

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/227689026>

Collective Indexing of Emotions in Images. A Study in Emotional Information Retrieval (EmIR)

Article in *Journal of the American Society for Information Science and Technology* · May 2009

DOI: 10.1002/asi.21043 · Source: DBLP

CITATIONS

85

READS

840

2 authors:



Stefanie Schmidt

Heinrich Heine University Düsseldorf

2 PUBLICATIONS 99 CITATIONS

SEE PROFILE



Wolfgang G. Stock

Heinrich Heine University Düsseldorf

182 PUBLICATIONS 3,240 CITATIONS

SEE PROFILE

Collective Indexing of Emotions in Images. A Study in Emotional Information Retrieval

Stefanie Schmidt and Wolfgang G. Stock

Heinrich-Heine-University Düsseldorf, Dept. for Information Science, Universitätsstr. 1, D-40225 Düsseldorf, Germany. E-mail: steff_schmidt@t-online.de; stock@phil-fak.uni-duesseldorf.de

Some documents provoke emotions in people viewing them. Will it be possible to describe emotions consistently and use this information in retrieval systems? We tested collective (statistically aggregated) emotion indexing using images as examples. Considering psychological results, basic emotions are *anger*, *disgust*, *fear*, *happiness*, and *sadness*. This study follows an approach developed by Lee and Neal (2007) for music emotion retrieval and applies scroll bars for tagging basic emotions and their intensities. A sample comprising 763 persons tagged emotions caused by images (retrieved from www.Flickr.com) applying scroll bars and (linguistic) tags. Using SPSS, we performed descriptive statistics and correlation analysis. For more than half of the images, the test persons have clear emotion favorites. There are prototypical images for given emotions. The document-specific consistency of tagging using a scroll bar is, for some images, very high. Most of the (most commonly used) linguistic tags are on the basic level (in the sense of Rosch's basic level theory). The distributions of the linguistic tags in our examples follow an inverse power-law. Hence, it seems possible to apply collective image emotion tagging to image information systems and to present a new search option for basic emotions. This article is one of the first steps in the research area of emotional information retrieval (EmIR).

Introduction

A Problem

Consider the following situation: A user wants to find images about a group of people that give him a feeling of *happiness*. He intends to use the documents for a marketing action. Another user tries to retrieve documents about human feet that lead to feelings of *disgust*. She is going to use the documents as illustrations of body parts provoking disgust. Documents that are "emotional-laden" (Newhagen, 1998,

p. 265) and provoke emotions in the observing persons can be

- images (Jørgensen, 1999; Jørgensen, 2003; Wang & Wang, 2005),
- videos (Salway & Graham, 2003),
- music (Hu & Downie, 2007; Kalbach, 2002; Li & Ogiwara, 2003), and
- some kinds of texts (e.g., novels, poems, Web pages or Weblogs; Goetz et al., 1993; Ni, Xue, Ling, Yu, & Yang, 2007; McKechnie, Ross, & Rothbauer, 2007; Rubin, Stanton, & Liddy, 2004; Yanbe, Jatowt, Nakamura, & Tanaka, 2007).

In this article, we describe problems (and possible solutions) concerning indexing and retrieval of emotions based upon images.

This paper is one of the first steps in the new research area of emotional information retrieval (EmIR), especially in the theoretical and empirical foundations of EmIR.

Emotional Information

Whether or not there is an "affective revolution in information science" (Nahl, 2007, p. 23), the analysis of indexing and retrieving of documents that cause emotions seems to be a very interesting research task in information science. "Emotion is one of the core factors in music information retrieval" (MIR; Lee & Neal, 2007, p. 732). Or, as Juslin and Västfjäll (2008, p. 559) put it: "Research indicates that people value music primarily because of the emotion it evokes." If we are going to know more about "user feelings about Web pages," then it is possible to create "a sentiment-aware search" (Yanbe et al., 2007, p. 111). The same holds true for image and video information retrieval. "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" (Hastings, Iyer, Neal, Rorissa, & Yoon, 2007, p. 1026). Corinne Jørgensen, in her famous book on image retrieval, claims that "an 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*

Received August 20, 2008; revised December 11, 2008; accepted January 8, 2009

© 2009 ASIS&T • Published online 20 February 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asi.21043

[italics added] characteristics” (Jørgensen, 2003, p. 3). Findings from a study by Kipp (2006) suggest “that users relate information to time related tasks, activities and *their own emotional reactions* [italics added].” Some news photographers interviewed by Diane Neal (2006) showed interest in searching images by emotion. In an image indexing experiment, especially subject novices, applied emotive or interpretative terms to the images, while subject experts did not (Beaudoin, 2008). In affective computing (Picard, 1997; Paiva, Prada, & Picard, 2007), computing is seen in its relations to emotions: Affective computing “relates to, arises from, or influence emotions” (Picard, 1995, p. 1). Rosalind W. Picard, a pioneer of affective computing, states that information retrieval is an area where affective computing may be applied (Picard, 1995, p. 15). People viewing fear and disgust pictures (in a pain experiment) exhibit decreasing pain intensity and unpleasantness thresholds (Meagher, Arnau, & Rhudy, 2001). In marketing and advertising research, attention has been given to the way in which “media content can trigger the particular emotions and impulse buying behavior of the viewer/listener/surfer” (Adelaar, Chang, Lancendorfer, Lee, & Morimoto, 2003, p. 247). Emotions are quite important in brand perception and purchase intents (Morris & Brown, 1998). To sum up, it seems to be necessary to classify emotion as a standard attribute of images. So, “image data . . . include perceptual, ‘semantic’, and ‘emotional’ attributes generated by a variety of cognitive, affective, or interpretive responses to the . . . images” (Jørgensen, 1998, p. 164).

Image Information Indexing and Retrieval

Is it possible to index emotions adequately so that users can retrieve documents that cause specific emotions? This is a task of either image indexing and retrieval (Rasmussen, 1997) or visual (i.e., image and video) information retrieval (Enser, 2008a; Enser, 2008b; Gupta & Jain, 1997). According to Corinne Jørgensen (2007), there are, up to the present, three ways to index images: “As we see now, in addition to content-based (using computer algorithms to describe low-level features) and concept-based (using human intelligence to assign higher-level descriptors) indexing, . . . there is now a third alternative available for providing access to images, that of social tagging or cooperative indexing.”

The first way is to follow the basic principles of *content-based information retrieval* (Stock, 2007, chap. 31) and try to find hints on emotions by computational analysis of images (Cho, 2004; Cho & Lee, 2001; Wang, Chen, Wang, & Zhang, 2003; Wang & Wang, 2005; Wang, Yu, & Jiang, 2006). It is very difficult to identify emotions automatically only by color distributions, texture, and shape. Pattern recognition of facial expressions of human emotions seems to be possible (Duthoit, Szttynda, Lal, Jap, & Agbinya, 2008), but there are huge amounts of emotional images besides human faces. Shneiderman, Bederson, and Drucker (2006, p. 69) stated: “Successful retrieval is based largely on attaching appropriate annotations to each image and collection since automated image content analysis is still limited.”

The second way of image indexing is *concept-based information retrieval*. Here, an indexer makes use of knowledge organization systems (KOS, such as nomenclatures, classification systems, or thesauri; Stock & Stock, 2008) and translates image-describing concepts into the language of the KOS. We have to presuppose that there is an agreement about the “right” image attribute (e.g., focal point, motion, orientation, shape, texture) and the “right” controlled vocabulary in the KOS. In addition, we have to presuppose that there is (at least to a high degree) inter-indexer and intra-indexer consistency, for, in practice, there is only one single indexer who handles the image description. All presuppositions are very awkward in information practice; inter-indexer consistency, especially, is very low (Beaudoin, 2008; Markey, 1984)—“the output of the indexing process seemed to be quite inconsistent” (Markkula & Sormunen, 2000, p. 273). We learn from Goodrum (2000, p. 64) that “manual indexing suffers from low term agreement across indexers, . . . and between indexers and user queries.” “There is evidence that current systems for image access often fail the user,” Jørgensen (1998, p. 162) adds. Although it is theoretically possible to create emotional concepts in KOS (Rorissa & Hastings, 2004), it has not been broadly applied in information practice, for there is a lack of experiences with emotional categories. “The influence of the photograph’s emotional tone on categorization has not been discussed much in previous studies,” Laine-Hernandez and Westman (2006) report. There is another, more practical problem: It seems to be impossible to index all the billions of images on the World Wide Web by professional indexers, for “manual assignment of textual attributes is both time consuming and costly” (Goodrum, p. 64).

So, we will try to follow the third way: the way of *social tagging*. Here, it is possible to apply folksonomy tagging (Dye, 2006; Furnas et al., 2006; Golder & Huberman, 2006; Gordon-Murnane, 2006; Guy & Tonkin, 2006; Kroski, 2008; Mathes, 2004; Neal, 2007; Noruzi, 2006; Peters & Stock, 2007; Peters & Stock, 2008; Peterson, 2006; Smith, 2008; Spiteri, 2006; Spiteri, 2007; Trant, 2006) as it is done, for example, in Flickr (Beaudoin, 2007; Kennedy, Naaman, Ahern, Nair, & Rattenbury, 2007; Van House, 2007; van Zwol, 2007). Folksonomy is a method of indexing documents using uncontrolled terms. There are no indexing rules, and there is no authority that controls the terminology and the taggers. Referring to Vander Wal (2005) there are two types of folksonomies: broad and narrow folksonomies. In a broad folksonomy (applied in, e.g., Del.icio.us), many different people may tag the same document using the same, similar, or completely different tags. Here, it is possible to calculate document-specific tag distributions. In a narrow folksonomy (applied in, e.g., Flickr), only the creator of the document (and, in some services, certain groups of other people) may tag the document, but a tag must not be indexed repeatedly. Folksonomies represent an authentic use of language, allow multiple interpretations, and are cheap methods of indexing. Besides such benefits, folksonomies have some underpinnings, a lack of precision (resulting from the absence

of controlled vocabulary), language merging, tags that do not identify aboutness (but planned actions as *to read*), spam-tags, user-specific tags, and other misleading keywords. In image tagging, there is a further problem: Taggers do not separate the different semantic levels of ofness and aboutness (Shatford, 1986) and isness (Ingwersen, 2002); they merge tags from all levels into one tag cloud (Peters & Stock, 2007). Kipp (2006) observed (in text indexing) that linguistic tags “such as *cool* or *fun* do not appear to add anything to the subject classification of an item and would not seem to be good candidates for search terms for information retrieval.” Furthermore, it is an open problem whether taggers are able to index intensities of emotions consistently.

All approaches of image indexing and retrieval (first, content-based, using algorithms only; second, concept-based using KOS; third, social tagging using folksonomies) are not free from problems. In addition to the third way of social tagging, we will follow a *fourth approach* recently developed by Hyuk-Jin Lee and Diane Neal (2007) for emotional music information retrieval: the application of *scroll bars* for tagging basic emotions and their intensities.

What is an Emotion? What is a Basic Emotion?

It seems very difficult to find a consensual definition of “emotion” (Kleinginna & Kleinginna, 1981). We adopt our working definition of “emotion” from Izard (1991, p. 14): “An emotion is experienced as a feeling that motivates, organizes, and guides perception, thought, and action.” In psychology, there are some approaches to define “fundamental” or “basic” emotions (see the list given by Ortoni & Turner, 1990, p. 316; see also Izard, 1992). According to Scherer (2005), emotions comprise five components: subjective feeling, cognition, motor expression, action tendencies, and neurological processes. We have to distinguish between the quality of an emotion and its quantity (intensity). To define the qualities, we will follow the cognitive theory of emotion by Oatley & Johnson-Laird (1987) and Power and Dalgleish (1997). Izard (1991, p. 49) presents some criteria to determine which emotions are “basic” or “fundamental”: “1. Fundamental emotions have distinct and specific neural substrates. 2. Fundamental emotions have a distinct and specific configuration of facial movements or facial expressions. 3. Fundamental emotions possess a distinct and specific feeling that achieves awareness. 4. Fundamental emotions were derived through evolutionary-biological processes. 5. Fundamental emotions have organizing and motivational properties that serve adaptive functions.” For Power and Dalgleish and many others, basic emotions are *anger*, *disgust*, *fear*, *happiness*, and *sadness* because all of them do not require propositional content (Oatley & Johnson-Laird). Power and Dalgleish (p. 150) present short definitions of the five basic emotions:

Sadness: Loss or failure (actual or possible) of valued role or goal,

Happiness: Successful move towards or completion of a valued role or goal,

Anger: Blocking or frustration of a role or goal through perceived other agent,

Fear: Physical or social threat to self or valued role or goal,

Disgust: Elimination or distancing from person, object, or idea repulsive to the self and to valued roles and goals.”

These five emotions play important roles in the neuroanatomy of emotions (Phan, Wager, Taylor, & Liberzon, 2002). For Izard (1991, p. 49), there are some more basic emotions such as interest, surprise, contempt, shame, guilt, and shyness. Jörgensen adds only one of the additional emotions to our five basic emotions, namely, surprise (Jörgensen, 2003, p. 28). In our present study, we followed Power and Dalgleish (1997) and worked only with five emotions. Follow-up studies have to consider Izard’s (1991) extension of the set of basic emotions.

Jörgensen (2003, p. 232) defines “emotion in images.” Emotion refers “to specific affective states or mental activity or states of being experienced or seeming to be experienced by the humans or animals in the picture.” Maybe, there are more emotional-laden images, e.g., a grey landscape in fall (without any human or animal) that provokes sadness. Or imagine a photo showing smiling hooligans (smiling is a sign of happiness) knocking other fans around. This photo provokes by no means happiness for all viewers, but probably anger or disgust. So, we define emotion in images broader and change the leading point from an image-based point of view to a viewer-based conception. All images are to be considered as emotional-laden, if they provoke emotions in the viewers, independent from the specific content of the picture.

In psychological research, there are some studies on evoked emotions while viewing images (Ekman & Friesen, 1998; Wild, Erb, & Bartels, 2001). The used images were photos of faces, which showed emotional affects. We learn from the study by Wild et al. (2001) that feelings of happiness or sadness are significantly, specifically, and repeatedly evoked in the viewer and that stronger expressions evoke more emotion.

Abstraction Levels of Tags, Prototypes, and Cognitive Reference Points

User-generated tags can be located on different levels of abstraction. In empirical psychological studies, Eleanor Rosch and her colleagues (Mervis & Rosch, 1981; Rosch, 1975a; Rosch, 1975b; Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976) could show that there are three abstraction levels and that the middle level—the basic level—is most frequently used because here “the information value of attribute clusters is maximized” (Mervis & Rosch, p. 92). Concerning Rosch, the three levels of abstraction are called subordinate, basic, and superordinate:

Suppose that basic objects (e.g., *chair*, *car*) are the most inclusive level at which there are attributes common to all or most members of the category. Then total cue validities are maximized at that level of abstraction at which basic objects are categorized. That is, categories one level more abstract will

be superordinate categories (e.g., *furniture*, *vehicle*) whose members share only a few attributes among each other. Categories below the basic level will be subordinate categories (e.g., *kitchen chair*, *sports car*) which are also bundles of predictable attributes and functions, but contain many attributes which overlap with other categories (for example, *kitchen chair* shares most of its attributes with other kinds of chairs). (Rosch et al., 1976, p. 385)

Basic level concepts are most useful in human life:

Universally, basic object categories should be the basic classifications made during perception, the first learned and the first named by children, and the most codable, most coded, and most necessary in the language of any people (Rosch et al., 1976, p. 435)

In the context of image indexing, Rorissa (2008, p. 1744) states: "Among these three levels, people, in general, prefer to use the basic level when they are thinking about the world around them." Here, one research question arises: Are in fact most of the user-created tags that describe images basic-level terms? A problem in practical research is arising: What are the conditions in which a given term can be assigned to the basic level (or superordinate or subordinate)? Green (2005) applies WordNet (Fellbaum, 1998), which is a lexical database for English terms that arranges the concepts (nouns and verbs, but not adjectives) in a hierarchical way. She is able to show that there is some regularity for the indication of the abstraction level of a noun, e.g., "if a word is longer than 15 characters, it is unlikely that it names a basic level category; if a lexical unit is a phrase, . . . it is unlikely that it is names a basic level category; . . . if a concept has more than four levels beneath it, it probably names a superordinate level category; if a concept has no children, it is unlikely to be basic level, but is much more likely to be a subordinate concept" (Green, 2005). Rorissa and Iyer (2008, p. 1392) add the following criteria besides WordNet: "terms for which one can think of several manifestations (examples)" for superordinate concepts; "convey information on the parts and components of objects" for basic concepts; and "describes a particular kind of a basic level object" for subordinate concepts. Green's, Rorissa's, and Iyer's analyses do not deal with emotional concepts, so further research is necessary to adjust their findings to emotions.

For a given concept, there are some objects that can be seen as prototypes. A prototype is the clearest case or a very good example for a concept. Prototypical members of a category are "those with most attributes in common with other members of that category and are those with least attributes with other categories" (Rosch & Mervis, 1975, p. 576). For example, a robin is a prototype for the concept *bird*, while a penguin is not (penguins fail to fly, which is common for most birds; Jørgensen, 2003, p. 40). So, we meet the problem relating to image tagging: Given a basic emotion, do prototypical images exist? Is it possible to measure prototypicality of images relative to fundamental emotions?

If there are indeed good fitting images or even prototypical images, then it would be an interesting analysis to separate

those cognitive reference points (Rosch, 1975b) that lead the users to tag a certain term or to adjust the scroll bar on a high value for the intensity of a basic emotion. Eleanor Rosch (1975b, p. 545) asks: "Do other natural categories, less obvious than colors, lines, or numbers, have stimuli which serve as reference points?" An example of a reference point is a landmark that people use to navigate through cities. Here, another research area becomes visible: Which are the landmarks of happiness, sadness, etc., on prototypical images? If we are able to find such cognitive reference points, these points can be used as a base for content-based image retrieval. But this research is far behind the study presented here and has to be considered elsewhere.

Methods

Empirical Tests of Collective Indexing of Emotions and Expressing Their Intensities With Scroll Bars

Is the "collective intelligence" of users able to produce consistent indexing results? Can image information systems apply those results to add a new retrieval option of searching emotions? In an empirical study, we tested the indexing of emotions by (linguistic) tags and by scroll bar "tags." To maximize the rate of participants, we constructed an online survey technically based on PHP and HTML. More than 700 subjects participated in our experiments in emotional image tagging.

To determine the intensity of the five basic emotions, we apply a scroll bar to the test set. On a scale from 0 (*there is no intensity of the given emotion quality; there is no evoked emotion*) up to 10 (*very high intensity*), the test persons move the scroll bar to the point of their perceived intensity of emotion (Figure 1).

Test Bed

Twenty-five images were manually selected from Flickr by using *Emotionen* (emotions) and *Emotion* (emotion) as well as the German words of the five fundamental emotions as search terms. Inside this hit set, we looked for images that most clearly represent the particular basic emotion. Furthermore, again manually, we chose five images with quite neutral emotional content as control items. So our test bed comprised a total of 30 images. For three examples, see Figures 2, 4, and 6. For every image, the test persons had the task to move the five scroll bars concerning their experienced feelings and to tag the image with words. They were free to choose their linguistic tags to describe the content (what they see in the picture) or their impressions and associations (what kind of emotion this picture generates). The online survey picked up personal data (such as gender and age), linguistic tags, and scroll bar adjustment (for every image).

Participants

There was a pretest with members of the staff of Düsseldorf's information science department. The actual

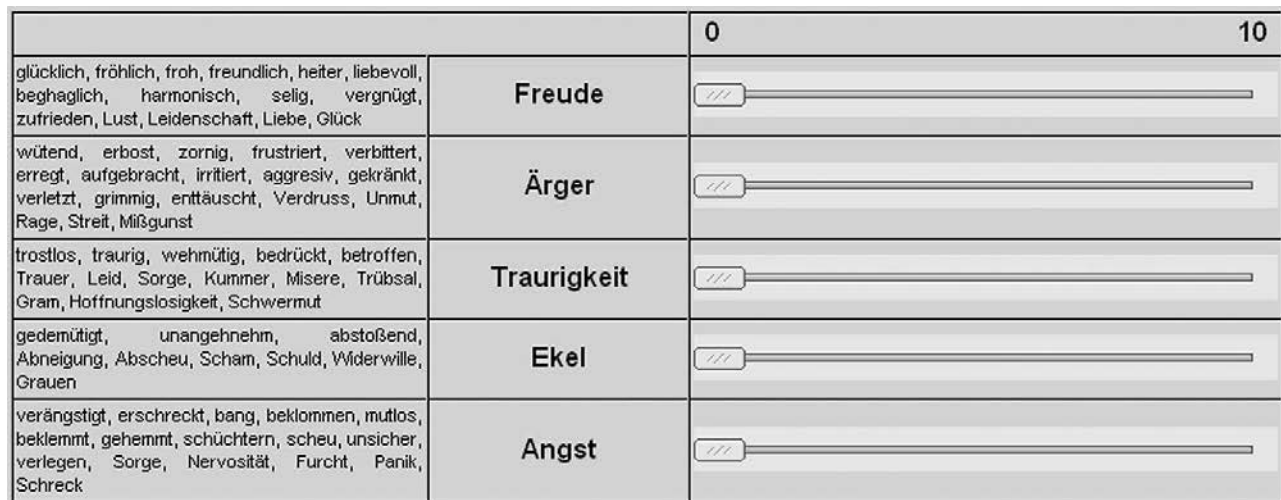


FIG. 1. Scroll bars for tagging intensities of emotions. (happiness [Freude], anger [Ärger], sadness [Traurigkeit], disgust [Ekel], fear [Angst]).

survey was conducted online from January 21 to March 1, 2008; participants were students of the Heinrich-Heine-University Düsseldorf. In all, 1,781 persons participated in the online test, of which 763 persons finished the scroll bar tagging of all 30 images. Only these 763 participants formed our sample. The average age of these test persons was 24 years, 63.6% were female and 36.4% male.

Data Collection and Data Analysis

Prior to importing the results into the database, the adjusted position of the graphically presented scroll bar was converted into numerical values. The empirical basis for the emotions' intensities comprised 22,891 items (scroll bar values). All information was processed in a database and the statistical analysis was performed by means of the data mining and statistical analysis software SPSS. Methods for data analysis followed descriptive statistics (mean, standard deviation) as well as analyses of correlations (Pearson, two-tailed). Linguistic tags were categorized manually (parts of speech: nouns, adjectives, verbs, etc.) and analyzed by frequency. For sample images, we analyzed the abstraction level of the tags the users annotated to these pictures. Using Rosch's basic level theory and the coding schemes by Green (2005) and by Rorissa and Iyer (2008), the tags were sorted into the three known classes: superordinate (a term that is more generic and is defined with less attributes than basic level terms), basic (a term on the medium level), and subordinate (a term that is more specific than basic level terms and that is defined with more properties). The coding was done by the authors. This work was not without problems. Though there are coding schemes, it is yet, for some cases, difficult to determine the semantic level of a term exactly. We tried to solve the problems by discussions. Adjectives were another problem because in WordNet this word form does not include the hierarchy relation (Miller, 1998). So, we additionally looked in WordNet for the derivative noun (e.g., *happiness* and *happy*)

to determine the hierarchical level of the tag. In case of doubt, we coded basic level.

Results

Consistency of Emotional Votes

For 17 out of the 30 images, the test persons had clear emotion "favorites." We want to exhibit this result using two cases (images 2 and 8; see Figures 2 and 4). In Figure 3, we see the basic emotions of our test persons and their intensities tagged by scroll bar concerning image 2. The mean for *happiness* is 7.9 (with a standard deviation [SD] of 2.6). There is a distance of 7.6 points to the next mentioned emotion (*fear*; mean: 0.3; SD: 0.9). In Figure 5, there is a similarly clear favorite emotion, which is caused by image 8. Here, the arithmetic average for *disgust* is 7.4 (SD: 3.1), the distance is up to 6.5 points (to *sadness*; mean: 0.9; SD: 2.0). For reasons of comparison, we add an image with no clear emotional effect (image 12 shown in Figures 6 and 7). The mean for the top-voted emotion is only 1.9 (SD: 2.6), and the figures for the other emotions are very low as well.

To measure the consistency of emotional tagging, we use the standard deviation of the mean for the five basic emotions (Table 1). The lower the value for the standard deviation, the higher is the consistency of votes. What we see is a gradation of consistency from *disgust* (the best consistency) to *happiness* (the lowest consistency). Concerning our results, people obviously agree more on *disgust*, *anger*, and *fear* than on *sadness* and *happiness*. There is an additional but only anecdotal result (Table 1, third and fourth column): The standard deviation of the results of female test persons are—for all five basic emotions—marginally higher than the males' values. Men seem to be a little bit more consistent in emotional tagging than women.

We were not very surprised by the high consistency of the users' votes for some images. Greisdorf and O'Connor (2002) report similar results. In an experimental study, users were



FIG. 2. Image 2. Source: Flickr. Uploaded by Grenzgänger@Eland.

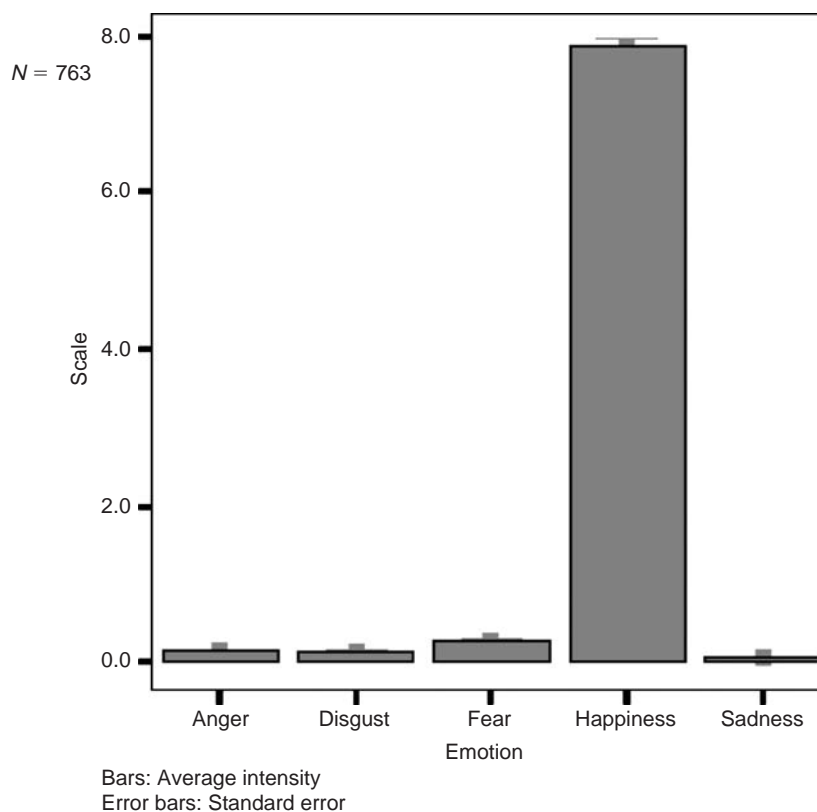


FIG. 3. Emotions and their intensities for image 2. $N = 763$.

instructed to treat given words as topic queries in a search for appropriately matching queries. For a special image, “users significantly responded to [this image] as ‘beautiful’ and ‘happy’” (Greisdorf & O’Connor, p. 15).

We must not confuse two results. On the one hand, the consistent adjustment of the scroll bar is an indicator for the accuracy of the tag. Here, the scroll bar is a tool for the determination of a consistency value. On the other hand,



FIG. 4. Image 8. Source: Flickr. Uploaded by status6.

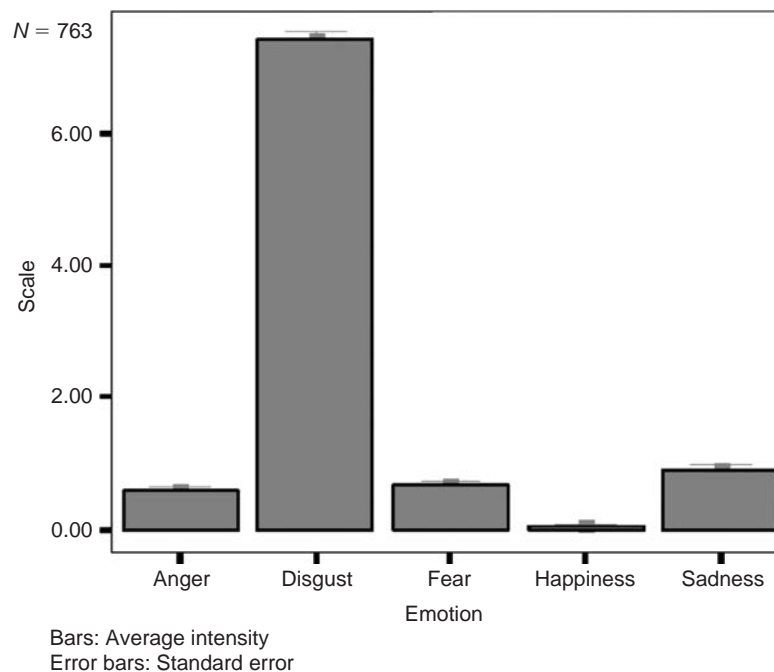


FIG. 5. Emotions and their intensities for image 8. $N = 763$.

the mean scroll bar adjustment is a measure of the emotion's intensity of the given image, which can be used for retrieval purposes.

Correlations Between Basic Emotions

Some pairs of basic emotions show clear correlations (Pearson; Table 2). The two main results are as follows:

- When *happiness* rises, all other emotions decline.
- *Fear* correlates positively with *sadness* and with *anger*.

The correlations are well-known phenomena in psychology. Izard discusses the “sadness–fear bind” (Izard, 1991, pp. 197, 306); and *fear* and *anger* are called the two stress-emotions, which are both answers to perceptions of risk (Lerner & Keltner, 2001).



FIG. 6. Image 12. Source: Flickr. Uploaded by frankierolalala.

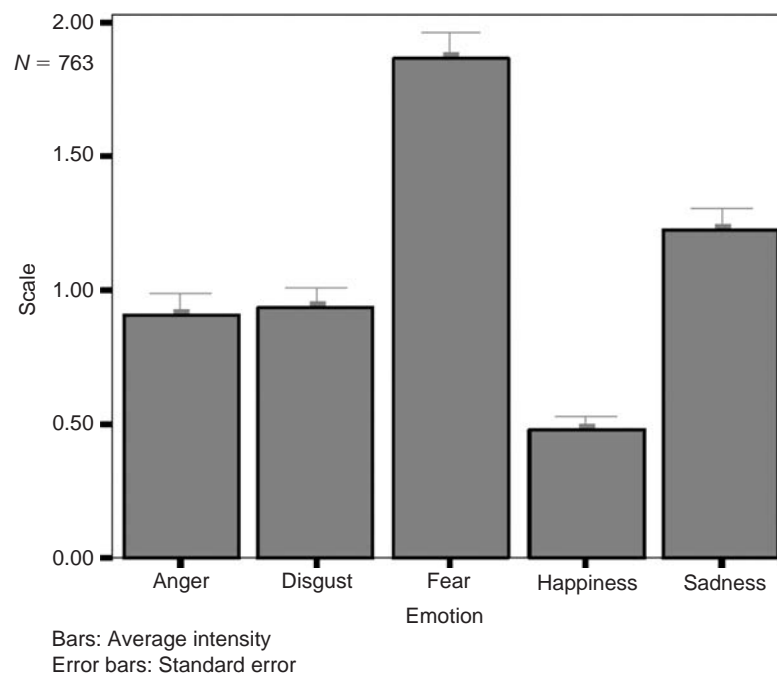


FIG. 7. Emotions and their intensities for image 12. $N = 763$.

TABLE 1. Uniformity of emotional votes.

Basic emotion	SD	Females	Males
Disgust	2.23	2.26	2.18
Anger	2.46	2.45	2.39
Fear	2.74	2.85	2.53
Sadness	3.07	3.17	2.87
Happiness	3.31	3.35	3.23

Note. Standard deviation of the mean of the emotions' intensity. $N = 22,891$ scroll bar tags, 763 participants.

TABLE 2. Correlations between the tested basic emotions.

	Happiness	Anger	Sadness	Disgust	Fear
Happiness	1				
Anger	-0.25	1			
Sadness	-0.32	+0.13	1		
Disgust	-0.18	+0.06	-0.07	1	
Fear	-0.26	+0.24	+0.30	+0.06	1

Note. $N = 22,891$ scroll bar tags, 763 participants. All correlations are significant on the 0.0001-niveau (Pearson, two-tailed).

For music emotion indexing, Lee and Neal (2007, p. 737) found that *happiness* negatively correlates with the other basic emotions (confirmed by our study). Their results demonstrate that *anger* shows a positive correlation with *fear* and *disgust*. (In our study, there is a strong correlation between *anger* and *fear* as well, but nearly no correlation with *disgust*.) Lee and Neal found positive correlations between *sadness* and *fear* (confirmed) and between *sadness* and *disgust* (not confirmed). In music emotion indexing, *fear* was positively correlated with *anger* (confirmed), *sadness* (confirmed), and *disgust* (not confirmed). There is a broad consensus between the correlation results of the two studies. The main difference lies in the role of *disgust*, which, in our study, shows no clear correlations with *anger*, *sadness*, and *fear*. Perhaps *disgust* in music indexing is different from *disgust* in image indexing. *Disgust* is not a preferred feeling in music; composers seldom write pieces of music that sound disgusting.

Emotion-Specific Prototype Images

In Table 3, we see an overview of all favorite emotions caused by the 30 test images. A basic emotion is called "favorite" to a single image if the scroll bar shows an arithmetic average of four or more for the emotion's intensity. (This threshold value is currently arbitrary and is in need of further empirical research.) There was no favorite emotion for 13 images; the test persons could not agree on the emotions caused by these images. Three images caused two favorite emotions (*fear* and *sadness* twice and *fear* and *anger* once—these couples are known for high correlation; see Table 2). The remaining 13 images were tagged with only one favorite basic emotion. For all images with favorite emotions, we calculated the distance to the next ranked emotion. In some

TABLE 3. Intensity of the favorite basic emotion for all test images.

Image No.	Basic emotion	Mean (SD)	Distance
1	Sadness	7.4 (2.7)	5.4
2	Happiness	7.9 (2.6)	7.6
3	Anger	5.4 (3.7)	
4	Fear	4.9 (3.6)	3.4
5	Disgust	5.7 (3.4)	3.7
6	None		
7	Fear	4.6 (3.2)	
8	Sadness	4.5 (3.1)	2.1
9	Sadness	5.6 (3.2)	4.2
10	Disgust	7.4 (3.1)	6.5
11	Anger	4.3 (3.3)	2.4
12	None		
13	None		
14	Happiness	6.3 (3.1)	5.8
15	Happiness	7.7 (2.8)	7.4
16	Happiness	7.1 (2.9)	6.9
17	None		
18	Happiness	4.7 (3.5)	1.5
19	None		
20	None		
21	Happiness	6.8 (3.4)	6.5
22	Sadness	4.1 (3.3)	1.7
23	Fear	5.1 (3.8)	
24	Sadness	4.1 (3.4)	3.1
25	None		
26	None		
27	None		
28	Sadness	4.5 (3.4)	2.5
29	Sadness	5.5 (3.5)	4.4
30	None		

Note. All emotions with an intensity > 4 .

Distance: distance to the intensity of the next ranked emotion. $N = 763$ (for each image).

cases, the distance is very high, e.g., image 2 (distance: 7.6 points on the scroll bar), image 14 (7.4 points), image 15 (6.9 points), image 8, and image 20 (6.5 points each). In such cases, the intensity of the given basic emotion and the distance to other emotions is consistently high, so we can use such tags as means of information retrieval. Such images seem to be prototypes for the given emotion. Of course, there are lots of images that do not affect emotions (e.g., Figure 6). In such cases, there is no possibility to display such an image as an answer to a request on emotions. Based upon an empirical investigation, Jørgensen (1999, p. 352) reports similar results. For some images, there were many terms describing affective attributes (e.g., 37.6% for her photo *Teach*), and for some others, there were few affective terms (e.g., only 1.0% for Jørgensen's photo called *Jungle*). But even images of scientific diagrams are occasionally described by terms of emotion (Brunskill & Jørgensen, 2002, p. 375).

If we have identified prototypical images, then the next question arises: Is it possible to extract special landmarks as cognitive entry points (in the sense of Rosch) into the pictures? Looking again to our example, image 2 (Figure 2),

we realize that there are some smiling faces. From psychological studies (e.g., Wild et al., 2001), we know that some facial expressions are able to provoke consistently specific basic emotions. “Happy faces, in comparison to sad faces, evoked significantly more happiness, were perceived as more pleasant and evoked less sadness, anger and fear” (Wild et al., p. 114). But, it is very probable that there are more types of cognitive entry points besides facial expressions (e.g., the nails on image 8). Within the limited frame of our study we are not able to give an answer. It seems necessary for future research to bridge the gaps between collective image indexing and retrieval and content-based approaches, for known entry points (identified on the basis of collective tagging) play a major role for the automated content-based indexing.

The Abstraction Level of Image Tags

How did the participants perform linguistic tagging? In Tables 4–6, you see the most commonly used tags for our three sample pictures 2, 8, and 12. Nearly half of the test persons indexed image 2 with *happiness*—in the same way they adjusted the scroll bar. There are some more terms in the semantic field of *happiness*, like *exultation* or *fun*, and there is one “mysterious” tag, namely, *soccer*, for on the image you cannot see a football game. We argued that the intensity of the

TABLE 4. Most commonly used tags for image 2.

Rank	Tag	% ^a	Level
1	Freude (happiness)	46.8	basic
2	Spannung (tension)	8.3	basic
3	Jubel (exultation)	4.6	basic
4	fröhlich (happy)	4.4	basic
4	Spaß (fun)	4.4	basic
6	Glück (luckiness)	3.5	basic
7	ausgelassen (frolic)	3.2	subordinate
8	Gemeinschaft (community)	2.4	basic
9	euphorisch (euphoric)	2.0	subordinate
9	Fußball (soccer)	2.0	subordinate
9	heiter (cheerful)	2.0	subordinate

^aPercentage of all users who have tagged this image, $N = 592$.

TABLE 5. Most commonly used tags for image 8.

Rank	Tag	Percent ^a	Level
1	Ekel (disgust)	37.8	basic
2	abstoßend (repulsive)	11.9	basic
3	ekelig/eklig (disgusting)	6.5	basic
4	ekelhaft (disgustful)	4.9	basic
5	ungepflegt (untended)	2.9	basic
6	widerlich (unsavory)	2.7	basic
7	unangenehm (unpleasant)	2.5	basic
7	alt (old)	2.5	basic
9	Abscheu (abhorrence)	2.2	subordinate
10	Alter (seniority)	1.6	subordinate
10	Mitleid (pity)	1.6	subordinate

^aPercentage of all users who have tagged this image, $N = 511$.

TABLE 6. Most commonly used tags for image no. 12.

Rank	Tag	Percent ^a	Level
1	Angst (fear)	7.3	basic
2	Abwehr (defense)	3.9	basic
3	gruselig (creepy)	2.4	subordinate
4	Kunst (art)	2.2	superordinate
5	Ablehnung (refusal)	1.9	basic
5	seltsam (strange)	1.9	basic
7	abstoßend (repulsive)	1.7	basic
7	Distanz (emotional distance)	1.7	subordinate
7	lustig (funny)	1.7	basic
7	unheimlich (weird)	1.7	basic

^aPercentage of all users who have tagged this image, $N = 463$.

emotions is clearly visible on the scroll bar. That is true, but for some cases, it seems to be possible to get analogous results with words. We see (in Table 4) the tag *happy* (a basic level concept) and the further tags *euphoric* and *cheerful*, which are subordinates to *happy* and which express intensities of *happiness*.

More than one third of the participants tagged image 8 with *disgust*, as the majority did on the scroll bar. A lot of the taggers used adjectives similar to *disgust*, namely, *repulsive*, *disgusting*, and *disgustful*.

Analogous to scroll bar adjusting, the test persons did not find clear favorite terms to describe image 12. Only 7.3% of the participants, who actually indexed this picture, attached *fear* and 3.9% selected *defense* to describe the picture’s content.

The test persons and the results could be biased, for the taggers saw the scroll bars. It is open how different the subjects would tag when they do not see the scroll bars of the five basic emotions.

The distributions of the linguistic tags in our three examples follow an inverse power-law (Newman, 2005)—a left-skewed shape with a long tail (Figure 8), which is one of the known distributions in informetrics (Stock, 2006). Such a distribution has the form

$$f(x) = C/x^a,$$

where C is a constant, x is the rank, and a is a value ranging normally from 1 to 2. For our prototypic images (2 and 8), the exponent a has values greater than 2, for the indifferent image (12) a has a value smaller than 1. The higher the value of the exponent a in the power-law formula, the higher is the probability that the image has prototype function concerning a given emotion.

The top 10 tags of the three examples are either substantives or adjectives; there are no verbs. Most of the (most commonly used) tags seem to be on the basic level according to Rosch’s theory. This finding is similar to the results of Rorissa (2008, p. 1747). He found that 63% of tags describing single pictures are basic-level concepts, 28% are on the subordinate, and 9% on the superordinate level.

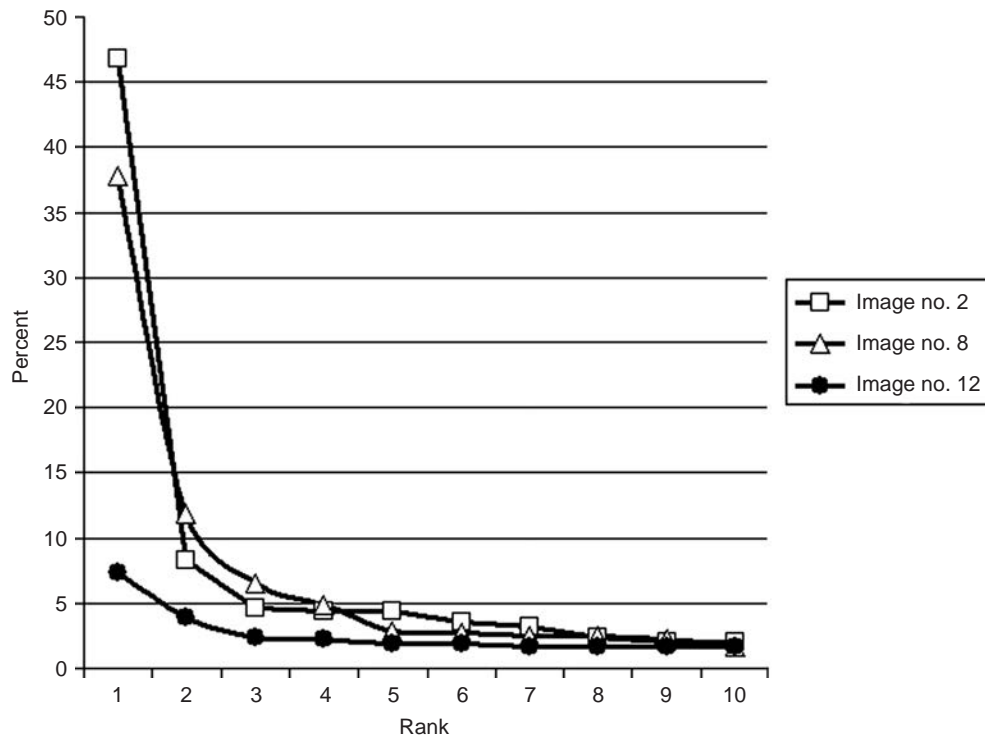


FIG. 8. Relative frequency (%) of the first 10 tags attached to images 2, 8, and 12.

Discussion

Collective Social Tagging and Statistical Aggregation

There may be different approaches to indexing an image like image 8 (Figure 4). A person who uses *disgust* may have a completely different feeling than a person who attaches *pity*. Is there really *one* collaborative description of a document using a folksonomy? Or are there many personal descriptions (of very different personomies) that are only statistically aggregated? This is an important point made by Diane Neal (personal communication, summer 2008): “Most of us are likely disgusted by the person’s feet in Figure 4, but would our response to it not be different if we knew those feet belonged to our aging, ill mother?” So, is there a difference between “collective” tagging, which means all users participate and are not aware of what others do, and “collaborative” tagging in which the system develops shared semantic meanings, which means that the users work together and they (or the systems) find consistent meanings? So far, our study captures only collective tagging (for the difference between collaborative and collective, see Gruber, 2008; Vander Wal, 2008), using statistical means to aggregate the data. But even those statistical calculations have to be handled with caution. “Merely looking at frequencies of tag use is a too simplistic measure. . . e.g., one cannot simply state that a definition is incorrect only because it is hardly used” (Spyns, de Moor, Vandenbussche, & Meersman, 2006, p. 745). Therefore, we cannot state that an image is not emotional-laden if only some (or none) users have tagged it. And we cannot state

that all people feel the same when they tag an image with, e.g., *disgust*.

Emotional Image Retrieval

“In making image systems accessible online, there should be appropriate access points so that users can perform searches on their own” (Choi & Rasmussen, 2003, p. 498). One of these access points is emotion (Choi & Rasmussen, p. 499). Besides ofness (objects of an image), aboutness (interpretation of the image; Shatford, 1986; Stock & Stock, 2008, pp. 31–36, 161), and isness (description of formal attributes; Ingwersen, 2002), there seems to be a further semantic level of image documents: emotiveness (emotions caused by an image). In our opinion, it is possible to retrieve the “unretrievable in electronic imaging systems”—the emotion (Jørgensen, 1999, p. 348). Imagine a picture showing a lucky scene in front of the Eiffel Tower: This is a photo of a tower; it is about a visit to the Eiffel Tower; it is by photographer X; it is exhibited in gallery Y; and, finally, it is emotive in terms of happiness.

The application of emotional tags in image information retrieval systems meets the following three requirements:

- adequate consistent collective indexing
- adequate retrieval algorithms
- adequate retrieval interface for emotional searches

If there are enough tagging users *and* if there is indeed an emotion caused by the image, our results show that it is

possible to get consistent indexing results using the scroll bar tagging method. However, we have to consider that more than 700 test persons participated in our test. It is open whether such amounts of tagging users per document can be reached in real-life situations (maybe lower amounts of scroll bar tags lead to the same consistent results as well; but this is actually an open research question). A possible way to get many scroll bar tags per image is to present the scroll bars to every user who accessed the given image document. (Did this image evoke any emotion? Are you emotionally concerned by the content of this image? If yes, please adjust the scroll bars of the basic emotions.)

The retrieval algorithm has to consider a first threshold value for the intensity of the basic emotion (say, 4 or more on a scale up to 10) and a second threshold value for the distance to the next ranked emotion (say, more than one standard deviation of the mean of the given emotion). The basic emotion that satisfies both presuppositions will be assigned to the image document just like a controlled term. The retrieval interface has to present the new search option for instance via a pull-down menu with a list of all searchable emotions.

Let us go back to the problematic situation of our users in the introduction. Now, our first user applies *people* as keyword and marks *happiness* as basic emotion, and our second user will search for *body parts* or *feet* and marks *disgust* in the emotion's menu. The image information system will be able to find precisely such images that satisfy the topical search argument and that cause the desired emotions.

Further Research

Our study is limited to image indexing and retrieval. The scope has to be broadened to all kinds of documents that are able to achieve emotions, namely, images, videos, music, and textual documents. We tested users' tagging behavior based on a small sample of only 30 images retrieved from Flickr. Our study is descriptive, and this article presents only three case studies. We are in need of much more experimental data. We worked with a set of five basic emotions; follow-up studies could add more emotions, e.g., surprise or shame. The empirical determination of the abstraction level of emotive concepts is in need of further clarification. What are the necessary and what are the sufficient conditions that a given tag is superordinate, basic, or subordinate? Given a sufficiently large set of emotional tags, is it possible to create a knowledge organization system for emotional indexing and retrieval (e.g., an emotional thesaurus)? We could show that people tag emotions via scroll bar more or less consistently. Is the same true for tagging emotions using words? As an anecdotal by-product, we comprise hints that women tag slightly less consistently than men. Is there indeed a gender-specific tagging behavior? And does it matter to emotional information retrieval (EmIR)? We were able to show that there are emotion-specific prototypical images, but we failed to explain why. Furthermore, we failed to present the cognitive reference points of the prototype images. Which are the landmarks of basic emotions on prototypical images that we can apply

in content-based image retrieval? For information retrieval, the most important questions arise: Are users really in need of search systems that can perform searches for emotion? Will users accept EmIR? Besides the problems, we hope that we could show that there is a fascinating research area in computer and information science: EmIR.

Acknowledgments

We would like to thank Diane Neal (North Carolina Central University), Abebe Rorissa (University at Albany, State University of New York), Isabella Peters and Katrin Weller (both of Heinrich-Heine University Düsseldorf), and the anonymous referees for valuable discussions and insightful comments.

References

- Adelaar, T., Chang, S., Lancendorfer, K.M., Lee, B., & Morimoto, M. (2003). Effects of media formats on emotions and impulse buying intend. *Journal of Information Technology*, 18, 247–266.
- Beaudoin, J. (2007). Flickr image tagging: Patterns made visible. *Bulletin of the American Society for Information Science and Technology*, 34(1), 26–29.
- Beaudoin, J. (2008). A preliminary investigation of image indexing: The influence of domain knowledge, indexer experience and image characteristics. In *Proceedings of the 19th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*, Columbus, Ohio.
- Brunskill, J., & Jörgensen, C. (2002). Image attributes: A study of scientific diagrams. In *Proceedings of the 65th Annual Meeting of the American Society for Information Science and Technology* (Vol. 39). Information, Connections, and Community (pp. 365–375).
- Cho, S.B. (2004). Emotional image and musical information retrieval with interactive genetic algorithm. *Proceedings of the IEEE*, 92(4), 702–711.
- Cho, S.B., & Lee, J.Y. (2001). Wavelet encoding for interactive genetic algorithm in emotional image retrieval. In C. Hsu (Ed.), *Advanced signal processing technology by soft computing* (pp. 225–239). Singapore: World Scientific.
- Choi, Y., & Rasmussen, E.M. (2003). Searching for images: The analysis of users' queries for image retrieval in American history. *Journal of the American Society for Information Science and Technology*, 54(6), 498–511.
- Duthoit, C.J., Szynda, T., Lal, S.K.L., Jap, B.T., & Agbinya, J.I. (2008). 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)* (article no. 5). Brussels: Institute for Computer Sciences, Social-Informatics and Telecommunication Engineering.
- Dye, J. (2006). Folksonomy: A game of high-tech (and high-stakes) tag. *E-Content*, 29(3), 38–43.
- Ekman, P., & Friesen, W.V. (1998). Constants across culture in the face and emotion. In J.M. Jenkins, K. Oatley, & N.L. Stein (Eds.), *Human emotions. A Reader* (pp. 63–72). Oxford: Blackwell.
- Enser, P.G.B. (2008a). Visual image retrieval. *Annual Review of Information Science and Technology*, 42, 3–42.
- Enser, P.G.B. (2008b). The evolution of visual information retrieval. *Journal of Information Science*, 34(4), 531–546.
- Fellbaum, C. (Ed.). (1998). *WordNet. An electronic lexical database*. Cambridge, MA: MIT Press.
- Furnas, G.W., Fake, C., von Ahn, L., Schachter, J., Golder, S., Fox, K., et al. (2006). Why do tagging systems work? In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (pp. 36–39). New York: ACM.

- Goetz, E.T., Sadoski, M., Stowe, M.L., Fetsco, T.G., & Kemp, S.G. (1993). Imagery and emotional response in reading literary texts: Quantitative and qualitative analyses. *Poetics*, 22, 35–49.
- Golder, S.A., & Huberman, B.A. (2006). Usage patterns of collaborative tagging systems. *Journal of Information Science*, 32(2), 198–208.
- Gordon-Murnane, L. (2006). Social bookmarking, folksonomies, and Web 2.0 tools. *Searcher—The Magazine for Database Professionals*, 14(6), 26–38.
- Guy, M., & Tonkin, E. (2006). Folksonomies: Tidying up tags? *D-Lib Magazine*, 12(1).
- Goodrum, A.A. (2000). Image information retrieval: An overview of current research. *Informing Science*, 3(2), 63–67.
- Green, R. (2005). Vocabulary alignment via basic level concepts. Final Report, 2003 OCLC/ALISE Library and Information Science Research Grant Project. Dublin, OH: OCLC.
- Greisdorf, H., & O'Connor, B. (2002). Modelling what users see when they look at images: A cognitive viewpoint. *Journal of Documentation*, 58(1), 6–29.
- Gruber, T. (2008). Collective knowledge systems: Where the social Web meets the semantic Web. *Journal of Web Semantics*, 6(1), 4–13.
- Hastings, S.K., Iyer, H., Neal, D., Rorissa, A., & Yoon, J.W. (2007). Social computing, folksonomies, and image tagging: Reports from the research front. In *Proceedings of the 70th Annual Meeting of the American Society for Information Science and Technology*, 45, 1026–1029.
- Hu, X., & Downie, J.S. (2007). Exploring mood metadata: Relationships with genre, artist and usage metadata. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*. Vienna: Austrian Computer Society.
- Ingwersen, P. (2002). Cognitive perspectives of document representation. In *CoLIS 4. 4th International Conference on Conceptions of Library and Information Science* (pp. 285–300). Greenwood Village: Libraries Unlimited.
- Izard, C.E. (1991). *The psychology of emotions*. New York: Plenum Press.
- Izard, C.E. (1992). Basic emotions, relations among emotions, and emotion-cognition relations. *Psychological Review*, 99(3), 561–565.
- Jørgensen, C. (1998). Attributes of images in describing tasks. *Information Processing & Management*, 34(2/3), 161–174.
- Jørgensen, C. (1999). Retrieving the unretrievable in electronic imaging systems: Emotions, themes, and stories. In *Human Vision and Electronic Imaging IV. Proceedings of SPIE*; 3644 (pp. 348–355).
- Jørgensen, C. (2003). *Image retrieval. Theory and research*. Lanham, MD: Scarecrow.
- Jørgensen, C. (2007). Image access, the semantic gap, and social tagging as a paradigm shift. In *Proceedings of the 18th Workshop of the American Society for Information Science and Technology Special Interest Group in Classification Research*, Milwaukee, WI.
- Juslin, P.N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences*, 31(5), 559–575.
- Kalbach, J. (2002). Classifying emotion for information retrieval. *Notes / Music Library Association*, 59(2), 408–411.
- Kennedy, L., Naaman, M., Ahern, S., Nair, R., & Rattenbury, T. (2007). How Flickr helps us make sense of the world: Context and content in community-contributed media collections. In *Proceedings of the 15th International Conference on Multimedia* (pp. 631–640). New York: ACM.
- Kipp, M.E.I. (2006, March). @toread and cool: Tagging for time, task and emotion. Paper presented at the 8th Information Architecture Summit, Las Vegas, NV.
- Kleinginna, P.R., & Kleinginna, A.M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and Emotion*, 5, 345–379.
- Kroski, E. (2008). *Web 2.0 for Librarians and Information Professionals*. New York: Neal-Schuman.
- Laine-Hernandez, M., & Westman, S. (2006). Image semantics in the description and categorization of journalistic photographs. In *Proceedings of the 69th Annual Meeting of the American Society for Information and Technology* (Vol. 44). *Information Realities: Shaping the Digital Future of All*.
- Lee, H.J., & Neal, D. (2007). 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* (Vol. 45). *Joining Research and Practice: Social Computing and Information Science* (pp. 732–741).
- Lerner, J.S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, 81(1), 146–159.
- Li, T., & Ogiwara, M. (2003). Detecting emotion in music. In *Proceedings of the 4th International Conference on Music Information Retrieval (ISMIR 2003)*. Baltimore, MD: Johns Hopkins University.
- Markey, K. (1984). Interindexer consistency tests: A literature review and report of a test of consistency in indexing visual materials. *Library & Information Science Research*, 6, 155–177.
- Markkula, M., & Sormunen, E. (2000). End-user searching challenges indexing practices in the digital newspaper photo archive. *Information Retrieval*, 1(4), 259–285.
- Mathes, A. (2004). *Folksonomies—Cooperative classification and communication through shared metadata*. Urbana, University of Illinois Urbana-Campaign / Graduate School of Library and Information Science.
- McKechnie, L.E.F., Ross, C.S., & Rothbauer, P. (2007). Affective dimensions of information seeking in the context of reading. In D. Nahl & D. Bilal (Eds.), *Information and emotion. The Emergent Affective Paradigm in Information Behavior Research and Theory* (pp. 187–195). Medford, NJ: Information Today.
- Meagher, M.W., Arnau, R.C., & Rhudy, J.L. (2001). Pain and emotion: Effects of affective picture modulation. *Psychosomatic Medicine*, 63, 79–90.
- Mervis, C.B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, 32, 89–115.
- Miller, K.J. (1998). Modifiers in WordNet. In C. Fellbaum (Ed.), *WordNet. An electronic lexical database* (pp. 23–46). Cambridge, MA: MIT Press.
- Morris, J.D., & Boone, M.A. (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.
- Nahl, D. (2007). The centrality of the affective in information behavior. In D. Nahl & D. Bilal (Eds.), *Information and emotion. The Emergent Affective Paradigm in Information Behavior Research and Theory* (pp. 3–37). Medford, NJ: Information Today.
- Neal, D. (2006). *News photography image retrieval practices: Locus of control in two contexts*. Unpublished doctoral dissertation, Denton, University of North Texas.
- Neal, D. (2007). Folksonomy and image tagging: Seeing the future? *Bulletin of the American Society for Information Science and Technology*, 34(1), 7–11.
- Newhagen, J.E. (1998). TV news images that induce anger, fear, and disgust: Effects on approach-avoidance and memory. *Journal of Broadcasting & Electronic Media*, 42(2), 265–276.
- Newman, M.E.J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, 46(5), 323–351.
- Ni, X., Xue, G.R., Ling, X., Yu, Y., & Yang, Q. (2007). Exploring the Weblog space by detecting informative and affective articles. In *Proceedings of the 16th International Conference on World Wide Web* (pp. 281–290). New York: ACM.
- Noruzi, A. (2006). Folksonomies: (un)controlled vocabulary? *Knowledge Organization*, 33(4), 199–203.
- Oatley, K., & Johnson-Laird, P.N. (1987). Towards a cognitive theory of emotions. *Cognition & Emotion*, 1, 29–50.
- Ortoni, A., & Turner, T.J. (1990). What's basic about basic emotions? *Psychological Review*, 97(3), 315–331.
- Paiva, A., Prada, R., & Picard, R.W., Eds. (2007). *Affective computing and intelligent interaction*. Berlin / Heidelberg, Germany: Springer.
- Peters, I., & Stock, W.G. (2007). Folksonomies and information retrieval. In *Proceedings of the 70th Annual Meeting of the American Society for Information Science and Technology*, 45, 1510–1542.
- Peters, I., & Stock, W.G. (2008). Folksonomies in Wissensrepräsentation und Information Retrieval [Folksonomies in knowledge representation and information retrieval]. *Information—Wissenschaft und Praxis*, 59(2), 77–90.

- Peterson, E. (2006). Beneath the metadata. Some philosophical problems with folksonomies. *D-Lib Magazine*, 12(11).
- Phan, K.L., Wager, T., Taylor, S.F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: A meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, 16(2), 331–348.
- Picard, R.W. (1995). Affective computing (Technical Report No. 321). Retrieved February 10, 2009, from <http://affect.media.mit.edu/pdfs/95.picard.pdf>. MIT Media Lab: Affecting Computing Group.
- Picard, R.W. (1997). *Affective Computing*. Cambridge, MA: MIT Press.
- Power, M., & Dalgleish, T. (1997). *Cognition and emotion. From order to disorder*. Hove, East Sussex, UK: Psychology Press.
- Rasmussen, E.M. (1997). Indexing images. *Annual Review of Information Science and Technology*, 32, 169–196.
- Rorissa, A. (2008). User-generated descriptions of individual images versus labels of groups of images: A comparison using basic level theory. *Information Processing & Management*, 44(5), 1741–1753.
- Rorissa, A., & Hastings, S.K. (2004). Free sorting of images: Attributes used for categorization. In *Proceedings of the 66th Annual Meeting of the American Society for Information Science and Technology*, 41, 360–366.
- Rorissa, A., & Iyer, H. (2008). Theories of cognition and image categorization: What category labels reveal about basic level theory. *Journal of the American Society for Information Science and Technology*, 59(9), 1383–1392.
- Rosch, E. (1975a). Cognitive representations of semantic categories. *Journal of Experimental Psychology—General*, 104(3), 192–233.
- Rosch, E. (1975b). Cognitive reference points. *Cognitive Psychology*, 7(4), 532–547.
- Rosch, E., & Mervis, C.B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7(4), 573–605.
- Rosch, E., Mervis, C.B., Gray, W.D., Johnson, D.M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8(3), 382–439.
- Rubin, V.L., Stanton, J.M., & Liddy, E.D. (2004). Discerning emotions in texts. In *Exploring Attitude and Affect in Text: Theories and Applications*. AAAI Spring Symposium Series, Stanford University, March 22–24.
- Salway, A., & Graham, M. (2003). Extracting information about emotions in films. In *Proceedings of the 11th ACM International Conference on Multimedia* (pp. 299–302). New York: ACM.
- Scherer, K.R. (2005). What are emotions? And how can they be measured? *Social Science Information*, 44(4), 695–729.
- Shatford, S. (1986). Analyzing the subject of a picture. A theoretical approach. *Cataloguing and Classification Quarterly*, 6(3), 39–62.
- Shneiderman, B., Bederson, B.B., & Drucker, S.M. (2006). Find that photo! Interface strategies to annotate, browse, and share. *Communications of the ACM*, 49(4), 69–71.
- Smith, G. (2008). *Tagging. People-powered metadata for the social Web*. Berkeley, CA: New Riders.
- Spiteri, L.F. (2006). The use of folksonomies in public library catalogues. *The Serials Librarian*, 51(2), 75–89.
- Spiteri, L.F. (2007). Structure and form of folksonomy tags: The road to the public library catalogue. *Webology*, 4(2), article 41.
- Spyns, P., de Moor, A., Vandenbussche, J., & Meersman, R. (2006). From folkologies to ontologies: How the twain meet. *Lecture Notes in Computer Science*, 4275, 738–755.
- Stock, W.G. (2006). On relevance distributions. *Journal of the American Society for Information Science and Technology*, 57(8), 1126–1129.
- Stock, W.G. (2007). *Information Retrieval*. Munich, Germany: Oldenbourg.
- Stock, W.G., & Stock, M. (2008). *Wissensrepräsentation [Knowledge representation]*. Munich, Germany: Oldenbourg.
- Trant, J. (2006). Exploring the potential for social tagging and folksonomy in art museums: Proof of concept. *New Review of Hypermedia and Multimedia*, 12(1), 83–105.
- Van House, N. (2007). Flickr and public image-sharing: Distant closeness and photo exhibition. Conference on Human Factors in Computing Systems. CHI '07 Extended Abstracts on Human Factors in Computing Systems (pp. 2717–2722). New York: ACM.
- van Zwol, R. (2007). Flickr: Who is looking? In *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 184–190). Washington, DC: IEEE Computer Society.
- Vander Wal, T. (2005). Explaining and Showing Broad and Narrow Folksonomies. Retrieved November 22, 2008, from <http://www.vanderwal.net/random/entrysel.php?blog=1635>
- Vander Wal, T. (2008). Welcome to the matrix! In Gaiser, B., Hample, T., & Panke, S. (Eds.), *Good tags—bad tags. Social Tagging in der Wissensorganisation* (pp. 7–9). Münster, Germany: Waxmann.
- Wang, S., Chen, E., Wang, X., & Zhang, Z. (2003). Research and implementation of a content-based emotional image retrieval. In *Proceedings of the 2nd International Conference on Active Media Technology* (pp. 293–302).
- Wang, S., & Wang, X. (2005). Emotion semantics image retrieval: A brief overview. *Lecture Notes in Computer Science*, 3784, 490–497.
- Wang, W.N., Yu, Y.L., & Jiang, S.M. (2006). Image retrieval by emotional semantics: A study of emotional space and feature extraction. In *IEEE International Conference on Systems, Man and Cybernetics (SMC '06)*, Vol. 4 (pp. 3534–3539).
- Wild, B., Erb, M., & Bartels, M. (2001). Are emotions contagious? Evoked emotions while viewing emotionally expressive faces: Quality, quantity, time course and gender differences. *Psychiatry Research*, 102(2), 109–124.
- Yanbe, Y., Jatowt, A., Nakamura, S., & Tanaka, K. (2007). Can social bookmarking enhance search in the Web? In *Proceedings of the 7th ACM/IEEE-CS Joint Conference on Digital Libraries* (pp. 107–116). New York: ACM.