Kathrin Knautz

Chapter 14. Emotion felt and depicted: Consequences for multimedia retrieval

Abstract: The great increase of multimedia documents in collaborative services like YouTube, Flickr and Last.fm brings with it the challenge of indexing and retrieving these documents. One aspect which has come into the focus of information science research is capturing emotions in videos, pictures and music. The goal is to identify emotions and enable the retrieval of documents based on these emotions. Research in this area is, however, still in the early stages and an application of the concept in the World Wide Web has yet to be implemented. Other fields, like psychology, film and musicology, on the other hand, pay great attention to the emotional description of films, images and music. In order to attain satisfying results regarding the capture of the emotional impact of multimedia documents and make these captured emotions retrievable, an interdisciplinary approach is necessary. The following chapter provides a look at the fundamentals of emotion research and applies them to the identification of emotions in multimedia documents. In the process, it becomes clear that it is necessary to distinguish between emotions *depicted* in the medium and emotions *felt* by the person viewing the medium. This has lasting consequences for indexing emotions and the retrieval of emotional content in multimedia documents, since it requires us to differentiate between these two types of emotions at both stages. We are using the search engine MEMOSE to exemplify how such a differentiation between depicted and felt emotions can be implemented.

Keywords: Emotional information retrieval, multimedia documents, image, video, music, felt emotion, depicted emotion, cognitive theory, appraisal process, commotion, tagging, scroll bar, Media EMOtion Search; MEMOSE

Kathrin Knautz, Research Associate and PhD Student, Department of Information Science, Heinrich-Heine-University Düsseldorf, Kathrin.Knautz@uni-duesseldorf.de

1 Introduction

Emotions are central phenomena of human life, since they occur frequently and are often connected to personally significant events (Lazarus, 1991). They are a

vital part of human communication and interaction. People take an active part in the emotions of others by trying to understand what caused the emotion and by helping to overcome certain emotion. Plutchik (1962), an important emotion researcher, writes: "The emotions have always been of central concern to men. In every endeavour, in every major human enterprise, the emotions are somehow involved" (Plutchik, 1962, p. 3). It is thus not surprising that researchers from various fields show an interest in emotions. Especially in the field of psychology, emotion research has a long tradition. Attaining a definition of the concept of an emotion is, however, rather difficult. Fehr and Russell (1984) write very aptly: "Everyone knows what an emotion is, until one is asked to give a definition. Then, it seems, no one knows" (1984, p. 464).

We all seem to know which emotions are typical for ourselves and can easily list examples. The difficulty, however, lies in formulating necessary and sufficient conditions for emotions. In section 2.1, we will illustrate how definitions vary based on different psychological orientations. Since no consensus has been reached for a common definition, we will also elucidate the alternative practice of using a so-called working definition.

Because of the difficulties in defining what an emotion is, some researchers use the features of emotions (2.2) to approach the problem. Psychologists use various models for this, which can be combined. The approaches range from describing emotions through dimensions like valence and intensity to clustering procedures which assign similar emotion words like aversion, contempt and disgust to the same class. Concerning the general delimitation of emotions, moods, feeling states and affect, we would like to refer to the paper by Otto, Euler and Mandl (2000).

The applications of emotion-psychological research results are wide-ranging. While emotional aspects play a vital role in music therapy (Davis, Gfeller, & Thaut, 1992), for example, emotion-psychological research results are also used in advertising and marketing, since "media content can trigger the particular emotions and impulse buying behaviour of the viewer/listener/surfer" (Adelaar, Chang, Langendorfer, & Morimoto, 2003, p. 247) and emotions are quite important in brand perception and purchase intents (Morris & Boone, 1998).

Web users also feel a need for emotional documents. Thanks to the emergence of collaborative services like YouTube (videos), Last.fm (music) and Flickr (pictures), we have massive amounts of user-generated content. In order to make these resources retrievable, the content is indexed via tags by the users of the service. Hastings, Iyer, Neal, Rorissa, and Yoon (2007) write the following about collaborative services and indexing via tags: "Image and video sharing services such as Flickr and YouTube pose new challenges in multimedia information indexing and retrieval and demand dynamic set of solutions" (p. 1026).

One of these challenges is the identification of the emotional content of the documents. Recent studies (Bischoff, Firan, Nejdl, & Paiu, 2008, 2010) show that if we assign categories like time, place or type to the various tags, the categories of tags that are assigned differ greatly from the categories of tags that are searched, especially in the area of opinions/qualities. 24 % of Flickr queries, for example, consist of emotional tags, but only 7 % of resources are indexed with emotional tags (Bischoff et al., 2010). We are thus confronted with the challenge of indexing emotions in multimedia documents and developing a system that facilitates their retrieval.

In order to solve this challenge, a closer examination of the media documents is necessary, because not all emotions are equal. An analysis of how emotions are formed when we see something (3.1) and what relation exists between felt emotions and depicted emotions (3.2) provides important clues regarding the indexation of media documents. In order to elucidate the emotional media impact, we will introduce the *component-process-model* by Scherer (1984) and illustrate its application to the process of watching/listening to videos, pictures and music. In these so-called appraisal models, emotions are the result of the subjective appraisal of a situation.

We will then integrate the emotion-psychological insights won from these considerations into a new approach to the emotional indexation of videos, pictures and music in a collaborative Web 2.0 environment (4.1). Using the search engine MEMOSE (Media Emotion Search) as an example, we will illustrate how a concrete realization of this approach might look (4.2).

2 Theoretical aspects of emotions

Below, we want to provide a first insight into the concept of emotions. To this end, we will first illuminate various possible definitions. It should be said that, as a result of the differing theoretic trends and changing scientific paradigms, psychological emotion researchers tend to prefer using an (incomplete) working definition. In the following, we will further consider the question of which emotions exist and how they can be differentiated.

2.1 The definition of emotion

Emotions play a vital role in social interaction and thus in the creation of social order. A definition of the term is, however, rather difficult. Emotion research has

a long history, but an exact, uniform definition for the concept of an emotion has yet to be produced. Depending on their research focus, researchers emphasize different aspects in their definitions. I will mention some of the most popular and influential definitions exemplarily.

Watson (1919, p. 165) said that "[a]n emotion is a hereditary pattern-reaction involving profound changes of the bodily mechanism as a whole, but particularly of the visceral and glandular systems". In his definition, Watson doesn't differentiate between an emotion and its reaction patter, which he understands to be inborn. William James (1884) says that "the bodily changes follow directly the PERCEPTION of the exciting fact, and that our feeling of the same changes as they occur IS the emotion" (p. 189). According to James, an emotion is thus a state which we experience, a bodily reaction that follows the perception of a specific stimulus.

In his two factor theory, Schachter (1964) also assumes that feeling bodily changes is necessary in order to experience emotions. He amends, however, that this experiencing state emerges from the integration of a feeling of arousal and a context-dependent cognition that matches the arousal. Bernard Weiner (1986), on the other hand, does not recognize an influence of physical arousal on emotions. According to Weiner, emotions are experiencing states with a positive or negative quality. They are caused by cognitive assessments that have motivational and informational consequences. Plutchik (1980) argues that the concept of emotion can be described through a combination of biological (ultimate), psychological, and physiological (proximate) explanations in his psycho-evolutionary emotion theory. Plutchik sees emotion as a syndrome that was created through natural selection in evolution, and is characterized by an experiencing state, physiological reaction, cognition, stimuli to act and observable behaviour.

These explanations and definitions show how different prevailing points of view have led to various emotion theories. These definitions only constitute a small part of the definitions of emotion found in the emotion-psychological specialized literature. Some notable newer approaches view emotions from a neurological point of view as neuro-physiological reactions (LeDoux, 1995) or as social constructs (Gergen, 1991) from the point of view of social constructionism.

An exact definition accepted by all thus does not exist. According to Otto, Euler and Mandl (2000), such a definition would require an exhaustive exploration of the subject area, something that has yet to be achieved in emotion research. Meyer, Schützwohl and Reisenzein (2001) point out that an exact definition of emotions is not the prerequisite, but rather the result of scientific analysis. Thus, research uses a so-called working definition, which should be as uncontroversial as possible and accepted by many researchers. It serves to describe a phenomenon and delimit the research area.

Kleinginna and Kleinginna (1981), for example, examined 100 statements and definitions from relevant specialized works, dictionaries and introductory texts and came up with the following working definition:

> Emotion is a complex set of interactions among subjective and objective factors, mediated by neural hormonal systems, which can (a) give rise to affective experiences such as feelings of arousal, pleasure/displeasure; (b) generate cognitive processes such as emotionally relevant perceptual effects, appraisals, labelling processes; (c) activate widespread physiological adjustments to the arousing conditions; and (d) lead to behaviour that is often, but not always, expressive, goal directed, and adaptive. (p. 355)

With this definition, Kleinginna and Kleinginna try to do justice to and combine all traditional, significant aspects of the different psychological theories. To summarize, according to this definition, an emotion is a complex pattern that is characterized by changes. Physiological arousal, feelings, cognitive processes and behaviour patterns are parts of this pattern, which occurs in situations that are significant for the individual. Notable newer working definitions put a greater emphasis on the cognitive component. As a central process, cognitive assessment is important not only for the creation of emotions, but also for continuous and recursive information processing. In the context of his Component Process Model (CPM), Klaus Scherer (1993) defines an emotion as

> an episode of temporary synchronisation of all major subsystems of organismic functioning represented by five components (cognition, physiological regulation, motivation, motor expression, and monitoring/feeling) in response to the evaluation of an external or internal stimulus event as relevant to central concerns of the organism. (p. 4)

According to this working definition, emotions function as continuous intermediaries between the organism and its environment. Cognitive processes play a vital role here, since they assess events in the environment and judge their significance for the individual.

Meyer, Reisenzein, and Schützwohl (2001) present another working definition. They define emotion as temporally dated, concrete occurrences which have certain attributes. In their view, emotions are current psychic states and thus distinguishable from other concepts like dispositions and moods. Emotions are also directed towards a specific – not necessarily existing – object with respect to the objects that cause said emotions. The authors further say that emotions reveal themselves as a so-called reaction triad consisting of subjective, behavioural and physiological aspects. They also have a specific quality, intensity and duration.

2.2 Classification of emotions

If we want to analyse emotions or use emotional aspects in the context of multimedia retrieval, it raises the question of which emotions exist and how they can be delimited. In the field of psychology, we can find three approaches to answer this question. Dimensional models try to quantitatively assess the features of emotions and map them to axes with respect to dimensions. When we analyse content emotionally based on a dimensional model, we can use either a 2-dimensional model, which models the arousal and valence of emotions, or a 3-dimensional model as proposed by Russell and Mehrabian (1977). Mehrabian's (1995) 3-dimensional model is very similar to the behaviouristic model proposed by Wundt (1906), although it is not based on Wundt's work. The three dimensions used in the (P-A-D-) model are as follows:

– pleasure and displeasure (valence; analogous to Wundt's *Lust* and *Unlust*)
– arousal and non-arousal (intensity; analogous to Wundt's *Erregung* and *Beruhigung*)
– dominance and submissiveness (dominance; approximately similar to Wundt's *Spannung* (tension) and *Lösung* (relaxation))

Whereas Wundt used his model to describe the progression of feelings, Mehrabian developed his dimensional model in order to categorize and group emotions. Arifin and Cheung (2007) hold a similar view of dimensional models: "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" (p. 69). We can see that the goal of dimensional models is to use a space defined by valence, dominance and intensity in order to identify and sort emotions.

Another method to identify and distinguish between emotions is the use of classes (category model). This method sorts emotion words according to similarity and uses statistics to group similar words into one category. This proves helpful for finding word fields that contain similar emotion words (Schmidt-Atzert & Ströhme, 1983). Another approach to categorize emotions which is sometimes used by psychologists is to reduce the number of emotions to a small, fixed number. These emotions are called fundamental emotions or base emotions. Advocates of the base emotion theory use psychological and/or (evolutionary) biological arguments for the existence of such fundamental emotions (Ortony & Turner, 1990). However, due to the differing research foci of emotion researchers, no consensus can be found for their number (Table 1). Furthermore, the existence of base resp. secondary emotions has been increasingly questioned in recent years (i.e. Ortony & Turner, 1990; Meyer, Reisenzein, & Schützwohl, 2003).

Fundamental Emotion	Basis for Inclusion	Reference
rage and terror, anxiety, joy	hardwired	Gray (1982)
expectancy, fear, rage, panic	hardwired	Panksepp (1982)
pain, pleasure	density of neural firing	Mowrer (1960)
anger, interest, contempt, disgust, distress, fear, joy, shame, surprise	unlearned emotional states	Tomkins (1984)
acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise	relation to adaptive biological processes	Plutchik (1980)
anger, aversion, courage, dejection, desire, despair, fear, hate, hope, love, sadness	relation to action tendencies	Arnold (1960)
anger, disgust, anxiety, happiness, sadness	do not require propositional content	Oatley & Johnson-Laird (1987)
anger, disgust, fear, joy, sadness, surprise	universal facial expressions	Ekman, Friesen, & Ellsworth (1982)
anger, disgust, elation, fear, subjection, tender-emotion, wonder	relation to instincts	McDougall (1908/1960; 1926)
anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise	hardwired	Izard (1971)
desire, happiness, interest, surprise, wonder, sorrow	forms of action readiness	Frijda (1986)
fear, grief, love, rage	bodily involvement	James (1884)
fear, love, rage	hardwired	Watson (1930)
happiness, sadness	attribution independent	Weiner & Graham (1984)

Table 1. Basic emotions (Ortony & Turner, 1990, p. 316).

3 Felt emotions and depicted emotions

Emotions can be found in all kinds of multimedia documents. However, the emotions depicted in media are not always the emotions that are aroused by said media. This raises the question of how these two types of emotions – depicted and felt – are related. We will begin our approach to the topic with an overview of the various emotion-psychological theories. These theories try to explain how emotions are formed from various points of view. Knowing how emotions are formed also means being able to influence people through depicted emotions. The influence emotions depicted in multimedia documents have on people shall be shown using a so-called appraisal model.

3.1 Emotions felt – theories in emergence

Different theories of emotion research try to explain how emotions are formed. We differentiate between

– behaviouristic theories of emotion;
– physiological-cognitive theories of emotion;
– evolution-psychological theories of emotion;
– attributional theories of emotion.

At this point we only want to provide a brief insight into the emergence of emotions from these diverse viewpoints.

(Classical) behaviourism

Behaviouristic approaches emphasize the behaviour aspect and the conditions that trigger emotions. Emotions are either seen as inborn (fear, anger and love) or as conditioned reaction patterns to specific stimuli (Watson, 1930). The most popular representative of this theory, Watson (1913), eschews any form of introspection. According to him, intra-subjective experience aspects are not scientifically acceptable, since they are not accessible to various independent observers.

Only inter-subjectively observable and measurable variables and stimuli can be used in a methodological approach. A felt emotion in the behaviourist theory is therefore only present if it is measurable; for example, through physiological reactions (perspiration, skin temperature, blood flow indicators). Voluntary disclosures in the form of utterances or questionnaires are used only as rough indications and must be complemented by accompanying independent behaviour monitoring and physiological measurements.

Specifically, Watson sees emotions as hereditary reaction patterns that occur when the triggering stimulus is presented (Watson, 1919). As basic reaction patterns, he postulates fear, anger, and love, which are modified by learning experience and stimulus substitution. Through classical conditioning, we acquire our own characteristic emotional reaction patterns in relation to environmental stimuli (e.g. Jones, 1924; Watson & Rayner, 1920).

Cognitive-physiological theories

In cognitive and physiological theories, it is postulated that emotions are determined through the interaction between physiological and cognitive changes.

Through the (in)direct perception of this change, people develop and adapt their emotions (James, 1884; Lange & Kurella, 1887). An emotion is here characterized as an experiencing state, the emergence of which is dependent on subjective feelings and bodily reactions. James (1884) describes the relationship as follows: "We feel sad because we cry, angry because we strike, afraid because we tremble, and neither we cry, strike, nor tremble because we are sorry, angry, or fearful, as the case may be" (p. 190). According to James, physiological reaction patterns (e.g. heartbeat, pulse rate), are reflexively triggered by the perception of an external stimulus (e.g. the appearance of a beautiful woman) specific to a certain emotion. The conscious perception of these physiological (motoric) changes is the emotion.

In summary we can say that, in contrast to Watson, a strong focus is on the subjective aspect. Physical changes are not the consequence, but the cause of the emotion. Due to much criticism James revised his theory ten years later. The new version held that not simply the perception of an object, but the assessment of the situation as a whole causes an emotional reaction. The belief that the origin of emotions is largely physiological, which has been held by a majority of cognitive-(physiological) theorists in the beginning, was gradually revised (e.g. Cannon, 1927; Marañon, 1924).

Theories of the 60s and 70s (Schachter, 1964; Valins, 1966; Zillmann, 1978) are therefore more strongly influenced by the cognitive component. Thus, although Schachter (1964) retains some of the basic assumptions of James, the perception of physiological changes is not sufficient for him and furthermore does not condition the emotion quality. Rather, cognition takes hold at this point. It conditions the quality of the emotion through an emotion-relevant assessment of the situation and causal attribution. The intensity of the emotion is determined by physiological arousal. Emotions can therefore be viewed in Schachter's theory as post-cognitive phenomena. Other studies such as studies with paraplegics (Hohmann, 1966) indicate that physiological arousal is not necessary for the emergence of emotions.

For this reason, recent research deals with so-called cognitive appraisal theories. These assume that emotions are a result of the interpretation and explanation of an incident. The lack of physical arousal is here acceptable (e.g. Frijda, 1986; Lazarus, 1995; Scherer, 1988, 1998a). Figure 1 demonstrates the emergence of emotions according to Scherer (1998b) in a simplified form. In the example, a person assesses the relevance of a perceived stimulus (e.g. an event or object) for their needs. This appraisal process includes aspects like the novelty of the stimulus, valence, purposefulness, certainty about the consequences, etc. Thus, the subjectively assessed importance of the situation is critical for triggering the emotion. The result of this process leads to a specific reaction pattern, which may be marked by physiological reactions, motor expressions, etc.
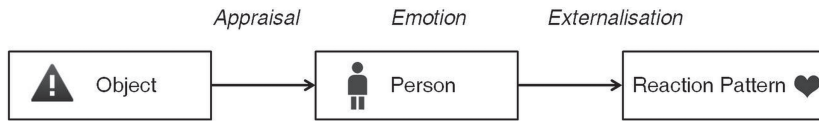
Figure 1. Emergence of emotions (according to Scherer (1998b)).

It has become increasingly established in research that emotions are based primarily on cognitions, appraisals and assessments. It is important, however, to conceptualize emotion and cognition in an appropriate manner. In the context of cognitive theories of emotion, attributional theories of emotion present a different concept by trying to accomplish this conceptualization.

Attributional theories of emotion

Attribution/attributional theories describe how humans try to understand and control their environment using causal attribution. According to these approaches, emotions are reactions to the results of these actions (e.g. Arnold, 1960; Weiner & Graham, 1984). Important suggestions regarding the causal explanation go back to Fritz Heider (1958). While attribution theories primarily deal with the formation of attributions, attributional theories are primarily concerned with the effects of already formed attributions.

According to Bernard Weiner (1986), a prominent representative of attributional theories, the emergence of emotions is a sequential cognitive information process. Before an emotion can be experienced, three cognitive steps must be completed upon the perception of the emotion-triggering event. In a first step, an assessment is carried out regarding the relation of the event to the achievement of objectives (event-driven emotions like happy, pleased, or satisfied). A second step is to ascribe the event to an originating factor like the person's own ability, effort or chance (attribution-specific emotions like surprise). Mapping the originating factor to dimensions is the last step (dimension-dependent emotions such as pride, shame, helplessness). These dimensions form
– personal dependencies: am I the cause or are other people the cause?
– stability: how long-lasting is the cause/origin? Talent or effort?
– controllability of the cause: controllable (effort), uncontrollable (aptitude).

Emotions ultimately result from the interaction of causal attribution and evaluation.

Evolution-psychological theories

Evolution-psychological theories emphasize the evolutionarily conditioned adaptive functions of emotions, such as their survival function. Besides proximate processes and their distal emergence, evolution-psychological approaches describe mainly ultimate (biological) functions. They go back to Darwin (1872), whose main concern was to demonstrate the evolutionary development of emotions. Darwin sees the evolutionary advantage of emotions in the fitness improvement (better perception, manipulation possibilities, effect on others). According to Darwin, emotions are innate mental states that occur automatically, depending on the situation. Cognitive assessments cause the emotional expression.

His studies were continued by many emotion researchers, who usually represent the most famous formulation of the evolutionary psychological position: the theory of discrete basic emotions (McDougall, 1926; Plutchik, 1980; Izard, 1971, 1977, 1994; Ekman et al. 1982; Tomkins 1962, 1963).

Among the most popular approaches is the psycho-evolutionary syndrome theory by American psychologist Robert Plutchik (1980). He postulates the existence of inherited dispositions towards eight fitness-enhancing adaptive behaviour patterns. These eight dispositions are the foundation of all other emotions (dyads or triads), and cannot be traced back to more fundamental emotions. Emotions are thus biologically and psychologically fundamental. Figure 2 portrays a simplified exemplary reaction sequence for the emergence of the emotion "fear".
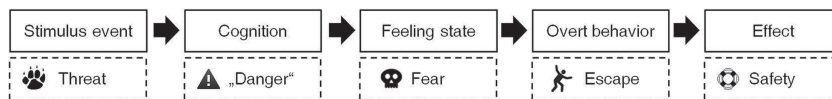

Figure 2. Chain reaction for the emotion "fear" (adapted from Plutchik, 1980).

The reaction sequence starts with a perceived stimulus which causes a threat, e.g., the emergence of a bear in a forest. The cognitive assessment of this event comes to the conclusion that a threat emanates from this animal and, through physiological reactions (increased autonomic activity), leads to the emotional state of "fear". The next step in Plutchik's sequential model is the activation of an action impulse, like escaping from the bear. This observable behaviour has the biological function of protection against threats.

3.2 Depicted emotions –
   emotions triggered by multimedia documents

How, then, do depicted emotions in multimedia documents influence people? Studies show that a significant number of emotions are caused by the representation of events in various media (Scherer, Wallbott & Summerfield, 1986; Cantor & Nathanson, 1996).

The question of the effect of media on the behaviour of recipients is as old as media itself. The fundamental paradigm of research on the emotional effects of media is the stimulus-response model. In this model, the outgoing stimuli of a communicator (media content) cause a reaction (response) in the recipient and possibly lead to emotional changes. Recipients are viewed as primarily passive – the motives that made them turn to the medium in the first place are neglected.

Modern models of media effects research are audience-centred approaches, which place the recipients and their media selection at the front. Thus, the *uses and gratification* approach (Katz & Foulkes, 1962), for example, tries to find out motives of media use.

The use of a medium is thus directed by the expected utility and in how far it will satisfy certain needs (Schenk, 1987). Emotional media impact – mood-management – goes back to Zillmann (1988). His theory on media selection is a specialization of the *uses and gratification* approach. The basic assumption is the innate desire for a positive emotional state that occurs through an average stimulation. An experience-based media selection occurs based on individual unconscious motives. Through the selection of media content, emotions arise in the recipient.

Regarding the question of how emotions caused in response to real events and emotions caused in response to media-mediated events differ, we would like to refer the reader to further reading relating to concepts such as *Sense of Reality*, *Law of Apparent Reality*, or *Perceived Reality* (Ortony, Clore & Collins, 1988; Frijda, 1988; Rothmund, Schreier & Groeben, 2001).

But how exactly are felt emotions and emotions depicted in multimedia documents related? For an explanation of media effects in the emotional sphere, appraisal models can be used (Mangold, Unz, & Winterhoff-Spurk, 2001; Schwab, 2001). In these models emotions are the result of a subjective assessment of the situation (see 3.1, Cognitive theories of emotion). With their peculiarities, component models integrate the triad of emotions derived from Izard's (1977) three component theory: the neurophysiological component, the subjective experience and the motor-expressive component. For the next steps, the *component-process-model* by Scherer (1984, 2001b) is used and tested for its use in a multimedia context.

In his *component-process-model*, Scherer postulates five subsystems, which are involved in the formation of an emotion (Table 2). These include

– cognitive processes (appraisal) due to the valuation of objects;
– physical reactions, produced in the neuroendocrine, autonomic and somatic nervous systems;
– motivational changes brought about by the appraisal process;
– facial and vocal expression and
– subjectively experienced emotional state.

Scherer explains the relations between the systems as follows: "The central assumption of the componential pattern theory is that the different organismic subsystems are highly interdependent and that changes in one subsystem will tend to elicit related changes in other subsystems" (Scherer, 2001b, p. 106).

Since a complete discussion of all components with reference to videos, music and images is not possible at this point, we will exemplarily discuss the appraisal component in greater detail below, since it is primarily responsible for triggering and differentiating between emotions. The cognitive component is intertwined with all other components through feedback loops (Brosch & Scherer, 2009) and is fundamental in order to illuminate the relationship between felt and depicted emotions.

Emotion function	Organismic subsystem	Emotion component
Evaluation of objects and events	Information processing (CNS)	Cognitive component (appraisal)
System regulation	Support (CNS, NES, ANS)	Neurophysiological component (bodily symptoms)
Preparation and direction of action	Executive (CNS)	Motivational component (action tendencies)
Communication of reaction and behavioural intention	Action (SNS)	Motor expression component (facial and vocal expression)
Monitoring of internal state and organism–environment interaction	Monitor (CNS)	Subjective feeling component (emotional experience)

Note: CNS = central nervous system; NES = neuro-endocrine system; ANS = auto-nomic nervous system; SNS = somatic nervous system.

Table 2. Relationships between organismic subsystems and the functions and components of emotion (Scherer, 2005, p. 698).

The appraisal component is used for information processing and includes a sub-jectively-assessed relevancy of the situation. Evaluation criteria (stimulus evalu-ation checks (SECs)) are relevance, implications, coping potential and normative significance (Scherer, 2001b). The following questions motivate these appraisals:

– relevancy: How relevant is this event for me? Does it affect me/my social reference group?
– implications: What are the implications or consequences of this event, how do they affect my well-being and my immediate or long-term goals?
– coping potential: How well can I deal with these consequences or adjust myself to them?
– normative significance: How important is this event in relation to my self-concept and social norms and values? (Brosch & Scherer, 2009, p. 197 [translated])

These information processing steps affect the individual sub-processes through complex multiple feedback and feed-forward processes. The result of this appraisal process is an emotion which is characterized by physiological symp-toms and movements ("motor-expressive movements") in the face, body and voice (Figure 3). When the mutual influence the subsystems exert on each other subsides, the emotional episode ends, according to Scherer (1984). Crucial for triggering the emotion is the subjective nature of the event for the actual motiva-tion of the organism.
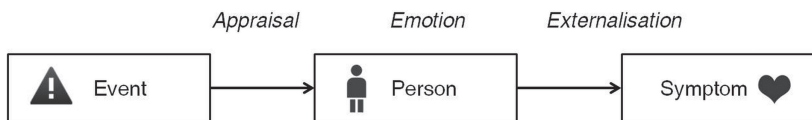

Figure 3. Emergence of emotions (Scherer, 1998b).

Emotions can also arise in a person, A, when s/he sees another person, B, experiencing an event that is relevant to B. These emotions which we feel when observing other people's emotions (depicted emotions) are called commotions by Scherer (1998b). According to the general *commotion model* (Scherer, 1998b; Scherer & Zentner, 2001), these emotions arise in the recipient via induction (appraisal process), empathy or emotional contagion. How exactly these emo-tions and commotions emerge when watching/looking at media is described in Figure 4.
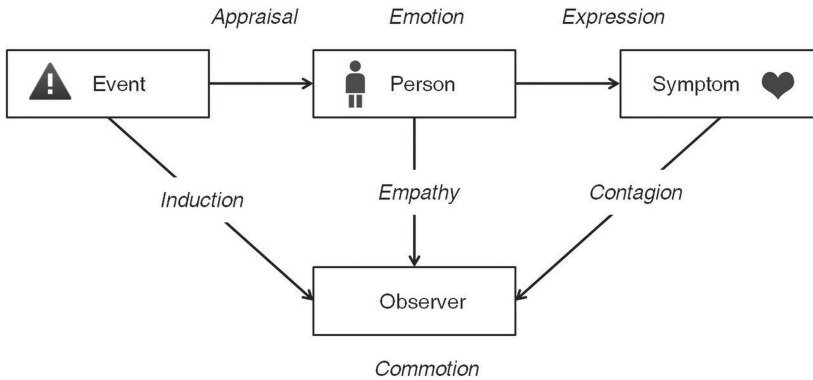
Figure 4. Model of 'normal' emotion induction via appraisal (upper part) and mediated 'commotion' due to empathy or other mechanisms of emotional communication (lower part) (Scherer & Zentner, 2001, p. 366).

Film reception

Videos and films are among those media through which an emotional impact can be achieved (Gross & Levenson, 1995). The origin of such considerations can be found as early as in antiquity. Aristotle was concerned with the effect depicted emotions have on the audience. According to Aristotle, the effect of tragedy is arousal and cleansing of the affective states of fear and compassion (catharsis). The depiction of emotions by the actors was supposed to evoke the same emotions in the audience. Scherer (2001a) formulates three possible origins of emotion emergence when viewing television reports, videos, etc. on the basis of the content displayed (see Figure 5): induction, empathy and contagion.

In real situations, the process of induction is equal to the appraisal process. In this process, a person assesses the relevancy of a perceived stimulus (e.g. event or object) for their goals or needs based on the SECs. In contrast, in the media environment, a fictional idea which the actor is trying to express takes the place of the object. The commotion is formed through a virtual appraisal. The emotional reaction of the audience is a direct result of an evaluation in regards to their own goals or values (induction). A good example is the outraged reaction (depicted emotion) of a person accused of murder in a crime series. The emotional reaction of the audience (felt reaction) is a subjective assessment of the situation. At this point an important distinction must be made. The felt emotion may coincide one depicted. In this case anger, for example, because in his evaluation process the viewer has come to the conclusion that the accused person is

innocent. However, the felt emotion may also be diametrically opposed. If the alleged murderer is a person negatively evaluated by the viewer, the felt emotion may be (malicious) joy.

Another possible cause for emotions in the viewer is empathic responses. This is the case if the recipient is not in any way personally affected, but can still assess the situation with regard to goal relevancy thanks to the actor/sender (empathy). An identification with the actor is also possible if the observer is aware that the emotion is only acted and thus not real. Here too, we must distinguish between depicted and felt emotions. Empathic responses can lead to symmetric commotions (the same emotion as the one of the sender, e.g., if the sender is likeable) or asymmetric commotions (different emotion as the sender, e.g. if the sender is dislikeable).

Emotional contagion is a third mechanism that causes emotions through the behaviour of the sender. This process is divided into two parts and arises when observing strong motor-expressive reactions. It includes, for example, motor mimicry, as when people yawn. The peripheral feedback of the emotional expression serves as a second process step. By acting out emotions via expression signals (e.g. gestures) the emotion is further strengthened. Hatfield, Cacioppo, and Rapson (1992, p. 153) define emotional contagion as a "tendency to automatically mimic and synchronize facial expressions, vocalizations, postures, and movements with those of another person and, consequently, to converge emotionally." Emotional contagion is thus based not on the appraisal of certain situations or the comparison with one's own goals. Rather, it is a mostly unconscious motoric reaction to the perceived behaviour of others. When emotional contagion happens, the depicted and felt emotions match. Studies show that emotional contagion is independent of empathy and identification with the sender (Hatfield et al., 1992).
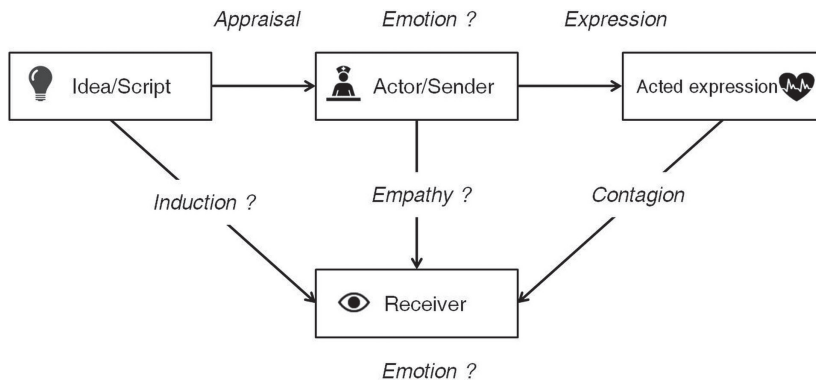
Figure 5. Emotion elicitation based on watching a fictional person's emotional expression in a media context (Scherer, 2001a, p. 138)

It can be said that an emotion may arise for various reasons (induction, empathy, contagion), and a distinction between depicted and felt emotions is fundamental for any analysis.

A starting point to figure out which emotions are conveyed by films in general can be found in the various genre terms, since the denomination of the genre is often reflected by the strongest triggered emotion (Gross & Levenson, 1995). Wirth and Schramm (2005) give an overview of many genre-specific studies regarding the impact of emotions in films. Research regarding the emotions generated by advertising are also numerous (e.g. Bagozzi, Gopinath, & Neyer, 1999; Edell & Burke, 1987). These studies find, for example, that any type of advertising evokes emotions (Zeitlin & Westwood, 1986) and that the product functions as a mediator (Holbrook & Batra, 1987; Mitchell & Olson, 1981).

This raises the question of which cinematic techniques can be used to trigger or reinforce emotions when viewing. A comprehensive meta-study regarding the effect of vocal expression and music performance has been compiled by Juslin and Laukka (2003). In addition to these acoustic parameters, visual features such as colour, texture, shape, or camera motion constitute another entry point. Apart from the extraction of content-based features, current information science studies use category models (e.g. Salway & Graham, 2003), dimensional models (e.g. Hanjalic & Xu, 2005; Soleymani, Chanel, Kierkels, & Pun, 2008) or hybrid models (e.g. Knautz, Neal, Schmidt, Siebenlist, & Stock, 2011) to identify emotions. Amongst other things, accompanying text-statistical methods (e.g. Chan & Jones, 2005) or the use of Hidden Markov models (e.g. Kang, 2003; Xu, Jin, Luo, & Duan, 2008) can also provide good results. Other methods, particularly for identifying felt emotions, include neuro-physiological approaches, observer and

self-report methods. Lopatovska (2011) as well as Lopatovska and Arapakis (2011) have compiled an extensive summary of these methods.

Picture reception

Just like films and videos, pictures can also induce emotions. Jörgensen (2003) claims that "[a]n image attribute is … not limited to purely visual characteristics, but includes other cognitive, affective, or interpretative responses to the image such as those describing spatial, semantic, or emotional characteristics" (p. 39). It can also be said that "[a]ll images are to be considered as emotional-laden, if they provoke emotions in the viewers, independent from the specific content of the picture" (Schmidt & Stock, 2009, p. 865).

But what is the relation between the depicted content and the evoked emotion? Scherer's (2001b) emergence mechanisms may come into play here as well. Viewing a picture can also entail an emotional reaction to the illustrated event. If this reaction is a direct result of the assessment of the situation with regard to one's own values, the emotion is created via induction. The relevance and assessment of the perceived image can cause symmetrical and asymmetrical reaction patterns. Depending on the assessed importance of events for the current motivation of the viewer, a war picture (see Figure 6 (a)), for example, may cause positive or negative emotions in the appraisal process. For example, if the viewer is a war veteran and associates terrible memories with the picture, likely reactions are anger, sadness, or aversion. Conversely, he might just as well be reminded of his youth and old comrades and feel longing or joy. The emotion felt may therefore vary quite a bit. Apart from differing assessments of the picture, the reaction may also be either symmetric (the same emotion is depicted and felt) or asymmetric (the opposite emotion is felt) just as when watching cinematic material.

Likewise, it is possible to feel empathy. This is the case when a situation can be understood from personal experience. Figure 6 (b) shows two boys whose relationship can be regarded as very amicable due to their posture, gestures and facial expressions. The depiction of friendship certainly has the potential to elicit empathy. Empathic reactions can, however, also lead to contrary commotions. Imagine a tragic situation (e.g. a mishap of a child with a bucket of water) that you can understand very well, but that would make you laugh.

The emotions depicted in pictures may well have a contagious effect, as seen in figure 6 (c). It shows a sad child, the head is bowed deeply. Due to his inability to help this child, the negative emotion, e.g. sadness, can be transmitted to

the viewer. Through the perception of this expressive behavior, emotions may be transferred between people (Le Bon, 1985).

Regarding the indexing and retrieval of emotions in pictures, low-level visual features are usually extracted and combined with content-based methods such as vector space approaches (Wang, Chen, Wang, & Zhang, 2003; Wang, Yu, & Zhang, 2004), neural networks (Dellandrea, Liu & Chen, 2010; Kim, Shin, Kim, Kim, & Shin, 2009), machine learning (Feng, Lesot, & Detyniecki, 2010), or fuzzy rules (Kim, Kim, Koo, Jeong, & Kim, 2005). Additionally, dimensional models (Hanjalic, 2006), category models or hybrid models (Knautz, Siebenlist, & Stock, 2010) often form the basis of these approaches. Interesting research approaches can also be found regarding the interaction of photographic documents and associated text (titles, tags, etc.) (Neal, 2010a, 2010b).

(a) Induction ?   (b) Empathy?   (c) Contagion ?

Figure 6. Possible images for emotion elicitation based on Scherer (2001).

Music

Music is present in every culture and plays an important role. It can elicit intense reactions (Gabrielsson, 2001; Gabrielsson & Lindström, 2001). According to Sloboda and Juslin (2001), the greatest power of music is to represent and express emotions. Unlike psychology, there is a long tradition of musical emotion research in the musical sciences (Budd, 1985). This research contains very few psychological assumptions. One reason for this could be the confusing picture that has arisen due to the many emotion-psychological theories and postulates (Sloboda & Juslin, 2001). Sloboda and Juslin pointed out that there is a need for multidisciplinary research, particularly with regard to a psychological approach.

But what is the connection between the communication of emotions and the emotions felt? At this point, too, the appraisal approach by Scherer is applicable (Figure 7). According to Scherer & Zentner (2001), this approach constitutes

a suitable model to reflect the personal significance of an event on the basis of a number of criteria in the music context.

In this process, the musical expression of a piece of music is received. The listener rates the intention on the basis of the application of specific combinations of musical parameters (Juslin & Laukka, 2003) and evaluates them in terms of relevance, implications, coping potential and normative significance. In addition to the emotion induction via appraisal, an induction can also arise during the reception of music on the basis of memories of certain situations (recall from memory) (Lang, 1979; Lang, Kozal, Miller, Levin, & MacLean, 1990; Scherer, 2004).

Just like videos and pictures, music can also evoke empathic responses. Imagine that you see an enthusiastic musician filled with emotions. Although you are not personally affected, you can still understand the situation. Also, as with pictures and videos, a positive emotion on the part of the observer requires sympathy (Scherer & Zentner, 2001). With regard to emotions mediated by music, emotional contagion can arise via the so-called peripheral route, according to Scherer (1998a, 2004). These are direct effects on the somatic and autonomic nervous system. As an example of this phenomenon, Scherer cites certain rhythms that generate emotions by activating the neurophysiological component and feedback mechanisms with the sub-systems.

Emotional music is a growing area of research in psychology and musicology. The results from this research are for example used in advertising and marketing (Bagozzi, Gopinath, & Neyer, 1999; Edell & Burke, 1987) as well as in music therapy. The applications in music therapy range from the treatment of patients with clinical depression (Hsu & Lai, 2004; Lai, 1999) or geriatric psychiatric patients (Clair & Memmott, 2008; Lee, Chan & Mok, 2010; Short, 1995) to the treatment of neurological diseases in the field of rehabilitation and procedures for cancer patients (Burns, 2001).

Information science and computer science researchers are also concerned with emotional music research. As with the indexing and retrieval of images, approaches start with the extraction of low-level features. The results of the meta-study compiled by Juslin and Laukka (2003) could present a very good starting point. They evaluated which acoustic parameters can communicate emotions in which way. Their results suggest that there exists emotion-specific acoustic information which is used for the communication of emotions through individual vowels and musical expression (e.g. loud and/or fast music communicates anger; high pitch and low high-frequency energy (tone colour, timbre) transport fear). These transmitted emotions, particularly anger and grief, are easily perceived by the recipients thanks to the musical parameters. We'd like to point out once more that the perceived emotions may differ from the felt emotions. In this context, Sloboda and Juslin (2005) postulate that a greater consensus can be

reached in identifying depicted emotions than in those that are felt. A comprehensive summary regarding emotional reactions to music in relation to the other subsystems has been given by Juslin and Västfjäll (2008).

In addition to the extraction of low-level features, similar procedures as for indexing and retrieving videos and images are applied. These are, among many others, vector-based approaches (Li & Ogihara, 2003), fuzzy logic (Yang, Liu, & Chen, 2006), software agents (Yang & Lee, 2004), regression (Yang, Lin, Su, & Chen, 2008), histograms (Li & Ogihara, 2006) and ontologies (Han, Rhon, Jun, & Hwang, 2010).
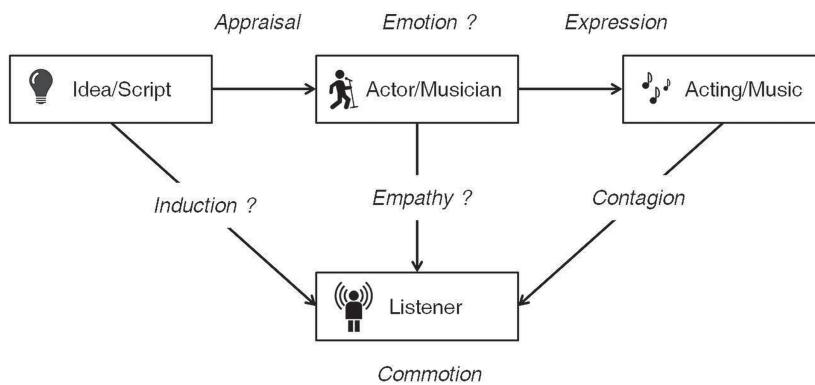

Figure 7. Emotion elicitation based in musical context (Scherer & Zentner, 2001, p. 370).

4 Consequences for multimedia indexing and retrieval

Below, it will be shown how the identified models, designs and research results can be applied to a multimedia search engine with emotional retrieval options. To this end, we will first discuss the approach, in which the indexing component in particular plays an important role. The second part of this section shows the actual implementation of the approach in an emotional search engine for multimedia documents. The search engine presented here is a project of the Heinrich-Heine University Düsseldorf. The author of this article participates in its development and implementation. Further information regarding the search engine can be found in Siebenlist & Knautz (this volume) as well as in other publications (Knautz et al., 2010; Knautz et al., 2011).

4.1 Approaches to emotional indexing on the World Wide Web

Concerning the indexing of multimedia documents, research distinguishes between two approaches. The automatic, content-based indexing of resources is neither time- nor labour-intensive. For images, colour, texture, and shape are taken into consideration. For videos, scenes that have shots with object and camera movement (panning or zooming), lighting, and cutting frequency also matter. Music is indexed according to pitch, rhythm, harmony, timbre and – if existent – the lyrics. However, the information that can be deduced from these low-level features is limited and therefore only partly suitable for an analysis of the semantic and emotional content of the Web. Concept-based solutions are considered promising alternatives to content-based approaches. Concept-based retrieval works with terms (concepts) that are currently primarily intellectually assigned, although, in principle, they could be deduced from the content. One way of achieving this is through knowledge organization systems (e.g. thesauri) and professional indexing. Given a taxonomy for emotions, it would at least theoretically be possible to intellectually assign controlled vocabulary about feelings toward documents such as videos. However, this method is highly dependent on the respective indexer. Additionally, there is the practical problem of having professional indexers evaluate the billions of pictures and videos available.

Using user-generated tags, i.e. social tagging or cooperative indexing – as can be found in various Web 2.0 services – has the distinct advantage of drastically reducing the time and manpower required (Peters, 2009). In broad folksonomies with a sufficiently large number of tagging users consistency problems have been found to be negligible (Knautz et al., 2011).

But how can users collaboratively index multimedia documents? As specified in section 1.3, three approaches for classifying emotions can be found in psychology (dimension model, category model, base emotions). Similar studies in emotional music retrieval (Lee & Neal, 2007) and in emotional image retrieval (Schmidt & Stock, 2009) work with a fixed set of five base emotions: sadness, happiness, anger, fear and disgust. The set of base emotions has been enriched on the basis of psychological specialized literature. For this project, the following base emotions have been selected: sadness, anger, fear, shame, disgust, surprise, longing, joy and love. Irrespective of these base emotions proposed in emotion psychology, we have included a further aspect in the form of humour resp. wit in order to do justice to this media component. The indexing of base emotions can be readily combined with the dimension model by considering the intensity of emotions. Following the approaches of Lee and Neal (2007), Schmidt and Stock (2009), and Knautz et al. (2010), sliders have been implemented in the search engine.

A far-reaching consequence of the research discussed above is the need to explicitly distinguish between depicted and felt emotions in multimedia documents. This applies to both the indexing and the retrieval options. The concrete implementation of the presented approach is shown below.

4.2 Implementation: An emotional multimedia search engine

Since there is currently no satisfactory possibility of emotionally indexing videos, music and images on the basis of content- and concept-based approaches on the Web, a consideration of the different indexing approaches leads to the conclusion that the most appropriate method to index those resources on the Web is the use of a broad folksonomy. In this approach, many different people index the same documents with tags, guaranteeing in turn the findability of the resources. The search engine MEMOSE (Media Emotion Search) presented here has adapted this concept. In addition to offering a keyword search like other Web 2.0 services, it also offers ten emotional (base) terms (Figure 10). The user can select one or more emotions via check boxes (e.g. fear) and combine them with other non-emotion-related search arguments (e.g. dog).

Figure 8. MEMOSE search interface.

The result of the query is presented in two lists (Figure 11). This is a consequence of the above considerations regarding the relationship between depicted and felt emotions. On the first page of the respective results list, the four most emotional pictures are included. More results can be viewed by clicking the arrow at the

right edge. The results confirm the chosen approach of differentiating between depicted and felt emotions. The first results list includes photos that were indexed by queried tag and the queried emotion. It is clear, however, that in these pictures the emotion fear is depicted. The second results list, on the other hand, shows pictures that trigger the emotion when viewed. The results in the two lists are completely different and a distinction is necessary. Video and music search results can be accessed via tabs.
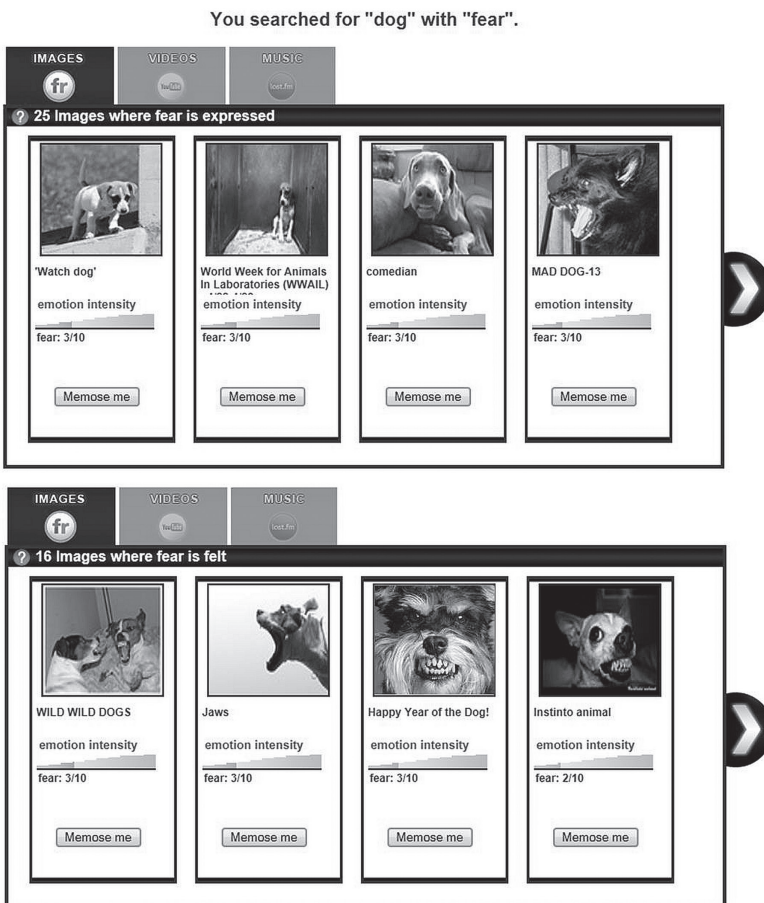

Figure 9. Search results (shown and felt) for "fear and dog".

The indexing of the multimedia files can be done in two different ways. First, a user can index existing multimedia documents. He can do this by clicking the "Memose me" button below a search result (Figure 9). Second, he can index

media he has uploaded himself. This can be done either by entering a URL from Flickr, YouTube or Last.fm or using MEMOSE's own uploader. In the first case, the tags are imported from the respective service; in the second case, the user adds them himself (Figure 12). The next step, the emotional indexing, is the same in both variants (Figure 13). To realize the retrieval options described above, the indexing tool includes three aspects:

1. general indexing with one or more emotions from a fixed set of emotions,
2. assessing the intensity of the emotion(s) on a scale of 0 to 10,
3. the distinction between depicted and felt emotions.

These three aspects are a consequence of the considerations in this chapter. During indexing, a distinction must be made between depicted and felt emotions in order to obtain relevant retrieval results. As an emotion can be felt or communicated with varying degrees of intensity, including the intensity makes for a useful ranking parameter. The restriction to a fixed set of emotions is a good approach on the Web due to the variety of emotions and the possibility of paraphrasing. In a future version it is planned to identify the emotion word fields (emotion clusters) and make them available to users via overlays over the individual emotions.

Figure 10: Multimedia uploader

Figure 11: Indexing tool.

5 Concluding remarks

The aim of this chapter was to present the peculiarities of emotions in relation to multimedia documents and to point out the consequences in regard to indexing and retrieving these documents. For this purpose, a small insight into emotion-psychological foundations was given. We found that there are many definitions of emotion and modern research makes use of working definitions. The variety of definitions is due to the amount of emotion-psychological theories with different emphases. We also gave a brief overview of how individual research orientations explain the origin of emotions.

In this context, studies show that a considerable number of emotional experiences are due to the representation of events in various media. To explain emotional media effects, we made use of an appraisal model – the *component-process-model* by Scherer (1984, 2001b) – in which emotions are the result of a subjective situation assessment. The subjective importance of the event for the current motivation of the subject is crucial for triggering the emotion.

According to the *commotion model* by Scherer (Scherer & Zentner, 2001), emotions can, however, also arise when looking at depicted emotions via induction (appraisal process), empathy or emotional contagion. Scherer's model was examined in relation to video, music, and pictures, and it was shown that the emotions felt not only emerge in different ways, but also that they can be com-

pletely different from the emotions depicted. This aspect is fundamental for any system that tries to make emotional content searchable. Which parameters (low-level features) cause or strengthen exactly which type of emotion in which media could not be discussed in detail in this chapter for lack of space. However, we referenced the relevant literature at the appropriate place.

Using results from the fields of psychology, musicology and information science, a model for indexing and retrieval of emotional content in multimedia documents was designed. Besides the use of ten base emotions, the dimension model was used for the intensity setting. The valence is, in our view, given by the concept of each base emotion. The distinction between felt emotions and depicted emotions in both in the indexing and the retrieval form a centrepiece of the approach and implementation. In this way, the specific requirements of the media component are satisfied based on the appraisal model.

MEMOSE is a project of the Heinrich-Heine-University Düsseldorf (Germany) and is currently still a prototype, with new insights being continually integrated into the system. We therefore do not claim that the system is perfect. Nevertheless, MEMOSE is a first look at how different emotion-psychological approaches can be used for the indexing and retrieval of multimedia documents.

References

Adelaar, T., Chang, S., Langendorfer, K. L., & Morimoto, M. (2003). Effects of media formats on emotions and impulse. *Journal of Information Technology, 18*, 247–266.

Arifin, S., & Cheung, P. (2007). A computation method for video segmentation utilizing the pleasure-arousal-dominance emotional information. *Proceedings of the 16th ACM International Conference on Multimedia* (p. 68).

Arnold, M. (1960). *Emotion and personality*. New York: Columbia University Press.

Bagozzi, R., Gopinath, M., & Nyer, P. (1999). The role of emotions in marketing. *Journal of the Academy of Marketing Science, 27*(2), 59–70.

Bischoff, K., Firan, C., Nejdl, W., & Paiu, R. (2008). Can all tags be used for search? *Proceedings of the 17th ACM Conference on Information and Knowledge Mining* (pp. 193- 202).

Bischoff, K., Firan, C., Nejdl, W., & Paiu, R. (2010). Bridging the gap between tagging and querying vocabularies: Analyses and applications for enhancing multimedia IR. *Journal of Web Semantics, 8*(2–3), 97–109.

Brosch, T., & Scherer, K. (2009). A case for the component process model as theoretical starting point for experimental emotion research. In V. Brandstätter & J. H. Otto (Eds.), *Handbuch der Allgemeinen Psychologie: Motivation und Emotion* (pp. 446–456). Göttingen: Hogrefe.

Budd, M. (1985). Understanding music. *Proceedings of the Aristotelian Society*, *59*, 233–248.

Burns, D. (2001). The effect of the bonny method of guided imagery and music on the mood and life quality of cancer patients. *Journal of Music Therapy*, *38*(1), 51–65.

Cannon, W. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *American Journal of Psychology, 100*(3–4), 567–586.

Cantor, J., & Nathanson, A. (1996). Children's fright reactions to television news. *Journal of Communication, 46*(4), 139–152.

Chan, C., & Jones, G. (2005). Affect-based indexing and retrieval of films. *Proceedings of the 13th Annual ACM International Conference on Multimedia* (pp. 427–430).

Clair, A., & Memmott, J. (2008). *Therapeutic uses of music with older adults.* Silver Spring, MD: American Music Therapy Association.

Darwin, C. (1872). *The expression of the emotions in man and animals.* London: John Murray.

Davis, W., Gfeller, K., & Thaut, M. (1992). *An introduction to music therapy: Theory and practice.* Dubuque, IA: Wm. C. Brown Publishers.

Dellandrea, E., Liu, N., & Chen, L. (2010). Classification of affective semantics in images based on discrete and dimensional models of emotions. *Proceedings of the IEEE International Workshop on Content-Based Multimedia Indexing (CBMI)* (pp. 1–6).

Edell, J., & Burke, M. (1987). The power of feelings in understanding advertising effects. *Journal of Consumer Research, 38*, 421–433.

Ekman, P., Friesen, W., & Ellsworth, P. (1982). What emotion categories or dimensions can observers judge from facial behavior. In P. Ekman (Ed.), *Emotion in the human face* (pp. 39–55). New York: Cambridge University Press.

Fehr, B., & Russell, A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General, 113*(3), 464–486.

Feng, H., Lesot, M., & Detyniecki, M. (2010). Association rules to discover color-emotion relationships based on social tagging. *Proceedings of Knowledge-Based and Intelligent Information and Engineering Systems* (pp. 544–553).

Frijda, N. (1986). *The emotions.* New York: Cambridge University Press.

Frijda, N. (1988). The laws of emotion. *American Psychologist, 43*, 349–358.

Gabrielsson, A. (2001). Emotions in strong experience with music. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 431–449). New York: Oxford University Press.

Gabrielsson, A., & Lindström, E. (2001). The influence of musical structure on emotional expression. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 223–248). New York: Oxford University Press.

Gergen, K. (1991). *Toward transformation in social knowledge.* London: Sage.

Gray, J. (1982). *The neuropsychology of anxiety: An inquiry into the functions of the septo-hippocampal system.* Oxford: Oxford University Press.

Gross, J., & Levenson, R. (1995). Emotion elicitation using films. *Cognition and Emotion, 9*(1), 87–108.

Han, B., Rhon, S., Jun, S., & Hwang, E. (2010). Music emotion classification and context- based music recommendation. *Multimedia Tools and Applications, 47*(3), 433–460.

Hanjalic, A. (2006). Extracting moods from pictures and sounds. *IEEE Signal Processing Magazine, 23*(2), 90–100.

Hanjalic, A., & Xu, l.-Q. (2005). Affective video content representation and modeling. *IEEE Transactions on Multimedia, 7*(1), 143–154.

Hastings, S., Iyer, H., Neal, D., Rorissa, A., & Yoon, J. (2007). Social computing, folksonomies, and image tagging: Reports from the research front. *Proceedings of the 70th Annual Meeting of the American Society for Information Science and Technology* (pp. 1026–1029).

Hatfield, E., Cacioppo, J., & Rapson, R. (1992). Primitive emotional contagion. In M. S. Clark (Ed.), *Emotion and social behavior* (pp. 151–177). Newbury Park: Sage.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Hohmann, G. (1966). Some effects of spinal cord lesions on experienced emotional feelings. *Psychophysiology, 3*, 143–156.

Holbrook, M., & Batra, R. (1987). Assessing the role of emotions as mediators of consumer responses to advertising. *Journal of Consumer Research, 14*, 404–420.

Hsu, W., & Lai, H. (2004). Effects of music on major depression in psychiatric inpatients. *Archives of Psychiatric Nursing, 18*(5), 193–199.

Izard, C. E. (1971). *The face of emotion.* New York: Appleton-Century-Crofts.

Izard, C. E. (1977). *Human emotions.* New York: Plenum Press.

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

James, W. (1884). What is an emotion? *Mind, 9*(34), 188–205.

Jones, M. (1924). A laboratory study of fear: The case of Peter. *The Pedagogical Seminary, 31*, 308–315.

Jörgensen, C. (2003). *Image retrieval: Theory and research.* Lanham, MD: Scarecrow Press.

Juslin, P., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770–814.

Juslin, P., & Västfjall, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*, 559–621.

Kang, H.-B. (2003). Affective content detection using HMMs. *Proceedings of ACM Multimedia* (pp. 259–263).

Katz, E., & Foulkes, D. (1962). On the use of the mass media as "escape" – clarification of a concept. *Public Opinion Quarterly, 26*, 377–388.

Kim, E., Kim, S., Koo, H., Jeong, K., & Kim, J. (2005). Motion-based textile indexing using colors and texture. *Proceedings of Fuzzy Systems and Knowledge Discovery* (pp. 1077–1080).

Kim, Y., Shin, Y., Kim, S., Kim, E., & Shin, H. (2009). EBIR: Emotion-based image retrieval. *Digest of Technical Papers, International Conference on Consumer Electronics* (pp. 1–2).

Kleinginna, P., & Kleinginna, A. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotions, 5*(4), 345–379.

Knautz, K., Neal, D., Schmidt, S., Siebenlist, T., & Stock, W. (2011). Finding emotional- laden resources on the World Wide Web. *Information, 2*(1), 217–246.

Knautz, K., Siebenlist, T., & Stock, W. (2010). MEMOSE. Search engine for emotions in multimedia documents. *Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval* (pp. 791–792).

Lai, Y. (1999). Effects of music listening on depressed women in Taiwan. *Issues in Mental Health Nursing, 20*, 229–246.

Lang, P. (1979). A bio-informational theory of emotional imagery. *Psychophysiology, 16*, 495–512.

Lang, P., Kozal, M., Miller, G., Levin, D., & MacLean, A. (1990). Emotional imagery: Conceptual structure and pattern of somato-visceral response. *Psychophysiology, 17*, 179–192.

Lange, C., & Kurella, H. (1887). *Ueber Gemüthsbewegungen: Eine Psycho-physiologische Studie.* T. Thomas.

Lazarus, R. (1991). *Emotion and adaptation.* New York: Oxford University Press.

Lazarus, R. (1995). Vexing research problems inherent in cognitive-mediational theories of emotion – and some solutions. *Psychological Inquiry, 6*, 183–196.

Le Bon, G. (1895). *Psychologie des Foules.* Paris: P.U.F.

LeDoux, J. (1995). Emotion: Clues from the brain. *Annual Review of Psychology, 46*, 209–235.

Lee, H., & Neal, D. (2007). Towards Web 2.0 music information retrieval: Utilizing emotion-based, user-assigned descriptors. *Proceedings of the 70th Annual Meeting of theAmerican Society for Information Science and Technology* (pp. 732–741).

Lee, Y., Chan, M., & Mok, E. (2010). Effectiveness of music intervention on the quality of life of older people. *Journal of Advanced Nursing, 66*(12), 677–2687.

Li, T., & Ogihara, M. (2003). Detecting emotion in music. *Proceedings of the 4th International Symposium on Music Information Retrieval* (pp. 239–240).

Li, T., & Ogihara, M. (2006). Toward intelligent music information retrieval. *IEEE Transactions on Multimedia, 8*(3), 564–574.

Lopatovska, I. (2011). Researching emotion: Challenges and solutions. *Proceedings of the 2011 iConference (iConference '11)* (pp. 225–229).

Lopatovska, I., & Arapakis, I. (2011). Theories, methods and current research on emotions in library and information science, information retrieval and human-computer interaction. *Information Processing and Management, 47*(4), 575–592.

Mangold, R., Unz, D., & Winterhoff-Spurk, P. (2001). Zur Erklärung emotionaler Medienwirkungen. Leistungsfähigkeit, empirische Überprüfung und Fortentwicklung theoretischer Ansätze. In *Theoretische Perspektiven der Rezeptionsforschung* (pp. 163–180). München: Reinhard Fischer.

Marañon, G. (1924). Contribution à l'étude de l'action émotive de l'adrenaline. *Revue Francaise d'Endocrinologie, 2*, 301–325.

McDougall, W. (1908 / 1960). *An introduction to social psychology.* Boston: Methuen.

McDougall, W. (1926). *An outline of abnormal psychology.* Boston: Luce.

Mehrabian, A. (1995). Framework for a comprehensive description and measurement. *Genetic, Social, and General Psychology Monographs, 121*, 339–361.

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

Meyer, W., Reisenzein, R., & Schützwohl, A. (2003). *Einführung in die Emotionspsychologie. Band 2: Evolutionspsychologische Emotionstheorien.* Bern: Huber.

Mitchell, A., & Olson, J. (1981). Are product attribute beliefs the only mediator of advertising effects on brand attitude? *Journal of Marketing Research, 18*, 318–332.

Morris, J., & Boone, M. (1998). The effects of music on emotional response, brand attitude, and purchase intent in an emotional advertising condition. *Advances in Consumer Research, 25*, 518–526.

Mowrer, O. (1960). *Learning theory and behavior.* New York: Wiley.

Neal, D. M. (2010a). *Breaking in and out of the silos: What makes for a happy photograph cluster?* Paper presented at the 2010 Document Academy Conference (DOCAM '10), Denton, TX, USA.

Neal, D. M. (2010b). Emotion-based tags in photographic documents: The interplay of text, image, and social influence. *Canadian Journal of Information and Library Science, 34*, 329–353.

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

Ortony, A., Clore, G., & Collins, A. (1988). *The cognitive structure of emotions.* Cambridge: Cambridge University Press.

Ortony, A., & Turner, T. (1990). What's basic about basic emotions? *Psychological Review, 97*, 315–331.

Otto, J., Euler, H. M., & Mandl. H. (2000). Begriffsbestimmungen. In J. Otto, H. A. Euler, & H. Mandl (Eds.), *Handbuch Emotionspsychologie* (pp. 11–18). Weinheim: Beltz, PsychologieVerlagsUnion.

Panksepp, J. (1982). Toward a general psychobiological theory of emotions. *The Behavioral and Brain Sciences, 5*, 407–467.

Peters, I. (2009). *Folksonomies: Indexing and retrieval in Web 2.0*. Berlin: De Gruyter Saur.

Plutchik, R. (1962). *The emotions: Facts, theories, and a new model.* New York: Random House.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In R. Plutchik & H. Kellerman (Eds.), *Theories of emotion: Emotion, theory, research, and experience* (pp. 3–33). New York: Academic.

Rothmund, J., Schreier, M., & Groeben, N. (2001). Fernsehen und erlebte Wirklichkeit I: Ein kritischer Überblick über die Perceived Reality-Forschung. *Zeitschrift für Medienpsychologie, 13*(1), 23–41.

Russell, J., & Mehrabian, A. (1977). Evidence for a three-factor theory of emotions. *Journal of Research in Personality, 11*(3), 273–294.

Salway, A., & Graham, M. (2003). Extracting information about emotions in films. *Proceedings of the 11th ACM International Conference on Multimedia* (pp. 299–302). New York: ACM.

Schachter, S. (1964). The interaction of cognitive and physiological determinants of emotional state. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 49–79). New York: Academic Press.

Scherer, K. (1984). On the nature and function of emotion. A component process approach. In K. R. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 293–318). Hillsdale: Erlbaum.

Schenk, M. (1987). *Medienwirkungsforschung*. Tübingen: Siebeck.

Scherer, K. (1988). Criteria for emotion-antecedent appraisal: A review. In *Cognitive Perspectives on Emotion and Motivation* (pp. 89–126). Dordrecht: Kluwer.

Scherer, K. (1993). Neuroscience projections to current debates in emotion psychology. *Cognition and Emotion, 7*, 1–41.

Scherer, K. (1998a). Appraisal theories. In T. Dalgleish & M. J. Power (Eds.), *Handbook of cognition and emotion* (pp. 637–663). Amsterdam: Elsevier Science.

Scherer, K. (1998b). Emotionsprozesse im Medienkontext: Forschungsillustrationen und Zukunftsperspektiven. *Medienpsychologie, 10*(4), 276–293.

Scherer, K. (2001a). Emotional experience is subject to social and technological change: Extrapolating to the future. *Social Science Information, 40*(1), 125–151.

Scherer, K. (2001b). Appraisal considered as a process of multilevel sequential checking. In K. R. Scherer, A. Schorr, & T. Johnstone (Eds.), *Appraisal processes in emotion: Theory, methods, research* (pp. 92–120). Oxford: Oxford University Press.

Scherer, K. (2004). Which emotions can be induced by music? What are the underlying: Mechanisms [sic]? And how can we measure them? *Journal of New Music Research, 33*(3), 239–251.

Scherer, K. (2005). What are emotions? And how can they be measured? *Social Science Information, 44*(4), 695–729.

Scherer, K., Wallbott, H., & Summerfield, A. (1986). *Experiencing emotion: A crosscultural study.* Cambridge: Cambridge University Press.

Scherer, K., & Zentner, M. (2001). Emotional effects of music: Production rules. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 361–391). New York: Oxford University Press.

Schmidt, S., & Stock, W. G. (2009). Collective indexing of emotions in images. A study in emotional information retrieval. *Journal of the American Society for Information Science and Technology, 60*(5), 863–876.

Schmidt-Atzert, L., & Ströhme, W. (1983). Ein Beitrag zur Taxonomie der Emotionswörter. *Psychologische Beiträge, 2*, 126–141.

Schwab, F. (2001). Unterhaltungsrezeption als Gegenstand medienpsychologischer Emotionsforschung. *Zeitschrift für Medienpsychologie, 13*, 62–72.

Short, A. (1995). Insight-oriented music therapy with elderly residents. *The Australian Journal of Music Therapy, 6*, 4–18.

Sloboda, J. A., & Juslin, N. (2001). Psychological perspectives on music and emotion. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and emotion: Theory and research* (pp. 71–104). New York: Oxford University Press.

Sloboda, J., & Juslin, P. (2005). Affektive Prozesse: Emotionale und ästhetische Aspekte musikalischen Verhaltens. In *Allgemeine Musikpsychologie* (pp. 767–841). Göttingen: Hogrefe.

Soleymani, M., Chanel, G., Kierkels, J., & Pun, T. (2008). Affective ranking of movie scenes using physiological signals and content analysis. *Proceedings of the 2nd ACM Workshop on Multimedia Semantics* (pp. 32–39). New York: ACM.

Tomkins, S. (1962). *Affect, imagery, consciousness, Vol. I: The positive affects.* New York: Springer Publishing.

Tomkins, S. (1963). *Affect, imagery, consciousness, Vol. II: The negative affects.* New York: Springer Publishing.

Tomkins, S. (1984). Affect theory. In K. Scherer & P. Ekman (Eds.), *Approaches to emotion* (pp. 163–195). Hillsdale, NJ: Erlbaum.

Valins, S. (1966). Cognitive effects of false heart-rate feedback. *Journal of Personality and Social Psychology, 4*(4), 400–408.

Wang, S., Chen, E., Wang, X., & Zhang, Z. (2003). Research and implementation of a content-based emotional image retrieval model. *Proceedings of the Second International Conference on Active Media Technology* (pp. 293–302).

Wang, W.-N., Yu, Y.-L., & Zhang, J.-C. (2004). Image emotional classification: Static vs. dynamic. *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics* (pp. 6407–6411).

Watson, J. (1913). Psychology as the behaviorist views it. *Psychological Bulletin, 20*, 158–177.

Watson, J. (1919). A schematic outline of the emotions. *Psychological Review, 26*(3), 165–196.

Watson, J. B. (1930). *Behaviorism.* Chicago: University of Chicago Press.

Watson, J., & Rayner, R. (1920). Conditioned emotional reaction. *Journal of Experimental Psychology, 3*, 1–14.

Weiner, B. (1986). *An attributional theory of motivation and emotion.* New York: Springer.

Weiner, B., & Graham, S. (1984). An attributional approach to emotional development. In E. Izard, J. Kagan, & R. Zajonc (Eds.), *Emotions, cognition, and behavior* (pp. 167–191). New York: Cambridge University Press.

Wirth, W., & Schramm, H. (2005). Media and emotion. *Communication Research Trends*, *24*(3), 1–39.

Wundt, W. (1906). *Vorlesungen über die Menschen und Tierseele.* Hamburg: Voss.

Xu, M., Jin, J., Luo, & Duan, L. (2008). Hierarchical movie affective content analysis based on arousal and valence features. *Proceedings of the 16th ACM International Conference on Multimedia* (pp. 677–680).

Yang, Y., Liu, C., & Chen, H. (2006). Music emotion classification: A fuzzy approach. *Proceedings of the 14th International Conference on Multimedia* (pp. 81–84).

Yang, D., & Lee, W. (2004). Disambiguating music emotion using software agents. *Proceedings of the 5th International Conference on Music Information Retrieval* (pp. 52–58).

Yang, Y., Lin, Y., Su, Y., & Chen, H. (2008). A regression approach to music emotion recognition. *IEEE Transactions on Audio, Speech and Language Processing, 16*(2), 448–457.

Zeitlin, D., & Westwood, R. (1986). Measuring emotional response. *Journal of Advertising Research, 26*, 34–44.

Zillmann, D. (1978). Attribution and misattribution of excitatory reactions. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (pp. 335–370). Hillsdale, NJ: Erlbaum.

Zillmann, D. (1988). Mood management through communication choices. *American Behavioral Scientist, 31*(3), 327–341.