EmoPhoto: Identification of Emotions in Photos

Soraia Meneses Alarcão
Computer Science and Engineering
IST/UL, Lisbon, Portugal
E-mail: soraiamenesesalarcao@tecnico.ulisboa.pt

***Abstract*—Nowadays, with the development in digital photography and the easiness of taking pictures, the amount of images, both in the Internet and in each person's private collections, is significantly growing. Whenever we interact with our collection, for example, to search for an image of a specific event, the results displayed will always be the same. However, our emotional state is not always the same: sometimes we feel happy, and other times sad, therefore we are more receptive to some images than others. In the worst case, our mood can deteriorate, which, given the impact of the emotions in our daily life, will lead to a worse performance during cognitive tasks such as attention or problem-solving. Although it would be convenient to take advantage of the emotions that an image conveys, currently there is no way of knowing which emotions are associated with a given image.**

In order to identify the emotions in an image, as well as if it transmits negative, positive or neutral feelings, we will use different information: Valence and Arousal, and the content of the image. Given the lack of annotated datasets (with emotional information), we will also provide a new one with the emotional content for each image.

***Index Terms*—emotion recognition, emotions in images, fuzzy logic, content-based image retrieval, emotion-based image retrieval**

I. Introduction

Images are an increasingly important class of data, especially as computers become more usable, with greater memory and communication capacities [21]. Nowadays, with the development in digital photography and the increasing easiness of acquiring cameras and smartphones, taking pictures (and storing them) is a common task. With this massive growth in the amount of visual information available, the need to store and retrieve images in an efficient manner arises, leading to an increase of the importance of Content-based Image Retrieval (CBIR) systems [38]. However, these systems do not take into account high level features like human emotions associated with the images or the emotional state of the users. To overcome this, a new technique, Emotion-based Image Retrieval (EBIR) was proposed in order to extend CBIR systems through the use of human emotions besides common features [45], [41].

We can interact with and explore image collections in many ways. One possibility is through their content, such as Color, Shape, Texture and Line, or through associated information such as tags, data or Global Positioning System (GPS) information. Every time we search for something, for example for images from a specific day or event, the images we receive will always be the same. However, our emotional state is not always the same: sometimes we are happy, and other times sad or depressed. Therefore we are more receptive to some images than others, depending of the emotions perceived from the image. In the worst case, these results will make us feel even worse, which, given the importance of the emotions in our daily life, will lead to a significantly negative performance during cognitive tasks such as attention, creativity, memory, decision-making, judgment, learning or problem-solving.

Although it seems interesting to take advantage of the emotions that an image transmits, for example, by using them to explore a collection of images, currently there is no way of knowing which emotions are associated with a given image. In order to identify the emotional content present in an image, i.e., the emotions that would be triggered when viewing the image, as well as the corresponding category of those emotions (Negative, Positive or Neutral), we will follow two approaches: one using Valence and Arousal information, and the other one using the content of the image.

In research areas that use visual stimuli, there are some datasets available in which each image has information about their Valence and Arousal ratings, and datasets related to facial expressions that are labeled with the corresponding emotions that each facial expression represents. More generally, an image has multiple kinds of information such as date and time, or even GPS data, but none of these are helpful for giving the user any information about the emotional content of the image itself or the impression it will make on a viewer. To achieve this, we developed two recognizers. The first one is based on the Valence and Arousal information associated to images, while the second one is based on the visual content of the images, such as color, shape or texture.

In order to provide a new dataset annotated with the emotional content of each image, we performed a study with different subjects. Besides the creation of the new dataset, we also collected important information about what users thought about the experience, and what influenced the way they felt during the visualization of an image.

This document is organized as follows: in section II, we describe the importance of emotions, as well as how they can be represented. Along with it, we detail the previous works in the recognition of emotions from images, how to identify the emotional state of a user, and some research areas where these two topics are combined: EBIR and Recommendation Systems (RecSys). We also describe the relationship between emotions and the different visual characteristics of an image. Finally, we present the datasets that we used in our work: International Affective Picture System (IAPS), Geneva Affective PicturE

Database (GAPED) and Mikels. In section III, we describe the Fuzzy Logic Emotion Recognizer (FLER) and the corresponding experimental results we achieved, while in section IV, both Content-based Emotion Recognizer (CBER) and the experimental results obtained are described. We also present an analysis of the different possible combinations between the different types of features used: color, texture, composition, and shape. In section V, we present a new dataset that is annotated with information collected through experiments with users. In chapter VI, we present the evaluation of FLER and CBER using our new annotated dataset. Finally, in section VII we present the main conclusions we evolve from this work.

II. Context and Related Work

Emotions have been described as discrete and consistent responses to external or internal events with particular significance for the organism. The role of emotion in human cognition is essential, namely in rational decision-making, perception, and human interaction [34]. Research in psychology [18] has found that temporary changes in affect can also significantly impact performance during cognitive tasks such as memory, attention, learning, judgment, creativity, and problem-solving. Therefore, the importance (and need) of automatic emotion recognition has grown with the increasing role of human-computer interface applications [26].

There are two different perspectives towards emotion representation. The first one (categorial), indicates that basic emotions have evolved through natural selection. Plutchik [35] proposed eight basic emotions: Anger, Fear, Sadness, Disgust, Surprise, Curiosity, Acceptance, and Joy. All the other emotions can be formed by these basic ones. Ekman, following a Darwinian tradition, based his work in the relationship between facial expressions and emotions derived from a number of universal basic emotions: Anger, Disgust, Fear, Happiness, Sadness, and Surprise. In the second perspective (dimensional), which is based on cognition, the emotions are mapped into the Valence, Arousal and Dominance (VAD) dimensions. Valence goes from a very positive feeling to very negative, Arousal is also called activation and goes from states like sleepy to excited, and finally, Dominance that corresponds to the strength of the emotion [11], [26].

In order to extract emotions from an image, we need to understand how its contents affect the way emotions are perceived by users. This content can be facial expressions of the faces present in the images, color, shape or texture information.

The human face is one of the major "objects" in our daily lives that is used to provide information about the gender, attractiveness and age of a person, but also helps to identify the emotion that person is feeling. Underneath our skin, a large number of face muscles allow us to produce different configurations, that can be summarized as Action Unit (AU) [12] and are used to define the facial expressions of an emotion: Joy, Surprise, Anger, Sadness, Disgust and Fear.

To describe how humans perceive and classify facial expressions of an emotion, there are two types of models: the continuous and categorical. The continuous model explains how expressions of emotion can be seen at different intensities, whereas the categorical explains, among other findings, why the images in a morphing sequence between two emotions, like happiness and surprise, are perceived as either happy or surprise but not something in between. There have been developed models of the perception and classification of the six facial expressions of emotion. Initially, they used feature and shape-based algorithms, but, in the last two decades, appearance-based models have been used [31].

CBIR is a technique that uses a set of image features, such as Color, Shape or Texture, to search for images in large databases. Color is the most extensively used visual content for image retrieval since it is the basic constituent of images. It is also the result of interpretation in the brain of the perception of light in the human eye [10]. Research has shown that color is a good predictor for emotions in terms of saturation, brightness, and warmth and that the relationship between colors and human emotions has a strong influence on how we perceive our environment. The same happens for our perception of images, i.e, all of us are in some way emotionally affected when looking at a photograph or an image [41].

Saturation indicates chromatic purity, i.e., corresponds to the intensity of a pixel color. The purer the primary colors the more striking the scenery is to viewers [20]. Brightness corresponds to a subjective perception of the luminance in the pixel's color [10]. Regarding color temperature, warm colors tend to be associated with excitement and danger, while images dominated by cool colors tend to create cool, clamming, and gloomy moods [7]. Usually red is considered to be vibrant and exciting and is assumed to communicate happiness, dynamism, and power. Yellow is the most clear, cheerful, radiant and youthful color. Orange is the most dynamic color and resemble glory. The blue color is deep and may suggest gentleness, fairness, faithfulness, and virtue. Green should elicit calmness and relaxation. Purple sometimes communicates fear, while brown is associated with relaxing scenes. A sense of quietness and calmness can be conveyed by the use of complementary colors, while a sense of uneasiness can be evoked by the absence of contrasting hues and the presence of a single dominant color region.

Since perception of emotion in color is influenced by biological, individual and cultural factors [10], mapping low-level color features to emotions is a complex task which theories about the use of colors, cognitive models and involve cultural and anthropological backgrounds must be considered [30]. Given that colors can be used in different ways, we need effective methods to measure their occurrence in an image, such as Color Moments [38], Color Histogram [38], Fuzzy Histogram (for Dominant Colors) [2], Color Correlogram [38], and the Number of Colors [10].

Shape corresponds to an important criterion for matching objects based on their physical structure and profile [38]. These features can represent the spacial information that is not represented by Texture or Color, and contains all the geometrical information of an object in the image. Growing

evidence indicates that the underlying geometry of a visual image is an effective mechanism for conveying the affective meaning of a scene or object, even for very simple context-free geometric shapes. Objects containing non-representational images of sharp angles are less well liked. Abstract angular geometric patterns tend to be perceived as threatening, and circles and curvilinear forms are usually perceived as pleasant [25]. Usually, perceptual shape features are extracted through angles, lines and curves. The number of angles can be used to describe complexity. Regarding the lines, their directions can express different feelings. Strong vertical elements usually indicate high tensional states while horizontal ones are much more peaceful. Oblique lines could be associated with dynamism [13]. Lines with many different directions present chaos, confusion or action. The longer, thicker and more dominant the line, the stronger the induced psychological effect [30].

Texture is defined as all that is left after color and local shape have been considered. It is used to look for visual patterns, with properties of homogeneity that are not achieved by the presence of a single color, in images and how they are spatially defined. Texture similarity can be used to distinguishing between areas of images with similar color such as sky and sea. It is also important for emotional analysis of an image [13], and its use can change the way other features are perceived; for example, in the case of the emotion unpleasantness, the addition of texture changes the perception of the image's colors [27]. In some situations, a great deal of detail gives a sense of reality to a scene, and less detail implies more smoothing moods [7]. The following texture features, Tamura [38], and Gabor Transform [38] are intended to capture the granularity and repetitive patterns of surfaces in an image.

Harmonious composition is essential in a work of art and useful to analyze an image's character. In terms of composition, images with a simplistic composition and a well-focused center of interest are sometimes more pleasant than images with many different objects [20]. Nature scenes, such as forests or waterscapes, are strongly preferred when compared to urban scenes for population groups from different areas of the world [20]. The most popular and widely known Composition rule is the Rule of Thirds, that can be considered as a sloppy approximation to the 'golden ratio' (about 0.618) [20]. It states that the most important part of an image is not the center of the image but instead at the one third and two third lines (both horizontal and vertical), and their four intersections.

An important thing we need to keep in mind is that human perception of image similarity is subjective, semantic, and task-dependent. Besides that, each type of visual feature usually captures only one aspect of image property, and it is usually hard for the user to specify clearly how different aspects are combined. However, there is no single "best" feature that gives accurate results in any general setting, which means that, usually, a combination of features is needed to provide adequate retrieval results [38].

However, the low-level information used in CBIR systems does not sufficiently capture the semantic information that the user has in mind. In order to solve this, the EBIR systems could be used [22]. These systems are a subcategory of the CBIR that, besides the common features, also use emotions as a feature. The most of research in the area is focused on assigning image mood on the basis of eyes and lips arrangement, but colors, textures, composition and objects are also used to characterized the emotional content of an image, i.e., some expressive and perceptual features are extracted and then mapped into emotions. In some cases it was possible to identify whether a picture is gloomy or not, associate the visual content of an image to adjectives such as aggressive, elegant or calm, or pairs of emotions such as like-dislike or gaudy-plain. Besides the extraction of emotions from an image, there has been an increasing number of attempts to use emotions in different ways such as the increase of the quality of recommendation systems [43]. These systems help users find a small and relevant subset of multimedia items based on their preferences.

Several possibilities have been explored so far in order to induce emotional reactions, relying on different contexts. The most used method of emotion induction is through the presentation of emotionally salient material like pictures, audio or video, without explicitly asking for a personal contribution from the participant. If stimuli are relevant enough, an appraisal is automatically executed and will trigger reactions in other measurable components of emotion such as physiological responses, expressivity, action tendencies, and subjective feeling.

In the different areas of research based on visual stimulation, such as EBIR systems or psychological studies, reliable databases are important for the success of emotion induction. Regarding this, in 1997, the IAPS [24] database was introduced. It contains about 1182 images, and provides a set of normative emotional stimuli for experimental investigations of emotion and attention. Each picture of the database is plotted in terms of the mean Valence and Arousal rating. However, the extensive use of the same stimuli lowers the impact of the images since it increases the knowledge that participants have of the images. In order to increase the availability of visual emotion stimuli, in 2011, a new database called GAPED [8], was created. It contains 730 pictures, 121 representing positive emotions, 89 for the neutral emotions, and 520 for the negative emotions. Pictures were rated according to Valence, Arousal, and the congruence of the represented scene with internal (moral) and external (legal) norms. Even though research has shown that the IAPS is useful in the study of discrete emotions, the categorical structure of the IAPS has not been characterized thoroughly. In 2005, Mikels [33] collected descriptive emotional category data on subsets of the IAPS in an effort to identify images that elicit discrete emotions. This new dataset is composed of 238 images from IAPS: 133 negative and 187 positive, and was annotated with the following emotions: Amusement, Anger, Awe, Contentment, Disgust, Excitement, Fear, and Sadness.

In Table I, it is possible to see that only GAPED and Mikels provide information about the category of an emotion, i.e.,

negative, neutral or positive. Mikels also discriminates the emotions elicited by the images. IAPS does not provide any information about the emotional content of the images that compose the dataset, only V-A information.

	Total	Negative	Neutral	Positive	Emotions
IAPS	1182	N/A	N/A	N/A	No
GAPED	730	520	89	121	No
Mikels	330	133	N/A	187	Yes

TABLE I
COMPARISION BETWEEN IAPS, GAPED AND MIKELS DATASETS

Besides the IAPS and GAPED databases, in which each image was annotated with their Valence and Arousal ratings, there are other databases (typically related to facial expressions) that were labeled with the corresponding emotions, such as NimStim Face Stimulus Set[1], Pictures of Facial Affect (POFA)[2] or Karolinska Directed Emotional Faces (KDEF)[3].

It is possible to see that a lot of work has been done identifying the relationship between emotions and the different visual characteristics of an image, recognizing faces in images and analyzing the emotions that they transmit or even the new technique EBIR used to retrieve images based in emotion's features. However, there is no system for automatically identify the emotional content present in an image.

III. FUZZY LOGIC EMOTION RECOGNIZER

We propose a recognizer to classify an image with the universal emotions present in it and the corresponding category (Negative, Neutral and Positive), based on their V-A ratings using Fuzzy Logic.

A. The Recognizer

In order to map V-A ratings into emotion labels, we used the Circumplex Model of Affect (CMA) [36] which states that all affective states arise from cognitive interpretations of core neural sensations that are the product of two independent neurophysiological systems: Valence and Arousal. It is important to mention that there are a lot of variations of this model with no consensus among them.

To build our dataset for training and testing our recognizer, we used Mikels dataset [39][40][33]. To our purposes, we have made two assumptions: 1) we assume that Amusement, Awe, Contentment and Excitement correspond to the basic emotion Happiness, and 2) besides each isolated emotion, we also consider classes of emotions that often occur together.

According to the assumptions made, our initial dataset is composed by 1 image of Anger, Disgust and Fear (ADF), 6 images of Anger, Disgust and Sadness (ADS), 1 image of Anger and Fear (AF), 1 image of Anger and Sadness (AS), 31 images of Disgust (D), 25 images of Disgust and Fear (DF), 11 images of Disgust and Sadness (DS), 12 images of Fear (F), 3 images of Fear and Sadness (FS), 114 images of Happiness (Ha), and finally, 43 images of Sadness (S). Given that we removed the classes of emotions with fewer samples (less than 5), the resulting dataset includes: ADS, D, DF, DS, F, Ha and S.

For each image in the dataset, we started by normalizing the V-A values (ranging between -0.5 and 0.5). Then, we divided the Cartesian Space, using these values, in order to define each class of emotions, and as we can see in Figure 1 there was a huge confusion among the different classes.

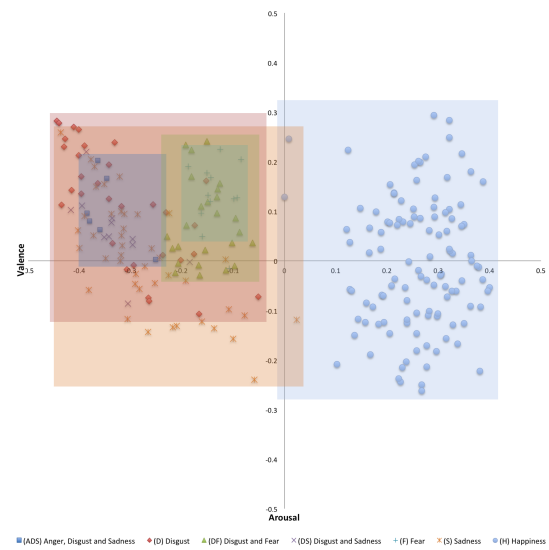

Fig. 1. Distribution of the images in terms of Valence and Arousal

In order to reduce the existing confusion, and considering the Circumplex Model of Affect, we used the Polar Coordinate System to represent each image in terms of Angle (see Equation 1) and Radius (see Equation 2), each computed using the V-A. Angle was used to identify the class of emotion for each image belongs to, while Radius was used to help reduce emotion confusion between images with similar angles.

$$Angle(valence, arousal) = \arctan(\frac{arousal}{valence}) \in [0°, 360°] \quad (1)$$

$$Radius(valence, arousal) = \sqrt[2]{valence^2 + arousal^2} \in [0, \frac{\sqrt[2]{2}}{2}] \quad (2)$$

Even with the use of Angle and Radius to describe each image, it still exists confusing among the different classes of emotions, so instead of using rigid intervals we decided to used Fuzzy Set Theory to describe each class of emotions, as well as the categories. A fuzzy set corresponds to a class of objects with a continuum Degree of Membership (DOM), where each set is characterized by a membership function, usually denoted as $\mu_A(x)$, which assigns to each object a DOM with a range between zero and one [46]. In our work we used the Product of Sigmoidal membership function (see Equation 3) and the Trapezoidal membership function (see Equation 4).

A sigmoidal function depends on two parameters: a and c. The first one controls the slope, while the second is the center of the function. Depending on the sign of the parameter a, the function is inherently open to the right or to the left. The

[1]http://www.macbrain.org/resources.htm
[2]http://www.paulekman.com/product/pictures-of-facial-affect-pofa/
[3]http://www.emotionlab.se/resources/kdef

Product of Sigmoidal membership function is given by:

$$psigmf(x : a_1, c_1, a_2, c_2) = \frac{1}{1 + e^{-a_1(x-c_1)}} \times \frac{1}{1 + e^{-a_2(x-c_2)}} \quad (3)$$

The trapezoidal curve depends on four scalar parameters a, b, c, and d (see Equation 4). The parameters a and d locate the "feet" of the trapezoid and the parameters b and c locate the "shoulders".

$$trapmf(x : a, b, c, d) = max(min(\frac{x-a}{b-a}, 1, \frac{d-x}{d-c}), 0) \quad (4)$$

Regarding the computation of the parameters for the membership functions, we started by using $mean$ and $stddev$ measures. In the case of the classes of emotions, for the Angle membership function, both measures were used for the slope parameters, i.e., a_1 and a_2. The parameters c_1 and c_2 correspond, respectively, to the lowest and highest value of the Angle for that subset of images. For the Radius membership function, b is the minimum, and c the maximum value of the Radius for that subset of images, while $a = b - \epsilon_1$ and $d = c + \epsilon_2$, with $\epsilon_1 = \epsilon_2 = 0.01$ (empirical value). In the case of the categories parameters, and since we used trapezoidal memberships for the Angles and the Radiuses, b and c parameters correspond to the lowest and highest value of the Angle/Radius for that subset of images; in the case of the parameters a and d, the only difference are the ϵ values that vary according to each category. For all the classes of emotions, we removed the outliers, i.e, images with angles or radius that were distant from the angles or radius of the majority of the images for the corresponding class.

Although fuzzy sets are commonly defined using only one dimension, they can be complemented with the use of cylindrical extensions. Given this, we used a two-dimensional membership function that is the result of the composition of the two membership functions mentioned above. For each category, we used Trapezoidal membership function, both for angle and radius. In the case of the classes of emotions, we used the Product of Sigmoidal membership function for the angle and the Trapezoidal membership function for the radius. Each image was annotated with the degree of membership for each possible category and class of emotions, and were also associated to the image the two dominant categories and the two dominant classes of emotions.

In Figure 2 there is a global view of the membership functions of all the classes of emotions for Angle, being possible to see the existing confusion between each of the classes of emotions. There is clearly a differentiation between the positive emotion Happiness, ($[0°, 95°] \cup [300°, 360°]$) and the negative emotions ($[120°, 280°]$). However, there is a lot of confusion among the negative emotions, being the main confusions between DF and F, between D, DF, ADS and DS, and finally between ADS, DS and S.

In Figure 3 there is a global view of the confusion between emotions regarding Radius. In this case, and contrary to what happened in the case of Angle, there is no clear differentiation between negative and positive emotions. As we can see, there

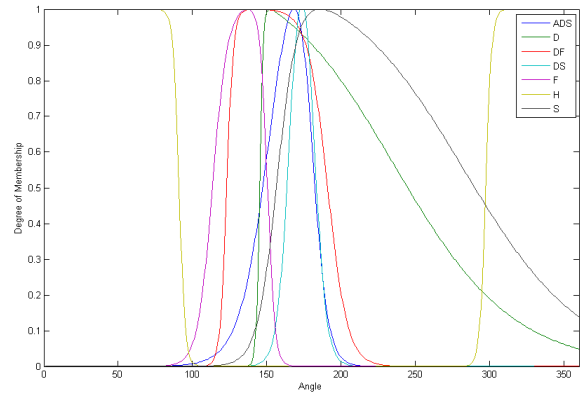

Fig. 2. Membership Functions of Angle for all classes of emotions

are no emotions in the proximity of the extremes (0 and 70), in fact the emotions are lying between 8 and 55. Almost all emotions are completely inside the D interval ([10, 55]). In some cases, such as F, which is completely inside DF, or S, which almost overlaps completely DF, the radius will not be particularly helpful. However, and considering the results for the angle, in the case of confusions between ADS, DS and S, the use of radius will be useful. So, the combination of the two attributes (Angle and Radius) allow us to better distinguish the emotions.

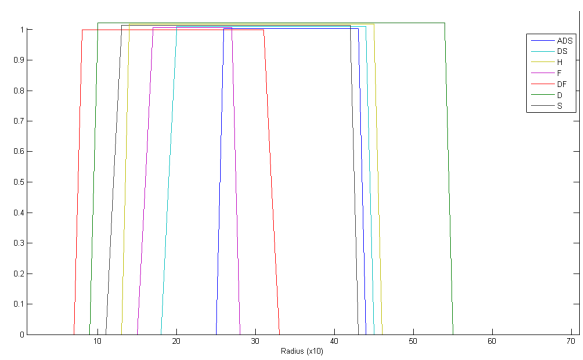

Fig. 3. Membership Functions of Radius for all classes of emotions

B. Experimental Results

For the proper evaluation of our model for the classes of emotions, and taking into account that "Clinicians and researchers have long noted the difficulty that people have in assessing, discerning, and describing their own emotions. This difficulty suggests that individuals do not experience, or recognize, emotions as isolated, discrete entities, but that they rather recognize emotions as ambiguous and overlapping experiences." as stated in [36], we consider that a result is correct if the expected class of emotion is present (totally or partially) in the result label; if it is not present, we consider the class of emotion with the biggest DOM as a confusion. For example, if the expected class of emotion was D, we considered correct results ADS, DS, D, DF or any combination of one of those with a second class of emotion.

As we can see in Table II, the best results were achieved for D, F and Ha. In the case of Ha, this result is due to the clear distinction between the angle values for Ha and the remaining emotions; for D we believe it is due to the big interval, both for angle and radius; DF shows the worst result, but this is expectable given that both the Angle and Radius intervals are overlapping with the majority of the emotions.

(%)	**ADS**	**D**	**DF**	**DS**	**F**	**S**	**H**
ADS	**83.33**					16.67	
D		**100**					
DF			**76**	8		16	
DS		9.09		**90.91**			
F					**100**		
S		4.65	4.65			**90.70**	
H							**100**

TABLE II
CONFUSION MATRIX FOR THE CLASSES OF EMOTIONS IN THE IAPS DATASET

For evaluating the model when it come to categories, we follow a similar approach as the one described above. As expected the results achieved when using the same set for training and test were 100% for negative and positive categories. With the use of GAPED as a testing set and with the addition of the neutral category (which did not exist in the training dataset), the results for the negative category became worse; however, this is mainly due to the existing confusion between the negative and neutral categories in the used dataset [8]. The positive category maintains an accuracy of 100%, while the neutral category obtained almost 99%. These results were achieved after the adjustment of the radius parameters for each category.

IV. CONTENT-BASED EMOTION RECOGNIZER

There is a general agreement on the fact that humans can perceive all levels of image features, from the primitive/syntactic to the highly semantic ones [37], and also that artists have been exploring the formal elements of art, such as lines, space, mass, light or color to express emotions [10]. We assumed that emotional content is characterized by the image Coloss, Texture and Shape. In order to acquire as much information as possible about an image, we will use different features regarding Color, Texture, Shape, Composition, among others. As stated in [37], [16], low-level image features can be easily extracted using computer vision methods, however they are no match for the information a human observer perceives. After the identification of the features that can be used to describe an image in terms of their emotional content, we will train different classifiers in order to understand which ones are the best features to describe an image according to their category. We used the following descriptor vector features: AutoColorCorrelogram (ACC) [19], Color Histogram (CH) [38], Color Moments (CM) [9], Number of Different Colors (NDC) [10], Opponent Histogram (OH) [44], Perceptual Fuzzy Color Histogram (PFCH) [1], [2], Perceptual Fuzzy Color Histogram with 3x3 Segmentation (PFCHS) [1], Reference Color Similarity (RCS) [23], Gabor (G) [32], Haralick (H) [17], Tamura (T) [42], Edge Histogram (EH) [3], Rule of Thirds (RT) [9], Color and Edge Directivity Descriptor (CEDD) [4], Fuzzy Color and Texture Histogram (FCTH) [6] and Joint Composite Descriptor (JCD) [5]. The majority of the features were extracted using jFeatureLib [14] and LIRE [28], [29], although PFCH, PFCHS, RT and NDC were implemented by us. Given this, we want to identify the combination of visual features that can match human perception as closely as possible regarding the positive or negative content of an image.

For the classification, we used Weka 3.7.11, a data mining software [15]. For the simple classifiers we used Naive Bayes (NB), Logistic (Log), John Platt's sequential minimal optimization algorithm for training a support vector classifier (SMO), C4.5 Decision Tree (algorithm from Weka) (J48), Random Forest (RF), and K-nearest neighbours (IBk). In the case of meta classifiers, i.e., classifiers based on other classifiers, we used LogitBoost (LB), RandomSubSpace (RSS), and Bagging (Bag). For the combination of classifiers we used Vote with the Average combination rule. Although one of the good practices of machine learning is to use normalized data, in our tests we did not find any difference in the results, so we kept the features unnormalized. For testing and training, we used Mikels dataset [39], [40]. The Positive images are the ones with the Happiness label, while the Negative ones correspond to ADS (6), D (31), DF (20), DS (11), F (12) and S (43) labels. We separate the data into a training and test set using K-fold Cross Validation with K = 5 [30].

We started by analyzing a set of classifiers, in order to understand which one learned best the relation between features and the given category of emotion. For these preliminary tests, we used all the features, but without any combination between them, and we do not consider the time required to build the model. With the used classifiers, we achieved average recognition rates between 52.75% and 56.62%. However, after the observation of the results, we were not able to choose only one classifier. Based on these relations (for each feature), we studied the following combinations of classifiers (using Vote classifier):

- V1: Vote(SMO+NB+LB+Log+Bag)
- V2: Vote(SMO+NB+LB+RF+RSS)
- V3: Is similar to Vote 2 (V2), but with default configurations for the LB and RSS classifier
- V4: Vote(SMO+NB+LB)
- V5: Vote(SMO+NB+Log)
- V6: Vote(SMO+NB)

Considering the preliminary results, we performed more tests using different combinations of features inside each feature type (Color, Composition, Texture, Shape and Joint) using the six vote classifiers. We consider the following combinations as the best ones: CH+CM+NDC+RCS, CH+OH+RCS, CH+CM+NDC+RCS+H, CH+OH+RCS+CEDD, CH+PFCH+RCS+H, CH+PFCHS+RCS+RT+H+T, CH+RCS+H+FCTH+JCD, and CH+PFCHS+RCS+RT+H,

with recognition rates above 68.00% for V2, and 66.50% for Vote 4 (V4).

In order to confirm if our selected combinations really discriminate an image in terms of their emotional content, we also trained the two classifiers V2 and V4 using a new dataset with images from GAPED. For the first tests, we used 121 Negative images (31 from Animal, 30 from Human, 30 from Snake, and 30 from Spider, chosen randomly) and 121 Positive images. For the tests using Positive and Negative categories in which the best combinations were CH+OH+ RCS+CEDD (70.25%) for Positive category, and CH+PFCH+RCS+H (82.11%) for Negative, in both cases using classifier V4.

Given the results achieved until now, and in order to select the best classifier and combination of features for our final recognizer, we created a new dataset of 468 images selected from both Mikels dataset and GAPED dataset. From each one we selected 121 Negative and 113 Positive images, giving us a total of 242 Negative images and 226 Positive images. We divided the dataset with $\frac{2}{3}$ for training (312 images), and the remaining for test (156 images). The best combination for the Negative category was OH+PFCHS+RCS+RT+H+T using classifier V4 (88.50%), while for the Positive it was CH+RCS+H+FCTH+JCD using classifier V2 (61.54%). For both the categories, and the one that we choose as the best overall combination was CH+CM+NDC+RCS that using classifier V2 achieved an average recognition rate of 72.44% (Negative: 87.18% and Positive: 57.69%).

Given that we performed all the tests using a small number of observations, and that in the majority of the tests the amount of features used for each image is considerably bigger than the number of observations, we consider the possibility of overfitting. Although we used cross-validation in all the tests we performed, which is helpful in reduce overfit in classifiers, we decided to verify if our final classifier suffers from overfitting. If our model was overfitting, when we use images for test that were not used to train the model, the classifier should perform considerably worst, however the recognition rates were similar as the ones using the training set. Additionally, we also tried reduce the number of features used, we also performed Principal Component Analysis (PCA), however the results achieved were that all the features used are important.

The recognizer uses a Vote classifier based on SMO, NB, LB, RF, and RSS, and is composed by CH, CM, NDC, and RCS features.

V. Dataset

In order to provide a new dataset annotated with the emotional content of each image, we performed a study with different subjects.

A. Image Selection

Concerning the creation of the dataset, we started by selecting the images, using the results of the recognizer developed in section III, from the IAPS, GAPED and Mikels datasets. From the first one we selected 86 images: 9 of A (Anger), ADS, D, DF, DS, F, Ha, Neutral (N), S and 5 of Surprise (Su). From GAPED we selected 76 images: 8 of A, ADS, D, DF, DS, F, Ha, N, S, and 4 of Su. Finally, from Mikels we selected 7 images: 1 for ADS, D, DF, DS, F, Ha and S. For each class of emotions, we selected images with the biggest DOM possible. The dataset contains multiple images with animals, such as snakes, spiders, dogs, sharks, horses, cats, tiger, among others. The remaining images include children, war scenarios, mutilation, poverty, diseases and death situations. It also include images from cirurgical procedures, as well as images of natural catastrophes, car accidents or fire.

For the experience, we divided the dataset into 4 subsets: DS0 to DS3. The first one contains 57 images, 20 from our subset of IAPS, 20 from our subset of GAPED, and all the images from our subset of Mikels. This dataset will be rated by all the participants. Dataset DS1 contains 40 images, while DS2 and DS3 contain 36 images each.

B. Description of the Experience

First, we started by explaining the purposes of the study and how it will be held. To ensure the willingness of the subject, regarding Negative images, we started by showing three images as examples of what can be expected. After that, the subject could decide to continue or not the study. If the subject decides to continue the study, s/he should fill the user's questionnaire with his/her personal information (age, gender, etc.), as well as the classification of their current emotional state (categories and emotions).

There are 7 different blocks with nearly 14 images each. Images were presented in a random order, i.e., each user will have a different sequence of images. Each image was shown to the user during 5 seconds. After looking at the image, the user should evaluate his/her current emotional state, and rate the image according to each of the universal emotions using a 5-Likert scale. When the user fills all the requested information for that image, the Next button appears and s/he can move on to the next image. Although in other study users usually have a limited time to respond, we decided not to do it. This way we allowed the user to spend as much time as needed, without feeling pressured to respond or even stressed out. In order to relax and avoid user fatigue, we provided a 30 seconds interval during which only a black screen was displayed.

C. Pilot Tests

In order to verify and validate if our procedure has any error and also if it is completely clear to the subjects, we performed two preliminary tests with different subjects. The first one was a 27 year old man, that performed the test in Portuguese. The second subject was an 18 year old female, that also preferred to take the test in Portuguese. Regarding an image that was duplicated, none of the subjects had any doubt or detected any error in our application for collecting their emotional information. An interesting aspect of the performed tests was the sensitivity to the negative images. The first subject considered the majority of the images very violent, while the second one considered them almost neutral, and in some cases she enjoyed the negative content. These

preliminary results demonstrated how subjective the emotional content of an image can be.

D. Results

We conducted 60 tests: 26 with females and 34 with males, with 70% of them belonging to the 18-29 age group. None of the users had participated in a study using IAPS or GAPED database. In fact, the overwhelming majority had no knowledge about these databases. Regarding their current emotional state, 31 participants classify it as Neutral, 25 as Positive, and only 4 as Negative. The majority of the participants were feeling moderately Happy or moderately Neutral, both with a Median of 3, in the beginning of the tests. Given the number of participants in our tests, each image of DS0 was rated by 60 participants, while each image of DS1, DS2 and DS3 was rated by 20 participants.

We compared the achieved results, concerning categories, between our dataset and the GAPED/Mikels datasets for each of the images of our dataset, in order to obtain the agreement between them. In the case of the images from Mikels dataset the agreement was 100% for the Positive image, where in the case of the Negative there is confusion between the Negative and Neutral categories. For the GAPED dataset we analyzed 76 images (33 Negative, 9 Positives, and 34 Neutral). For the Neutral and Positive categories the achieved agreement was 100% for each, while in the case of the Negative, similarly to what happens for Mikels dataset, there is confusion between the Negative and Neutral categories.

During each session, participants were encouraged to share their opinions/comments about the experience. More than 40% of participants indicated some type of difficulty in understanding the content of some of the images, leading to confusion about their feelings. The majority identified the lack of context as the main reason for this. For example, some users did not understand if an animal in front of a car will be hit by it or not. In this case there is confusion between feeling negative if the animal is hit, and neutral/positive otherwise. Besides this concrete example, one of these users explained that if he sees an image of a hideous act that is made based on religious fanaticism, he feels disgusted and angry, but if it is due to necessity (poverty, to get food, etc.) he only felt sadness.

Some of the users (2) indicated that there are too many emotions to rate. However, other users (8) suggest that there should be an option such as confusion or anxiety, because they consider that some images do not correspond to any of the available emotions. Moreover, another user considered that Happiness is not enough to discriminate the positive feelings of some images, such as cuteness. In the case of surprise, some users (5) claimed that it is subjective, difficult to understand and difficult to elicit from an image. There seem to be some exceptions to this, such as images with unexpected content like a lamp or stairs. However, two users considered surprise as one of the most common emotions in the beginning of the test, but that tends to disappear during the test. In the case of anger, two users explicitly indicated us that none of the images was able to trigger that emotion. In the case of the neutral emotion, and given the existence of the neutral category, four users did not comprehend the use of the emotion suggesting that a rate of “3” in all the other emotions will be equivalent to "feeling neutral"; one user suggested the use of indifference instead of neutral.

Regarding the personal taste of the users, some of them do not appreciate spiders (4), snakes (3) or aquatic animals (1), but some of them consider images with these animals beautiful because of the colors in them. However, the opposite is also true (some users appreciate snakes (4) or spiders (3)). Two users hate needles, one user hates hunting and another is afraid of “heights”, i.e, he reported that he felt fear from an image in which he thinks that should feel happy. In the contrary, two users identified that a specific image should be considered as “negative”, but since they enjoy the content in it (fire and cirurgical instruments), they felt positive and happy. Finally, one user also noticed that in a image with a couple in which the woman is pregnant, usually this scenario would be neutral to him, but as his sister is pregnant, he feels happy because he remembers his sister.

One of the users was particularly happy in the beginning of the test, and reported us he did not feel affected by the images. However, after viewing various negative images, he said that his emotional state was getting worse. In fact, more users (4) stated that sequential negative images, for example 3, negatively affects the emotional state more than, for example, one negative, one neutral, one negative. The same happens for a positive image, the user feels positive, but he is also influenced by the negative images, so he did not feel so happy as he “should”. However, two users justified that, given the extensive amount of sequential negative images, they tend to rate a positive image with a higher value. Finally, some users mentioned that the emotional content of the last image also interferes in the way they were feeling at that moment.

Concerning the design of the test, six users considered it very long, i.e., with too many images, and two other users suggested that should have been more positive images. A large number of users also reported that the test had too many images of snakes (18) or spiders (7). With so many images of snakes/spiders, the users (6) reported that they got used to them, and stopped feeling afraid or disgusted. To avoid this, some users (3) suggested the use of salamanders, grasshoppers, scorpions or maggots. In the case of the pause screen, seven users considered it very long. At least one user appreciated the pause screen, and suggested the use of a timer to indicate the time left for resting. Finally, some users (6) explained that it was complicated to analyse what they were feeling, given that it was very subjective, and also difficult to rate from 1 to 5; two of them gave the example that they would only give a rating of 5 in extreme cases, such as if they started crying or laughing out loud.

The existing comments as well as the reported inconsistencies represent a minority of participants (10%). The remaining participants did the experiment as it should be, and their responses were aligned with the emotions that were supposed to be transmitted by the images.

VI. Evaluation

Each image of the new dataset was classified by the two recognizers: FLER and CBER. Concerning the categories, each image was annotated with the dominant category using CBER and FLER; in the later each image was annotated with up to two dominant categories. In the case of the emotions, only FLER was used to annotate the image with the most dominant emotions.

A. Fuzzy Logic Emotion Recognizer

From our dataset we used 21 Positive, 67 Neutral and 81 Negative images. In the case of the Negative category, the achieved recognition rate was 100%, while in the Positive category, it achieved almost 86%. For the Neutral category, the achieved results was considerably worst (only 28%). When we compared these results with the ones achieved using only GAPED dataset the Negative and Positive categories achieved good results, but the Neutral category achieved a poor result; it decreases from almost 99% to 28%. This result can be explained by the lack of agreement between the results from our users and the previous classification of the images from the GAPED dataset, as well as the existing confusion between Negative and Neutral category for the GAPED.

Concerning the classification in terms of the emotions that an image conveys, we considered that a given emotion is present in the image if the median of the values assigned by users to that emotion was ≥ 2.0. Considering this, from the 169 images that compose our dataset, almost 23% did not have any emotion associated. From the 131 images with emotions associated, there were no images with the emotions Anger or Surprise. For the remaining Negative emotions, we had 18 images of Sadness, 8 of Fear, and 5 images associated with Disgust. In the case of Happiness there were 17 images, while for Neutral we had 36. In the case of combinations of emotions in the same image, we had: DS (8), AS (7), HaN (4), DF(3), FSu (2), NS (2), AD (1), AF (1), DSu (1), FS (1), ADS (7), AFS (3), ADF(2), DFN (1), DFSu (1), FHaN (1), HaNSu (1), and ADSSu (1).

To check if the emotion identified by our recognizer was correct, we assumed that a result is considered correct, if at least one of the emotions for our dataset is present in the emotions identified by the recognizer. For example if an image has the emotions ADS from the dataset, all the following emotions, from the recognizer, will be considered correct: A, D, S, AD, AS, or DS. Given this, our FLER achieved a success rate of 68.70%.

B. Content-Based Emotion Recognizer

For this evaluation, and given that CBER only classifies an image in terms of being Negative or Positive, we did not consider the Neutral images of our dataset; therefore we used 21 Positive and 81 Negative images. For the negative category, we achieved a classification rate of 76.54%, and of 52.38% to the positive category. If we compare these results with the ones obtained earlier, in both cases there was a decrease in the recognition rates. Although this can be justified by the use of only one category, for each image, given that even across our users, in many cases and for different reasons, there is no consensus about which feeling each image transmits.

VII. Main Results and Contributions

With the completion of this work, we achieved the following results:

- A **Fuzzy recognizer** with a classification rate of 100% in the case of categories and 91.56% in the case of emotions, for Mikels dataset [33]; with the GAPED we achieved an average classification rate of 95.59% for the categories. Using our dataset, we achieved a success rate of 68.70% for emotions. In the case of categories, we achieved 100% for Negative category, 88% for the Positive and 28% for the Neutral.
- A recognizer based on the **content of the images**, that has a recognition rate of 87.18% for the Negative category, and 57.69% for the Positive, using a dataset of images selected both from IAPS and from GAPED datasets. Using our dataset, we achieved a recognition rate of 76.54% for the Negative category and 53.28% for the Positive.
- A **new dataset** of 169 images from IAPS and GAPED annotated with the **dominant categories and emotions**, according to what people felt while viewing each image.

VIII. Conclusions and Future Work

Although there is a lot of work done in understanding the content of an image, the majority of this work did not specifically focus on the emotions or categories that an image conveys. With our work we provided two recognizers of the emotional content of an image: categories and emotions, using the Valence and Arousal information from the image, or the visual content of the image. This allow us to increase the number of images annotated with their emotions without the need of manual classification, reducing both the subjectivity of the classification and the extensive use of the same stimuli.

From the experimental evaluation of the developed recognizers, we can establish new guidelines for the work to be done in the future. Concerning FLER, we used a the distribution of the images, according to each class of emotion, that is not balanced, and in the majority of the cases, there are a small number of images of each. Given this, we consider important to use more annotated images to adjust the Fuzzy Sets for each class of emotions, and consequently the Fuzzy Sets for each of the categories.

Considering the results obtained throughout this work in the case of the categories, the next possible step is to join the two recognizers into one. If a particular image, provided as input to the "new" recognizer has information about their Valence and Arousal values, a weighting system should be use between the values of DOM (assigned by FLER) and the estimated probability (assigned by the CBER) in order to classify the image. Otherwise, it should be used only CBER classification. Further, we suggest to complement the new dataset with data collected using an Brain-Computer Interfaces (BCI) device

(e.g., Emotiv). This way, each image will have the emotion felt, and the emotion reported by the users.

REFERENCES

[1] JC Amante. *Colorido : Identificação da Cor Dominante de Fotografias*. PhD thesis, 2011.

[2] JC Amante and MJ Fonseca. Fuzzy Color Space Segmentation to Identify the Same Dominant Colors as Users. *DMS*, 2012.

[3] Shih-Fu Chang, T Sikora, and A Purl. Overview of the MPEG-7 standard. *Circuits and Systems for Video Technology, IEEE Transactions on*, 11(6):688–695, June 2001.

[4] SA Chatzichristofis and YS Boutalis. CEDD: color and edge directivity descriptor. A compact descriptor for image indexing and retrieval. *Computer Vision Systems*, pages 312–322, 2008.

[5] Savvas A Chatzichristofis, Y S Boutalis, and Mathias Lux. Selection of the proper compact composite descriptor for improving content based image retrieval. In B Zagar, editor, *Signal Processing, Pattern Recognition and Applications (SPPRA 2009)*, page 0, Calgary, Canada, February 2009. ACTA Press.

[6] Savvas A Chatzichristofis and Yiannis S Boutalis. FCTH: Fuzzy Color and Texture Histogram - A Low Level Feature for Accurate Image Retrieval. In *Proceedings of the 2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, WIAMIS '08, pages 191–196, Washington, DC, USA, 2008. IEEE Computer Society.

[7] Chin-han Chen, MF Weng, SK Jeng, and YY Chuang. Emotion-based music visualization using photos. *Advances in Multimedia Modeling*, 4903:358–368, 2008.

[8] Elise S Dan-Glauser and Klaus R Scherer. The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance. *Behavior research methods*, 43(2):468–77, June 2011.

[9] Ritendra Datta, Dhiraj Joshi, Jia Li, and James Z Wang. Studying Aesthetics in Photographic Images Using a Computational Approach. In *Proceedings of the 9th European Conference on Computer Vision - Volume Part III*, ECCV'06, pages 288–301, Berlin, Heidelberg, 2006. Springer-Verlag.

[10] CM de Melo and Jonathan Gratch. Evolving expression of emotions through color in virtual humans using genetic algorithms. *Proceedings of the 1st International Conference on Computational Creativity ({ICCC-X)}*, 2010.

[11] Paul Ekman. *Basic emotions*, chapter 3, pages 45–60. John Wiley & Sons Ltd, New York, 1999.

[12] Paul Ekman and Wallace Friesen. Pictures of Facial Affect. *Consulting Psychologists Press, Palo Alto, CA*, 1976.

[13] Syntyche Gbèhounou, François Lecellier, Christine Fernandez-maloigne, and U M R Cnrs. Extraction of emotional impact in colour images. *6th European Conference on Colour in Graphics, Imaging and Vision*, 2012.

[14] Franz Graf. JFeatureLib, 2012.

[15] Mark Hall, Eibe Frank, and Geoffrey Holmes. The WEKA Data Mining Software: An Update. *ACM SIGKDD*, 11(1):10–18, 2009.

[16] Alan Hanjalic. Extracting Moods from Pictures and Sounds. *IEEE SIGNAL PROCESSING MAGAZINE*, (March 2006):90–100, 2006.

[17] R Haralick, K Shanmugam, and I Dinstein. Texture Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 3(6), 1973.

[18] Lane Harrison, Drew Skau, and Steven Franconeri. Influencing visual judgment through affective priming. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2949–2958, 2013.

[19] Jing Huang, S Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih. Image Indexing Using Color Correlograms. *1997 IEEE Conference on Computer Vision and Pattern Recognition*, 0:762, 1997.

[20] Dhiraj Joshi, Ritendra Datta, Elena Fedorovskaya, Quang-tuan Luong, James Z Wang, Li Jia, and Jiebo Luo. Aesthetics and Emotions in Images [A computational perspective]. *IEEE Signal Processing Magazine*, (SEPTEMBER 2011):94–115, 2011.

[21] A Khokher and R Talwar. Content-based image retrieval: state of the art and challenges. *(IJAEST) INTERNATIONAL JOURNAL OF ADVANCED ENGINEERING SCIENCES AND TECHNOLOGIES*, 9(2):207–211, 2011.

[22] Youngrae Kim, So-jung Kim, and Eun Yi Kim. EBIR: Emotion-based image retrieval. *2009 Digest of Technical Papers International Conference on Consumer Electronics*, pages 1–2, January 2009.

[23] Hans-Peter Kriegel, Erich Schubert, and Arthur Zimek. Evaluation of Multiple Clustering Solutions. In *MultiClust@ECML/PKDD*, pages 55–66, 2011.

[24] P.J. Lang, M.M. Bradley, and B.N. Cuthbert. International affective picture system (IAPS): Affective ratings of pictures and instruction manual. *NIMH Center for the Study of Emotion and Attention*, 1997.

[25] Christine L. Larson, Joel Aronoff, and Elizabeth L. Steuer. Simple geometric shapes are implicitly associated with affective value. *Motivation and Emotion*, 36(3):404–413, October 2011.

[26] Yisi Liu, Olga Sourina, and MK Nguyen. Real-time EEG-based emotion recognition and its applications. *Transactions on computational science XII*, 2011.

[27] Marcel P. Lucassen, Theo Gevers, and Arjan Gijsenij. Texture affects color emotion. *Color Research & Application*, 36(6):426–436, December 2011.

[28] Mathias Lux and Savvas A Chatzichristofis. Lire: Lucene Image Retrieval: An Extensible Java CBIR Library. In *Proceedings of the 16th ACM International Conference on Multimedia*, MM '08, pages 1085–1088, New York, NY, USA, 2008. ACM.

[29] Mathias Lux and Oge Marques. *Visual Information Retrieval Using Java and LIRE*. Synthesis Lectures on Information Concepts, Retrieval, and Services. Morgan & Claypool Publishers, 2013.

[30] Jana Machajdik and Allan Hanbury. Affective image classification using features inspired by psychology and art theory. *Proceedings of the international conference on Multimedia - MM '10*, page 83, 2010.

[31] Aleix Martinez and Shichuan Du. A Model of the Perception of Facial Expressions of Emotion by Humans: Research Overview and Perspectives. *Journal of Machine Learning Research : JMLR*, 13:1589–1608, May 2012.

[32] S Marčelja. Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am.*, 70(11):1297–1300, November 1980.

[33] Joseph a Mikels, Barbara L Fredrickson, Gregory R Larkin, Casey M Lindberg, Sam J Maglio, and Patricia a Reuter-Lorenz. Emotional category data on images from the International Affective Picture System. *Behavior research methods*, 37(4):626–30, November 2005.

[34] W.R. Picard. Affective Computing, 1995.

[35] R. Plutchik. The nature of Emotions. *Am. Sci.*, 89(4):344–350, 2001.

[36] Jonathan Posner, James a Russell, and Bradley S Peterson. The circumplex model of affect: an integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology*, 17(3):715–34, January 2005.

[37] Thomas Rorissa, Abebe; Clough, Paul; Deselaers. Exploring the Relationship Between Feature. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY*, 59(5):770–784, 2008.

[38] DS Shete and MS Chavan. Content Based Image Retrieval: Review. *International Journal of Emerging Technology and Advanced Enginnering*, 2(9):85–90, 2012.

[39] A. Smith. A new set of norms. *Behavior Research Methods, Instruments, and Computers*, (3x(x), xxx-xxx), 2004.

[40] A. Smith. Smith2004norms.txt. *Retrieved October 2, 2004 from Psychonomic Society Web Archieve: http://www.psychonomic.org/ARCHIEVE/*, 2004.

[41] Martin Solli. *Color Emotions in Large Scale Content Based Image Indexing*. PhD thesis, 2011.

[42] H Tamura, S Mori, and T Yamawaki. Texture features corresponding to visual perception. *IEEE Transactions on Systems, Man and Cybernetics*, 8(6), 1978.

[43] M Tkalčič, A Kosir, and J Tasic. Affective recommender systems: the role of emotions in recommender systems. *Decisions@RecSys*, 2011.

[44] Koen E a van de Sande, Theo Gevers, and Cees G M Snoek. Evaluating color descriptors for object and scene recognition. *IEEE transactions on pattern analysis and machine intelligence*, 32(9):1582–96, September 2010.

[45] WN Wang and YL Yu. Image emotional semantic query based on color semantic description. *Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou*, (August):18–21, 2005.

[46] L. A. Zadeh. Fuzzy Sets*. *Information and Control*, 8:338–353, 1965.