Article

Finding Emotional-Laden Resources on the World Wide Web

Kathrin Knautz [1],*****, **Diane Rasmussen Neal** [2]**, Stefanie Schmidt** [3]**, Tobias Siebenlist** [1] **and Wolfgang G. Stock** [1]

[1] Department of Information Science, Heinrich-Heine-University, D-40225 Düsseldorf, Germany;
E-Mails: Tobias.Siebenlist@uni-duesseldorf.de (T.S.); stock@phil-fak.uni-duesseldorf.de (W.G.S.)

[2] North Campus Building, Room 240, Faculty of Information and Media Studies, The University of
Western Ontario, Ontario, N6A 5B7, Canada; E-Mail: dneal2@uwo.ca

[3] Dormagen, Germany; E-Mail: Steff_Schmidt@t-online.de

***** Author to whom correspondence should be addressed: E-Mail: Kathrin.Knautz@uni-duesseldorf.de.

Abstract: Some content in multimedia resources can depict or evoke certain emotions in users. The aim of Emotional Information Retrieval (EmIR) and of our research is to identify knowledge about emotional-laden documents and to use these findings in a new kind of World Wide Web information service that allows users to search and browse by emotion. Our prototype, called Media EMOtion SEarch (MEMOSE), is largely based on the results of research regarding emotive music pieces, images and videos. In order to index both evoked and depicted emotions in these three media types and to make them searchable, we work with a controlled vocabulary, slide controls to adjust the emotions' intensities, and broad folksonomies to identify and separate the correct resource-specific emotions. This separation of so-called power tags is based on a tag distribution which follows either an inverse power law (only one emotion was recognized) or an inverse-logistical shape (two or three emotions were recognized). Both distributions are well known in information science. MEMOSE consists of a tool for tagging basic emotions with the help of slide controls, a processing device to separate power tags, a retrieval component consisting of a search interface (for any topic in combination with one or more emotions) and a results screen. The latter shows two separately ranked lists of items for each media type (depicted and felt emotions), displaying thumbnails of resources, ranked by the mean values of intensity. In the evaluation of the MEMOSE prototype, study participants described our EmIR system as an enjoyable Web 2.0 service.

1. Introduction

Certain content in documents may represent or provoke emotions in their users. "Emotive" documents can be found as music, images and movies as well as text (e.g., poems, blog posts, *etc.*) The aim of Emotional Information Retrieval (EmIR) is to identify these emotions, to make them searchable, and thus to utilize them for retrieval in a variety of areas.

In this article, only a few of the many practical applications of this research are discussed. Emotional aspects play an important role in music therapy. There are connections between emotions expressed in the music and those in the patients which are being treated [1,2]. Adequate indexing of the emotions depicted in the music would help in tracking down appropriate songs and to improve therapies. Another application can be found in advertising; emotionally colored images and videos (including music) play a crucial role. The emotional content of images or videos has to fit the message of the advertisement [3]. Besides uses in therapy and advertising, private uses are possible as well. For example, someone may want to use an emotionally appropriate image for party invitations; someone else might be looking for emotionally appropriate music for a party. Through EmIR, the information needs of such users could be fulfilled satisfactorily.

A theoretical starting point of EmIR is "affective computing" [4], a concept related to human-computer interaction and research on information retrieval. Picard asks whether searching for emotions in retrieval systems is even possible. Are assessments of the users' feelings so consistent that they can be clearly assigned to documents and also made searchable? How can resources like images, videos and music become more accessible for retrieval systems? Information retrieval research distinguishes between two broad approaches:

- Concept-based information retrieval (with the aid of concepts and knowledge organization systems)
- Content-based information retrieval (only based on the low-level content of the document)

A time- and labor-independent approach is the automatic, content-based indexing of resources. For images, we can work with color, texture and shape. For videos, we can also work with scenes or shots, and with movement-like camera movement or recorded movement [5-7]. Content-based music is indexed according to pitch, rhythm, harmony, timbre, and—if available—lyrics. These components are all grouped under "low-level" features. Generally, these features describe metadata and basic attributes comprised of visual, auditory and textual components. But the information which can be deduced from these low-level features is limited, and therefore only partly suitable for an analysis of the semantic and emotional content. There are already approaches which try to extract emotional content from low-level features automatically. Some exemplary research projects are:

- MER (Music Emotion Recognition): [8-10],

- VER (Video Emotion Recognition): [11-13],
- IER (Image Emotion Recognition): [14-19].

Concept-based solutions are considered as promising alternatives to content-based approaches. In concept-based retrieval, users work with text-based terms such as keywords. In principle, these terms could even be extracted automatically from the content. Currently, they are primarily assigned intellectually. A first approach leads to knowledge organization systems (such as thesauri) and professional indexing. If there is a knowledge system for emotions, it would be (at least theoretically) possible to assign the feelings to the video intellectually. However, this method depends strongly on the indexer. A central problem of concept-based image, music and video retrieval is indexer inconsistency. This has been found to be particularly problematic in videos and images [20]. In a study investigating archivists' indexing of journalistic photographs contrasted with journalists' searching behaviors, "The output of the indexing process seemed to be quite inconsistent" [21]. Goodrum [22] reports that in the context of image retrieval, "manual indexing suffers from low term agreement across indexers (…), and between indexers and user queries." We do not yet have any experience with intellectual indexing of emotions. Ultimately, such an undertaking on the web exemplifies a practical problem: it is impossible for professional indexers to describe the millions of images, videos and pieces of music available online, most of which are user-generated.

The necessary time and manpower to describe all these documents argue in favor of user-generated tags (social tagging), as used in various Web 2.0 services [23]. Folksonomies are either "narrow" (only the author of the document may assign tags; for example, YouTube), "extended narrow" (authors and their friends may also attach tags, but every tag is only assigned once, e.g., Flickr) or "broad" (each user of the information service may assign tags, which can therefore be assigned several times to a document, e.g., Delicious or Last.fm).

There is an even greater problem with indexing inconsistency in narrow and extended narrow folksonomies. Only relatively few consistency problems occur in those tags that are assigned most often in broad folksonomies with a sufficiently large user base. However, the tags here do not come from any standardized vocabulary, but from the language of the user. Those users rarely use emotions to index resources.

Thus, our survey of the research ends with a rather negative result: at present, neither content- nor concept-based retrieval provides useful results for the emotional retrieval of images, videos and music on the internet.

2. Emotional Basics

In order to get a clear understanding of what "emotions" are, we can consult psychology discussions and research results. The research of emotions has a long history, but it has nonetheless so far proven impossible to agree on a common, exact definition of the concept of emotions. Although everyone knows what an "emotion" is, formulating a concrete terminology is rather difficult to accomplish [24]. Schmidt Atzert also concludes that no consensus has been reached concerning the meaning of "emotion" in [25].

According to Izard [26] a complete definition of "emotion" has to cover three aspects:

- the experience or conscious sensation of the emotion,

- the processes which are formed in the brain and the nervous system, as well as
- noticeable bodily reactions, especially facial expressions.

As early as 1981, Kleinginna and Kleinginna examined 100 statements and definitions from pertinent specialist books, dictionaries and introductory texts and tried to formulate a working definition: "Emotion is a complex pattern of changes which comprises physiological excitation, sensation, cognitive processes and behavior patterns that arise in response to a situation which the individual has perceived as important" [27].

Meyer, Reisenzein and Schützwohl also conduct their research on the basis of a working definition, since the definition itself is a central question within emotion research and is therefore not a prerequisite for scientific research of emotions, but rather its result [28]. But despite these factors, emotions can still be somewhat more closely characterized, according to Meyer, Reisenzein and Schützwohl in [28]: They are current mental states, and we can thus distinguish between them and other concepts like dispositions and moods. Emotions are further aimed at one or several specific—not necessarily actually existing—objects that cause said emotions. Another feature that the authors mention is that emotions show up in the so-called reaction triad consisting of subjective, behavioral, and physiological aspects. Emotions also have a certain quality, intensity and duration. The quality functions as a grouping feature, so that a situation like "*person A is happy*" can be categorized as belonging to the quality type *happiness*. The concrete occurrence of an emotion of a certain quality type is further defined by its duration as well as its intensity. According to Meyer, Reisenzein and Schützwohl, the strength of an emotion can be measured on a scale from low to medium to high intensity [28].

Different theories of emotion research with varying biases try to clarify how emotions are formed. Behavioristic approaches emphasize the preconditions that cause the emotion as well as the resulting behavior. Emotions are seen as innate (fear, anger and love) or develop as conditioned reaction patterns to certain stimuli [29].

Furthermore, there are evolutionary psychological approaches, which stress the adaptive functions of emotions (e.g., their survival function) that were in all likelihood caused by evolution. They can be traced back to Darwin [30], who wanted to prove the phylogenetic development of emotions, and have since been taken up by various emotion researchers [26,31-37].

Cognitive-physiological theories postulate that emotions are caused by the interaction between physiological changes. Humans develop and adapt their emotions through the (in) direct perception of those changes [38-40]. Attribution theories describe how humans try to understand and control their environment through the attribution of agency. According to these theories, emotions are reactions to action outcome [41,42].

In the course of our research, we had to consider which emotions should be included in the project and used for EmIR. Psychology researchers commonly reduce emotions to a small and relatively fixed number. These are called fundamental or basic emotions [43]. However, the exact number is still controversial, since the emotion researchers following differently oriented theories do not manage to reach a common agreement.

Current approaches to emotional music retrieval [44] as well as emotional image retrieval [45] work with five basic emotions: sadness, happiness, anger, fear, and disgust [46]. Following the

psychological literature, we decided to expand upon those basic emotions. The fundamental emotions for our project are thus: sadness, anger, fear, disgust, shame, surprise, desire, happiness and love. Although it is typically not included as a basic emotion in psychological literature, we have also added humor/wit in order to cover this highly media-related aspect.

We can take note of the following things:

1. The concept of emotions is difficult to define and thus no commonly accepted definition has yet been found. We try to approximate the meaning through the use of working definitions and features of emotions.
2. There are different approaches to categorize emotions in psychology which are founded on different psychological orientations and pursue differing goals based on differing assumptions.

3. Emotional Multimedia Indexing and Retrieval

To understand what EmIR and indexing means and to understand what kind of problems it entails, one has to be concerned with the basics of information retrieval. A retrieval system is "a system that is capable of storage, retrieval and maintenance of information" [47] and is designed to retrieve "information which may act as a supplement to human conscience and unconscious mental conditions in a given situation" [48]. Information indexing is a method of description and summarization of content in documents. It employs methods and tools of information filtering (e.g., folksonomy, nomenclature, classification system, thesaurus, ontology) and of summarization methods (abstracting and information extraction). Information encompasses all types of media and formats, which we are able to publish online. This includes textual content as well as audio-visual content (images, videos, music/sounds). In the Web 2.0 environment, we can find a multitude of user-generated content, e.g., images (Flickr), videos (YouTube), and music (Last.fm). Essentially, online information is multimedia-based content. However, current information retrieval systems are mostly only able to utilize metadata that appears in textual form, making search and retrieval of non-text information problematic.

Digital multimedia systems have the potential to offer users many possibilities for access modes. The inclusion of the emotional dimension in information indexing and retrieval is slowly becoming a central point of research. To this end, features extracted from the lower levels are often combined with new—intellectually facilitated or automated—methods in order to achieve an emotion-based annotation.

3.1. Emotional Image Indexing and Retrieval

Most existing approaches to emotional image and music retrieval and indexing have fallen into either the content-based or concept-based lines of thought. This research bifurcation [49] is unfortunate; not only does it allow for the best features of both approaches to work together [50], but each method exhibits its own weaknesses [45].

For these reasons, Lee and Neal [44] as well as Schmidt and Stock [45] advocate for a new approach that allows for the use of folksonomies as well as slide controls for "tagging basic emotions and their intensities". Their preliminary studies suggested that prototypical representations for basic emotions may exist, and that combining prototypes with scroll bars for indicating emotion intensity

may be a satisfactory method of retrieving documents by emotion. Tagging is a promising way to address the emotion-based subjectivity issues inherent in indexing non-text documents [44]. Knautz *et al.* stated in [51] that when people tagged images for emotion and intensity of said emotion for their study, they found it fun. This approach is further investigated in the current paper. While searching by emotion in digital image, musical, and video documents is a new concept to most people [51], the authors believe that it will be a welcomed method of retrieval once it is more thoroughly researched and implemented.

Although the "influence of the photograph's emotional tone on categorization has not been discussed much" [52], "images are to be considered as emotional-laden" [45]. Scholars from the art domain, such as Arnheim [53], have discussed the existence of emotional connotations in visual works, and information science scholars have observed the existence of emotions inherent in people's semantic descriptions of visual documents. For example, participants in Jörgensen's study [16] provided emotion-based descriptions such as "angry looking" and "worn out from too many years of stress". Conversational or uncontrolled responses to images in studies such as O'Connor, O'Connor, and Abbas [54], Greisdorf and O'Connor [55], and Lee and Neal [56] frequently included affective connotations. The difficulty, of course, is the highly subjective nature of a visual image, especially when higher levels of semantic understanding are used [57].

The emotional, subjective nature of a visual document becomes even more complex when the translation between text and image comes into play. While "pictures are not words and words are not pictures" [58], text and image definitely work together to convey the emotional meaning of an image [59,60]. Neal demonstrated that emotions within "photographic documents" containing emotion-based tags on Flickr are conveyed with an interplay between the image, the associated text (captions, titles, tags, *etc.*), and the social interaction provided by visitors to a Flickr photograph. Additionally, "happy" pictures are more common on Flickr than sad, angry, disgusting, or fearful pictures. Her model of text-based and image-based methods of emotion conveyance were tested in a follow-up qualitative study of pictures tagged with "happy" which demonstrated that photographs of smiling people and "symbolic" photographs (such as flowers, weddings, and natural beauty) may be the most common types of "happy" pictures on Flickr [61]. In this paper, she also suggests that serendipitous exploration, not just "search," may be needed for users to interact with emotions present in photographs. Yoon's small-scale study [62] provided promising results for the use of the language-independent Self-Assessment Manikin (SAM) in assigning emotional representations for images, despite subjectivity issues.

These points of variation noted, research noted in Jörgensen [63] suggests that affective responses to pictures are fairly universal, as demonstrated with psychological research such as studies involving the International Affective Response to Pictures (IAPS). Even if this is true, the significant increase in the number of non-text documents now available online makes it impossible for people, whether librarians or everyday users, to supply all the potential descriptors needed to represent associated emotions. For this reason, the trend in concept-based emotional image retrieval research is noteworthy. In most cases, this work involves tying low-level visual features such as lines, colors, and textures to their translations in art [64,65] and implementing these connections using computer algorithms. For example, "horizontal lines are always associated with static horizon and communicate calmness and relaxation; vertical lines are clear and direct and communicate dignity and eternality; slant lines are

unstable and communicate notable dynamism" [66]. Working from this perspective, Wang and Wang [67] note four issues in emotion-based semantic image retrieval: Performing feature extraction, determining users' emotions in relation to images, creating a usable model of emotions, and allowing users to personalize the model. These rules are widely noted in the content-based literature: "defining rules that capture visual meaning can be difficult" [68]. Dellandrea *et al.* [69] noted three issues in classifying emotion present in images: "emotion representation, image features used to represent emotions and classification schemes designed to handle the distinctive characteristics of emotions."

Methods of content-based emotional image retrieval implementation have included the Support Vector Machine approach [70,66], machine learning [71], fuzzy rules [72], neural networks [73,74,69], and evidence theory [69]. While the majority of these studies suggest promisingly effective results, the systems are very minimally tested by users before they are reported in research. Additionally, personal preferences and individual differences are not incorporated into the work.

3.2. Emotional Music Indexing and Retrieval

The existence of emotions inherent in listening to and creating music is undeniably widely experienced [75-77]. Based on the efficacy demonstrated in empirical research, music is frequently utilized for improving conditions associated with various illnesses [78,79], and is used often in advertising despite mixed research results [80]. Musicologists have identified various musical elements that convey particular emotions, such as the "descending tetrachord" element present in works such as "Dido's Lament" in Henry Purcell's opera "Dido and Aeneas" [81]. Hevner's frequently cited study [82] outlined sets of emotion-based adjectives as they relate to musical elements, such as major keys sounding "happy, merry, graceful, and playful," and dissonant music sounding "exciting, agitating, vigorous, and inclined toward sadness". However, "the notion of musical emotions remains controversial, and researchers have so far been unable to offer a satisfactory account of such emotions" [83]. Emotional response to music is a growing area of interest in psychology, computer science, and information science, although researchers in these complementary fields rarely cite each other.

Emotional music retrieval and indexing is a relatively new field that deserves much more research. Recent, modestly-sized information science studies involving tagging and music emotion have also concluded that it is difficult to agree on which emotions are represented in specific music. Lee and Neal [44] found that participants rating emotion present in music considered prototypical of basic emotions agreed more on "happy" music with a relatively high level of agreement, but music representative of other emotions represented little to no agreement. Neal *et al.* [84] received similar results, noting that harmonic, timbral, textual, and bibliographic facets of popular music may influence users' choice of emotion-based tags more than the pitch and temporal facets [85]. Hu and Downie [86] found relationships between artist and mood as well as genre and mood.

There has been little work to date on emotional music retrieval; most of what does exist has taken a content-based approach. Like content-based emotional image retrieval, content-based emotional music retrieval aims to link low-level musical features with commonly associated emotions. Huron [87] noted that while it is important to index musical mood, musical summarization and characterization are difficult tasks. Additionally, Chai and Vercoe [88] noted the importance of user modeling in music information retrieval, but also noted its difficulty, since different users have different responses to

music. Approaches to content-based emotional music retrieval have taken forms similar to its image counterpart. These include vector-based systems [89-91], music extractions with coefficient histograms [91], fuzzy logic [92], software agents [93], accuracy enhancement algorithms [94], interactive genetic algorithms [95], regression [96], emotional probability distribution [97], extraction by focusing on "significant" emotion expression [98], and emotion classification by ontology [99].

Lu and Tseng [100] present an interesting approach to music recommendation, combining "content-based, collaboration-based and emotion-based methods by computing the weights of the methods according to users' interests". Oliviera and Cardoso [101] present one of the few interdisciplinary studies that could be located on this topic, describing an affective music retrieval system based on a knowledge base grounded in music psychology research. Dunker *et al.* [102] outlined an automatic multimedia classification system; similar to the tagging studies described above, more consistency was found with images than with music. In addition to user subjectivity inherent in experiencing emotions while listening to music, the "multilabel classification problem" [91,103] signifies that more than one emotion can be attributed to one piece of music. Bischoff *et al.* [104] conducted a user study in Facebook to determine the quality of their mood-based automatically created tags, providing promising results. The use of both lyrics and audio in automatically extracted musical mood classifications may be more effective than audio alone [105]. Kim *et al.* [106] provide a thorough overview of automatic emotional music retrieval beyond what can be realistically provided in this article.

3.3. Emotional Video Indexing and Retrieval

In the field of emotion-based video indexing, the low-level features are comprised of the audio and visual realms. In this section, a few interesting research projects shall illustrate which methods may be used in order to obtain an emotional augmentation of content. Based on their approaches, the studies can be divided into two categories:

- Utilization of categorization models
- Utilization of dimension models

When using a categorization model, we work with preassigned emotion types with respect to basic categories. Thus Salway and Graham [12], for instance, use audio descriptions of films in order to extract the emotions of the actors. Based on the cognitive emotion theory, according to Ortony *et al.* [107], the narrative structures of 30 films were evaluated. Twenty-two emotion types, comprising a total of 679 tokens (semantically related keywords, grammatical variants, synonyms and hyponyms), were used for classification.

Besides the automatic extraction of the emotions, a validation test was performed in which participants performed an intellectual categorization for four films. Only the emotions that were actually present are shown in the figures, and they are augmented with temporal information. Salway and Graham concluded that all 30 films contained the emotion "fear" and many films also contained positive emotions towards the end of the films. The control test showed a relationship of 63%. The frequency of emotions allows, amongst other things, to draw conclusions concerning the type and climax of the film. Information drawn from dialogues, lighting and music was not incorporated and will—together with emotion intensity—form an additional exploratory focus in their continuing study.

When an emotional analysis of the contents is performed on the basis of a dimensional model, one can work with a two-dimensional model comprising arousal and valence of emotions, or with a three-dimensional model according to Russel and Mehrabian [108,109]. According to Mehrabian, the three-dimensional model is very similar to Wundt's behavioristic model [110], although it is not based on it. In detail, the (P-A-D) model represents the following dimensions:

- pleasure and displeasure (valence; analogous to Wundt's Lust and Unlust) = P
- arousal and nonarousal (intensity; analogous to Wundt's Erregung and Beruhigung) = A
- dominance and submissiveness (dominance; only roughly similar to Wundt's Spannung and Lösung) = D

While Wundt used his model to describe the course of an emotion, Mehrabian created his dimensional model in order to categorize and group emotions. Arifin and Cheung [111] have a similar understanding of the use of a dimensional model: "This model does not reduce emotions into a finite set, but attempts to find a finite set of underlying dimensions into which emotions can be decomposed." Thus, the goal is to identify and classify emotions in a space defined by intensity and valence.

Soleymani *et al.* [112] conducted a study to assess film scenes emotionally on the basis of a two-dimensional model. They combined the subjective impressions of the viewer (intensity and valence) with audio- and video-based low-level features and physiological examination (e.g., blood pressure, breathing, temperature, *etc.*) They used 64 scenes from eight movies of different genres (drama, horror, action, and comedy) which were shown to the participants in random order and with neutral intermediate elements. The subjective impressions of the participants were captured by the subjects themselves using sliders. Soleymani *et al.* found out that the subjective impressions sometimes differed greatly from the actual physical symptoms, which were consistent. All in all, the emotions that the participants felt fluctuated greatly and changed according to the participant's personality and mood. The extraction of multimedia features and their combination with physiological responses, however, is seen as an adequate method for the emotional analysis of video material by the authors based on their research results. Other studies which combine dimensional models with physiological examinations are, for example, Chanel, Ansari-Asl, and Pun [113], Chanel, Kronegg, Grandjean, and Pun [114] or Khalili and Moradi [115].

Hanjalic and Xu [116] also extracted low-level video- and audio-based features from videos (motion, shot length variation, sound energy and pitch) and dimensionally related these to valence and arousal of the user's emotion. In this way, they obtain emotion curves, which can be used to extract highlights in films and sports features (see also [117]).

Cui *et al.* [118] showed how the dimensional model of emotional indexing can be used for music videos. They envision a way to allow for more effective video retrieval than the traditional bibliographic search by using emotional indexing of videos to satisfy emotion-based user queries. They base their work on Zhangs *et al.* [119-121] in the area of emotion-based analysis of music videos. They extract emotion-based audio features and evaluate their correlation with intensity and valence. Additionally, participants rate some of the videos on a four point scale with respect to the two dimensions. Cui *et al.* could thus identify eleven features like short time energy, saturation and sound energy that correlated with the intensity. Eight features (e.g., tempo, short time energy and motion

intensity) correlated with the valence. Other studies which try to discover emotionally relevant audio features are, amongst others, Lu, Liu, and Zhang [122] and Ruxanda, Chua, Nanopoulos, and Jensen [123].

Chan and Jones [11] used 22 predefined emotion labels in their study and combined these with automatically extracted low-level audio features to index films emotionally. Chan and Jones used the results of Russell and Mehrabian for their work, and associated each measured point on their affect curve (see also [116]) with an emotion described by its valence and intensity. They used text statistics to index a film with its dominant emotions [124]. Based on his studies on the analysis of emotional highlights in films, Kang [125] worked with the intensity and valence model as well as with categorization. He trained two Hidden Markov Models (HMM) and was able to identify the emotions sadness, fear and joy in films [126]. Arifin and Cheung [111,127,128] showed how the dimensional model can be used to segment videos according to emotional concentration. For instance, in their study "A Computation Method for Video Segmentation Utilizing the Pleasure-Arousal-Dominance Emotional Information," they used an HMM in conjunction with a three-dimensional model to identify the emotions sadness, violence, neutral, fear, happiness and amusement.

Xu *et al.* [129] also used an HMM in their study to identify five predefined emotions, but they combined it with a fuzzy clustering algorithm for the analysis. The intensity levels were high, medium and low. The precision of emotion discovery was found to be higher in action and horror films than in dramas and comedies. Xu *et al.* concluded that fuzzy clustering is an appropriate method for identifying emotional intensity, and that a two step analysis offers a slight improvement in emotion identification.

The illustrated studies show that emotional indexing is not only possible in various ways, but also that research in the field has only just begun. Researchers currently work with a fixed set of emotions (categorization model) or on the basis of the intensity and valence of emotions (dimensional model). Some studies combine both models by trying to categorize emotions according to their dimensions. The extraction of appropriate low-level features relating to emotions forms an important starting point in many studies. Furthermore, the studies show that not only the emotions displayed in the video, but also the emotions that are experienced by the viewer, are objects for research. In this paper, we introduce an alternative way of indexing documents, which uses the theories and studies introduced in this section as a basis.

Our model (Figure 1) uses the knowledge gained from the different studies and consists of three blocks/constituents. Currently there is no satisfactory way to emotional index multimedia documents on a content- and concept-based level on the web. In conclusion, the different indexing approaches lead to the result that the use of broad folksonomies is the most appropriate method, with many different people indexing the same resource (Section 1).

In our indexing approach we use the results of research on emotional videos, since they include both aspects of image and music. This raises the question whether a categorical or dimensional-based approach should be used. We combine both models and use a solid set of ten emotions (category model) on the basis of emotion research (Section 2). Furthermore we give the opportunity to evaluate the intensity (model dimension) (Section 3).

Figure 1. Our model used for indexing.
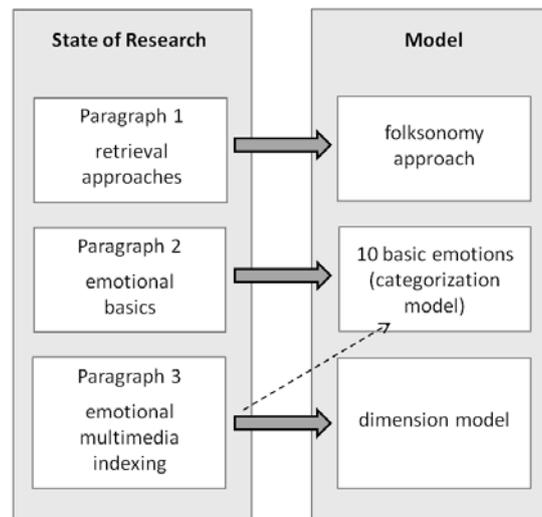
3.4. Slide Control Emotional Tagging: Consistent and Useful for a Search Engine?

If we want to construct a search engine for emotional-laden documents in the context of Web 2.0, we must be sure that users are able to tag basic emotions consistently, and that we are able to separate the appropriate emotions. The most useful method of indexing such documents on the web is to use a broad folksonomy, *i.e.* to let a multitude of different people index the same pieces of information (in this instance, images, videos, and music). The psychological studies previously described lead us to utilize ten basic emotions for this task. Furthermore, we know that emotions have intensities. Building on these three results, the following proved most workable in indexing instances of feeling:

- a controlled vocabulary for the ten basic emotions,
- slide controls to adjust the intensity per emotion (Scale: 0 [non-existent] to 10 [very strong],
- the broad folksonomy, thus letting different users tag the documents via the different slide control settings.

Previously, we conducted two empirical studies (with images in [45], and with videos in [51]) to answer three basic research questions:

- How consistent are user-oriented procedures for tagging emotions in multimedia documents?
- How many different users are needed for a stable distribution of the emotion tags? Are there stable distributions of the emotion tags at all?
- Can power tags corresponding to the emotions ("power emotions") be derived from the tag distributions?

Schmidt and Stock [45] selected 30 images from Flickr, which were presented to test persons in order to index them by type of emotion (disgust, anger, fear, sadness, happiness) and their intensity. 763 persons finished the slide control tagging of all images. In another study, Knautz *et al.* [51] replicated the experiment with videos and a set of nine basic emotions. A total of 776 people watched 20 films each (prepared YouTube documents) and tagged them via slide controls according to the different emotions. In contrast to Schmidt and Stock's study, Knautz *et al.* differentiated between

emotions expressed in the documents and emotions felt by the participants. To test the consistency of emotional votes, they calculated the standard deviation of the votes for all basic emotions per document, and then determined the arithmetic method of the standard deviations per emotion. The lower the value, the more consistently the subjects voted. Table 1 shows the results.

Table 1. Consistency of emotional voting (standard deviation).

Basic emotion	Images	Videos (expressed emotions)	Videos (felt emotions)
Love		0.88	0.66
Disgust	2.23	0.92	0.91
Sadness	3.07	1.11	1.15
Anger	2.46	1.21	1.15
Fun		1.46	1.24
Happiness	3.31	1.54	1.36
Desire		1.55	1.56
Fear	2.74	1.61	1.71
Surprise		1.94	1.75

Note: N = 776 participants (images); N = 763 participants (videos). Sources: [45,51].

What we see in image tagging is a gradation of consistency from disgust (with the best consistency) to happiness (with the lowest consistency). One finding of particular importance to these results is that people obviously agree more on disgust, anger, and fear than on sadness and happiness. In video tagging, both the expressed emotions and the experienced emotions showed relatively low values of standard deviation, and consequently high consistency values for the participants' estimations.

There is a very high consistency for the emotion of love in particular (with a median standard deviation of 0.88 for the expressed emotion and 0.66 for the experience), but disgust, sadness and anger are apparently also highly consensual. Surprise produced an interesting result as this emotion is (with a value of 1.94 for the expressed emotion) the most contentious of all in the video voting. Fun also leads to large differences in interpretation (the standard deviation being 1.75). We can answer the first research question in the positive: participants did indeed consistently index emotions in images and videos. This means that a central hurdle with regard to emotional retrieval has been overcome. It can be assumed that users allocate more or less the same basic emotions to the documents, with at least similar intensity.

The goal of the second research question was to find out how many users would be needed to lead to a stable distribution of the emotional tags. Afterwards, the number of slide control adjustments might change, but without any repercussions on the shape of the distribution curve. In order to find out how many people are needed to provide tagging stability, the numbers of participants for the respective documents (here, videos) are divided into eight blocks with n participants: 1st Block: n = 76, 2nd Block: n = 176 *etc.* up to the 8th Block: n = 776. For each of these blocks, the average slide control adjustment per emotion was calculated for each video. We will exemplify this procedure with the concrete example of video no. 20. The average slide control adjustment for anger (as an expressed basic emotion) was 6.3 for the first (randomly selected) 76 participants, 6.6 for 176 subjects, and so on, until the final value of 6.6 was reached with the number of 776 participants. The same

procedure was repeated for all other emotions. The evaluations make it clear that a stable distribution is reached as early as in the first block. Votes that break away from the arithmetic mean are "corrected" by the sheer number of users. This means that in emotional indexing practice, we do not require a large number of tagging users; if anything, the distribution of basic emotions "stands" with a few dozen users. But there is a "cold-start" problem: if there are only a few users providing tags, we are not able to separate the appropriate emotions.

There are three different forms of distribution that appear: *First*, the users find no predominant emotions at all, resulting in no distribution; *second*, there are Power Law distributions, where one emotion dominates all others; *third*, there are inverse-logistical distributions with several items heavily represented. The Power Law and the inverse-logistical distribution are typical curve progressions for rankings of informational objects [130].

Can we conclusively separate from among the distributions those basic emotions that provide the best emotional description of the video in question? To solve this problem, we draw on the concept of "power tags" [23,131]. "Power tags are tags that best describe the resource's content, or the platform's focal point of interest, according to Collective Intelligence … since they reflect the implicit consensus of the user community," as defined in [23]. We are interested only in power tags on the resource level. Peters [23] emphasizes that the restriction of a power tag search will increase that search's precision, as the document-specific "long tail" of tags will no longer be searched. As we have only ten basic emotions in total, it appears to make sense to regard all terms from the second-ranked one on down as the "long tail", and to cut them off accordingly, in the case of a Power Law distribution. Such a distribution follows the formula $f(x) = C/x^a$, where x is the rank position, C a constant and a a value between (roughly) 1 and 2. To make sure, we will assume a small value for a (at the moment, we calculate with 1). A curve is a Power Law precisely when the value of the item ranked second is equal to or less than half of the value of the item ranked first. In all other cases, the form of distribution is noted as "inverse-logistical" (this does not correspond to the prevalent opinion in the literature, which defines the turning point of the distribution as a threshold value [23], but is easily applicable). For inverse-logistical distributions, all emotions are noted whose intensity value is larger than/equal to the used threshold value (at present, we work with 4). Thus, up to three emotions could be indexed. Now we are able to answer the third research question. In the distribution of basic emotions according to a Power Law, exactly one emotion is filtered out, which is allocated to the resource as an emotional point of access. In the case of inverse-logistical distributions, several emotions (two or three) are attached to the resources. We must distinguish between expressed emotions (related to the documents) and experienced emotions (related to the users' feelings), as the users rate them differently as well.

4. Media EMOtion SEarch (MEMOSE)

MEMOSE is an attempt at comprising emotional-laden, user-generated image, music and video content within one information retrieval system. Based on the different types of EmIR, Knautz *et al.* [51] implemented a specialized search engine for fundamental emotions in several different kinds of emotional-laden documents called Media EMOtion SEarch (MEMOSE). The first approach covered only a limited set of images from Flickr. The further development and significant points of the system will be presented here. The retrieval of emotional-laden documents is the most important part of MEMOSE.

However, before these parts are described in detail, we will describe how data is put into MEMOSE, and how emotional values are assigned.

4.1. MEMOSE Indexing Sub-System

MEMOSE is a web application built upon a relational database, with PHP scripts for processing and visualizing data. As a basis for multimedia documents, MEMOSE can process data from Flickr (images), YouTube (videos) and Last.fm (music). To put a multimedia document into MEMOSE, the user just needs to know the URL of a document present in one of the three mentioned services.

Figure 2. Tagging an image with emotional tags.
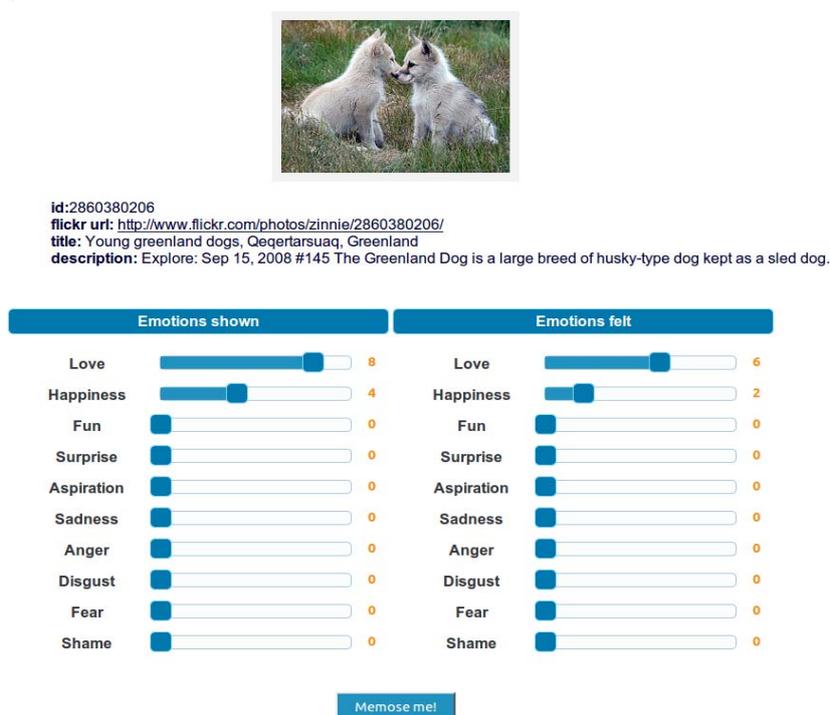
This URL is then checked for validity and processed further with help of the respective API in order to gather the title, tags, and other metadata. These data are stored in a relational database. The documents themselves are not stored in MEMOSE, but rather are accessed from their original websites. Because of this, a periodical check for availability is needed in order to ensure that there are no missing documents in the list of results. After adding a document, the user is asked to emotionally tag the respective document. For this purpose, we use slide controls to enable the user to value the document with emotions, as can be seen in Figure 2.

The user can see a preview of the document respective to its type: a thumbnail-size image, a video in an embedded player, or a music track in an embedded player. Below the document, the user can see the title of the document, and a group of 20 slide controls partitioned into two groups.

We differentiate between the emotion that is *expressed* (what is displayed) and *felt* (what the user feels while interacting with the document). Therefore, we display the ten basic emotions for each of these emotion types.

The range of every slide control contains values from 0 to 10 in order to get gradual results and to enable the user to be able to tag the document with a modifiable intensity. Since every user of MEMOSE can tag every multimedia document existing in the system, we use the approach of broad folksonomies. To avoid abuse of the system, each user can tag a document only once. A document can be retrieved only if a certain amount of users have tagged it. In order to encourage the users to tag documents which have not yet reached the necessary number of emotional tags, these documents are shown in random order with the request to tag these documents.

4.2. MEMOSE Retrieval Sub-System

The retrieval interface of MEMOSE consists of 10 checkboxes for each emotion, a text field for search terms, and a "Search" button. A screenshot of the search interface is shown in Figure 3.

Figure 3. The MEMOSE search interface.
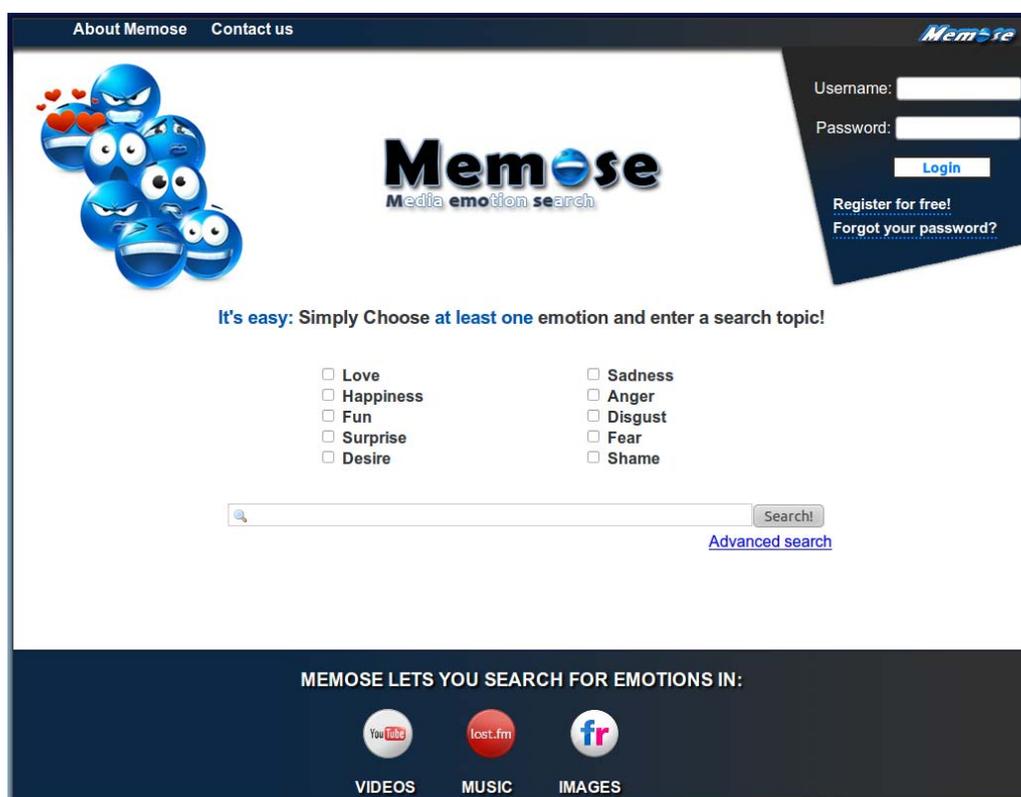
The user can search only for one or more emotions without entering search terms. This way the user will get the highest rated documents in descending order. If search terms are entered, these will be included in the query. The basis for search terms in a query are the tags retrieved from the mentioned web services. For every document, all assigned tags are stored in the MEMOSE database. The user's search terms are matched against these tags in order to get results for this query. As a supportive element, an auto-completion feature is used. The results are displayed in two tabbed slide controls, one for expressed and one for felt emotions. Every query is executed for all three media types. The three tabs correspond to the three media types. Inside the tabbed environment, four documents are displayed. Images are represented by thumbnails, and videos and music are represented by icons.

Below the title, the intensity of the chosen emotions and a button that leads to the tagging area for the respective document are displayed. The intensity of emotions is shown separately for every chosen emotion with a colorful scale as well as the arithmetic mean value of assigned emotional tags. An example for a result list is shown in Figure 4. By clicking the image or icons, an overlay opens and shows the document either as a larger image or in an embedded player together with the emotional scales and the "Memose me" button. Further results can be displayed using the arrow elements at the right end of the slider. Besides the normal search interface, there is another search interface for advanced users. With this advanced search interface, the user is able to choose which media types should be searched, and whether only felt, only expressed, or both options should be applied. Further text fields for a search for at least one of the given terms or to exclude terms from a query are available.

Figure 4. Search results in a tabbed slider environment.
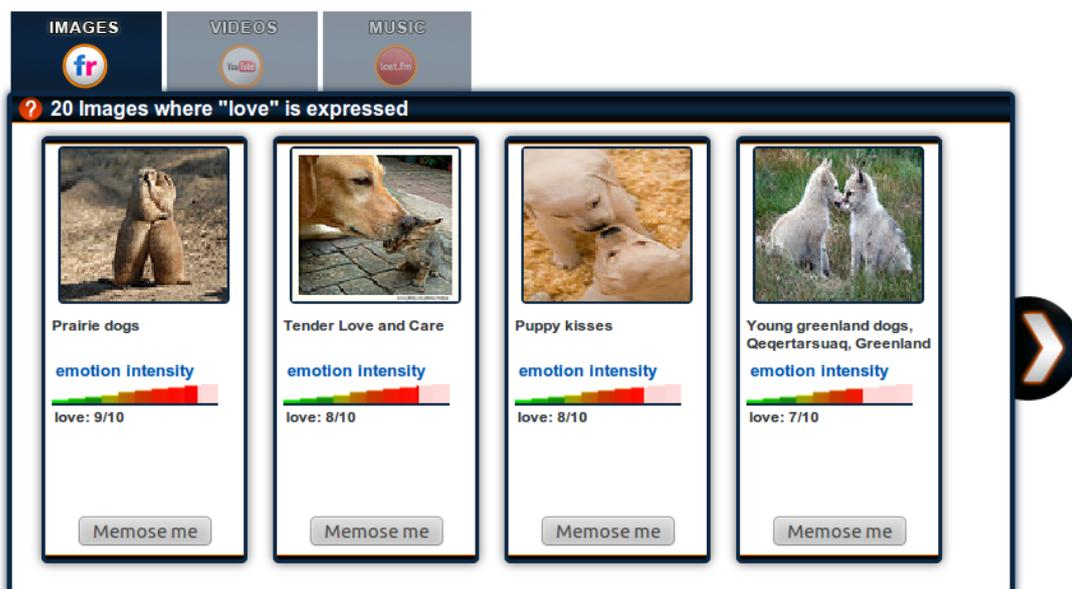
4.3. Evaluation

For the evaluation of MEMOSE, three different methods were utilized:

- IT-SERVQUAL [132-135];
- Customer Value Discovery [136];
- Usability [137].

For the IT-SERVQUAL method the participants were 50 undergraduate information science students. The same 50 participants were taken for the second method that acted as "users" and additional 7 graduate information science students who developed the MEMOSE prototype were taken as "developer" participants. The usability evaluation was divided into two parts. For a thinking-aloud test 13 regular Web 2.0 users aged 19 to 30 were used and for an online survey 74 participants with the same profile were taken.

As it does not appear useful to only work with "classical" indicators for a Web 2.0 service as MEMOSE, the evaluation is based on the tool box of evaluation methods that has been developed by

Knautz *et al.* [138]. The methods IT-SERVQUAL and Customer Value Discovery are applied to Web 2.0 services as well as the aspect of enjoyment. Knautz *et al.* [138] stated that fun is a very important usage aspect in Web 2.0 environments. Figure 5 describes the IT success model that was used for the evaluation. The following methods and models were developed using this model as a basis.

First, we utilized the SERVQUAL method. For this portion, two questionnaires with similar questions were designed in order to get information for the four dimensions of system quality, information quality, intent to use/perceived benefit, and user satisfaction. The basis for the contained questions was the information system success model from DeLone and McLean [139] and the revised version of this model from Jennex and Olfman [140], which have been combined into one model and used in our IT success model (see Figure 5).

Figure 5. IT success model used for evaluation.
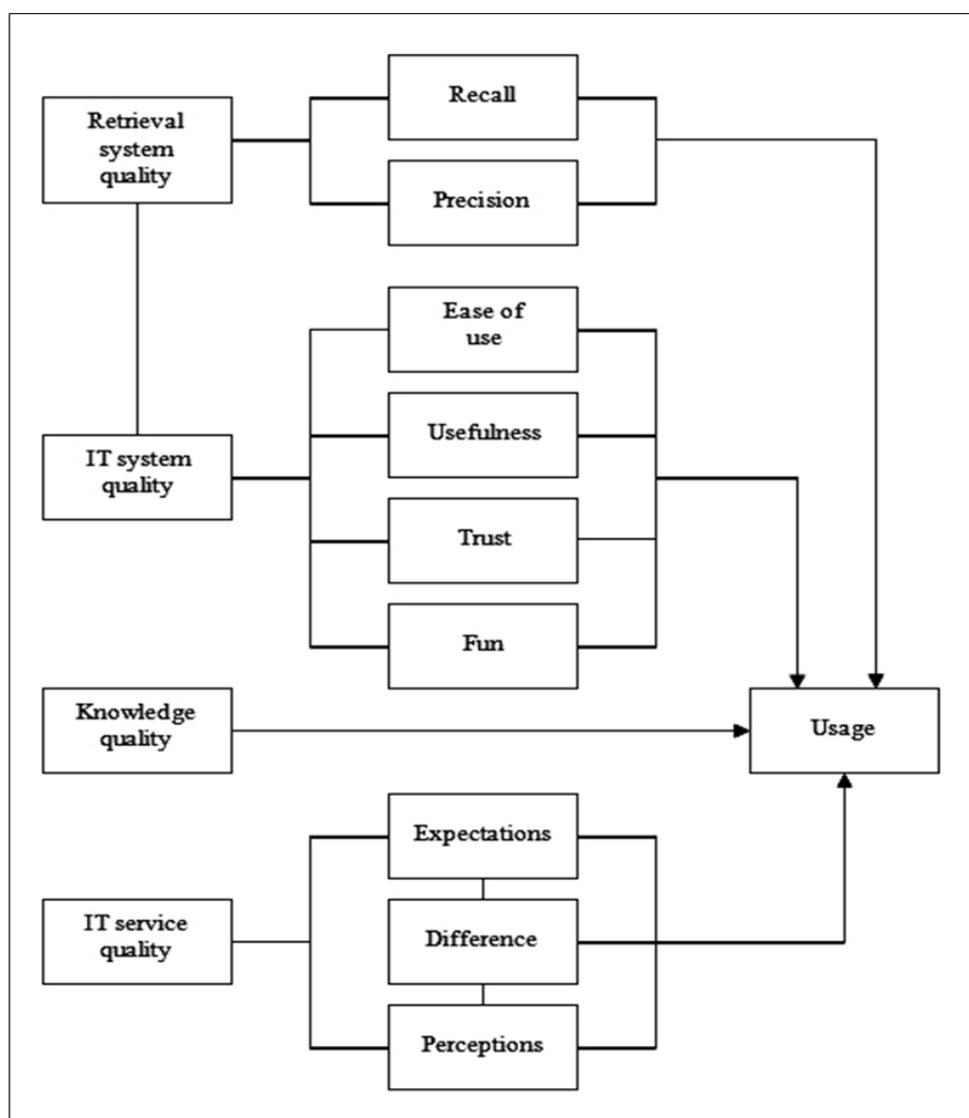
Both questionnaires had 22 statements, each containing a Likert scale with a range from 1 to 7. The first questionnaire presented statements about properties an optimal retrieval system should have with respect to ease of use, processing time, retrieval of relevant information, trust, stability, and so on. For example questions like "The search for emotions is an interesting feature for a retrieval system" or

"The search interface of an optimal retrieval system is easy to understand and easy to use". The questions on were changed on the second questionnaire so that the questions focused on MEMOSE. This questionnaire was given to 50 undergraduate information science students before they knew anything about MEMOSE and its purpose. Later, the participants were introduced to MEMOSE and were encouraged to use it. After having used MEMOSE for some time, the participants were asked to fill out the second questionnaire. This questionnaire contained the same kinds of statements as the first one, but the questions were aimed towards the MEMOSE prototype that the participants used. In order to get the overall rating for every question, the actual values for each question were summed, normalized by number of participants, and subtracted from the normalized sum of expected values. After doing this for every question, statements about MEMOSE's quality could be made. The results for the questions can be further aggregated regarding the four dimensions that should be analyzed, and this led to the following results.

For the dimension of system quality, the expected value is 5.6, not far from the actual value of 5.23. The users had high expectations regarding system quality, and the results for MEMOSE fulfill these expectations with only minor concerns. For the dimension of information quality, the expected value is 5.64, a little higher than the actual value of 4.83. This result shows that there is room for improvement in information quality; in other words, the users were not totally satisfied with the retrieved results. Regarding the dimension of intent to use/perceived benefit, we can draw a similar conclusion, with an expected value of 5.74 and an actual value of only 4.73.

The expectations for this dimension are high, and the actual condition is not able to meet this value, although it is above the mean value of 4. The last of the four analyzed dimensions is user satisfaction. For this dimension, we received an expected value of 4.56 and an actual value of 4.12. The results for this dimension are similar to the ones for the two dimensions before: the value of the actual condition is below the expected value but it is above the mean value of value, and the difference is rather small. One set of questions in which a considerable difference was detected is "I would use an optimal retrieval system again" *versus* "I would use this retrieval system again" with an expected value of 6.45 and an actual value of 3.36. This result clearly showed that the users were not pleased with the retrieval system as it existed at the time of this evaluation.

The Customer Value Discovery research method [136] was used to compare satisfaction regarding the main components of MEMOSE from two groups: users and developers. The "user" participants were the same as for the SERVQUAL method. The developers were seven graduate information science students who received the same questionnaire and were asked to fill it out in regards to their expectations of the users' evaluation results. The same questionnaire was used for both groups, and the groups did not know the other group's responses. The developers made a good estimate of how the users would rate the system. For 13 out of the 22 statements, the difference was below 1. The mean deviation is 0.79, whereas the maximum deviation is 1.96. Differences between 1 and 1.96 were found for the statements that applied to the operation of MEMOSE. The developers assigned lower values to both ease of use and relevant results, so they had lower expectations regarding these points. The participants rated these statements higher, and were more pleased with the search and the results they received. Overall, only small differences between the two groups could be measured, from which we concluded that the developers are aware of existing issues and suboptimal features. However, many of

the missing and suboptimal features could not be addressed because of the limited amount of time that was available to develop this system.

Furthermore, we administered a usability test with a task-based thinking-aloud protocol. In order to achieve results from quantitative and qualitative approaches, we combined three different methods. These methods are the concurrent thinking-aloud method (CTA), screen and voice capturing, and a survey after finishing the first two methods. Participants of the task-based thinking-aloud protocol included 13 regular Web 2.0 users aged 19 to 30 who were not information science students. For the online survey, there were 74 participants with the same profile.

The thinking-aloud participants were first introduced to what they should do, and it was made clear that the retrieval system is being tested, not the participants themselves. After that, the thinking-aloud test was performed. The screen and voice capturing was accomplished with the free software Wink. The participants got a small list of tasks they should do e.g., "Find all images with cats on them which are associated with disgust." In the testing room, only the participant and the instructor were present. After the thinking-aloud test was finished, the participants were asked to fill out a questionnaire containing 17 questions. The questions applied to the design of the system, the available search options, the retrieved results, the displayed emotional ratings, and to possible errors while using MEMOSE.

It was difficult for them to understand that several emotions could be chosen. The display of the emotions in the results was not universally recognized, so the ranking procedure was not clear to all participants. Regarding the free-text search, there was a problem with choosing the appropriate term; terms had to be entered as they existed in Flickr, YouTube, or last.fm. Neither a handling of singular and plural forms nor natural language processing could be utilized, although these would have proved useful. Additionally, a search for the term "pig" brought up results for pigs as well as for guinea pigs. This happened not because the term "pig" exists within the phrase "guinea pigs" but because the tag "pig" was given to the image. Since Web users are from countries all over the world, tags are not only present in English but in many other languages. Finally, the distinction between the sliders for expressed and felt emotions was not clear. The participants did not recognize the difference at first, and even after they saw the note regarding the meaning, it was not clear what they were expected to do. It became clear that another approach for the display of these two different groups of slide controls needed to be developed. There results were retrieved from the thinking-aloud test as well as from the online survey.

With about 60% the participants in the online survey were more pleased than the participants in the thinking-aloud test where 50% were pleased. Both groups of participants stated that the operation of MEMOSE was intuitive and easy to understand. The participants generally wondered why there was no option to get back to the starting page, where only the search interface without any results could be seen. 80% of the participants of the thinking-aloud test as well as of the online survey stated that the auto complete feature was very useful when typing in search terms. Regarding the interesting problem in the distinction of expressed and felt emotions, 50% of the participants in the thinking-aloud test found it reasonable to differentiate between these points after they understood the difference. In contrast a majority of 75% of the participants that took the online survey found this distinction reasonable. The remaining participants in both tests did not find it reasonable, mainly because they did not understand the distinction.

Generally speaking, the participants have high expectations for an optimal retrieval system that MEMOSE could not completely meet. The gap between the expected value and the actual value was (in the majority of cases) small. The actual value was above the mean value with a few exceptions, as previously noted. One reason for this difference includes the fact that the evaluated version of MEMOSE was the first iteration of a working prototype that had to be developed quickly. Also, the amount of available content was quite small, leading to similar results for different queries. The technical aspects were rated well, and the ability to search for emotions was described as an exciting task. However, it was not clear how this kind of retrieval system could be useful, and the majority of participants said they would not use it again. We suggest, however, that their hesitation might be due to MEMOSE's highly novel and therefore unfamiliar approach to searching; user acceptance may be higher after additional exposure to the concept of searching by emotion. A major usability issue was the display of the separated columns with the distinction between expressed and felt emotions. The participants stated clearly that this distinction was at first hard to understand. Regarding the small number of images in the database, the participants were not always satisfied with the retrieved results.

All of these results showed that MEMOSE was headed in the right direction, but some various issues prevented the participants from giving it a good rating. Based on these results, several changes were made to MEMOSE. First of all, a new design was established, making it easier to understand how the search works (see Figure 3 and Figure 6). The search options were moved to the middle of the screen.

Figure 6. First version of the MEMOSE search interface.

The options' positions no longer change when the results are displayed, as they did previously. The results display was a main area of criticism. In the evaluated version, there were two columns next to each other, with each column containing nine images, differentiated by the emotion shown or felt.

As it was difficult to understand this difference, our new approach put one result list below the other, and reduced the number of displayed images to four. The tabbed-slider environment that was presented in Section 3.2 replaced the previous results table (see Figure 7). Consequently, the difference between the two result lists was made clearer by a small amount of text that leads to additional text where the difference is explained. The search has been extended by an advanced search, where not

only emotions but also the media type and the emotion type (expressed *vs.* felt) can be selected, and where a search with the Boolean operators AND, OR and NOT is possible.

Figure 7. Retrieved results in a table environment in the first version of MEMOSE.

The measurement of retrieval system quality and knowledge quality are essential parts of an evaluation, but we have decided to perform such evaluation tasks at a later date because we primarily wanted to make sure that the basic dimensions of IT system quality and the IT service quality are accepted by the users. Thus not only the construction of the MEMOSE system but its evaluation as well is work in progress.

5. Conclusions

EmIR is a new research activity in computer science and information science. EmIR analyses the emotional content in all kinds of resources and tries to construct search engines for emotional-laden documents. In this article, we combined works from many fields. We reviewed applications of EmIR in music therapy, advertising, and private households. We noted that, at present, there is no useful system for the retrieval of emotions in images, videos and pieces of music. In order to develop an understanding of "emotion" we consulted psychological literature, and found that there is no commonly accepted definition of "emotion," but a working definition of "basic emotions" exists. We decided, based upon psychological research, to work with a set of ten basic emotions (sadness, anger, fear, disgust, shame, surprise, desire, happiness, love, and humor).

Our next step was an in-depth discussion of emotional indexing and retrieval of images, music and videos in computer science and in information science. It is possible to perform information retrieval using content-based (using low-level features, such as color or shape for image indexing) or concept-based (using concepts to index the resources' meaning) methods. Due to problems with both approaches, we decided to work with a new method, adding a controlled vocabulary (for the ten basic emotions), slide controls (to adjust the emotions' intensities) and broad folksonomies (so that multiple users can tag a resource emotionally).

With the aid of user tests, we found that users are able to tag basic emotions in a consistent way, that there are stable distributions of tags per document after a few dozen users tag it, and that algorithms exist to derive the "true" emotions from the tag distributions. Consequentially, we designed and coded a prototype of a web search engine for emotional content called Media EMOtion SEarch (MEMOSE). MEMOSE is made up of an indexing sub-system and a retrieval sub-system.

An evaluation of MEMOSE (done by means of IT-SERVQUAL, Customer Value Discovery and Usability) showed that MEMOSE's basic design is well-accepted by users, but also that there are some problems to be solved. For further research and development, it seems to be necessary to process the users' tags (stored by Flickr, YouTube and Last.fm) by means of natural language processing. At the beginning of every document's presence in the system, there is no tag distribution, so the system cannot attach an emotion. Therefore, we propose a combination of content-based and concept-based indexing, of low-level features (color, shape, *etc.*), and limited tagging from only a few users.

References

1. Gabrielsson, A. Emotion Perceived and Emotion Felt: Same or Different? *Musicae Sci.* **2002**, *6*, 123-147.
2. Schubert, E. Locus and Emotion: The Effect of Task Order and Age on Emotion Perceived and Emotion Felt in Response to Music. *J. Music Ther.* **2007**, *44*, 344-368.
3. Holbrook, M.; O'Shaughnessy, J. The Role of Emotion in Advertising. *Psychol. Market.* **1985**, *1*, 45-64.
4. Picard, R. *Affective Computing*; MIT: Cambridge, MA, USA, 1995.
5. Enser, P. The Evolution of Visual Information Retrieval. *J. Info. Sci.* **2008**, *34*, 531-546.
6. Enser, P. Visual Image Retrieval. *Annu. Rev. Inform. Sci.* **2008**, *42*, 3-42.
7. Smeaton, A. Techniques Used and Open Challenges to the Analysis, Indexing and Retrieval of Digital Video. *Inf. Syst.* **2007**, *32*, 545-559.
8. Han, B.; Rho, S.; Dannenberg, R.; Hwang, E. *SMERS:* Music Emotion Recognition Using Support Vector Regression. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, Kobe, Japan, October 2009; pp. 651-656.
9. Li, T.; Ogihara, M. Detecting Emotion in Music. *In Proceedings of the 4th International Society for Music Information Retrieval Conference (ISMIR 2003*), Baltimore, MD, USA, June 2003.
10. Yang, Y.; Su, Y.; Lin, Y.; Chen, H. Music Emotion Recognition: The Role of Individuality. In *Proceedings of the International Workshop on Human-Centered Multimedia*, Augsburg, Germany, September 2007; pp. 13-22.

11. Chan, C.; Jones, G. Affect-Based Indexing and Retrieval of Films. In *Proceedings of the 13th Annual ACM International Conference on Multimedia*, Singapore, November 2005; pp. 427-430.

12. Salway, A.; Graham, M. Extracting Information About Emotions in Films. In *Proceedings of the 11th ACM International Conference on Multimedia*, Berkeley, CA, USA, November 2003.

13. Xu, M.; Chia, L.; Jin, J. Affective Content Analysis in Comedy and Horror Videos by Audio Emotional Event Detection. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME 2005)*, Amsterdam, The Netherlands, July 2005; pp. 622-625.

14. Cho, S.; Lee, J. Wavelet Encoding for Interactive Genetic Algorithm in Emotional Image Retrieval. In *Advanced Signal Processing Technology by Soft Computing*; Hsu, C., Ed.; World Scientific: Singapore, 2001; pp. 225-239.

15. Duthoit, C.; Sztynda, T.; Lal, S.; Jap, B.; Agbinya, J. Optical Flow Image Analysis of Facial Expressions of Human Emotions—Forensic Applications. In *Proceedings of the 1st International Conference on Forensic Applications and Techniques in Telecommunication, Information, and Multimedia (e-Forensics 2008)*, Adelaide, Australia, January 2008.

16. Jörgensen, C. Retrieving the Unretrievable in Electronic Imaging Systems: Emotions, Themes, and Stories. In *Proceedings of SPIE—International Society for Optical Engineering*, San Jose, CA, USA, January 1999; pp. 348-355.

17. Wang, S.; Chen, E.; Wang, X.; Zhang, Z. Research and Implementation of a Content-Based Emotional Image Retrieval. In *Proceedings of the 2nd International Conference on Active Media Technology*, Chongqing, China, May 2003; pp. 293-302.

18. Wang, W.; Wang, X. Emotion Semantics Image Retrieval. A Brief Overview. *Lect. Notes Comput. Sci.* **2005**, *3783*, 490-497.

19. Wang, W.; Yu, Y.; Jiang, S. Image Retrieval by Emotional Semantics: A Study of Emotional Space and Feature Extraction. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '06)*, Taipei, Taiwan, October 2006; pp. 3534-3539.

20. Markey, K. Interindexer Consistency Tests: A Literature Review and Report of a Test of Consistency in Indexing Visual Materials. *Libr. Inf. Sci. Res.* **1984**, *6*, 155-177.

21. Markkula, M.; Sormunen, E. End-User Searching Challenges: Indexing Practices in the Digital Newspaper Photo Archive. *Inf. Retr.* **2000**, *1*, 259-285.

22. Goodrum, A. Image Information Retrieval: An Overview of Current Research. *Informing Sci.* **2000**, *3*, 63-67.

23. Peters, I. *Folksonomies: Indexing and Retrieval in Web 2.0*; De Gruyter Saur: Berlin, Germany, 2009.

24. Fehr, B.; Russell, A. Concept of Emotion Viewed from a Prototype Perspective. *J. Exp. Psychol.: General* **2005**, 113, 464-486.

25. Schmidt-Atzert, L. *Lehrbuch der Emotionspsychologie*; Kohlhammer: Stuttgart, Germany, 1996.

26. Izard, C. *Die Emotionen des Menschen. Eine Einführung in die Grundlagen der Emotionspsychologie*; Psychologie Verlags Union: Weinheim, Germany, 1994.

27. Kleinginna, P.; Kleinginna, A. A Categorized List of Emotion Definitions, with Suggestions for a Consensual Definition. *Motiv. Emot.* **1981**, *5*, 345-379.

28.   Meyer, W.; Reisenzein, R.; Schützwohl, A. *Einführung in die Emotionspsychologie. Band I: Die Emotionstheorien von Watson*, *James und Schachter*; Verlag Hans Huber: Bern, Switzerland, 2001.

29.   Watson, J.B. *Behaviorism*; University of Chicago Press: Chicago, IL, USA, 1930.

30.   Darwin, C. *The Expression of the Emotions in Man and Animals*; John Murray: London, UK, 1872.

31.   McDougall, W. *An Outline of Abnormal Psychology*; Luce*:* Boston, MA, USA, 1926.

32.   Plutchik, R. A General Psychoevolutionary Theory of Emotion. In *Emotion: Theory*, *Research*, *and Experience*, *Vol. 1: Theories of Emotion*; Plutchik, R., Kellerman, H., Eds.; Academic: New York, NY, USA, 2006; pp. 3-33.

33.   Izard, C.E. *The Face of Emotion*; Appleton-Century-Crofts: New York, NY, USA, 1971.

34.   Izard, C. *Human Emotions*; Plenum Press: New York, NY, USA, 1977.

35.   Ekman, P.; Friesen, W.; Ellsworth, P. What Emotion Categories or Dimensions Can Observers Judge from Facial Behavior. In *Emotion in the Human Face*; Ekman P., Ed.; Cambridge University Press: New York, NY, USA, 1982; pp. 39-55.

36.   Tomkins, S. *Affect*, *Imagery*, *Consciousness*, *Vol. I: The Positive Affects*; Springer Publishing: New York, NY, USA, 1962.

37.   Tomkins, S. *Affect*, *Imagery*, *Consciousness*, *Vol. II: The Negative Affects*; Springer Publishing: New York, NY, USA, 1963.

38.   James, W. What is an Emotion? *Mind* **1984**, *9*, 188-205.

39.   Panksepp, J. Toward a General Psychobiological Theory of Emotions. *Behav. Brain Sci.* **1982**, *5*, 407-467.

40.   Frijda, N. *The Emotions*; Cambridge University Press: New York, NY, USA, 1986.

41.   Arnold, M. *Emotion and Personality*; Columbia University Press: New York, NY, USA, 1960.

42.   Weiner, B.; Graham, S. An Attributional Approach to Emotional Development. In *Emotions*, *Cognition*, *and Behavior*; Izard, E., Kagan, J., Zajonc, R., Eds.; Cambridge University Press: New York, NY, USA, 1984; pp. 167-191.

43.   Ortony, A.; Turner, T.J. What's Basic About Basic Emotions? *Psychol. Rev.* **1990**, *97*, 315-331.

44.   Lee, H.; Neal, D. Towards Web 2.0 Music Information Retrieval: Utilizing Emotion-Based, User-Assigned Descriptors. In *Proceedings of the 70th Annual Meeting of the American Society for Information Science and Technology*, Milwaukee, WI, USA, October 2007; pp. 732-741.

45.   Schmidt, S.; Stock, W. Collective Indexing of Emotions in Images. A Study in Emotional Information Retrieval. *J. Am. Soc. Inf. Sci. Technol.* **2009**, *60*, 863-876.

46.   Power, M.J. The Structure of Emotion: An Empirical Comparison of Six Models. *Cognit. Emot.* **2006**, *20*, 694-713.

47.   Kowalski, G.; Maybury, M.T. *Information Storage and Retrieval Systems: Theory and Implementation*, 2nd ed.; Kluwer Academic: Boston, MA, USA, 2000.

48.   Ingwersen, P. *Information Retrieval Interaction*; Taylor Graham, London, UK, 1992.

49.   Persson, O. Image Indexing—A First Author Co-citation Map. 2002. Available at http://www.umu.se/inforsk/Imageindexing/imageindex.htm (Accessed on 2 March 2011)..

50.   Neal, D. News Photographers, Librarians, Tags, and Controlled Vocabularies: Balancing the Forces. *J. Lib. Metadata* **2008**, *8*, 199-219.

51. Knautz, K.; Siebenlist, T.; Stock, W.G. MEMOSE: Search Engine for Emotions in Multimedia Documents. In *Proceedings of the 33rd International ACM SIGIR Conference on Research and development in information retrieval*, Geneva, Switzerland, July 2010; pp. 791-792.

52. Laine−Hernandez, M.; Westman, S. Image Semantics in the Description and Categorization of Journalistic Photographs. In *Proceedings of the American Society for Information Science and Technology*, Austin, TX, USA, November 2006.

53. Arnheim, R. *The Power of the Center: A Study of Composition in the Visual Arts*, rewritten ed.; University of California Press: Berkeley and Los Angeles, CA, USA, 1983.

54. O'Connor, B.C.; O'Connor, M.K.; Abbas, J.M. User Reactions as Access Mechanism: An Exploration Based on Captions for Images. *J. Am. Soc. Inf. Sci.* **1999**, *50*, 681-697.

55. Greisdorf, H.; O'Connor, B. Modelling What Users See When They Look at Images: A Cognitive Viewpoint. *J. Doc.* **2002**, *58*, 6-29.

56. Lee, H.J.; Neal, D. A New Model for Semantic Photograph Description Combining Basic Levels and User-Assigned Descriptors. *J. Inf. Sci.* **2010**, *36*, 547-565.

57. Shatford, S. Analyzing the Subject of a Picture: A Theoretical Approach. *Cataloging Classification Q.* **1986**, *6*, 39-62.

58. O'Connor, B.C.; Wyatt, R.B. *Photo Provocations*; Scarecrow Press: Lanham, MD, USA, 2004.

59. Adelaar, T.; Chang, S.; Lancendorfer, K.M.; Lee, B.; Morimoto, M. Effects of Media Formats on Emotions and Impulse Buying Intent. *J. Inf. Technol.* **2003**, *18*, 247-266.

60. Neal, D.M. Breaking In and Out of the Silos: What Makes for a Happy Photograph Cluster? In *Proceedings of the 2010 Document Academy Conference (DOCAM '10)*, Denton, TX, USA, March 2010.

61. Neal, D.M. Emotion-Based Tags in Photographic Documents: The Interplay of Text, Image, and Social Influence. *Can. J. Infom. Libr. Sci.* **2010**, *34*, 329-353.

62. Yoon, J.W. Utilizing Quantitative Users' Reactions to Represent Affective Meanings of an Image. *J. Am. Soc. Inf. Sci. Technol.* **2010**, *61*, 1345-1359.

63. Jörgensen, C. *Image Retrieval: Theory and Research*; Scarecrow Press: Lanham, MD, USA, 2003.

64. Arnheim R. *The Power of the Center: A Study of Composition in the Visual Arts*, rewritten ed.; University of California Press: Berkeley and Los Angeles, CA, USA, 1983.

65. Itten J. *The Art of Color: The Subjective Experience and Objective Rationale of Color*; 1973 ed.; Van Nostrand Reinhold: New York, NY, USA.

66. Wang, W.-N.; Yu, Y.-L.; Zhang, J.-C. Image Emotional Classification: Static *vs.* Dynamic. In *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, The Hague, Netherlands, October 2004, pp. 6407-6411.

67. Wang, S.; Wang, X. Emotion Semantics Image Retrieval: An Brief Overview. In *Proceeding of 1st International Conference* on *Affective Computing and Intelligent Interaction*, Beijing, China, October 22-24, 2005; Springer: Amsterdam, The Netherlands, 2005; pp. 490-497.

68. Colombo, C.; Del Bimbo, A.; Pietro, P. Semantics in Visual Information Retrieval. *IEEE MultiMedia* **1999**, *6*, 38-53.

69. Dellandrea, E.; Liu, N.; Chen, L. Classification of Affective Semantics in Images Based on Discrete and Dimensional Models of Emotions. In *Proceedings of the IEEE International*

Workshop on Content-Based Multimedia Indexing (CBMI), Grenoble, France, June 2010; pp. 1-6.

70. Wang, S.; Chen, E.; Wang, X.; Zhang, Z. Research and Implementation of a Content-based Emotional Image Retrieval Model. In *Proceedings of the Second International Conference on Active Media Technology*, Chongqing, China, May 29-31, 2003; pp. 293-302.

71. Feng, H.; Lesot, M.J.; Detyniecki, M. Using Association Rules to Discover Color-Emotion Relationshiops Based on Social Tagging. In *Proceedings of Knowledge-Based and Intelligent Information and Engineering Systems*, Cardiff, Wales, UK, September 2010; pp. 544-553.

72. Kim, E.Y.; Kim, S.; Koo, H.; Jeong, K.; Kim, J. Emotion-Based Textile Indexing Using Colors and Texture. In *Proceedings of Fuzzy Systems and Knowledge Discovery*, Changsha, China, August 2005, pp. 1077-1080.

73. Kim, N.Y.; Shin, Y.; Kim, E.Y. Emotion-Based Textile Indexing Using Neural Networks. In *Proceedings of the 12th International Conference on Human-Computer Interaction: Intelligent Multimodal Interaction Environments*, Beijing, China, July 2007; pp. 349-357.

74. Kim, Y.; Shin, Y.; Kim, S.; Kim, E.Y.; Shin, H. EBIR: Emotion-Based Image Retrieval. In the Digest of Technical Papers, International Conference on Consumer Electronics, Las Vegas, NV, USA, May 2009.

75. Meyer, L.B. *Emotion and Meaning in Music*, University of Chicago Press: Chicago, IL, USA, 1954.

76. Juslin, P.N.; Sloboda, J.A. *Music and Emotion: Theory and Research*; Oxford University Press: New York, NY, USA. 2001.

77. Hallam S. Music Education: The Role of Affect. In *The Oxford Handbook of Music Psychology*; Hallam, S., Cross, I., Thaut, M., Eds.; Oxford University Press: Oxford, UK, 2009.

78. Davis, W.B.; Gfeller, K.E.; Thaut, M.H. *An Introduction to Music Therapy: Theory and Practice*; Wm. C. Brown Publishers: Dubuque, IA, USA, 1992.

79. Riley, P.; Alm, N.; Newell, A. An Interactive Tool to Promote Creativity in People with Dementia. *Comput. Hum .Behav.* **2009**, *25*, 599-608.

80. Morris, J.D.; Boone, M.A. The Effects of Music on Emotional Response, Brand Attitude, and Purchase Intent in an Emotional Advertising Condition. *Adv. Consum. Res.* **1998**, *25*, 518-526.

81. Rosand, E. The Descending Tetrachord: An Emblem of Lament. *Musical Q.* **1979**, *65*, 346-359.

82. Hevner, K. Experimental Studies of the Elements of Expression in Music. *Am. J. Psychol.* **1936**, *48*, 246-268.

83. Juslin, P.N.; Västfjäll, D. Emotional Responses to Music: The Need to Consider Underlying Mechanisms. *Behav. Brain Sci.* **2008**, *31*, 559-575.

84. Neal, D.; Campbell, A.; Neal, J.; Little, C.; Stroud-Mathews, A.; Hill, S.; Bouknight-Lyons, C. Musical Facets, Tags, and Emotion: Can We Agree? *Proceedings of the iConference (iSociety: Research, Education, Engagement)*, Chapel Hill, NC, USA, 2009.

85. Downie, J.S. Music Information Retrieval. *Ann. Rev. Inform. Sci. Tech.* **2003**, *37*, 295-340.

86. Hu, X.; Downie, J.S. Exploring Mood Metadata: Relationships with Genre, Artist and Usage Metadata. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR'07)*, Vienna, Austria, September 2007.

87. Huron, D. Perceptual and Cognitive Applications in Music Information Retrieval. In *Proceedings of the 1st International Symposium on Music Information Retrieval, Plymouth*, MA, USA, October 2000.

88. Chai, W.; Vercoe, B. Using User Models in Music Information Retrieval Systems. In *Proceedings of the 1st International Symposium on Music Information Retrieval*, Plymouth, MA, USA, October 2000.

89. Li, T.; Ogihara, M. Detecting Emotion in Music. In *Proceedings of the 4th International Symposium on Music Information Retrieval, Baltimore*, MD, USA, October 2003; pp. 239-240.

90. Wang, M.; Zhang, N.; Zhu, H. User-Adaptive Music Emotion Recognition. In *Proceedings of the 7th International Conference on Signal Processing*, Beijing, China, August-September 2004; pp. 1352-1355.

91. Li, T.; Ogihara, M. Toward Intelligent Music Information Retrieval. *IEEE Trans. Multimed.* **2006**, *8*, 564-574.

92. Yang, Y.H.; Liu, C.C.; Chen, H.H. Music Emotion Classification: A Fuzzy Approach. In *Proceedings of the 14th Annual ACM International Conference on Multimedia*, Santa Barbara, CA, USA, October 2006; pp. 81-84.

93. Yang, D.; Lee, W.S. Disambiguating Music Emotion Using Software Agents. In *Proceedings of the 5th International Conference on Music Information Retrieval (ISMIR '04)*, Barcelona, Spain, October 2004; pp. 52-58.

94. Liu, C.C.; Yang, Y.H.; Wu, P.H.; Chen, H.H. Detecting and Classifying Emotion in Popular music. In *Proceedings of the 9th Joint Conference on Information Sciences (JCIS/CVPRIP)*, Kaohsiung, Taiwan, China, October 2006; pp. 996-999.

95. Cho, S.B. Emotional Image and Musical Information Retrieval with Interactive Genetic Algorithm. *Proc. IEEE* **2004**, *92*, 702-711.

96. Yang, Y.H.; Lin, Y.C.; Su, Y.F.; Chen, H.H. A Regression Approach to Music Emotion Recognitions. *IEEE Tran. Audio Speech Lang. Proc.* **2008**, *16*, 448-457.

97. Wu, T.L.; Jeng, S.K. Probabilistic Estimation of a Novel Music Emotion Model. In *Proceedings of Advances in Multimedia Modeling*, Kyoto, Japan, January 2008; pp. 487-497.

98. Wu, T.; Jeng, S. Extraction of Segments of Significant Emotional Expressions in Music. In *Proceedings of the Workshop on Computer Music and Audio Technology*, Taipei, Taiwan, March 2006.

99. Han, B.; Rho, S.; Jun, S.; Hwang, E. Music Emotion Classification and Context-Based Music Recommendation. *Multimedia Tools Appl.* **2010**, *47*, 433-460.

100. Lu, C.C.; Tseng, V.S. A Novel Method for Personalized Music Recommendation. *Expert Syst. Appl.* **2009**, *36*, 10035-10044.

101. Oliveira, A.; Cardoso, A. Modeling Affective Content of Music: A Knowledge Base Approach. In *Proceedings of the Sound and Music Computing Conference*, Berlin, Germany, 2008.

102. Dunker, P.; Nowak, S.; Begau, A.; Lanz, C. Content-Based Mood Classification for Photos and Music: A Generic Multi-Modal Classification Framework and Evaluation Approach. In *Proceedings of the 1st ACM International Conference on Multimedia Information Retrieval*, Vancouver, Canada, October 2008; pp. 97-104.

103. Trohidis, K.; Tsoumakas, G.; Kalliris, G.; Vlahavas, I. Multi-label Classification of Music into Emotions. In *Proceedings of the 9th International Symposium on Music Information Retrieval*, Philadelphia, PA, USA, September 2008; pp. 325-330.

104. Bischoff, K.; Firan, C.S.; Nejdl, W.; Paiu, R. How Do You Feel about Dancing Queen? Deriving Mood & Theme Annotations from User Tags. In *Proceedings of the 9th ACM/IEEE-CS Joint Conference on Digital Libraries*, Austin, TX, USA, June 2009; pp. 285-294.

105. Hu, X.; Downie, J.S. Improving Mood Classification in Music Digital Libraries by Combining Lyrics and Audio. In *Proceedings of the 10th Annual Joint Conference on Digital Libraries*, Brisbane, Australia, June 2010; pp. 159-168.

106. Kim, Y.E.; Schmidt, E.M.; Migneco, R.; Morton, B.G.; Richardson, P.; Scott, J.; Speck, J.A.; Turnbull, D. Music Emotion Recognition: A State of the Art Review. In *Proceedings of the 11th International Society for Music Information Retrieval Conference*, Utrecht, Netherlands, August 2010; pp. 255-266.

107. Ortony, A.; Clore, G.L.; Collins, A. *The Cognitive Structure of Emotions*; Cambridge University Press: New York, NY, USA, 1988.

108. Russell, J.; Mehrabian, A. Evidence for a Three-Factor Theory of Emotions. *J. Res. Pers.* **1977**, *11*, 273-294.

109. Mehrabian, A. Framework for a Comprehensive Description and Measurement. *Genet. Soc. Gen. Psychol. Monogr.* **1995**, *121*, 339-361.

110. Wundt, W. *Vorlesungen über die Menschen und Tierseele*; Voss: Hamburg, Germany, 1906.

111. Arifin, S.; Cheung, P.A. Computation Method for Video Segmentation Utilizing the Pleasure-Arousal-Dominance Emotional Information. In *Proceedings of the 15th International Conference on Multimedia*, Augsburg, Bavaria, Germany, September 2007; pp. 68-77.

112. Soleymani, M.; Chanel, G.; Kierkels, J.; Pun, T. Affective Ranking of Movie Scenes Using Physiological Signals and Content Analysis. In *Proceedings of the 2nd ACM Workshop on Multimedia Semantics*, Vancouver, Canada, October 2008; pp. 32-39.

113. Chanel, G.; Ansari-Asl, K.; Pun, T. Valence-Arousal Evaluation Using Physiological Signals in an Emotion Recall. In *Proceedings of the IEEE SMC International Conference on Systems, Man and Cybernetics, Smart Cooperative Systems and Cybernetics: Advancing Knowledge and Security for Humanity*, Montréal, Canada, October 2007; pp. 2662-2667.

114. Chanel, G.; Kronegg, J.; Grandjean, D.; Pun, T. Emotion Assessment: Arousal Evaluation Using EEG's and Peripheral Physiological Signals. In *Proceedings of the International Workshop on Multimedia Content Representation, Classification and Security (MRCS)*, Istanbul, Turkey, September 2006; pp. 530-537.

115. Khalili, Z.; Moradi, M. Emotion Recognition System Using Brain and Peripheral Signals: Using Correlation Dimension to Improve the Results of EEG. In *Proceedings of the 2009 International Joint Conference on Neural Networks*, Atlanta, GA, USA, June 2009; pp. 1920-1924.

116. Hanjalic, A.; Xu, L.Q. Affective Video Content Representation and Modeling. *IEEE Trans. Multimedia* **2005**, *7*, 143-154.

117. Hanjalic, A. Extracting Moods from Pictures and Sounds: Towards Truly Personalized TV. *IEEE Signal Process. Mag.* **2006**, *23*, 90-100.

118. Cui, Y.; Jin, J.; Zhang, S.; Luo, S.; Tian, Q. Music Video Affective Understanding Using Feature Importance Analysis. In *Proceedings of the ACM International Conference on Image and Video Retrieval*, Xi'an, China, July 2010; pp. 213-219.

119. Zhang, S.; Huang, Q.; Tian, Q.; Jiang, S.; Gao, W. Affective mtv Analyses Based on Arousal and Valence Features. In *Proceedings of IEEE International Conference on Multimedia and Expo*, Hannover, Germany, June 2008; pp. 1369-1372.

120. Zhang, S.; Huang, Q.; Tian, Q.; Jiang, S.; Gao, W. i.mtv: An Integrated System for mtv Affective Analysis. In *Proceedings of the 16th ACM International Conference on Multimedia*, Vancouver, BC, Canada, October 2008; pp. 985-986.

121. Zhang, S.; Huang, Q.; Tian, Q.; Jiang, S.; Gao, W. Personalized mtv Affective Analysis Using User Profile. In *Proceedings of the 9th Pacific Rim Conference on Multimedia: Advances in Multimedia Processing*, Tainan, Taiwan, December 2008; pp. 327-337.

122. Lu, L.; Liu, D.; Zhang, H.J. Automatic Mood Detection and Tracking of Music Audio Signals. *IEEE Trans. Audio Speech Lang. Proc.* **2006**, *14*, 5-18.

123. Ruxanda, M.M.; Chua, B.Y.; Nanopoulos, A.; Jensen, C.S. Emotion-based Music Retrieval on a Well-Reduced Audio Feature Space. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Taipei, Taiwan, April 2009; pp. 181-184.

124. Luhn, H. The Automatic Creation of Literature Abstracts. *IBM J. Res. Dev.* **1958**, *2*, 159-165.

125. Kang, H.B. Analysis of Scene Context Related with Emotional Events. In *Proceedings of the 10th ACM International Conference on Multimedia*, Juan les Pins, France, December 2002; pp. 311-314.

126. Kang, H.-B. Affective Content Detection Using HMMs. In *Proceedings of the Eleventh ACM International Conference on Multimedia*, Berkeley, CA, USA, November 2003; pp. 259-263.

127. Arifin, S.; Cheung, P. User Attention Based Arousal Content Modeling. In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, Atlanta, GA, USA, 2006; pp. 433-436.

128. Arifin, S.; Cheung, P. Affective Level Video Segmentation by Utilizing the Pleasure-Arousal-Dominance Information. *IEEE Trans. Multimedia* **2008**, *10*, 1325-1341.

129. Xu, M.; Jin, J.; Luo, S; Duan, L. Hierarchical Movie Affective Content Analysis Based on Arousal and Valence Features. In *Proceedings of the 16th ACM International Conference on Multimedia*, Vancouver, BC, Canada, October 2008; pp. 677-680.

130. Stock, W.G. On Relevance Distributions. *J. Am. Soc. Inf. Sci. Technol.* **2006**, *57*, 1126-1129.

131. Peters, I.; Stock, W.G. "Power Tags" in Information Retrieval. *Library Hi. Tech.* **2010**, *28*, 81-93.

132. Parasuraman, A.; Zeithaml, V.A.; Berry, L.L. SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. *J. Retailing* **1988**, *64*, 12-40.

133. Pitt, L.F.; Watson, R.T.; Kavan, C.B. Service Quality: A Measure of Information Systems Effectiveness. *MIS Q.* **1995**, *19*, 173-187.

134. van Dyke, T.P.; Kappelman, L.A.; Prybutok, V.R. Measuring Information Systems Service Quality: Concerns on the Use of the SERVQUAL Questionnaire. *MIS Q.* **1997**, *21*, 195-208.

135. Pitt, L.F.; Watson, R.T.; Kavan, C.B. Measuring Information Systems Quality: Concerns for a Complete Canvas. *MIS Q.* **1997**, *21*, 209-221.

136. McKnight, S. Improving Customer Satisfaction: Changes as a Result of Customer Value Discovery. *Evid. Base. Lib. Info. Prac.* **2008**, *3*, 33-52.

137. Nielsen, J.; Tahir, M. *Homepage Usability: 50 Websites Deconstructed*; New Riders Publishing: Indianapolis, IN, USA, 2001.

138. Knautz, K.; Soubusta, S.; Stock, W.G. Tag Clusters as Information Retrieval Interfaces. In *Proceedings of the 43th Annual Hawaii International Conference on System Sciences (HICCS-43)*, Koloa, Kauai, Hawaii, January 5-8, 2010.

139. DeLone, W.H.; McLean, E.R. The DeLone and McLean Model of Information Systems Success: A Ten-Year Update. *J. Manag. Inform. Syst*. **2003**, *19*, 9-30.

140. Jennex, M.E.; Olfman, L. A Model of Knowledge Management Success. *Int. J. Knowl. Manag.* **2006**, *2*, 51-68.