over a neighbor, his timeline appears under the user's timeline (if the user already watched the movie), or the overall timeline is displayed instead (if the user did not watch the current movie). This feature provides the user with information about the sequence of emotions felt by neighbors in that particular movie. We also removed the director's perspective in this prototype, because it was too difficult to obtain such information for all the movies, and we substituted it with the other compatible user's timeline. But for its recognized value, we keep it in the model for future versions and whenever this information is available. The concept of neighbors and compatibility was thoroughly explained above in section 5.4.4.

We also removed the current emotion from this prototype because this time we would not test whether users liked to watch videos and view the current emotion at the same time. Therefore, we explored this notion of the current emotion in the following new functionalities of the movie's profile: the Explorer timeline and the Explore movie scenes, explained below.

2) The **Explore timeline** allows the user to explore movies by using mouse over and clicking actions, providing the possibility of viewing the respective movie scene. For example, the user has the possibility of viewing the sequence of

emotions felt and watch the corresponding movie scenes by further clicking on them. It is also possible to have two timelines displayed at the same time. For example, we may want to compare our most compatible neighbor timeline with our own timeline. In our opinion, this is especially useful to enable users to gain a more accurate insight about the similarity of that neighbor (see an example of such a comparison in Figure 43).

Figure 43 - Exploring timeline

3) The "**Explore Movie Scenes**" feature allows users to observe a non-sequential representation of the movie along with information about how many emotions were felt for each category with their corresponding intensity, beyond being able to watch those scenes. For each emotional scene represented by a circle, its size represents its dominance (as depicted in Figure 44).

4) In this prototype, we also provided the **share** possibility suggested by the first study's participants. For this purpose, we already included in every page header the future possibility of sharing emotion's classification by clicking in icons representing the most common social sharing services for movies (as illustrated in Figure 45).

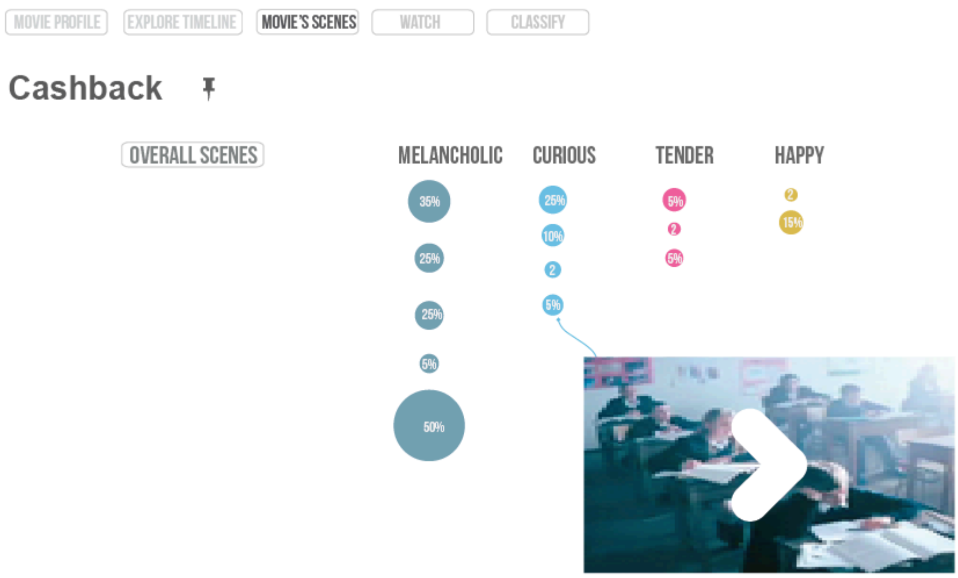

Figure 44 - iFelt scene exploration

Figure 45 - Sharing options

5) The **manual classification** of movies is also a new feature that we included in this prototype. This feature allows users to provide a manual classification of the emotion and its intensity (as shown in Figure 46). To classify movies, users are presented with a wheel, similar to the search wheel of emotions (SWE) described in detail below in section 5.7.2.1. Users then choose an emotion and its intensity by specifying the size of the circle, or by scrolling through an intensity measure.

In Figure 46 we present an example of the classification of the movie “Cashback” with “curious”, and an intensity of 30%, as the main emotion felt for this movie. The scroll user interface element allows for specifying the percentage of that emotion’s intensity, and the circle that corresponds to that percentage is highlighted automatically. This can also be accomplished the other way round, i.e., the user may choose a circle and the scroll position is automatically moved accordingly. Users must classify each movie with a minimum of two and a maximum of five emotions.

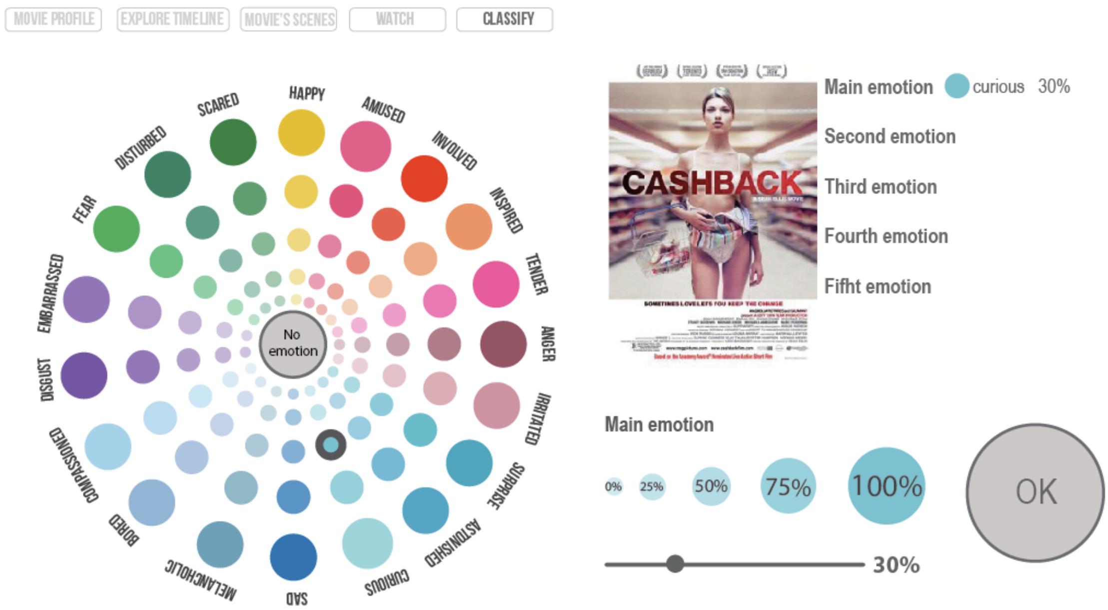

Figure 46 - Movie manual classification mechanism

5.7.1.4 User's Emotional Profile

For the user's emotional profile we preserved the representations used in the first prototype. However, we changed the way we represent the albums, and the MSW is displayed according to this new design; the new feature of the neighbors is also represented with direct access to their own profiles (as shown in Figure 47). Nevertheless, we have a context menu that systematically links to four pages: a) Albums, where users have organized their movies into albums (as shown Figure 48); b) Lists, which allow to access title lists (as shown in Figure 40); and 3) Favorites, which represent movies organized by albums per emotions.

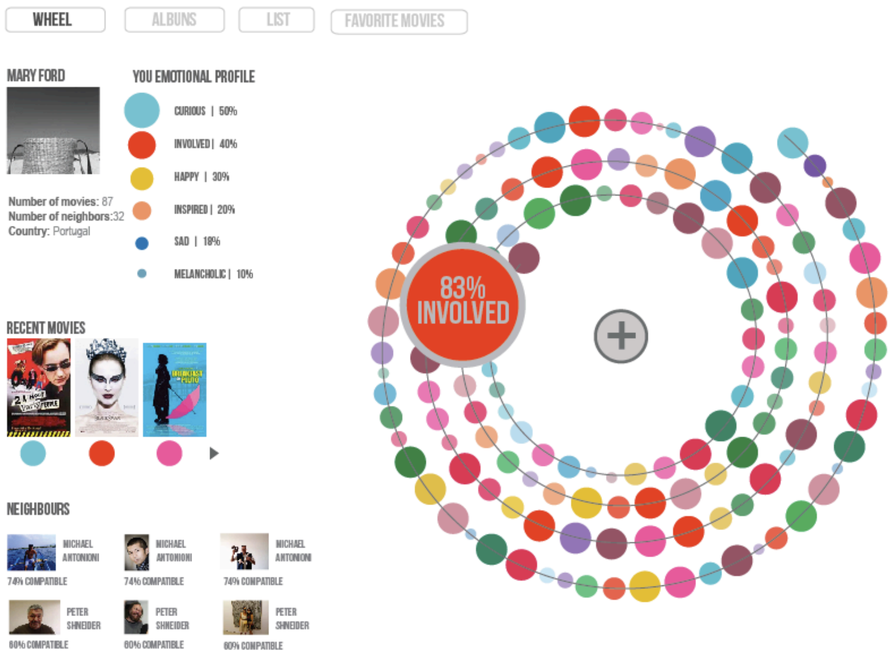

Figure 47 - User profile movie space wheel page

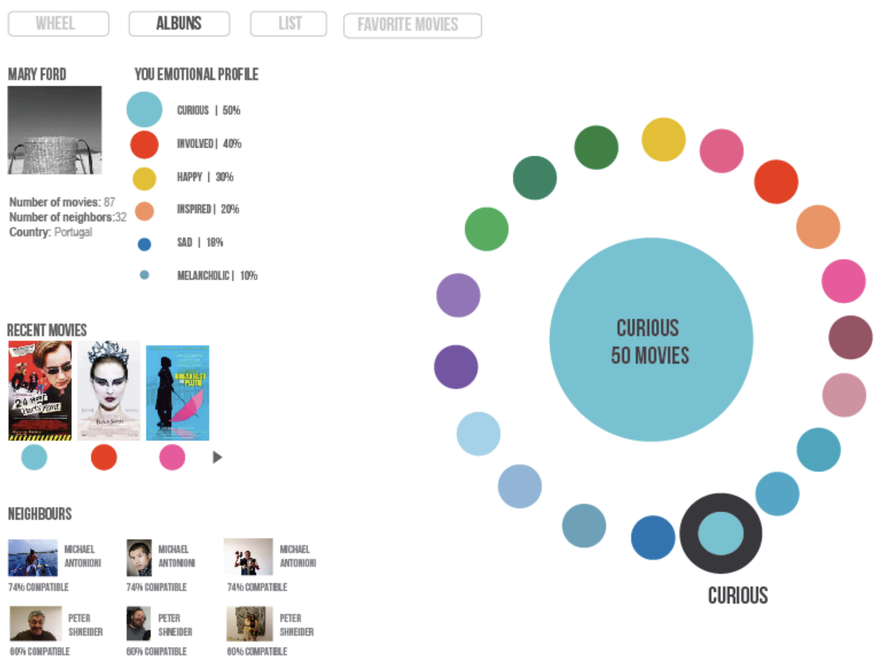

Figure 48 - User albums organization

5.7.2 Movie Search and Access by Emotions

In this second prototype we designed different mechanisms to search, access and find movies in our experimental system. Besides the typical search keywords, that we continue to provide in this prototype, we describe below the remainder mechanisms we introduced.

5.7.2.1 Search Wheel of Emotions

The SWE is based in the Geneve Emotional Wheel (GEW) and in studies around the wheel such as the ones reported by Caicedo & van Beuzekom (2006). The SWE serves two main purposes: a) to search, and b) to classify movies (as shown in Figure 49). As in GEW, we have a set of emotions organized in four dimensions, negative and positive, calm and excited. The group of emotions selected for our SWE was chosen among the Scherer list of emotions and those that, in our opinion, best describe users' feelings after watching a movie, after our study on the emotional impact of movies on viewers (Chambel, Oliveira & Martins, 2011).

Each emotion is represented with a color, and the circles that become smaller and less intensive represent the level of intensity a user can search for or classify. There is also the possibility of classifying a movie with no emotion or with another emotion (outside our list), but we did not explore these two additional options in this prototype. Every emotion besides the basic ones is a subset of the 6 basic emotions, because these are the basic ones and this also allows us to match this classification with the automatic classification we developed and described in Chapter 4. To search through the wheel, users may select the circles corresponding to the emotion they wish to find, specifying a maximum of two emotions. The result can be displayed as a wheel or as a list sorted by time or by intensity, according to the user's preferences. Figure 49 shows an example of a time ordered wheel that represents a sample search for "amused" plus "curious" emotions.

When users want to search for a set of emotions existing on movies, they may select the circle representing the emotion of the correspondent intensity. The movies represented in the wheel shown in Figure 47 are exactly the movies that were classified as "amused" or "curious" as both the main and the secondary emotion.

5.7.2.2 Movie's Title List

As in the first prototype, the movie's title list allows users to access movies based on their title in addition to their dominant emotion.

5.7.3 Browsing Large Quantities of Movies from an Affective Perspective

We now focus our presentation around the browsing features introduced in the second version of the system's prototype.

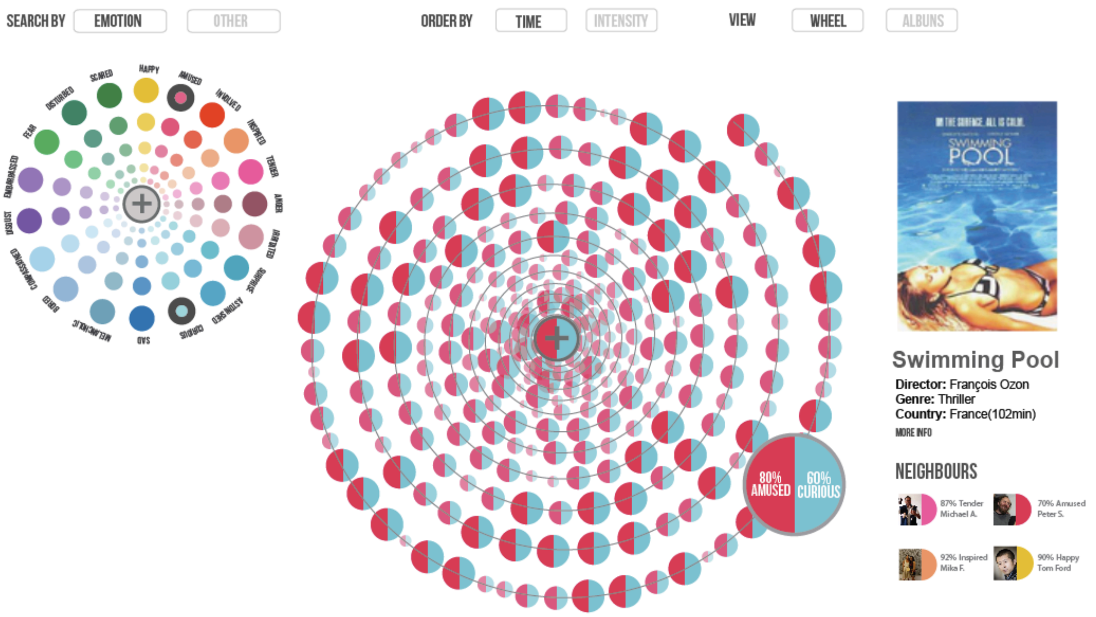

Figure 49 - Search wheel of emotions (SWE) based on GEW.
This example represents a search for curious and amused films.

In order to allow a huge collection of movies to be represented in a wheel, we adopted another way of organizing and displaying movies, effectively enabling an easier navigation among thousands of movies.

5.7.3.1 Movie Space Wheel

For this purpose, we created the movie space wheel (MSW) of emotions already presented in Figure 49 above. The MSW is capable of representing hundreds of movies in a limited display space and enables users to browse among an even larger number of movies by clicking in the plus signal provided for this purpose. Next to the MSW, we also display a scale (a timeline as shown in Figure 50) where users can observe the depth of the search for time or intensity purposes.

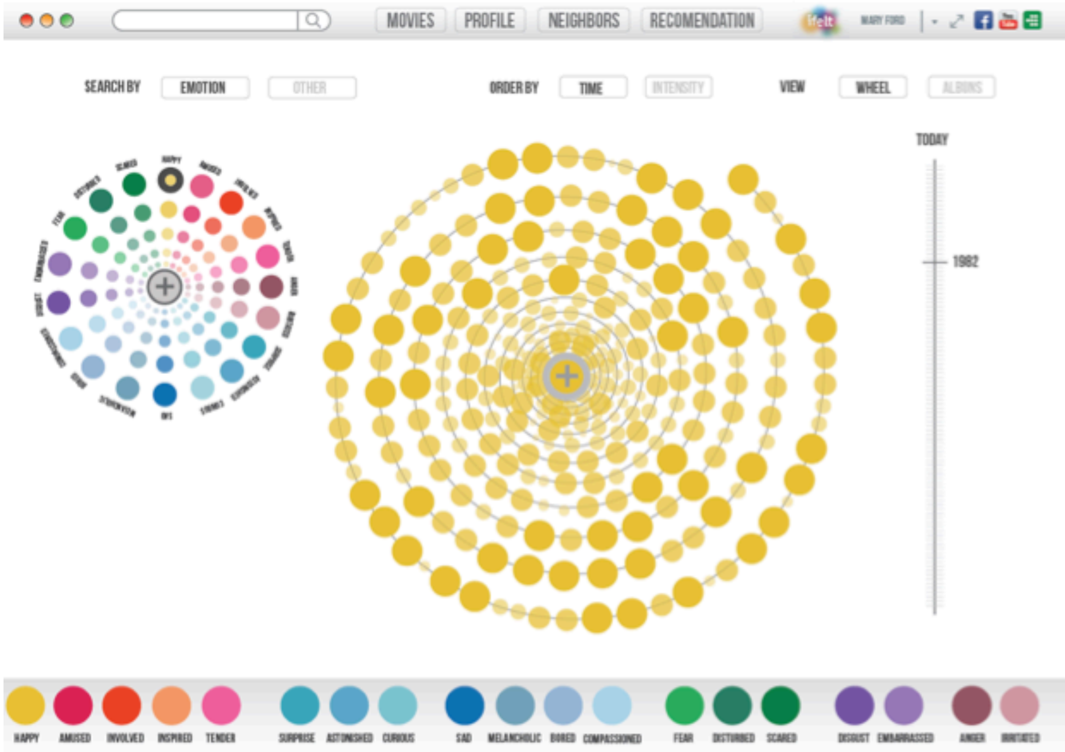

Figure 50 - Browsing over millions of movies

When users move the cursor over a circle, a summary of the information related with that particular movie is then displayed, as we described above.

5.7.3.2 Search Wheel of Emotions

The MSW is usually accompanied by a search wheel of emotions (SWE), or by a form containing general information about movies, according to users' preferences. For example, when a user is searching for the emotion "happy", the user may click in the corresponding circle and the MSW reacts by changing the set of displayed movies (as shown in Figure 51). As the user makes choices, the MSW is constantly changing according to the filters that are being applied, until the final result is displayed.

5.7.4 Movie Recommendation by Analyzing Neighbors

In order to help users find movies by their emotional profile, we propose a set of features that may probably improve the findings of emotionally compatible movies.

As stated earlier, user's emotional profiles are constructed over time by collecting and analyzing all the emotional user data detected for each movie. If we have many users' profiles analyzed along time, we then have the possibility of comparing such profiles, effectively creating "emotional neighbors".

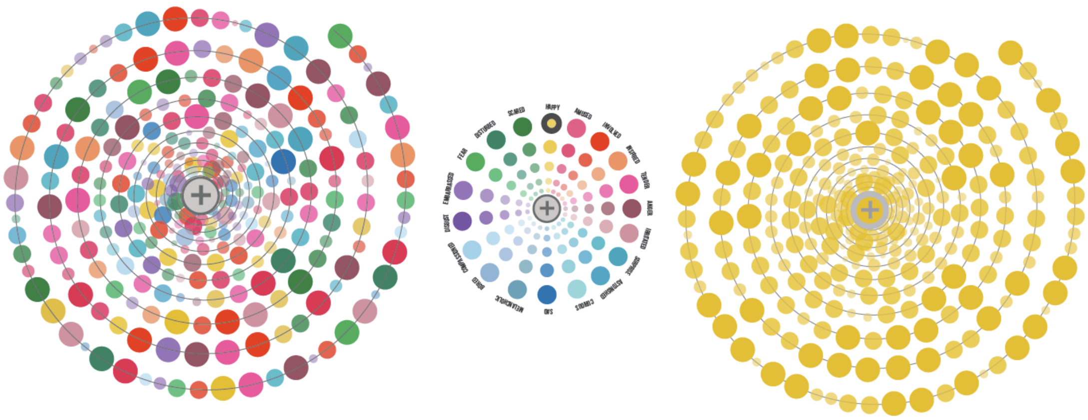

Figure 51 - Movie space wheel reaction to click on the happy circle of the SWE

These emotional neighbors are users that watched similar movies in their albums and possess similar emotional classifications. The use of features related to this concept of neighbors may help to improve the accuracy of finding movies by emotional profile. The features based on neighbors are described in the following sections.

5.7.4.1 Finding Neighbors

As neighbors represent the most similar users, there is a great amount of information readily available to find and access movies. In this second prototype, we propose to find and access the list of neighbors by accessing the neighbors' page (an example of which is shown in Figure 52).

We thus developed a feature to find neighbors (as shown in Figure 52) and to display their own lists. In this case, when the cursor is over a circle (such as the yellow circle depicted in Figure 53), it then becomes highlighted, and the number of neighbors that correspond to that particular emotion appears in the middle of the inside circle, along with additional information about the percentage of compatibility.

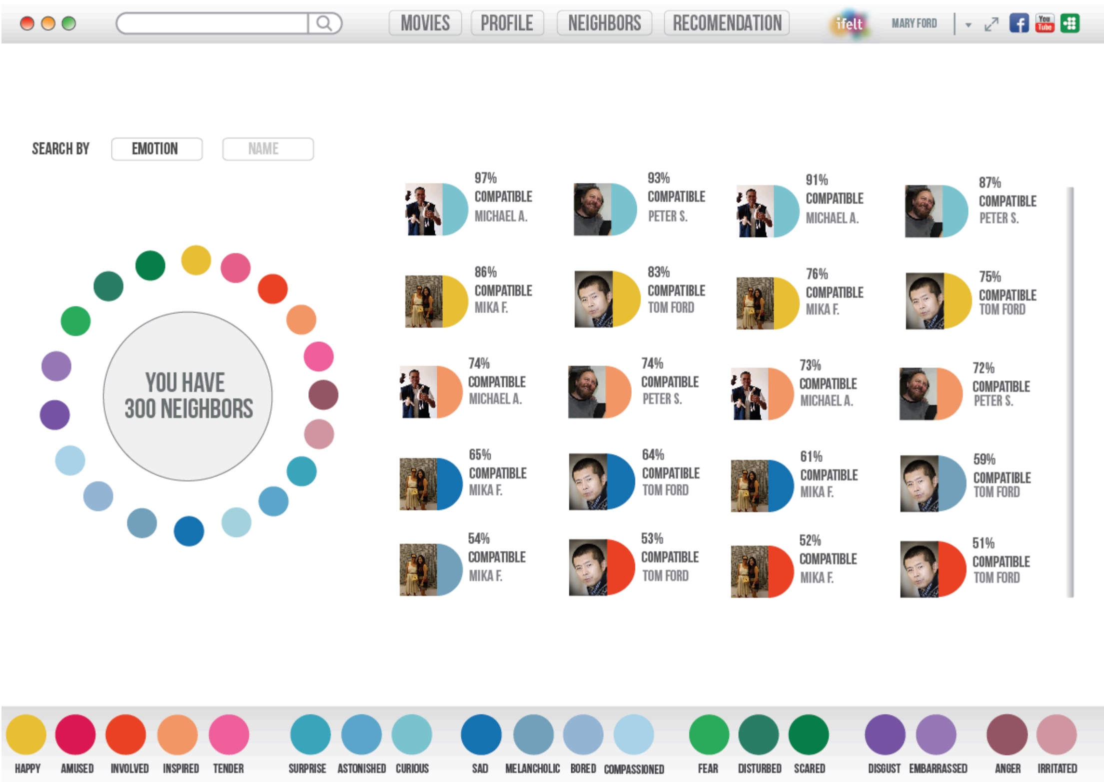

Figure 52 - Neighbors search by emotion or by name

When the mouse is clicked, the list of neighbors reacts accordingly and displays a list corresponding only to that emotion (Figure 53).

5.7.4.2 Neighbor's Profile

The neighbor's section on our system allows for finding users with similar profiles to the current user's profile, as well as to access their user's profile pages (see an example in Figure 54). In that page, every neighbor (if allowed by the owner) can view the list of movies and every classification pertaining to that particular user. To enhance the finding of new emotionally compatible movies, we introduced four lists of movies emotionally approximated to the one a user might be looking for. The first list contains the movies that the user didn't watch; the second one contains the movies emotionally similar, which help users to compare a user's compatibility with another one; the third list presents the less emotionally similar movies, which can also be a relevant aspect when comparing profiles; and the fourth list corresponds to the list of favorites of a particular user, which may be useful if a user is found to be very compatible with another one.

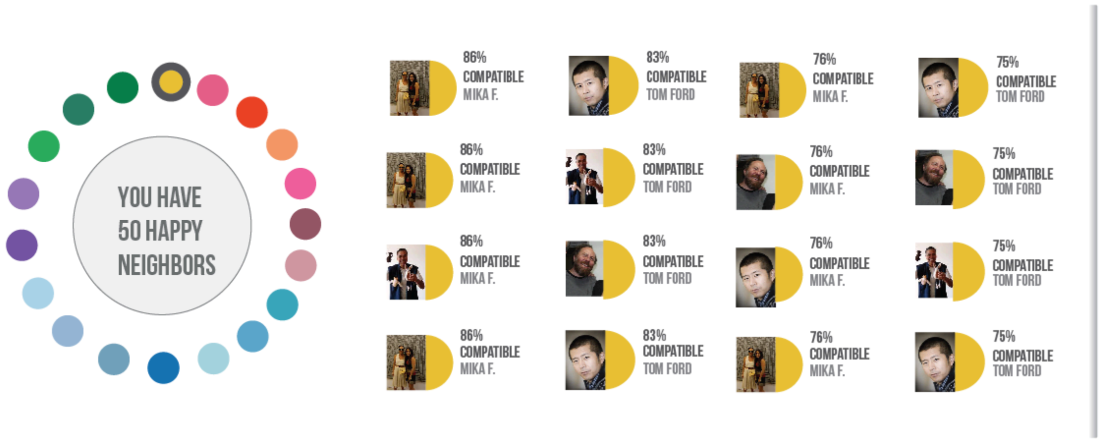

Figure 53 - iFelt - Happy neighbors

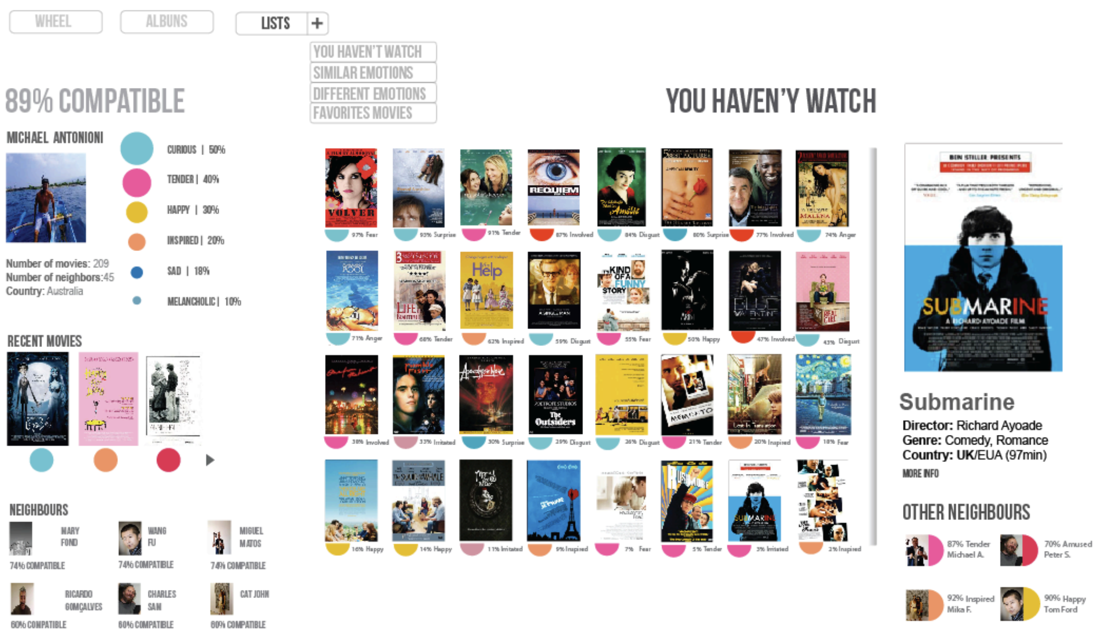

Figure 54 – Neighbor's profile page

5.7.4.3 Recommendation Feature Based on Neighbors

The purpose of this feature is either to recommend movies based on emotions or based on a specific movie (see Figure 55). In the first case, we provide the general search mechanism already described above in the section 5.7.2.

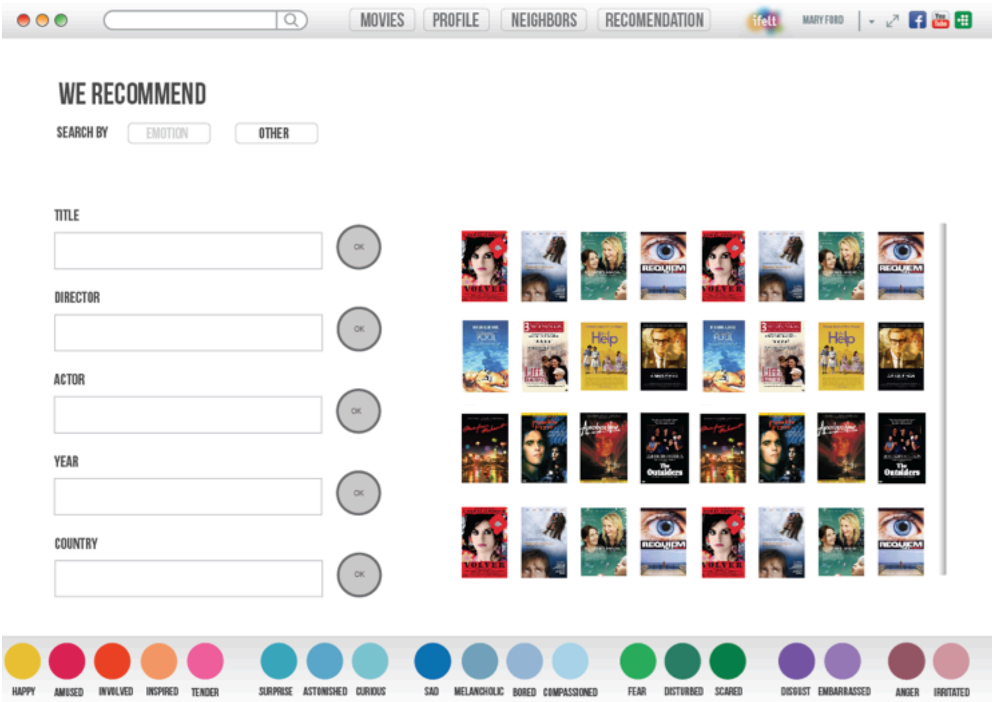

Figure 55 - iFelt emotional recommendation movies by general information

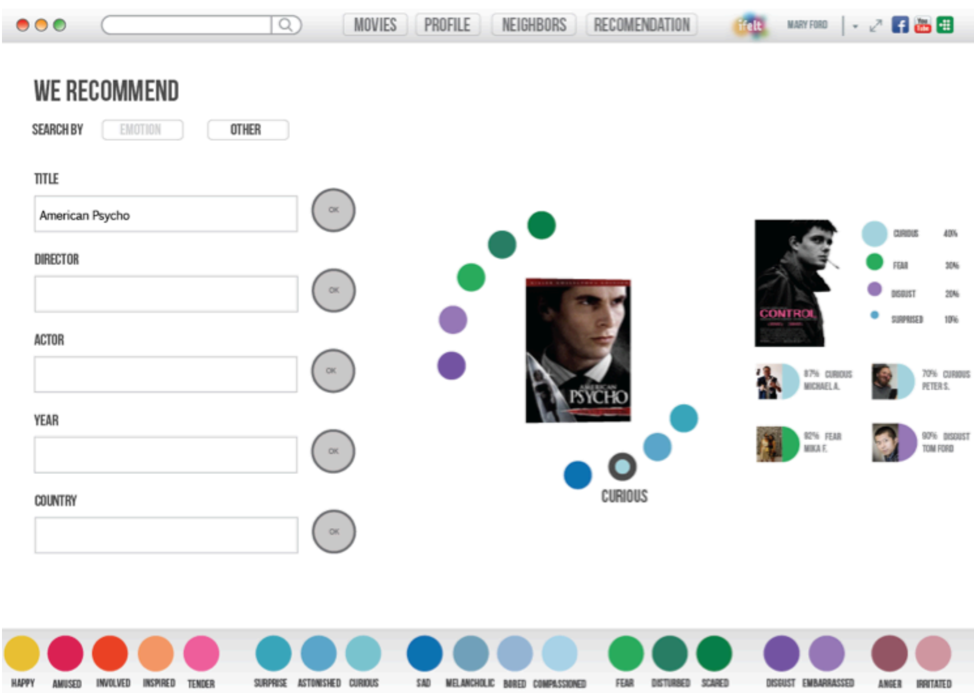

Figure 56 – Result of an emotional movie recommendation by movie title

In the second case (in Figure 56), we provide the option to ask the system to suggest films emotionally similar to a specific one by using common information about movies such as its title, director, actors, year or country. For example, let's suppose a user wishes to watch movies similar to "Kill Bill", which in our system is classified as 70% Amused, 20% Curious and 10% Disgust. The system will first search the neighbors' profiles for movies with the dominant emotion "Amused", second emotion "Curious" and third "Disgust". Then it would query the iFelt movies' database for movies that have a similar classification. For example, Figure 56 shows the result of a recommendation issued for movies similar to "American Psycho". As shown, the system presents a set of emotions related to this movie. If the user clicks on a circle representing an emotion (in this example "curious") a list of movies classified with that particular main emotion is then displayed.

Moreover, when the cursor is over a circle (as shown in Figure 52 over "Curious"), the system displays a sidebar with that movie information; an emotional abstract is also listed along with the neighbor and the emotion felt by that neighbor. This feature provides access to emotional information about movies that the user didn't yet watch.

The following list summarizes the features related to movie search and access by emotions that were described above:

- **Searching** features:
  a) Search by typing;
  b) Search the wheel of emotion;
  c) Search by typing general information.

- **Access movies profiles** features:
  a) Movie space wheel;
  b) User profile albums list;
  c) Neighbors profile lists;
  d) Recommendation of movies;

- **Access movie albums** features:
  a) User's profile albums list;
  b) Neighbors' profile albums;
  c) User's profile circles.
- F**ind movies** features:
  a) Exploring the wheel;
  b) User's profile suggestions;
  c) Neighbors' lists exploration;
  d) Searching for recommendations.

We now describe a second case study conducted to evaluate the usability of the proposals and features just introduced in case study II.

5.7.5 Usability Evaluation – Case Study II

The usability evaluation carried out in the second case study was different from the first case study. For the second case study, we implemented low or mid-fidelity prototypes with high-fidelity colored visual elements. This decision was made for two main reasons. First, the implementation of all the features suggested above would exceed the time available to develop this work. Second, the functionality of the system could be demonstrated through the use of low or mid-fidelity prototypes, as suggested by the literature on user interface usability evaluation (Synder, 2003), and especially if the interfaces resemble those of high-fidelity prototypes and a Wizard of Oz approach is adopted to simulate the system's behavior, as in our case. The prototypes used in this second usability study were already presented and described along section 5.7.

This case study was conducted by having users perform a set of twenty-four tasks prepared in order to assess the research questions identified above. In the beginning, participants were briefed about the testing method, specifically in what concerns the use of a pen to substitute the mouse and paper keyboard to type whatever they needed. They were also instructed to explore every element of the interfaces leisurely, because it could display animations that could help them accomplishing the proposed tasks.

5.7.5.1 Usability Goals

The usability goals for this second case study were similar to the ones already identified for the first case study. In brief, we intended to assess the system's interface usefulness, satisfaction and ease-of-use through the USE-based questionnaire, but also the fun and engagement factors, by asking users whether they enjoyed using the system at the end of each task, if it was fun to use it and if they would recommend the system to other people. We also assessed the perception of the presented information by asking whether participants understood the choices made for visually representing the information and if they would recommend another way to represent the information they used to perform the suggested tasks. Participants were also asked to provide global comments and suggestions (see section 5.5.1).

5.7.5.2 Usability Evaluation Methodology

As in the previous case study, we have also employed an observation-based methodology by annotating observation and comments using a pen and a paper, as well as interviews (a methodology similar to the one followed for the first study, see appendix A, where the interview guidelines may be found).

We briefed all participants, informing them about the context and objectives of the experiment, and explaining them the characteristics of the system that we intended to test, explicitly asking users to Think Aloud to help us observe how they thought about what they were doing, and to understand why they failed and why they hesitated in performing some tasks.

Following an identical procedure to the first study, participants watched two scenes from a movie and classified them to increase their awareness of the global concepts being evaluated. A questionnaire was again filled by participants to later help on the analysis of the results. Satisfaction and engaging parameters of the interface were again observed by the interviewer and fully annotated. At the end, participants were invited to answer a final questionnaire about their opinion of the system as a whole.

For this case study we recruited 8 computer literate participants (3 female, 5 male), with ages between 26 and 45 years old (average 37 years old), to perform the tasks included in this usability evaluation.

5.7.5.3 Results

In order to describe the evaluation performed in the context of the second case study, we itemize below the questions that were answered by the participants and present the results obtained for each of them. The mean (M) and the standard deviation (SD) obtained from the analysis of the USE results, based on 5 five point Likert scales, on the proposed tasks (Tn) and underlying questions are presented in tables 15 to 19, complemented with the presentation and discussion of the most relevant results and comments obtained, in addition to our own observations.

In order to analyze the perceptiveness and ease-of-use of the new information representation features we developed for this second case study, we repeated some of the tasks performed in the first case study to make sure our choices remained understandable by users and also to obtain participants' feedback on the new features.

Although in the first evaluation we wanted to test users' acceptance of new ways of representing movies by emotions and verify whether our representations and mechanisms were well accepted, in this second evaluation, we also wanted to further assess the relevance of some of the features that were not developed for the first prototype. These features included the more complex movie search/browse mechanisms, a different mechanism for the manual classification procedure, and the recommendation and neighborhood features which were completely left out in the first evaluation. Although we did not implement these features in code for the second experiment, we still wanted to assess their utility, engagement, and fun factors. In the following sections, we describe in detail the evaluation we performed and we present the results, their discussion and their comparison with the first case study results organized by feature.

5.7.5.3.1 Visual representation of movies by emotions

In order to analyze the usability of the visual representation and navigation offered by the movie space wheel (MSW), we asked participants to perform a set of tasks which main goal was to assess the perception of our visual representations (task T1) and how easy it was to understand the movie space wheel sorting and movie browsing mechanisms (tasks T2 and T3):

T1) Find the most happy movies in the collection.

T2) Find the most intense movies in this collection.

T3) Sort the movies space wheel in time and find the less intense movies.

Table 15 - USE evaluation of movie space (scale: 1-5)

USE	Usefulness		Satisfaction		Ease of use	
Features in	**M**	**SD**	**M**	**SD**	**M**	**SD**
T1	4.75	0.5	5	0	5	0
T2	5	0	5	0	5	0
T3	4.5	0.5	5	0	4.3	0.5

For this group of tasks, we observed that participants demonstrated a nice first impression when they first observed the user interface. The colors and the wheel had a positive impact on participants, and 80% of them stated that it was very interesting when asked to make a general assessment of the system. Regarding the perception of the movie space wheel, in task T1 the wheel was organized by intensity (order by intensity), as shown in Figure 38 above. We observed that 6 of the participants did not have any difficulty in choosing every yellow circle and obtain the additional information about the movie that is presented on mouse over. However, 2 participants took more time to explore the movie space wheel, but they performed the task in a quite reasonable time (less than a minute). In fact, when participants were exploring the interface, 80% observed the *order by* label and the corresponding highlighted intensity, and then turned their attention to the labels to look for the "happy" color, and, finally, they easily answered the question in task T1. We suppose that maybe it helped them the fact that the first view of the MSW was organized by intensity.

In what concerns the question in task T2, and given that the wheel was organized by intensity, we observed that participants answered it almost instantaneously, given that 100% of participants answered it without any delays. The system features (MSW) used to perform this task were perceived by participants as useful, satisfactory and easy to use (mean=5, SD=0).

Concerning task T3, most participants (80%) also easily obtained the answer by ordering the wheel by the time filter. However, to perform this task, some participants did not find the MSW useful to enable them to find the less intense movies, but found it useful to browse along the wheel. When they were asked about the usefulness of browsing the wheel, 90% of the participants stated that it was more useful to browse the wheel when it was ordered by time, than when it was ordered by intensity to search for intense movies. One participant stated that it was more useful to browse for recent and intense movies when order by time, and more useful to use the intensity order to search only for intense movies. We have also noticed that the labels provided in the user interface were considered one of the facilitator features in what concerns the recognition of emotions by color (the meaning of the chosen colors). Some of the comments provided about the ordered wheel included that “it was very helpful to understand the spatial distribution of the movies”.

5.7.5.3.2 Movie search and access by emotion

We then tested the movie’s profile features and the classification of a movie, and for the movie classification we asked the participants to imagine that that they could watch the full movie and, whenever asked to, its corresponding scenes. For this purpose, we selected the 9 tasks presented below in order to assess how easy it was to access the movies by name (task T4), the perception of the information provided by the system about emotion characteristics of movies (tasks T5, T6), the perception of the neighbors related information (tasks T7, T8, T11), the clarity of the symbol used to inform users that a movie belongs to their albums (task T9), the ease-of-use of the navigation links (tasks T10, T12, T13), the perception of the classification mechanism (task T10), and the global satisfaction of using the provided tools to explore emotions in movies (tasks T12, T13):

T4) Find the movie “Cashback”.
T5) Which is its main emotion?
T6) Which are the emotions that characterize this movie and their percentages?
T7) How did your neighbors feel about this particular movie?
T8) Does the system recommend any similar movie to this one? Which ones?
T9) Have you already watched this movie?
T10) Watch the movie and classify it emotionally with the following emotions and intensity: melancholic 4, curious 3 and tender 2. Share your emotions with

other people in a social network.

T11) How different is your classification from your best neighbor's classification, and when compared to all users' classification?

T12) Access the movie scenes. How many scenes are there and what are their intensities?

T13) Explore the timeline. What is the dominant emotion on the first scene?

At the beginning of this set of tasks, participants were informed that they should take their time in the exploration of the interface due to the high level of animation associated with the features they were about to use. In general, this set of tasks was considered very interesting to perform by participants, and it was also very useful in revealing the engagement and fun aspects associated with this system.

Table 16 - USE evaluation of movie space (scale: 1-5)

USE	Usefulness		Satisfaction		Ease of use	
Features in task	M	SD	M	SD	M	SD
T4	5	0	5	0	5	0
T5	5	0	5	0	5	0
T6	5	0	5	0	5	0
T7	5	0	5	0	5	0
T8	5	0	5	0	5	0
T9	4.3	0.5	3.1	0.4	1.9	1.4
T10	5	0	5	0	4.6	0
T11	5	0	5	0	4	0
T12	4.4	0.5	4.5	0.5	4.4	0.5
T13	5	0	5	0	5	0

In task T4, when asked to search for a specific movie, all participants used the search text field to search for the movie "*Cashback*". When asked for their comments about the method used to perform the search, 90% of the participants reported that given that they had the title of the movie, the more intuitive way to search for it was to locate it by using the text search feature.

Similarly, the features of the system used to perform tasks T4, T5, T6, T7 and T8 were also deemed as useful (mean=5, SD=0), satisfactory (mean=5, SD=0) and easy to use (mean=5, SD=0). These results can be justified by noting that every one of these tasks involved perceiving information that was explicitly given by the system on the movie's profile page. In fact this set of tasks, designed to test the perception of the information provided by the system, achieved the best results possible.

From task T8 to task T13, we observed that participants were visibly enjoying using the system's features associated with them, and felt increasingly engaged with the system than when they first started the experiment. In fact, in task 13, participants were curious, and even surprised, when they understood the underlying concept: the possibility of exploring the timeline, watching the scenes and comparing their classification with their neighbors' classifications were the most appreciated features, given that they were perceived by participants as very useful, satisfactory and easy to use, and also found it "very useful" and "cool".

Nevertheless, task T9 had the lowest results regarding the ease-of-use measure, mainly because most of the participants (80%) took a lot of time (approximately 1,5 minute) to verify whether the movie was already on their album or not at all. We observed that two participants could not even answer the question in task T9, which probably explains the poor results obtained for satisfaction (mean=3,1 SD=0.4) and ease-of-use (mean=1,9 SD=1,4). Regarding the usefulness of the corresponding feature (mean=4.3 SD=0.5), some participants stated that they would remember every movie they have ever watched, especially if it was emotionally significant.

In what concerns the manual classification mechanism (task T10), participants intuitively explored the scroll UI component to change the percentage, and while changing the scroll position, the tester updated the classification wheel accordingly to simulate this system feature. It was interesting to note that, after observing that change in the classification wheel, 80% of the participants began to use the wheel to perform this task, instead of using the scroll, and this was an important moment of the experiment, when they realized that the wheel also reveals intensity. In conclusion, the classification task gave participants a true insight about the representation of the emotions intensities in our system.

In task T11, we observed that participants took some time to discover how the mechanism worked, because they should do a mouse over action on the neighbor's photo in order to change the secondary timeline (as shown above in Figure 42), and this fact probably justifies the lower result obtained for the ease-of-use usability measure (mean=4, SD=0).

In what the task T12 is concerned, the usability measures were very good but not as high as the others. We suppose that the reason for this is that the paper prototype does not provide the necessary dynamic behavior required to use these features properly. In any case, participants found it useful to have the possibility of re-watching the movie organized by the emotional scenes. More relevant than the USE measures, the usability evaluation in this task gave us a better insight about the participants' perceptions about this feature, which was overall considered positive.

One achievement that proceeds from the results of the task T13 is related with the participants' understanding of the underlying concept of the emotional scenes. In fact, when we asked to identify the dominant emotion in the first scene of the movie, participants easily got the correct answer by looking at the label MIN, revealing that they clearly understood that each color corresponded to an interval of time, or a scene.

We tested the user's profile in what concerns its main features, such as the access to user's emotions statistics associated with movies, and the access to the emotions albums. As before, we performed this particular evaluation to ensure that the concepts we developed were easily understood in this new interface. Participants performed the following tasks:

T14) Access your profile. Which emotions best describe you as a movie watcher?
T15) Which are the last emotions you felt?
T16) Access your curious album. How many films do you have in this album?

Table 17 - USE evaluation of user emotional profile (scale: 1-5)

USE	Usefulness		Satisfaction		Ease of use	
Features in task	M	SD	M	SD	M	SD
T14	5	0	5	0	5	0
T15	4.5	0.5	5	0	5	0
T16	5	0	5	0	4.75	0.5

However, for this second case study, we have included the emotional albums access (task T16), a feature which was globally considered as useful (mean=5, SD=0.5), satisfactory (mean=5, SD=0.5) and easy to use (mean=4.75, SD=0.5). Regarding the ease-of-use measure results, 20% of the participants did not immediately explore the album wheel. When asked to justify this behavior, participants answered that it was different from the other wheel and thought it had no similar functionality.

5.7.5.3.3 Browsing large quantities of movies from an affective perspective

At this moment, it is important to recall that the assessment of the usability of movie search, access and browse by the emotional wheel constituted one of our main goals, given that the wheel is one of our key proposals that were developed for the second system interface. In essence, we wanted to create a useful and engaging search mechanism using both the movie space wheel (MSW) and the search wheel of emotions (SWE). Hence, we created a complementary mechanism for visually browsing movies by emotion. In order to test whether users might understand the visual representations we developed for exploring more than 250 movies, and their perception of the wheel-based browsing mechanism, we prepared the following set of search tasks (T17, T18, T19) to assess the usefulness of the movie space wheel (MSW) and the search emotions wheel (SWE):

T17) Search for happy movies and identify two examples.
T18) Search for curious and amused movies and identify the more intense ones.
T19) Search for curious and amused movies from the year 1982.

At the beginning of this set of tasks, participants were again informed that they should take their time in the exploration of the interface, due to the high level of animation involved in using the system's features associated with these tasks. After this briefing, the flow of the test was quite similar to that of a functional prototype because participants were aware of the length of the paper replacements and easily waited for some change when doing a mouse over (pen over) movement. Therefore, the results were very satisfactory. In fact, 90% of the participants found it very engaging to explore the movie space wheel (MSW), and felt a little discouraged for not having the possibility of using the dynamics of the actual mechanism in a concrete implementation of the system.

Table 18 – USE evaluation of search mechanisms (scale: 1-5)

USE	Usefulness		Satisfaction		Ease of use	
Features in task	**M**	**SD**	**M**	**SD**	**M**	**SD**
T17	5	0	5	0	5	0
T18	5	0	5	0	5	0
T19	4.4	0.5	4.6	0.5	4.5	0.5

When asked for the most happy movies (task T17), we observed that participants took some time to analyze the question due to the similarity to the first one (task T1). After reflecting about the difference between the two tasks, 90% of the participants selected the yellow circle of the search wheel of emotions. One person opted to perform the search through the search input text component. We believe that the fact that participants had already made a classification gave them the required insight that the smallest search wheel of emotions (SWE) could effectively filter the larger movie search wheel (MSW). As the experiment progressed forward, we observed an increased interest of participants to explore the wheel to see what happened. Thus, participants easily answered the question in task T14, providing the best possible results for every USE measure (mean=5, SD=0).

After performing this task, users easily performed task T18 as well, first selecting "curious" in the SWE (and the tester changed the MSW), and then selecting the "amused" circle in the SWE (and a dual colored circle MSW appeared as provided by the tester). Participants were surprised by the visual effect and 80% of them reported that it was indeed very useful without even being asked about it. The ease-of-use measure results obtained for this task was favored, probably, by the timing of the task. In fact participants were visibly comfortable in answering to the questions asked in the context of these tasks. For example, when asked to search for curious and amused movies from 1982 (task T19), participants stared at the MSW and, intuitively, clicked in the plus signal. Whenever they clicked on the sign, the depth timeline line dropped a little, and then they could directly select the line and push it to the bottom in order to accomplish the task. The ease-of-use was penalized because we did not have a year on the top of the line, and maybe that introduced difficulties in the task completion. But, in general, participants found the mechanism very intuitive after the first use. Globally, participants also found the search mechanism engaging and intuitive as almost every participant stated they would find it very fun to use the final system when it would be implemented.

5.7.5.3.4 Recommended movies by analyzing emotional neighbors

We now describe the assessment of the recommendation features of the iFelt system's interface. As stated above, the recommendation of movies was a new conceptual feature implemented in this second version of the system. The recommendation mechanisms described in section 5.4.4 are based on the neighbors, which are the most emotional compatible users in the system to the current user. In this context we wanted to test whether participants could infer the role of the neighbors feature by asking them to identify their more similar users in the system (tasks T20, T21), and if participants found the neighbors feature useful to search for new movies (tasks T21, T22). We also wanted to verify whether the recommendation of movies focused on emotions by movie title was also regarded as useful (tasks T23 and T24):

T20) Find your neighbors. Which users are more similar to you?
T21) Provide an example of two movies you haven't yet watched that you would probably like.
T22) Which are your "happy" neighbors?
T23) Ask for a recommendation for movies emotionally similar to "American Psyco".

T24) Which are the emotions that are similar to those in "American Psyco"?

Regarding the assessment of the neighbors' concept (task T20) participants found it globally satisfactory and useful. Some participants (40%) that normally use other recommendation systems really appreciated this feature because, in their opinion, it consisted of a very effective way to find new and compatible information about new movies to watch. However, the remaining participants had some initial difficulties in understanding the concept of finding neighbors, but when they understood more clearly the concept, they found it very interesting, satisfying (mean=5, SD=0), and also a very useful feature (mean=5, SD=0). It is interesting to note that participants used the neighbor link at the top of the page to access to the neighbors page and therefore complete the task, which is, probably, the reason why the ease-of-use measure result was the best possible (mean=5, SD=0). When asked to look for unwatched movies, that were already watched by their most similar user in the system (task T21), participants did not quite understand this question but, intuitively, clicked in the username of their most compatible user and accessed his or her page. At the user's page, participants took a moment to look for any information that might provide them with the answer to this task, and 90% of them clicked in the menu *lists*. Then, participants clicked in the "haven't watched list", and easily answered the question in task 21. Notably, one participant did not explore the link *lists*. In general, these features were perceived by participants as very useful (mean=5, SD=0).

Table 19 - USE evaluation of recommendation based on neighbors' concept (scale: 1-5)

USE	Usefulness		Satisfaction		Ease of use	
Features in task	**M**	**SD**	**M**	**SD**	**M**	**SD**
T20	5	0	5	0	5	0
T21	5	0	5	0	4.8	0.4
T22	4.1	0.4	4.5	0.5	4.1	0.4
T23	4.8	0.4	4.4	0.7	4.3	0.7
T24	5	0	4.6	0.5	4.5	0.5

At the end of the test, during the final phase related with the global assessment of the system, when asked for the best method to discover new movies in iFelt, 70% of the participants considered these lists as the best method to discover new movies. Regarding task T22, 80% of the participants did not understand this question. In fact, even without knowing exactly what the correct answer to that question was, they opted to explore the wheel. At this point, they clicked in the yellow circle, and found the functionality useful (mean=4.1, SD=0.4), satisfactory (mean=4.5, SD=0.5) and easy to use (mean=4.1, SD=0.4). Curiously, afterwards, one of the participants found this feature particularly "original". Although high, we consider that these results were not the best possible due to the fact that only 50% of the participants were used to recommendation systems, so they could not realize the real usefulness of such a feature. Moreover, the feature associated with finding emotion related neighbours was also globally found as a useful and considered to be a "cool" feature, having obtained a reasonable result for every usability measure.

To test the recommendation feature based on a movie title, we asked participants to perform task T23. The major part of the participants (90%) did not take much time to choose the recommendation link, and once in that page, they chose "Other" in the search option without any difficulty. When they typed in the keyboard "American Psycho", participants were visibly surprised with the visual result, and took some time to understand the concept. In fact, 90% of the participants, even hesitantly, still explored the circles, and when they did a mouse over action on a circle, the tester highlighted the correspondent circle and presented the movie information at the side of the page. In that moment, participants understood the idea and essentially found it fun. When asked if it was useful, 50% of participants, although having rated it with high results (mean=4.8, SD=0.4), were somewhat apprehensive about its usefulness. Regarding satisfaction (mean=4.4, SD=0.7) and ease-of-use (mean=4.3, SD=0.7) the results reflect this somewhat vague feature, probably due to experimenting a mid-fidelity prototype.

After task T23, the accomplishment of task T24 was easy for 80% of the participants, given that only two took some time to understand the concept, but still found the feature useful (mean=5, SD=0). We believe that the satisfaction and ease-of-use measures results were not higher due to the participants' difficulty in understanding the concept. Indeed, when asked to classify these measures, some participants stated that, maybe due to their lack of experience in this type of systems, they had some difficulties in grasping the utility of this feature, but that they also believed that it was easy to use after understanding what it might be used for.

5.8 Summary

In this chapter we presented the description of two interface proposals for the iFelt system and their respective assessment. In the first case study, we tested users' sensibility towards the exploration of movies based on emotional characteristics. The usability evaluation we have performed informed us about the usefulness of such a system, and also about the level of perception of the participants about some concepts we have created in this context, for example the movie emotional timeline. In the first case study we did not test every goal of this system that we defined and described in section 5.4, having tested mainly the visual representation of emotions, browsing and movie watching, in the movies space, the scenes spaces, the individual movie's profile, and the user's profile.

The second case study, although performed in mid-fidelity prototypes, tested the whole system's functionalities, encompassing the visual representation of emotions, movies search and access by emotions, browsing of large quantities of movies from an affective perspective and recommending movies by analyzing emotional neighbors.

In both studies, users participated in an experiment designed to observe their attitudes, awareness and preferences about the emotional impact of movies. The inquired viewers strongly agreed that watching a movie can fulfill one's soul or make one sad. They also reported that they quite often feel the need to watch movies, and that sometimes turn to movies to achieve a specific emotional state. For both case studies, we first described the user interface and then presented the usability results of a USE usability evaluation. Globally, the results we obtained confirm that participants found both solutions as very useful, satisfactory and easy to use in their main functionalities, regardless of the fact that some limitations in the first prototype were addressed in the second prototype with good results.

Chapter 6

Conclusions and Future Work

This thesis presented a comprehensive work on the exploration of emotions conveyed in videos from the users' impact perspective, spanning from the gathering of emotions acquired from users to their treatment and visual exploration. This work encompassed the comprehension and analysis of the current state of the art in different scientific fields, including relevant results from psychophysiological measures; physiological signal transformation by applying signal processing techniques; semantic representation and organization of emotional information; and the perceptive and useful visual representation for the access, browsing and exploration of movie emotions from the point of view of their viewers.

In this thesis, we developed a basis for understanding how to use automatic recognition of emotions felt by users while watching movies and how to apply visual and interactive mechanisms for representing acquired emotions. This is part of the set of novel proposals presented in Chapter 4 and Chapter 5. This work also contributes to understanding that the automatic recognition of emotions can be accomplished with low-cost methods. In fact, it may be carried out by using a pattern recognition system such as the one described in this thesis, using easily computed statistical features and quite reasonable recognition rate with only three biometric sensors.

The work described in this thesis uncovered that, until the present, there has been an absence of works, to our best knowledge, about the visual exploration of emotions conveyed on movies based on the automatic emotion recognition from users or, in other words, using the users' emotional impact. This thesis also uncovered that, regarding the emotion-based access and navigation of movie collections, users prefer to use a search input textbox when searching for specific movies, but prefer, and feel more engaged, to search and explore movie collections using the movie space wheel when exploring the emotional properties of video. Also, users thought it was useful to browse and explore the movie collections with a spatial order. Besides, the engagement and fun associated with the emotional exploration of movies by scenes, and the comparison with other users, probably helped participants to understand the whole concept underneath the interface.

Based on the results of the work described in this thesis, we semantically organized -

using Emotional ML, a W3C standard for emotion applications - psychophysiological information acquired from a user while watching a movie together with other user related and movie related information, allowing for the indexation and the subsequent search for emotional information of movies from the point of view of their viewers.

In this chapter, we revisit the results we obtained by performing two case studies in the context of the initial research hypothesis. The thesis concludes with the presentation of the major contributions of this work and a number of suggestions of future directions for continuing this work.

6.1 Revisiting the Hypothesis

The initial motivations for this thesis were some recent developments reported in the literature, such as studies based on emotional theories, and the fact that emotions effectively influence the cognitive process, creativity, concentration and health. Thus, given the importance of emotions on cognitive and engagement aspects of the interaction, the HCI research community encouraged studies regarding the use of emotional aspects in digital multimedia and interaction. On one side, video, and more specifically movies, constitute media which are particularly rich in emotional information. On the other hand, due to the recent technological developments that facilitated the access to huge collections of videos online, and due to the fact that video is increasingly used over the Internet as an entertainment media, it was therefore deemed appropriate to apply methods and techniques of affective computing in both the presentation and access of movies emotional information.

These motivations served as the basis for the following hypothesis that guided the research work detailed in this thesis:

The emotional classification of videos, acquired from users, automatically based on physiological information when watching films and manually through user labeling, has the potential to enrich the ***exploration*** *of emotional information of videos, both by enriching the set of video search criteria and by allowing for a more perceptive and useful visual representation for browsing video collections.*

In general, this hypothesis has been supported through the experimental studies described in Chapter 4 and Chapter 5. The automatic recognition of emotions while watching movies, through physiological data, indeed provided individual emotional information about users while they were watching a specific movie. The usability assessment on the usefulness, satisfaction and ease-of-use of the proposed user interface representations and exploration mechanisms, validated the hypothesis that more perceptive, satisfactory and useful visual representations influenced positively the exploration of emotional information in movies.

To analyze in detail the thesis conclusions, which constitute in large steps the conclusions of the steps we took to test our hypothesis, we describe these conclusions in terms of their correspondent research achievements.

6.2 Research Achievements

The achievements of the research work described in this thesis contributed to enrich current understanding in three key areas:

- How to automatically recognize emotions felt by users;
- How to explore visual representations of emotions to enrich the access to video collections, both in terms of searching and browsing;
- How to use current W3C standards to semantically structure the emotional information related with movies collections.

6.2.1 Automatic Recognition of Emotions from Users

Concerning the emotion recognition and classification part of our work, we developed a novel emotional recognition system that seems capable of automatically recognizing a set of human emotional states using psychophysiological measures and pattern recognition techniques based on discriminant analysis and support vector machine classifiers. In order to test the performance of our system, we suggested and described a novel emotion elicitation scheme, based on emotions induced on viewers by watching selected movie scenes, producing a moderate degree of confidence in collected, emotionally relevant biosignals. Discrete state recognition via physiological signal analysis, using pattern recognition and signal processing, was shown to be reasonably accurate. A correct average recognition rate of 69% was achieved using sequential forward selection and Fisher dimensionality reduction, coupled with a Linear Discriminant Analysis classifier.

However, there are still a number of limitations in our method. In fact, recognition greatly depends on the training phase, i.e., the more emotionally accurate the initial scenes are, the more accurate would be the resulting recognition. Also, the larger the quantity of the training scenes, the better the recognition process will perform. But, it is not very effective to overload users with emotional scenes because users become fatigued, undermining the recognition process, as we observed. A possible solution could be a system that occasionally performs a training phase. Another possible solution, and a more challenging one, would be to learn from the automatic recognition of users while watching movies, asking users, now and then, to confirm the recognized emotion, allowing the learning in this fashion. Another open challenge related to the automatic recognition of emotions is how to optimize the algorithm in what regards the pattern matching behavior, trying to enhance its respective accuracy.

Although the study we performed had a number of limitations, the results we obtained were sufficient to show that easily computed statistical features could be effective. This is promising in what concerns real-time classification systems, allowing moving towards an emotional interaction system. Our results also revealed that few features can achieve pretty good results. Even though physiological sensors are invasive, recent technological advances are resulting in the development of wearable and less intrusive sensors, which make our recognition engine useful in assessing human emotional states during human-computer interactions and further validates the use of movies as powerful emotional triggers.

Thus, these findings support envisioning the real-time classification of discrete emotional states, which further allows the subsequent recommendation of movies based on the actual emotional state of the user, and would also facilitate the adjustment of interfaces according to users' emotions and based on emotional regulation theories. Other possible applications of the work we developed on emotion recognition include multimedia content classification and user interaction mechanisms by developing emotional aware applications that react in accordance to users' emotions. Another interesting possibility, despite its complexity, is the automatic creation of emotional scenes, if rules for scene definition are previously defined.

6.2.2 Exploring Visual Representations of Emotions to Enrich the Access to Movie Collections

We designed two proposals of a web-based video user interface system for the classification, access, exploration and visualization of movies based on an emotional paradigm, and presented two user studies that were carried out with the objective of evaluating a number of suggested functionalities, usability and perceived utility. The main limitation of both studies was the use of prototypes that were not fully implemented versions of the system, making impossible to test its efficacy (as a complete system with information that actually corresponded to the viewers testing the system). Nevertheless, we were still able to evaluate a number of usability aspects concerning the interaction with such a system and measure its usefulness, satisfaction and ease-of-use, as well as the engaging aspect of the resulting user experience.

The first study was conducted to perform the evaluation of our suggested representations of emotional aspects of movies (circles, colors), and users' and movies' profiles representations. The second study covered improvements on the prototype used in the first study and introduced and assessed new visual representations for some mechanisms to browse, classify, and search video collections containing large numbers of movies, as well as the visual representation of a new concept that we described as emotional neighbors.

From the first study, based on a high-fidelity prototype, our first conclusion, based on users' feedback, was that the experimental system, iFelt, was generally perceived as useful, satisfactory and easy to use by the participants in the experiment. Regarding the access and exploration of movie related information, we concluded that participants preferred to use the "Movie Title List" to search for a specific movie, while to search or explore emotional properties of movies, most of the participants tended to opt for the "Movies Emotional Wheel". The exploration of emotional timelines was considered useful, especially in their ability to allow the comparison of users' own emotional classifications with directors' and all users' perspectives We did not include the directors' perspective on the second prototype because identifying the director's intentions, scene by scene, was found extremely difficult to obtain, although it was found very useful and interesting in many scenarios, and kept in our model for when we have access to that information (Oliveira, Martins & Chambel, 2011).

From the second study, based on a mid-fidelity prototype, we concluded that exploring, browsing, classifying and searching movies by emotions in our prototype was also generally perceived as useful, satisfactory and easy to use, now that we increased the number of emotions addressed, in a wider model of emotional impact, as suggested by our study (Chambel, Oliveira & Martins, 2011), and that we also explored the visualization of very large amounts of movies. Regarding the exploration of emotions on movies, users have shown interest in exploring emotions through the movie space wheel (MSW) feature, because it made them more aware of the movies representation by the colored circles, due to the presence of the emotion labels. In fact such awareness resulted in a more pleasant experience and an understanding of the usefulness of the whole system by the participants. In what concerns the perceptiveness of our representations, the usability evaluation provided very satisfactory results, given that participants provided correct answers in all the suggested tasks, leading to the conclusion that they understood correctly the information representation mechanisms we were testing.

In the second study the emotional scenes were easily understood through the interpretation of the additional timeline, and were perceived as a very useful mechanism by all the participants, mainly because it enabled them to revisit a movie and directly choose any scene such as, for example, the happiest scenes. The feature “explore timeline” was one of the most important mechanisms of the interface, in what concerns engagement and fun properties. This feature also helped to understand a concept that was difficult to understand in the first prototype: the emotional scenes. The possibility of exploring timelines was visibly the most engaging feature for all second study’s participants, even higher than the engagement found in the first study’s participants, mainly because users found the neighbors’ feature very useful. In fact, the neighbors concept was regarded as a powerful way to discover new emotionally interesting movies, because the recommendation is based on the comparison of similar users’ emotional profiles, which was found a reliable way to recommend new movies to watch from the point of view of most of the second study’s participants.

We have also demonstrated that our visual representation proposals were found by test participants to be appropriated to communicate emotions on movies, i.e., participants easily understood that colored circles constitute an effective way to represent movies by emotion, and that colored timelines can effectively represent a movie timeline divided into emotional scenes. In what concerns the movie space wheel representation, we observed that it was also clearly understood by our test participants and was regarded as a perceptive mechanism for the time and intensity organization of movies' emotions.

6.2.3 Semantic Structure of Movies' Emotional Information Based on Standards

This thesis also contributed with the design and development of an Emotion ML based structure for the emotional information manipulated in the experimental system. In this way, we demonstrated the possibility of structuring the emotional information manipulated in our system based on a W3C standard. This fact allows us to conclude that our emotional information might be universally understandable by other systems, given that such a system has the advantage of being open to external communication with other systems that might require, or need, to use the emotional information acquired and stored in our system. We believe that the structure we proposed may provide emotional meaning and organization for movies, both as a collection and as media linked to a user, as well as by providing machine understandable meaning to a user profile in what concerns a person's psychological profile, constructed while the person is watching a movie. The possibility of sharing such information also provides a small contribution to the information retrieval area, specifically in order to enhance search mechanisms by adding new search criteria.

6.3 Future Work

This thesis, as an exploratory study, uncovered a number of interesting issues and results regarding the future exploration of emotions elicited by watching movies. In fact, this work suggests a number of new challenges such as the ones described below.

Regarding the emotion recognition process, there is the need to improve the recognition rates by changing some scenes. Also of interest is to investigate to what extent films can be automatically divided into emotional segments with our recognition engine. Given that this engine is based on easily computed statistical features, an interesting avenue for future research would be to implement real-time classification of discrete emotional states. Our classification method could also be improved by adding arousal/valence mappings from biosignals, and additional emotions, for multimedia content classification.

In what concerns the visual representation of emotions and system features, we are already implementing the second interface and the next step would be to improve and extend the system we described to include a number of additional features, including: extending the concept of video summaries to present movies in chosen emotional perspectives and preferences enriching the system with additional search criteria other than just selecting scenes with one chosen emotion; summarizing, searching or recommending movies based on users current emotional states or other previously defined emotional criteria; to enable finding movies by example, with emotional timelines similar to the timeline of a given movie.

Another interesting feature to investigate and develop would be to include support for historical emotional information gathered along time, so people may witness the evolution of users' emotional reactions to movies over time, and compare them to other perspectives, including the actors and directors involved, across several movies genres.

We also intend to make all the information gathered in our system more available, or visible, on the World Wide Web as a shared and recommender environment based on the emotional classification of movies. Such an endeavor might prove useful for the general public, as well as for professionals, such as directors and actors.

Finally, iFelt is currently focused in movies and the web environment, but this same approach might prove to be useful and interesting to be explored with other types of digital videos, as is the case of advertisement videos that typically aim to induce specific emotional reactions in the viewers. It might also prove to be interesting for 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.

References

Ahlberg, C., & Truvé, S. (1995). Tight coupling: Guiding user actions in a direct manipulation retrieval system. In M. A. R. Kirby, A. J. Dix, and J. E. Finlay (Eds.). *People and Computers X: Proceedings of HCI'9.* New York: Cambridge University Press.

Arapakis, I., Konstas, I., & Jose, J. (2009). Using facial expressions and peripheral physiological signals as implicit indicatior of topical relevance. In *Proceedings of the Seventeen ACM Conference Multimedia.* New York: ACM.

Axelrod, L. & Hone. K. (2006). *Multimodal Expressions: E-Motional and Non-Verbal Behaviour.* In G. Ghinea & S. Y. Chen (Eds.), *Digital Multimedia perception and design*. New York: IGI Global.

Bestiario, Videosphere (2008, May). Retrieved from http://www.bestiario.org/research/videosphere/

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning.* Cambridge, UK: Springer.

Blythe, M., Monk, A., Overbeeke, K., & Wright, P. (Eds.). (2003). Funology: From Usability to Enjoyment. Boston: Kluwer, Academic Publisher.

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, *25*, 49-59.

Bulterman, D. C. A., & Rutledge, L. (2008). SMIL 2.0: Interactive multimedia for web and mobile devices. Amsterdam: Springer.

Caciedo, D., Van Beuzekom (2006). “How do you Feel?” An assessment of existing tools for the measurement of emotions and their application in consumer products research. Delft University of Technology. Unpublished. Manuscript.

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think.* San Diego: Academic Press.

Chambel, T., Oliveira, E., & Martins, P. (2011). Being happy, healthy and whole watching movies that affect our emotions. *Affective Computing and Intelligent Interaction*, Volume 1, 35-45.

Christie, I. C., & Friedman, B. H. (2004). Autonomic specificity of discrete emotion and dimensions of affective space: A multivariate approach. *International Journal of Psychophysiology*, *51*, 143-153.

Damasio, A. (1995). *Descartes' error*. New York: Harper-Collins.

Darwin, Charles (1998 /1872). The Expression of the Emotions in Man and Animals. 3rd ed. New York : Oxford University Press, Inc..

De Carolis, B., De Rosis, F., Carofiglio, V., Pelachaud, C., & Poggi, I. (2001). Interactive information presentation by an embodied animated agent. *International Workshop on Information Presentation and Natural Multimodal Dialogue*, 14-15.

Descartes, R. (1989/1649). The passions of the soul. *The Philosophical Writings of Descartes*. (originally published in 1649). Indiana: Hackett Publishers.

Desmet P.M.A., Porcelijn, R., & van Dijk, M. (2007). Emotional design; application of a research based design approach. *Journal of Knowledge, Technology & Policy*, 20, 141-155.

Desmet, P. (2003). Measuring emotion; development and application of an instrument to measure emotional responses to products. In: M.A. Blythe, A.F. Monk, K. Overbeeke, & P.C. Wright (Eds.), *Funology: from Usability to Enjoyment (pp. 111-123)*. Dordrecht: Kluwer Academic Publishers.

Desmet, P. (2002). *Designing emotions.* Delft University of Technology. Unpublished dissertation. The Netherlands.

Devillers, L., Vidrascu, L., & Lamel, L. (2005). Challenges in real-life emotion annotation and machine learning based detection. *Neural Net*, *18*, 407-22. doi:10.1016/j.neunet.2005.03.007

Dix, A., Finlay, J., & Abowd, G. D. (2004). *Human-Computer Interaction (3rd Ed)* Essex, UK: Prentice-Hall.

Doherty, J., Girgensohn, A., Helfman, J., Shipman, F., & Wilcox, L. (2003). Detail-On-Demand hypervideo. *Proceedings of the Eleventh ACM International Conference on Multimedia*, 601. New York: ACM.

Douglas-Cowie, E., Campbell, N., Cowie, R., & Roach, P. (2003). Emotional speech: Towards a new generation of databases. *Speech Communication*, *40*, 33-60.

Douglas-Cowie, E., Cowie, R., Sneddon, I., Cox, C., Lowry, O., Mcrorie, M., .Batliner, A. (2007). The HUMAINE database: Addressing the collection and annotation of naturalistic and induced emotional data. *Affective Computing and Intelligent Interaction* (pp. 488-500). Amsterdam: Springer.

Duchenne, G. B., & Cuthbertson, R. A. (1990). *The Mechanism of Human Facial Expression.* Cambridge: Cambridge University Press.

Duda, R. O., Hart, P. E., & Stork, D. G. (2001). *Pattern Classification.* New York: Wiley.

Ekman, P. (1992). Are there basic emotions? *Psychological Review*, *99*, 550-553.

Ekman, P. (1994). All emotions are basic. In P. Ekman & R. Davidson (Eds.), *The Nature of Emotion.* Oxford: Oxford University Press.

Ekman, P. (1999). Basic emotions. T. Dalgleish & M. Power (Eds.), *Handbook of Cognition and Emotion (pp. 45-60)*. Sussex, UK: Wiley.

Ekman, P. (2004). *Emotions Revealed: Understanding Faces and Feelings*. New York: Times Books.

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous system activity distinguishes among emotions. *Science, 221,* 1208-1210.

doi:10.1126/science.6612338

E*MMA: Extensible Multimodal Annotation Markup Language.* (2010, May). *EMMA: Extensible Multimodal Annotation Markup Language.* Retrieved from http://www.w3.org/TR/emma/#s3.2

Fehr, B., & Russell, J. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology, 113,* 464-486.

Fontaine, J. R. J., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological Science*, *18*, 1050.

Friedman, B. H., & Thayer, J. F. (1998). Autonomic balance revisited: Panic anxiety and heart rate variability. *Journal of Psychosomatic Research*, *44*, 133-151.

Frijda, N. H. (1987). Emotion, cognitive structure, and action tendency. *Cognition & Emotion*, *1*, 115-143.

Gebhard, P. (2005). ALMA: A layered model of affect. *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, 29-36. New York: ACM Digital Library.

Godbole, N., Srinivasaiah, M., & Skiena, S. (2007). Large-Scale sentiment analysis for news and blogs. *Proceedings of the International Conference on Weblogs and Social Media (ICWSM)*, 219-222.

Gomez, P., & Danuser, B. (2004). Affective and physiological responses to environmental noises and music. *International Journal of Psychophysiology*, *53*, 91-103.

Gomez, P., Stahel, W. A., & Danuser, B. (2004). Respiratory responses during affective picture viewing. *Biological Psychology*, *67*, 359-373.

Greenwald, M. K., Cook, E. W., & Lang, P. J. (1989). Affective judgment and psychophysiological response: Dimensional covariation in the evaluation of pictorial stimuli. *Journal of Psychophysiology*, *3*, 51-64.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & Emotion*, *9*, 87-108.

Hanjalic, A., & Xu, L. Q. (2005). Affective video content representation and modeling. *Multimedia, IEEE Transactions on 7*, 143 - 154.

Harris, & Kamvar. (2009). We feel fine: An almanac of human emotion. *We Feel Fine: An Almanac of Human Emotion.* New York: Scribner Books.

Hassenzahl, M. (2005). The quality of interactive products: Hedonic needs, emotions and experience. In C. Ghaoui (Ed.), *Encyclopedia of Human Computer Interaction (pp.226-272). Idea Group.*

Hauptmann, A. (2005). Lessons for the future from a decade of informedia video analysis research. In Proceedings of the 4th international conference on Image and Video Retrieval (CIVR'05), Wee-Kheng Leow, Michael S. Lew, Tat-Seng Chua, Wei-Ying Ma, and Lekha Chaisorn (Eds.). Springer-Verlag, Berlin, Heidelberg, 1-10. DOI=10.1007/11526346_1 http://dx.doi.org/10.1007/11526346_1.

Human, B. (2007). Human-Computer interaction in the year 2020. Cambridge: Microsoft Research.

Huppert, F. (2006a). Positive emotions and cognition: Developmental, neuroscience and health perspectives. *In J. P. Forgas (Ed), Affect in social thinking and behavior (pp. 245-252).* New York: Psychology Press.

Isen, A. M., Daubman, K. A., & Nowicki, G. P. (1998). Positive affect facilitates creative problem solving. *Human Emotions: A Reader*, 288-297.

Junghöfer, M., Bradley, M. M., Elbert, T. R., & Lang, P. J. (2001). Fleeting images: A new look at early emotion discrimination. *Psychophysiology*, *38*, 175-178.

Kang, H. B. (2003). Affective content detection using hmms. *Proceedings of the Eleventh ACM International Conference on Multimedia*, 259-262.

Kim, J., & André, E. (2008). Emotion recognition based on physiological changes in music listening. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2067-2083.

Koren, Y., Bell, R., and Volinsky, C. (2009). Matrix factorization techniques for recommender systems. Computer, 42(8), 30–37.

Kreibig, S. D., Wilhelm, F. H., Roth, W. T., & Gross, J. J. (2007). Cardiovascular, electrodermal, and respiratory response patterns to fear- and sadness-inducing films. *Psychophysiology*, *44*, 787-806. doi:10.1111/j.1469-8986.2007.00550.x

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: Chicago University Press.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International affective picture system (IAPS): Technical manual and affective ratings. *NIMH Center for the Study of Emotion and Attention.*

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry*, *44*, 1248-1263.

Lange, C. G., & James, W. (1922). The emotions (Vol. 4). Williams & Wilkins.

Langlois, T., Chambel, T., Oliveira, E., Carvalho, P., Marques, G., & Falcão, A. (2010). VIRUS: Video information retrieval using subtitles. *Proceedings of the 14Th International Academic Mindtrek Conference: Envisioning Future Media Environments*, 197-200.

Lazarus, R. S. (1984). On the primacy of cognition. *American Psychologist, 39*, 124-129.

LeDoux, J. (1996). *The emotional brain.* New York: Touchstone Ed.

Levenson, R. W., Ekman, P., & Friesen, W. V. (1990). Voluntary facial action generates emotion-specific autonomic nervous system activity. *Psychophysiology*, *27*, 363-384.

Lew, M., Sebe, N., Djeraba, C., & Jain, R. (2006). Content-Based multimedia information retrieval: State of the art and challenges. *Transactions on Multimedia Computing, Communications, and Applications TOMCCAP*, *2*.

Lichtenstein, A., Oehme, A., Kupschick, S., & Jürgensohn, T. (2008). Comparing two emotion models for deriving affective states from physiological data. *Affect and Emotion in Human-Computer Interaction. LNCS.*

Luneski, A., & Bamidis, P. D. (2007). Towards an emotion specification method: Representing emotional physiological signals. *Twentieth IEEE International Symposium on Computer-Based Medical Systems, 2007. CBMS'07*, 363-370.

Lund, A. M. (2001). Measuring usability with the USE questionnaire. Usability & UserExperience (the STC Usability SIG Newsletter), 8(2).

Maaoui, C., Pruski, A., & Abdat, F. (2008). Emotion recognition for human-machine communication. *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*, 1210 - 1215. doi:10.1109/IROS.2008.4650870

Madden, M. (2009). *The Audience for Online Video-Sharing Sites Shoots Up.* Online Report, retrieved (May 2010) from www.pewinternet.org

Malamed, C. (2009). *Visual Language for Designers: Principles for Creating Graphics That People Understand.* New York: Rockport Publishers.

Marston, A., Hart, J., Hileman, C., & Faunce, W. (1984). Toward the laboratory study of sadness and crying. *The American Journal of Psychology*, *97*, 127-131.

Martin, M. (1985). Induction of depressed mood in the laboratory. *The American Journal of Psychology*, *98*, 635-639.

Martinho, J., & Chambel, T. (2009). Colorsinmotion: Interactive visualization and exploration of video spaces. In Proceeding of Academic MindTRek'2009: Everyday life in the ubiquitous Era, in cooperation with ACM SIGCHI and SIGMM, ambient and Ubiquitous Media Track, pp. 190-197, Tampere, Finland.

Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion*, *23*, 209-237. doi:10.1080/02699930802204677

Metz, C. (1974). *Film language: a semiotics of the cinema*. New York: Oxford University Press

Money, A. G., & Agius, H. (2009). Analysing user physiological responses for affective video summarization. *Displays*, *30*, 59-70.

Münsterberg, H. (1970). *The Film: A Psychological Study.* New York: Dover Edition.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *Journal of the Acoustical Society of America, 93,* 1097-1108.

Neisser, U. (1976). *Cognition and Reality: Principles and Implications of Cognitive Psychology.*Oxford: Oxford University Press.

Nielsen, Jakob (1993). Usability Eengineering. San Francisco: Morgan Kaufmann.

Norman, D. A. (2004). *Emotional design: Why we love (or hate) everyday things.* New York: Basic Books.

Oatley, K., & Johnson-Laird, P. (1987). Towards a cognitive theory of emotions. *Cognition & Emotion, 1,* 29-50.

Oatley, K., Keltner, D., & Jenkins, J. M. (2006). *Understanding Emotions*. Oxford: Blackwell Publishers.

Oliveira, E., Benovoy, M., Ribeiro, N., & Chambel, T. (2011). Towards Emotional Interaction: Using Movies to Automatically Learn Users' Emotional States. In *Proceedings of Interact'2011*: "13th IFIP TC13 International Conference on Human-Computer Interaction", pp.152-161, Lisbon, Portugal, September 5-9, 2011.

Oliveira, E., Martins, P., & Chambel, T. (2011). iFelt: Accessing Movies Through Our Emotions. In *Proceedings of EuroITV'2011: "9th International Conference on Interactive TV and Video: Ubiquitous TV"*, in cooperation with ACM SIGWEB, SIGMM & SIGCHI, pp.105-114, Lisbon, Portugal, June 29-July 1, 2011. (Best Paper Award).

Oliveira, E., Ribeiro, N., & Chambel, T. (2010). Towards Enhanced Video Access and Recommendation Through Emotions. In *"Brain, Body and Bytes: Psychophysiological User Interaction" Workshop, at ACM CHI'2010*, Atlanta, GA, USA, Apr 10-15, 2010.

Oppenheim, A.V., and R.W. Schafer. Discrete-Time Signal Processing. Englewood Cliffs, NJ: Prentice Hall, 1989. Pgs. 311-312.

Pantic, M., & Rothkrantz, L. J. M. (1999). An expert system for multiple emotional classification of facial expressions. *IEEE International Conference on Tools with Artificial Intelligence.*

Perez, S. (2008). The best tools for visualization. Retrieved (June, 2008) from http://www.readwriteweb.com/archives/the_best_tools_for_visualization.php

Philippot, P. (1993). Inducing and assessing differentiated emotion-feeling states in the laboratory. *Cognition & Emotion*, *7*, 171-193.

Philippot, P., Baeyens, C., Douilliez, C., & Francart, B. (2004). Cognitive regulation of emotion: Application to clinical disorders. *The Regulation of Emotion*, 71-97.

Picard, R. (1999). Affective computing for HCI. *In Proceedings of HCI International (the 8th International Conference on Human-Computer Interaction) on Human-Computer Interaction: Ergonomics and User Interfaces-Volume I (pp.* 829-833). Hillsdale, NJ: Erlbaum Associates.

Picard, R. W. (1997). *Affective Computing.* Massachusetts: MIT Press.

Picard, R. W., & Wexelblat, A. (2002). Future interfaces: Social and emotional. *CHI 2002. Conference on Human Factors in Computing Systems*, 698-699.

Picard, R. W., Vyzas, E., & Healey, J. (2001). Toward machine emotional intelligence: Analysis of affective physiological state. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *23*(10), 1175-1191.

Pilászy, I., & Tikk, D. (2009). Recommending new movies: Even a few ratings are more valuable than metadata. In *Proceedings of the third ACM conference on* *recommender systems* (pp. 93-100)

Plantinga, C.R. (2009). *Moving viewers: American film and the spectator's experience*. Univ of California Press.

Plutchik, R. (1980). *Emotion: A psychoevolutionary synthesis*. New York: Harper.

Pollatos, O., Herbert, B. M., Matthias, E., & Schandry, R. (2007). Heart rate response after emotional picture presentation is modulated by interoceptive awareness. *International Journal of Psychophysiology*, *63*, 117-124.

Purcell. (2010). *The State of Online Video.* Retrieved (June, 2010) from http://www.pewinternet.org/Reports/2010/State-of-Online-Video.aspx

Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006a). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *Int J Psychophysiol*, *61*, 5-18. doi:10.1016/j.ijpsycho.2005.10.024

Rangayyan, R. M. (2002). *Biomedical Signal Analysis: A Case-Study Approach.*New York: Wiley.

Rao, R. K. A., & VIKRAM, K. Y. (2001). Decreased chaos and increased nonlinearity of heart rate time series in patients with panic disorder. *Autonomic Neuroscience: Basic & Clinical*, *88*, 99-108.

Reeves, B., & Nass, C. (1996). *How People Treat Computers, Television, and New Media Like Real People and Places.* Cambridge: Cambridge University Press.

Rimm-Kaufman, S. E., & Kagan, J. (1996). The psychological significance of changes in skin temperature. *Motivation and Emotion*, *20*, 63-78.

Rocha, T., & Chambel, T. (2008). Videospace: A 3D video experience. *Proceedings of Artech. 4th International Conference on Digital Arts, pp.305-310, Portuguese Catholic University, Porto, Portugal, November 7-8.*

Rottenberg, J., Ray, R. R., & Gross, J. J. (2007). Emotion elicitation using films. In J. A. Coan & J. J. B. Allen (eds.), *The Handbook of Emotion Elicitation and*

Assessment (pp. 9-28). London: Oxford University Press.

Russell, J. A. (1980). A circumplex model of emotion. *Journal of Personality and Social Psychology*, *39*, 1161-1178.

Scherer, K. R. (1999). Appraisal theories. In T. Dalgleish, & M. Power (Eds.). *Handbook of Cognition and Emotion* (pp. 637–663). Chichester: Wiley

Scherer, K. R. (2000). Psychological models of emotion. In J. Borod (Ed.). *The neuropsychology of emotion* (pp. 137–162). Oxford: Oxford University Press.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*, 695-729. doi:10.1177/0539018405058216

Scherer, K. R. & Zentner, K. R. (2001). Emotional effects of music: production rules. In P. N. Juslin & J. A. Sloboda (Eds). Music and emotion: Theory and research (pp. 361–392). Oxford: Oxford University Press.

Schroder, M., Baggia, P., Burkhardt, F., Pelachaud, C., Peter, C., & Zovato, E. (2011). Emotionml--An upcoming standard for representing emotions and related states. *Affective Computing and Intelligent Interaction (ACII), 1*, 316-325

Schroder, M., Devillers, L., Karpouzis, K., Martin, J., Pelachaud, C., Peter, C., Wilson, I. (2007). What should a generic emotion markup language be able to represent? *Lecture Notes in Computer Science*, *4738*, 440.

Schroder, M., Pirker, H., & Lamolle, M. (2006). First suggestions for an emotion annotation and representation language. *Proceedings of LREC*, 88-92.

Schröder, M., Wilson, I., Jarrold, W., Evans, D., Pelachaud, C., Zovato, E. and Karpouzis, K. (2008). What is most important for an Emotion Markup Language? *Proc. Third Workshop Emotion and Computing, KI 2008, Kaiserslautern, Germany*.

Schupp, H. T., Cuthbert, B. N., Bradley, M. M., Cacioppo, J. T., Ito, T., & Lang, P. J. (2000). Affective picture processing: The late positive potential is modulated by motivational relevance. *Psychophysiology*, *37*, 257-261.

Sharp, H., Rogers, Y., & Preece, J. (2007). *Interaction Design: Beyond Human-Computer Interaction.* Sussex: Wiley.

Shneiderman, B. (2004). Designing for fun: How to make user interfaces more fun. ACM Interactions 11, 5, 48-50.

Smeaton, A. F., & Rothwell, S. (2009). Biometric responses to music-rich segments in films: The cdvplex. *Biometric Responses to Music-Rich Segments in Films: The Cdvplex*, 162-168. doi:10.1109/CBMI.2009.21

Snyder, C. (2003). *Paper prototyping: The fast and easy way to design and refine user interfaces.* San Francisco: Morgan Kaufmann Publishers.

Snoek, & Worring. (2008). Concept-Based video retrieval. *Information Retrieval*, *2*, 215–322. doi:10.1561/1500000014

Soleymani, M.; Chanel, G.; Kierkels, J.; Pun, T. (2008/a), "Affective Characterization of Movie Scenes Based on Multimedia Content Analysis and User's Physiological Emotional Responses," *Multimedia, 2008. ISM 2008*. pp.228-235, 15-17.

Soleymani, M. S., Chanel, C. G., Kierkels, J. K., & Pun, T. P. (2009). Affective characterization of movie scenes based on content analysis and physiological changes. " *Int'l J. Semantic Computing,* v. 3, 2, 235-254.

Soleymani, M., Chanel, G., Kierkels, J., and Pun, T. (2008/b). Affective ranking of movie scenes using physiological signals and content analysis. In *Proceedings of the 2nd ACM workshop on Multimedia semantics* (MS '08). ACM, New York, NY, USA, 32-39.

Steinbeis, N., Koelsch, S., & Sloboda, J. A. (2006). The role of harmonic expectancy violations in musical emotions: Evidence from subjective, physiological, and neural responses. *Journal of Cognitive Neuroscience*, *18*, 1380-1393.

Thomas, N., Mulligan, M. & Wiramihardja, L. (2010). *The Future of Online Video. Online report retrieved (June, 2010) from www.forrester.com/rb/Research/future_of_online_video/q/id/53765/t/2*

Ware, C. (2004). *Information Visualization: Perception for Design.* San Francisco: Morgan Kaufmann Publisher.

Westermann, R., Spies, K., Stahl, G., & Hesse, F. W. (1996). Relative effectiveness and validity of mood induction procedures: A meta-analysis. *European Journal of Social Psychology*, *26*, 557-580.

William, J. (1950). *The principles of psychology: volume 2.* New York: Dover publications (First published in 1890)

Winkielman, P. & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation elicits positive affect. *Journal of Personality and Social Psychology, 81*, 989-1000.

Winters, A. (2005). Perceptions of body posture and emotion: A question of methodology. *The New School Psychology Bulletin*, *3*, 35-45.

Virzi, Robert A. (1992). Refining the test phase of usability evaluation: how many subjects is enough? Human Factors, 34(4), 457 - 468.

Xu, M., Luo, S., Jin, J. S., & Park, M. (2009). Affective content analysis by mid-level representation in multiple modalities. *Proceedings of the First International Conference on Internet Multimedia Computing and Service*, 201-207.

Yoo, H. -W., & Cho, S. -B. (2007). Video scene retrieval with interactive genetic algorithm. *Multimedia Tools Applications*, *34*, 317-336. doi:10.1007/s11042-007-0109-8

Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist, 35*, 151-175.

Zajonc, R. B. (1984). On the primacy of affect. *American Psychologist, 39*, 117-123.

Zhang, J. (2008). *Visualization for Information Retrieval.* Berlin: Springer-Verlag.

Zwaag, M., & Broek, E. (2010). *Guidelines for Biosignal-Driven HCI.* CHI Workshop on Brain, Body and Bytes: Psychophysiological User Interaction.

Appendix A

Interview Guidelines

Part 1

[User personal information and cinema habbits]

Dados do Entrevistador e do Entrevistado

Número da Entrevista:_____________________________________

Entrevistador:____

Como é feito o registo:_____________________________________

Sobre o entrevistado:

Idade:________

Sexo: F ☐ M ☐

Profissão: ____

Já viu filmes online? Sim ☐ Não ☐

Com que frequência vê filmes online?____

Que tipo de serviços online costuma utilizar Imdb ☐ mubi ☐ Vimeo ☐

Google ☐ outros?___________________________________

Gosta de cinema? Sim ☐ Não ☐

Que tipo de filmes gosta:

Cómicos ☐ Acção ☐ Drama ☐ Romance ☐ Terror ☐ Fantasia ☐ Animação ☐ Ficção científica ☐ Musicais ☐

Prefere os filmes típicos de Hollywood, ou filmes alternativos (europeus, chineses, japoneses, africanos, medio-oriente)?

Já chorou em filmes? Sim ☐ Não ☐

Costuma emocionar-se? Sim ☐ Não ☐

O filme que mais gostou? _______________________________________

Com que frequência vê filmes?___________________________________

Como descreve esse filme, em 3 palavras?

_____________ _______________ ______________

Conhece algum realizador? Tem algum preferido? ____________________

Conhece algum actor? Tem algum preferido? ________________________

Há algum tipo de filme que não goste de ver?________________________

Onde preferencialmente vê filmes?________________________________

Que temas prefere: política, histórias de amor, histórias de guerra, filmes reais, biografias?____

Part 2

[Tester briefing text, to inform user about the experience purpose]

Explicação sobre o sistema e o que se vai avaliar

O sistema iFelt é na sua essência um arquivo de filmes que incorpora um player e gere a informação emocional dos utilizadores, associada aos filmes. A aplicação armazena filmes e classifica-os por emoções (alegria, medo, tristeza, raiva e nojo) e dimensões (positivo, negativo, calmo, stressante) e permitindo pesquisar, navegar e visualizar de acordo com estas perspectivas.

(Para o avaliado) Pretendemos com esta experiência avaliar a usabilidade do sistema.

(Para nós) Com esta experiência pretende-se avaliar a utilidade, satisfação, facilidade de uso: acesso:pesquisa/navegação e percepção/visualização do sistema iFelt.

Appendix B

Xml implementation: iFelt Schema.xsd

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs = "http://www.w3.org/2001/XMLSchema"
 xmlns:iFelt = "/Users/oliveira/Dropbox/_Tese/XMLImplementation/iFelt.xsd">

<xsd:element name="MovieInformation" type="xsd:string"/>
                <xsd:complexType>
                    <xsd:sequence>
                        <xsd:element name="MovieId" type="xsd:id" use="required"/>
                        <xsd:element name="Title" type="xsd:string" use="required"/>
                        <xsd:element name="Director" type="xsd:string" use="required"/>
                        <xsd:element name="Country" type="xsd:string" use="required"/>
                        <xsd:element name="Year" type="xsd:string" use="required"/>
                        <xs:element name="Actors" maxOccurs="unbounded">
                          <xs:complexType>
                            <xs:sequence>
                              <xs:element name="name" type="xs:string"/>
                    </xs:sequence>
                </xs:complexType>
                </xs:element>
            </xsd:sequence>
                </xsd:complexType>
<xsd:complexType name="UserInformation" type="xsd:string"/>
                <xsd:complexType>
                    <xsd:sequence>
                        <xsd:element name="UserID" type="xsd:id" use="required"/>
                        <xsd:element name="Name" type="xsd:string" use="required"/>
                        <xsd:element name="NumberMovies" type="xsd:integer" use="required"/>
                        <xsd:element name="NumberNeighbors" type="xsd:integer" use="required"/>
                        <xsd:element name="Country" type="xsd:string" use="required"/>
                    </xsd:sequence>
    </xsd:complexType>
<xsd:complexType name="UserMovie">
        <xsd:annotation>
            <xsd:documentation>This information represents users classification on a specific
            movie</xsd:documentation>
        </xsd:annotation>
                <xsd:complexType>
                    <xsd:sequence>
                        <xsd:element name="UserID" type="iFelt:UserInformation"/>
                        <xsd:element name="MovieID" type="iFelt:MovieInformation"/>
                </xsd:sequence>
                </xsd:complexType>
</xsd:complexType>
<xsd:complexType name="ListUsers">
        <xsd:annotation>
            <xsd:documentation>This information represents a list of users</xsd:documentation>
        </xsd:annotation>
                <xsd:complexType>
                    <xsd:sequence>
                        <xsd:element name="UserID" type="iFelt:UserInformation" minOccurs="0"
                        maxOccurs="unbounded"/>
                </xsd:sequence>
                </xsd:complexType>
</xsd:complexType>
<xsd:complexType name="ListMovies">
        <xsd:annotation>
            <xsd:documentation>This information represents a list of movies</xsd:documentation>
        </xsd:annotation>
                <xsd:complexType>
                    <xsd:sequence>
                        <xsd:element name="MovieID" type="iFelt:MovieInformation"/>
                </xsd:sequence>
                </xsd:complexType>
</xsd:complexType>
<xsd:complexType name="LastEmotions">
        <xsd:annotation>
            <xsd:documentation>This information represents the emotions felt by a user in the
            last three movies watched</xsd:documentation>
        </xsd:annotation>
                <xsd:complexType>
                    <xsd:sequence>
                        <xsd:element name="Position" type="xsd:integer"/>
                        <xsd:element name="Emotion" type="xsd:string" use="required"/>
                        <xsd:element name="Title" type="xsd:string" use="required"/>
                </xsd:sequence>
                </xsd:complexType>
</xsd:complexType>

    <xsd:element name="Prefences" type="xsd:integer"/>

</xs:schema>
```

User.xml


```xml
<?xml version="1.0" encoding="UTF-8"?>
<emotionml xmlns="http://www.w3.org/2009/10/emotionml"
    category-set="/Users/oliveira/Dropbox/_Tese/XMLImplementation/vocabulary#iFeltEmotions"
    dimension-set="http://www.w3.org/TR/emotion-voc/xml#scherer-dimension"
    xmlns:iFelt="/Users/oliveira/Dropbox/_Tese/XMLImplementation/iFelt.xsd">


<info>
  <iFelt:UserInformation>
  <iFelt:UserID>XXXXXX</iFelt:UserID>
  <iFelt:name>MaryFond</iFelt:name>
  <iFelt:NumberMovies>123</iFelt:NumberMovies>
  <iFelt:NumberNeighbors>34</iFelt:NumberNeighbors>
  <iFelt:Country>Portugal</iFelt:Country>
  </iFelt:UserInformation>
</info>

<info>
  <iFelt:LastEmotions>
    <iFelt:Position>1</iFelt:Position>
    <iFelt>Curious</iFelt:Emotion>
  <iFelt:MovieTitle>24 Hour Party People</iFelt:MovieTitle>
  </iFelt:LastEmotions>
    <iFelt:LastEmotions>
    <iFelt:Position>2</iFelt:Position>
    <iFelt:Emotion>Involved</iFelt:Emotion>
  <iFelt:MovieTitle>Black Swan</iFelt:MovieTitle>
  </iFelt:LastEmotions>
    <iFelt:LastEmotions>
    <iFelt:Position>3</iFelt:Position>
    <iFelt:Emotion>Tender</iFelt:Emotion>
  <iFelt:MovieTitle>Breakfast in Pluto</iFelt:MovieTitle>
  </iFelt:LastEmotions>
</info>


<info>
  <iFelt:ListMovies>
    <iFelt:MovieID>MMMMMM</iFelt:MovieID>
    <iFelt:MovieID>XXXXX1</iFelt:MovieID>
    <iFelt:MovieID>XXXXX2</iFelt:MovieID>
    <iFelt:MovieID>XXXXX3</iFelt:MovieID>
    <iFelt:MovieID>XXXXX4</iFelt:MovieID>
  </iFelt:ListMovies>
</info>

<emotion>
        <category name="curious" value="0.5"/>
</emotion>

<emotion>
        <category name="involved" value="0.4"/>
</emotion>

<emotion>
        <category name="happy" value="0.3"/>
</emotion>

<emotion>
        <category name="inspired" value="0.2"/>
</emotion>

<emotion>
        <category name="sad" value="0.18"/>
</emotion>

<emotion>
        <category name="melancholic" value="0.1"/>
</emotion>


<emotion expressed-through="text">
<dimension name="arousal" value="0.3"/>
  </emotion>


</emotionml>
```

Movie.xml

```xml
<?xml version="1.0" encoding="UTF-8"?>
<emotionml xmlns="http://www.w3.org/2009/10/emotionml"
           category-set="http://www.w3.org/TR/emotion-voc/xml#big6"
           appraisal-set="http://www.w3.org/TR/emotion-voc/xml#scherer-appraisals"
           iFelt-set="/Users/oliveira/Dropbox/_Tese/XMLImplementation/vocabulary.xml"
           xmlns:iFelt="/Users/oliveira/Dropbox/_Tese/XMLImplementation/iFelt.xsd">

<info>
    <iFelt:MovieInformation>
        <iFelt:MovieID>MMMMMM</iFelt:MovieID>
        <iFelt:Title>Cashback</iFelt:Title>
        <iFelt:Director>Sean Ellis</iFelt:Director>
        <iFelt:Year>2009</iFelt:Year>
        <iFelt:Actors>
            <iFelt:Actor>Tom B</iFelt:Actor>
            <iFelt:Actor>Ann A</iFelt:Actor>
        </iFelt:Actors>
    </iFelt:MovieInformation>
</info>

<info>
    <iFelt:ListUsers>
        <iFelt:UserID>XXXXXX</iFelt:UserID>
        <iFelt:UserID>XXXXX0</iFelt:UserID>
        <iFelt:UserID>XXXXX1</iFelt:UserID>
        <iFelt:UserID>XXXXX2</iFelt:UserID>
        <iFelt:UserID>XXXXX3</iFelt:UserID>
        <iFelt:UserID>XXXXX4</iFelt:UserID>
    </iFelt:ListUsers>
</info>


<emotion expressed-through="text">
        <category name="curious" value="0.5"/>
</emotion>

<emotion expressed-through="biosignal">
        <category name="happy"/>
</emotion>

<emotion expressed-through="text">
        <dimension name="arousal" value="0.3"/>
</emotion>


<emotion expressed-through="text">
        <category name="involved" value="0.4"/>
</emotion>

<emotion expressed-through="text">
        <category name="happy" value="0.3"/>
</emotion>

<emotion expressed-through="text">
        <iFelt name="inspired" value="0.2"/>
</emotion>

<emotion expressed-through="text">
        <category name="sad" value="0.18"/>
</emotion>

<emotion expressed-through="text">
        <category name="melancholic" value="0.1"/>
</emotion>


<emotion expressed-through="text">
<dimension name="arousal" value="0.3"/>
</emotion>


</emotionml>
```

User-Movie.xml

```xml
<?xml version="1.0" encoding="UTF-8"?>
<emotionml xmlns="http://www.w3.org/2009/10/emotionml"
    category-set="http://www.w3.org/TR/emotion-voc/xml#everyday-categories"
           appraisal-set="http://www.w3.org/TR/emotion-voc/xml#scherer-appraisals"
           dimension-set="http://www.w3.org/TR/emotion-voc/xml#pad-dimensions"
           xmlns:ifelt="/Users/oliveira/Dropbox/_Tese/XMLImplementation/iFelt.xsd">


<info>
    <iFelt:UserInformation>
    <iFelt:UserID>XXXXXX</iFelt:UserID>
        <iFelt:name>MaryFond</iFelt:name>
        <iFelt:NumberMovies>MaryFond</iFelt:NumberMovies>
        <iFelt:NumberNeighbors>MaryFond</iFelt:NumberNeighbors>
        <iFelt:Country>MaryFond</iFelt:Country>
    </iFelt:UserInformation>
</info>

<info>
    <iFelt:MovieInformation>
    <iFelt:MovieID>MMMMMM</iFelt:MovieID>
        <iFelt:name>Cashback</iFelt:name>
        <iFelt:Director>Sean Ellis</iFelt:Director>
        <iFelt:Year>2009</iFelt:Year>
        <iFelt:Actors>
            <iFelt:Actor>Tom B</iFelt:Actor>
            <iFelt:Actor>Ann A</iFelt:Actor>
        </iFelt:Actors>
    </iFelt:MovieInformation>
</info>


<emotion  expressed-through="biosignal">
<reference uri="movieId.avi#t=105,110"/>
<category name="happy"/>
</emotion>

<emotion  expressed-through="biosignal">
<reference uri="movieId.avi#t=115,120"/>
<category name="curious"/>
</emotion>

<emotion expressed-through="text">
<category name="happy" value="0.7"/>
</emotion>

<emotion expressed-through="text">
<category name="curious" value="0.3"/>
</emotion>

<emotion expressed-through="text">
<dimension name="arousal" value="0.3"/>
</emotion>

<emotion expressed-through="biosignal">
<dimension name="valence" value="0.9"/>
</emotion>


</emotionml>
```

Vocabulary.xml

```xml
<vocabulary type="category" id="iFeltEmotions">
        <item name="happy" />
        <item name="amused"/>
        <item name="involved"/>
        <item name="inspired"/>
        <item name="tender"/>
        <item name="surprised"/>
        <item name="astonished" />
        <item name="curious"/>
        <item name="melancholic"/>
        <item name="bored"/>
        <item name="compassioned"/>
        <item name="fear"/>
        <item name="disturbed" />
        <item name="amused"/>
        <item name="scared"/>
        <item name="disgust"/>
        <item name="embarrased"/>
        <item name="anger"/>
        <item name="irritated" />
</vocabulary>
```