

On Shape and the Computability of Emotions

Xin Lu Poonam Suryanarayan Reginald B. Adams, Jr.

Jia Li Michelle G. Newman James Z. Wang*

The Pennsylvania State University, University Park, Pennsylvania
{xinlu, pzs126, regadams, jiali, mgn1, jwang}@psu.edu

ABSTRACT

We investigated how shape features in natural images influence emotions aroused in human beings. Shapes and their characteristics such as roundness, angularity, simplicity, and complexity have been postulated to affect the emotional responses of human beings in the field of visual arts and psychology. However, no prior research has modeled the dimensionality of emotions aroused by roundness and angularity. Our contributions include an in-depth statistical analysis to understand the relationship between shapes and emotions. Through experimental results on the International Affective Picture System (IAPS) dataset we provide evidence for the significance of roundness-angularity and simplicity-complexity on predicting emotional content in images. We combine our shape features with other state-of-the-art features to show a gain in prediction and classification accuracy. We model emotions from a dimensional perspective in order to predict valence and arousal ratings which have advantages over modeling the traditional discrete emotional categories. Finally, we distinguish images with strong emotional content from emotionally neutral images with high accuracy.

Categories and Subject Descriptors

H.3.1 [Information Storage and Retrieval]: Content analysis and indexing; I.4.7 [Image Processing and Computer Vision]: Feature measurement

General Terms

Algorithms, Experimentation, Human Factors

*J. Li and J. Z. Wang are also affiliated with the National Science Foundation. This material is based upon work supported by the Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Foundation.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

MM'12, October 29-November 2, 2012, Nara, Japan.

Copyright 2012 ACM 978-1-4503-1089-5/12/10 ...\$15.00.

Keywords

Human Emotion, Psychology, Shape Features

1. INTRODUCTION

The study of human visual preferences and the emotions imparted by various works of art and natural images has long been an active topic of research in the field of visual arts and psychology. A computational perspective to this problem has interested many researchers and resulted in articles on modeling the emotional and aesthetic content in images [10, 11, 13]. However, there is a wide gap between what humans can perceive and feel and what can be explained using current computational image features. Bridging this gap is considered the “holy grail” of computer vision and the multimedia community. There have been many psychological theories suggesting a link between human affective responses and the low-level features in images apart from the semantic content. In this work, we try to extend our understanding of some of the low-level features which have not been explored in the study of visual affect through extensive statistical analyses.

In contrast to prior studies on image aesthetics, which intended to estimate the level of visual appeal [10], we try to leverage some of the psychological studies on characteristics of shapes and their effect on human emotions. These studies indicate that roundness and complexity of shapes are fundamental to understanding emotions.

- **Roundness** - Studies [4, 21] indicate that geometric properties of visual displays convey emotions like anger and happiness. Bar *et al.* [5] confirm the hypothesis that curved contours lead to positive feelings and that sharp transitions in contours trigger a negative bias.
- **Complexity of shapes** - As enumerated in various works of art, humans visually prefer simplicity. Any stimulus pattern is always perceived in the most simplistic structural setting. Though the perception of simplicity is partially subjective to individual experiences, it can also be highly affected by two objective factors, parsimony and orderliness. Parsimony refers to the minimalistic structures that are used in a given representation, whereas orderliness refers to the simplest way of organizing these structures [3].

These findings provide an intuitive understanding of the low-level image features that motivate the affective response, but the small scale of studies from which the inferences have been drawn makes the results less convincing. In order



Figure 1: Example images from IAPS (The International Affective Picture System) dataset [15]. Images with positive affect from left to right, and high arousal from bottom to top.

to make a fair comparison of observations, psychologists created the standard International Affective Picture System (IAPS) [15] dataset by obtaining user ratings on three basic dimensions of affect, namely *valence*, *arousal*, and *dominance* (Figure 1). However, the computational work on the IAPS dataset to understand the visual factors that affect emotions has been preliminary. Researchers [9, 11, 18, 23, 25, 26] investigated factors such as color, texture, composition, and simple semantics to understand emotions, but have not quantitatively addressed the effect of perceptual shapes. The study that did explore shapes by Zhang *et al.* [27] predicted emotions evoked by viewing abstract art images through low-level features like color, shape, and texture. However, this work only handles abstract images, and focused on the representation of textures with little accountability of shape.

The current work is an attempt to systematically investigate how perceptual shapes contribute to emotions aroused from images through modeling the visual properties of roundness, angularity and simplicity using shapes. Unlike edges or boundaries, shapes are influenced by the context and the surrounding shapes influence the perception of any individual shape [3]. To model these shapes in the images, the proposed framework statistically analyzes the line segments and curves extracted from strong continuous contours. Investigating the quantitative relationship between perceptual shapes and emotions aroused from images is non-trivial. First, emotions aroused by images are subjective. Thus, individuals may not have the same response to a given image, making the representation of shapes in complex images highly challenging. Second, images are not composed of simple and regular shapes, making it difficult to model the complexity existing in natural images [3].

Leveraging the proposed shape features, the current work attempts to automatically distinguish the images with strong emotional content from emotionally neutral images. In psychology, emotionally neutral images refer to images which evoke very weak or no emotions in humans.

Also, the current study models emotions from a non-categorical or discrete emotional perspective. In previous work, emotions were distinctly classified into categories like anger, fear, disgust, amusement, awe, and contentment,

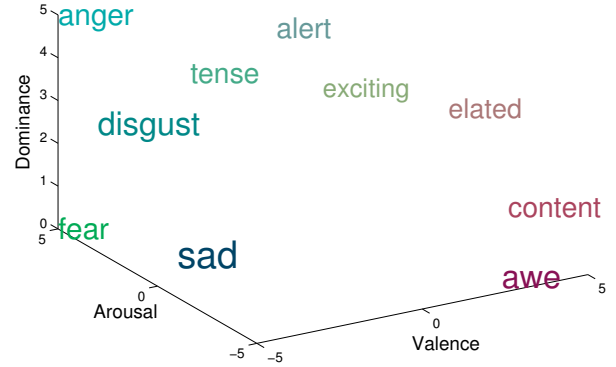


Figure 2: Dimensional representation of emotions and the location of categorical emotions in these dimensions (Valence, Arousal, and Dominance).

among others. This paper is, to our knowledge, the first to predict emotions aroused from images by adopting a dimensional representation (Figure 2). Valence represents the positive or negative aspect of human emotions, where common emotions, like joy and happiness, are positive, whereas anger and fear are negative. Arousal describes the human physiological state of being reactive to stimuli. A higher value of arousal indicates higher excitation. Dominance represents the controlling nature of the emotion. For instance, anger can be more controlling than fear. Researchers [2, 12, 28] have investigated the emotional content of videos through the dimensional approach. Their emphasis was on the accommodation of the change in features over time rather than low-level feature improvement. However, static images, with less information, are often more challenging to interpret. Low-level features need to be punctuated.

This work adopts the dimensional approaches of emotion motivated by recent studies in psychology, which argued for the strengths of dimensional approaches. According to Bradley and Lang [6], categorized emotions do not provide a one-to-one relationship between the content and emotion of an image since participants perceive different emotions in the same image. This highlights the utility of a dimensional approach, which controls for the intercorrelated nature of human emotions aroused by images. From the perspective of neuroscience studies, it has been demonstrated that the dimensional approach is more consistent with how the brain is organized to process emotions at their most basic level [14, 17]. Dimensional approaches also allow the separation of images with strong emotional content from images with weak emotional content.

In summary, our **main contributions** are:

- We systematically investigate the correlation between visual shapes and emotions aroused from images.
- We quantitatively model the concepts of roundness-angularity and simplicity-complexity from the perspective of shapes using a dimensional approach.
- We distinguish images with strong emotional content from those with weak emotional content.

The rest of the paper is organized as follows, Section 2 provides a summary of previous work. Section 3 introduces some definitions and themes which recur throughout the paper. The overall framework followed by details of the

perceptual shape descriptors are described in Section 4. Experimental results and in-depth analyses are presented in Section 5. We conclude in Section 6.

2. RELATED WORK

Previous work [11, 26, 18] predicted emotions aroused by images mainly through training classifiers on visual features to distinguish categorical emotions, such as happiness, anger, and sad. Low-level stimuli such as color and composition have been widely used in computational modeling of emotions. Affective concepts were modeled using color palettes, which showed that the bag of colors and Fisher vectors (i.e., higher order statistics about the distribution of local descriptors) were effective [9]. Zhang *et al.* [27] characterized shape through Zernike features, edge statistics features, object statistics, and Gabor filters. Emotion-histogram and bag-of-emotion features were used to classify emotions by Solli *et al.* [24]. These emotion metrics were extracted based on the findings from psychophysiological experiments indicating that emotions can be represented through homogeneous emotion regions and transitions among them.

The first work that comprehensively modeled categorical emotions, Machajdik and Hanbury [18] used color, texture, composition, content, and semantic level features such as number of faces to model eight discrete emotional categories. Besides the eight basic emotions, to model categorized emotions, adjectives or word pairs were used to represent human emotions. The earliest work based on the Kansei system employs 23 word pairs (e.g., like-dislike, warm-cool, cheerful-gloomy) to establish the emotional space [23]. Along the same lines, researchers enumerated more word pairs to reach a universal, distinctive, and comprehensive representation of emotions in Wang *et al.* [25]. Yet, the aforementioned approaches of emotion representation ignore the interrelationship among types of emotions.

3. CONCEPT INTERPRETATION

This work captures emotions evoked by images by leveraging shape descriptors. Shapes in images are difficult to capture, mainly due to the perceptual and merging boundaries of objects which are often not easy to differentiate using even state-of-the-art segmentation or contour extraction algorithms. In contemporary computer vision literature [7, 20], there are a number of statistical representations of shape through characteristics like the straightness, sinuosity, linearity, circularity, elongation, orientation, symmetry, and the mass of a curve. We chose roundness-angularity and simplicity-complexity characteristics because they have been found previously by psychologists to influence the affect of human beings through controlled human subject studies. Symmetry is also known to effect emotion and aesthetics of images [22]. However, quantifying symmetry in natural images is challenging.

To make it more convenient to introduce the shape features proposed, this section defines the four terms used: line segments, angles, continuous lines, and curves. The framework for extracting perceptual shapes through lines and curves is derived from [8]. The contours are extracted using the algorithm in [1], which used color, texture, and brightness of each image for contour extraction. The extracted contours are of different intensities and indicate the algorithm’s confidence on the presence of edges.

Considering the temporal resolution of our vision system, we adopted a threshold of 40%. Example results are presented in Figures 3, 4, 5, and 6. Pixels with an intensity higher than 40% are treated equally, which results in the binary contour map presented in the second column. The last three columns show the line segments, continuous lines, and curves.

Line segments - Line segments refer to short straight lines generated by fitting nearby pixels. We generated line segments from each image to capture its structure. From the structure of the image, we propose to interpret the simplicity-complexity. We extracted locally optimized line segments by connecting neighboring pixels from the contours extracted from the image [16].

Angles - Angles in the image are obtained by calculating angles between each of any two intersecting line segments extracted previously. According to Julian Hochberg’s theory [3], the number of angles and the number of different angles in an image can be effectively used to describe its simplicity-complexity. The distribution of angles also indicates the degree of angularity of the image. A high number of acute angles makes an image more angular.

Continuous lines - Continuous lines are generated by connecting intersecting line segments having the same orientations with a small margin of error. Line segments of inconsistent orientations can be categorized as either corner points or points of inflexion. Corner points, shown in Figure 7(a), refer to angles that are lower than 90 degrees. Inflexion points, shown in Figure 7(b), refer to the midpoint of two angles with opposite orientations. Continuous lines and the degree of curving can be used to interpret the complexity of the image.

Curves - Curves are a subset of continuous lines, the collection of which are employed to measure the roundness of an image. To achieve this, we consider each curve as a section of an ellipse, thus we use ellipses to fit continuous lines. Fitted curves are represented by parameters of its corresponding ellipses.

4. CAPTURING EMOTION FROM SHAPE

For decades, numerous theories have been promoted that are focused on the relationship between emotions and the visual characteristics of simplicity, complexity, roundness, and angularity. Despite these theories, researchers have yet to resolve how to model these relationships quantitatively. In this section, we propose to use shape features to capture those visual characteristics. By identifying the link between shape features and emotions, we are able to determine the relationship between the aforementioned visual characteristics and emotions.

We now present the details of the proposed shape features: line segments, angles, continuous lines, and curves. A total of 219 shape features are summarized in Table 1.

4.1 Line segments

Psychologists and artists have claimed that the simplicity-complexity of an image is determined not only by lines or curves, but also by its overall structure and support [3]. Based on this idea, we employed line segments extracted from images to capture their structure. Particularly, we used the orientation, length, and mass of line segments to determine the complexity of the images.

Orientation - To capture an overall orientation, we employed statistical measures of minimum (min), maximum

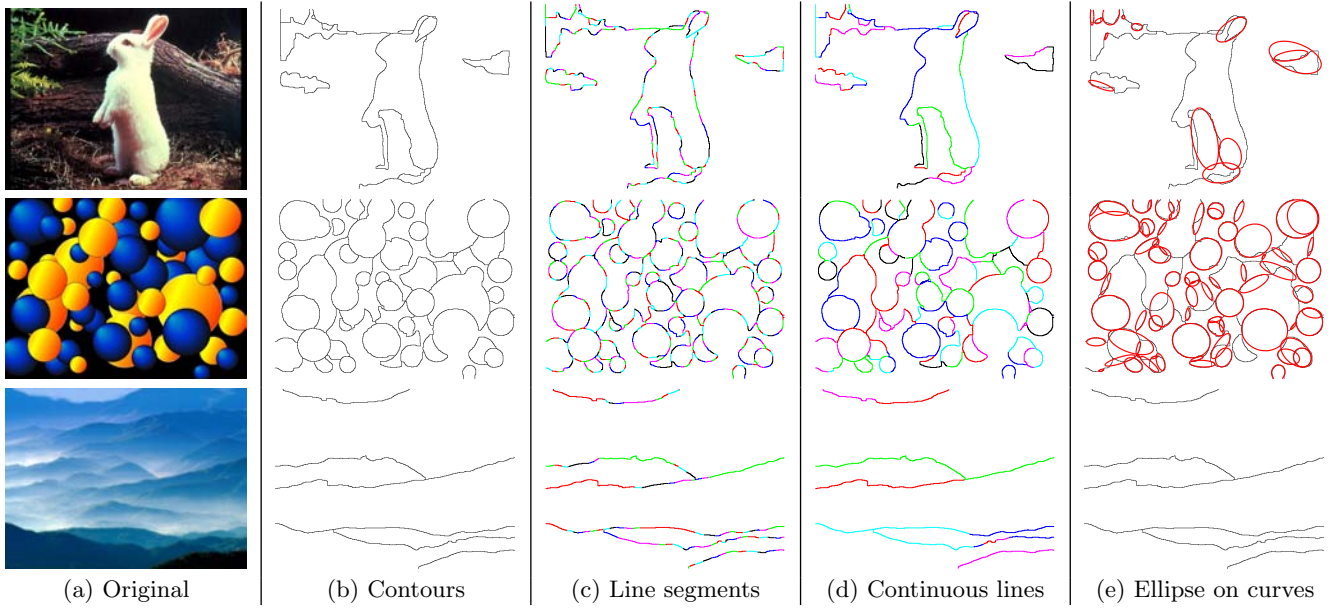


Figure 3: Perceptual shapes of images with *high valance*.

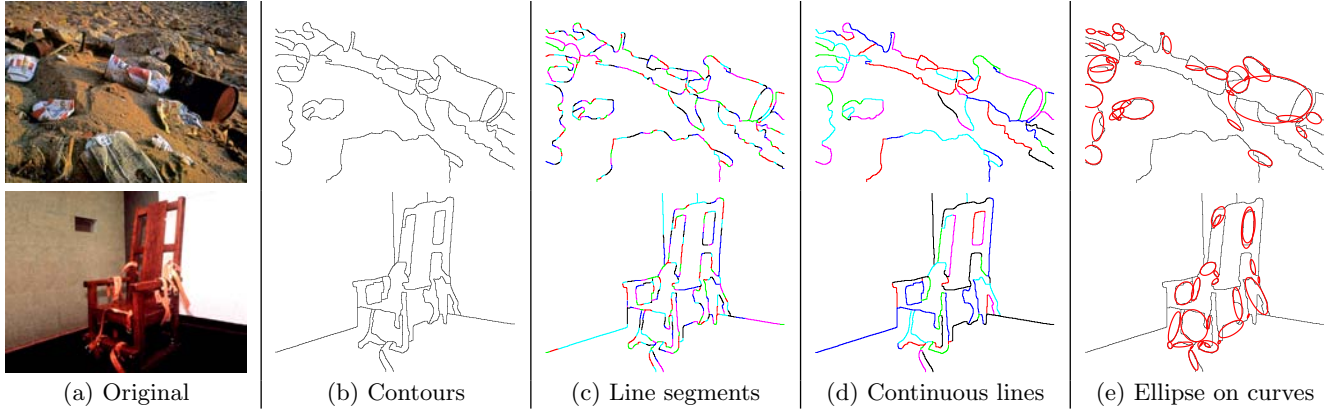


Figure 4: Perceptual shapes of images with *low valance*.

Table 1: Summary of shape features.

Category	Short Name	#
Line Segments	Orientation	60
	Length	11
	Mass of the image	4
Continuous Lines	Degree of curving	14
	Length span	9
	Line count	4
	Mass of continuous lines	4
Angles	Angle count	3
	Angular metrics	35
Curves	Fitness	14
	Circularity	17
	Area	8
	Orientation	14
	Mass of curves	4
	Top round curves	18

(max), 0.75 quantile, 0.25 quantile, the difference between 0.75 quantile and 0.25 quantile, the difference between max and min, sum, total number, median, mean, and standard deviation (we will later refer to these as {statistical measures}), and entropy. We experimented with both 6- and 18-bin histograms. The unique orientations were measured based on the two histograms to capture the simplicity-complexity of the image.

Among all line segments, horizontal lines and vertical lines are known [3] to be static and to represent the feelings of calm and stability within the image. Horizontal lines suggest peace and calm, whereas vertical lines indicate strength. To capture the emotions evoked by these characteristics, we counted the number of horizontal lines and vertical lines through an 18-bin histogram. The orientation θ , of horizontal lines fall within $0^\circ < \theta < 10^\circ$ or $170^\circ < \theta < 180^\circ$, and $80^\circ < \theta < 100^\circ$ for vertical lines.

Length - The length of line segments reflects the simplicity of images. Images with simple structure might use long lines to fit contours, whereas complex contours have shorter lines. We characterized the length distribution by calculating the {statistical measures} of lengths of line segments within the image.

Mass of the image - The centroid of line segments may indicate associated relationships among line segments within the visual design [3]. Hence, we calculate the mean and standard deviation of the x and y coordinates of the line segments to find the mass of each image.

Some of the example images and their features are presented in Figures 8 and 9. Figure 8 presents the ten lowest mean values of the length of line segments. The first row shows the original images, the second row shows the

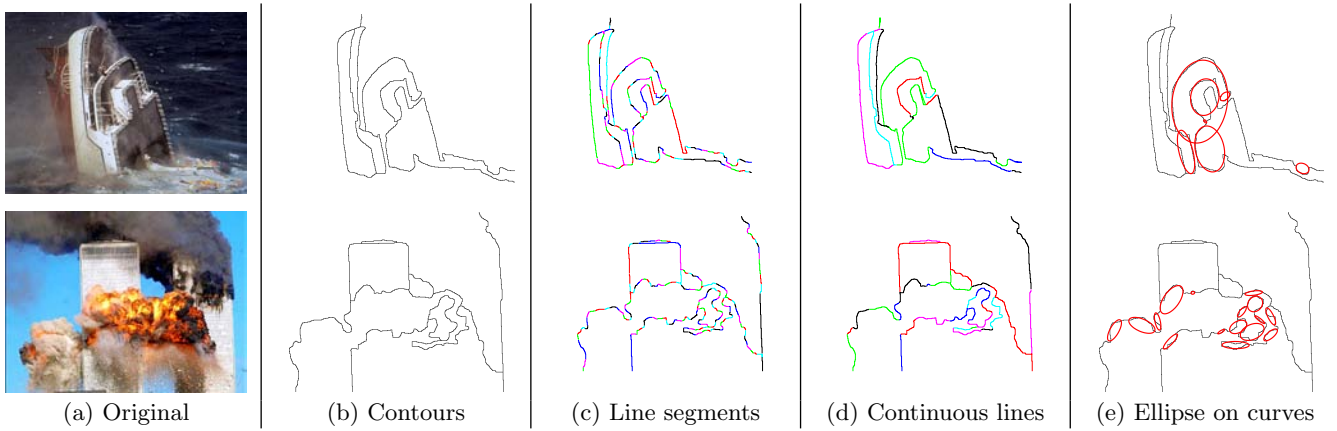


Figure 5: Perceptual shapes of images with *high arousal*.

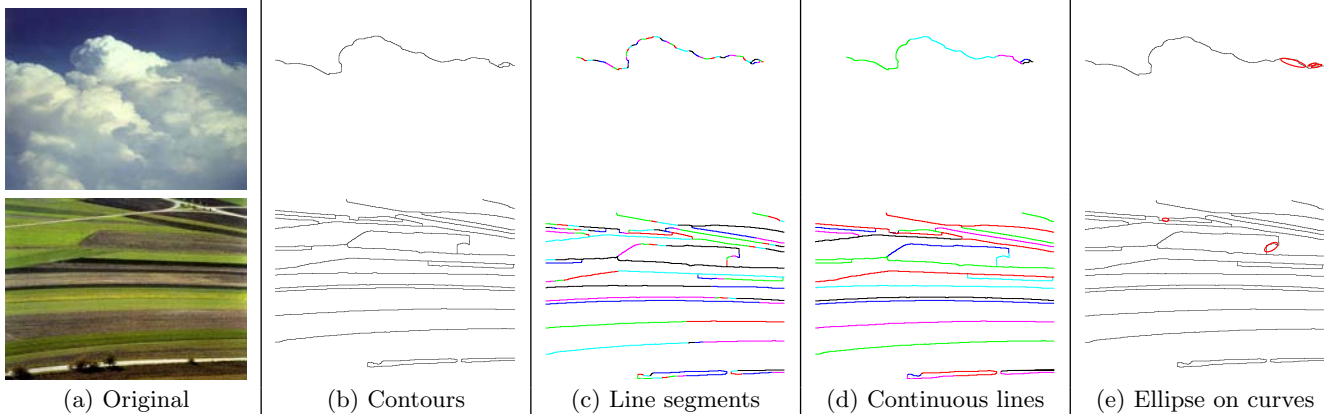


Figure 6: Perceptual shapes of images with *low arousal*.

line segments extracted from these images and the third row shows the 18-bin histogram for line segments in the images. The 18 bins refer to the number of line segments with an orientation of $[-90 + 10(i - 1), -90 + 10i)$ degrees where $i \in \{1, 2, \dots, 18\}$. Similarly, Figure 9 presents the ten highest mean values of the length of line segments.

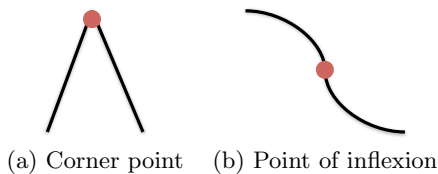


Figure 7: The corner point and point of inflexion.

These two figures indicate that the length or the orientation cannot be examined separately to determine the simplicity-complexity of the image. Lower mean values of the length of line segments might refer to either simple images such as the first four images in Figure 8 or highly complex images such as the last four images in that figure. The histogram of the orientation of line segments helps us to distinguish the complex images from simple images by examining variation of values in each bin.

4.2 Angles

Angles are important elements in analyzing the simplicity-complexity and the angularity of an image. We capture the visual characteristics from angles through two perspectives.

- **Angle count** - We first calculate the two quantitative

features claimed by Julian Hochberg, who has attempted to define simplicity (he used the value-laden term “figural goodness”) via information theory: “The smaller the amount of information needed to define a given organization as compared to the other alternatives, the more likely that the figure will be so perceived” [3]. Hence this minimal information structure is captured using the number of angles and the percentage of unique angles in the image.

- **Angular metrics** - We use the {statistical measures} to extract angular metrics. We also calculate the 6- and 18-bin histograms on angles and their entropies.

Some of the example images and features are presented in Figures 10 and 11. Images with lowest and highest number of angles are shown along with their corresponding contours in Figure 10. These examples show promising relationships between angular features and simplicity-complexity of the image. Example results for the histogram of angles in the image are presented in Figure 11. The 18 bins refer to the number of line segments with an orientation in $[10(i - 1), 10i)$ degrees where $i \in \{1, 2, \dots, 18\}$.

4.3 Continuous lines

We attempt to capture the degree of curvature from continuous lines, which has implications for the simplicity-complexity of images. We also calculated the number of continuous lines, which is the third quantitative feature specified by Julian Hochberg [3]. For continuous lines, open/closeness are factors affecting the simplicity-complexity of an image. In the following, we focus on the

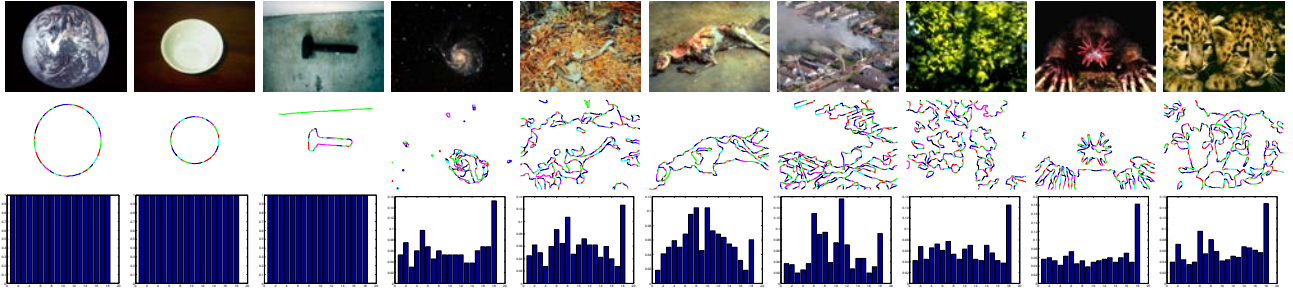


Figure 8: Images with *low* mean value of the length of line segments and their associated orientation histograms. The first row is the original images; the second row shows the line segments; and the third row shows the 18-bin histogram for line segments in the images.

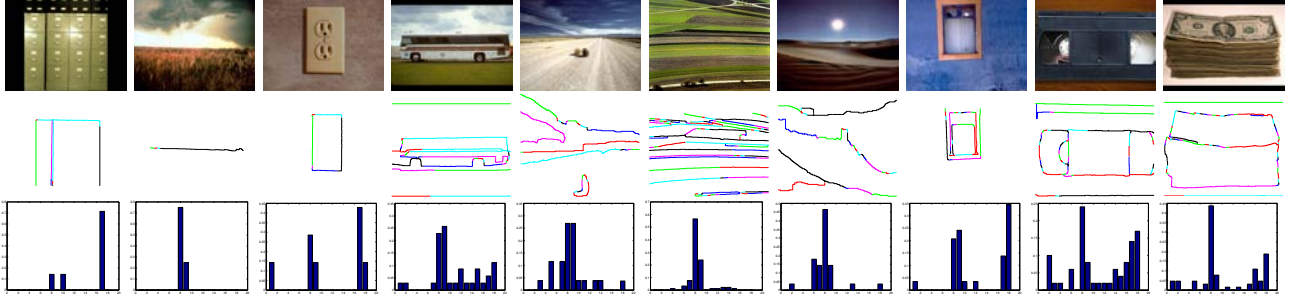


Figure 9: Images with *high* mean value of the length of line segments and their associated orientation histograms. The first row is the original images; the second row shows the line segments; and the third row shows the 18-bin histogram for line segments in the images.

calculation of the degree of curving, the length span value, and the number of open lines and closed lines. The length span refers to the highest Euclidean distance among all pairs of points on the continuous lines.

$$\text{Length Span}(l) = \max_{p_i \in l, p_j \in l} \text{EuclideanDist}(p_i, p_j), \quad (1)$$

where $\{p_1, p_2, \dots, p_N\}$ are the points on continuous line l .

- **Degree of curving** - We calculated the degree of curving of each line as

$$\text{Degree of Curving}(l) = \text{Length Span}(l)/N, \quad (2)$$

where N is the number of points on continuous line l .

To capture the statistical characteristics of contiguous lines in the image, we calculated the {statistical measures}. We also generated a 5-bin histogram on the degree of curving of all continuous lines (Figures 12 and 13).

- **Length span** - We used {statistical measures} for the length span of all continuous lines.
- **Line count** - We counted the total number of continuous lines, the total number of open lines, and the total number of closed lines in the image.

4.4 Curves

We used the nature of curves to model the roundness of images. For each curve, we calculated the extent of fit to an ellipse as well as the parameters of the ellipse such as its area, circularity, and mass of curves. The curve features are explained in detail below.

- **Fitness, area, circularity** - The fitness of an ellipse refers to the overlap between the proposed ellipse and the curves in the image. The area of the fitted ellipse

Table 2: Average number of curves in terms of the value of fitness in positive and negative images.

	(0.8, 1]	(0.6, 0.8]	(0.4, 0.6]	(0.2, 0.4]
Positive imgs	2.12	9.33	5.7	2.68
Negative imgs	1.42	7.5	5.02	2.73

Table 3: Average number of curves in terms of the value of circularity in positive and negative images.

	(0.8, 1]	(0.6, 0.8]	(0.4, 0.6]	(0.2, 0.4]
Positive imgs	0.96	2.56	5.1	11.2
Negative imgs	0.73	2.19	4	9.75

is also calculated. The circularity is represented by the ratio of the minor and major axes of the ellipses. The angular orientation of the ellipse is also measured. For each of the measures, we used the {statistical measures} and entropies of the histograms as the features to depict the roundness of the image.

- **Mass of curves** - We used the mean value and standard deviation of (x, y) coordinates to describe the mass of curves.
- **Top round curves** - To make full use of the discovered curves and to depict roundness, we included the fitness, area, circularity, and mass of curves for each of the top three curves.

To examine the relationship between curves and positive-negative images, we calculated the average number of curves in terms of values of circularity and fitness on positive images (i.e., the value is higher than 6 in the dimension of valance) and negative images (i.e., the value is lower than 4.5 in the dimension of valance).

The results are shown in Tables 2 and 3. Positive images

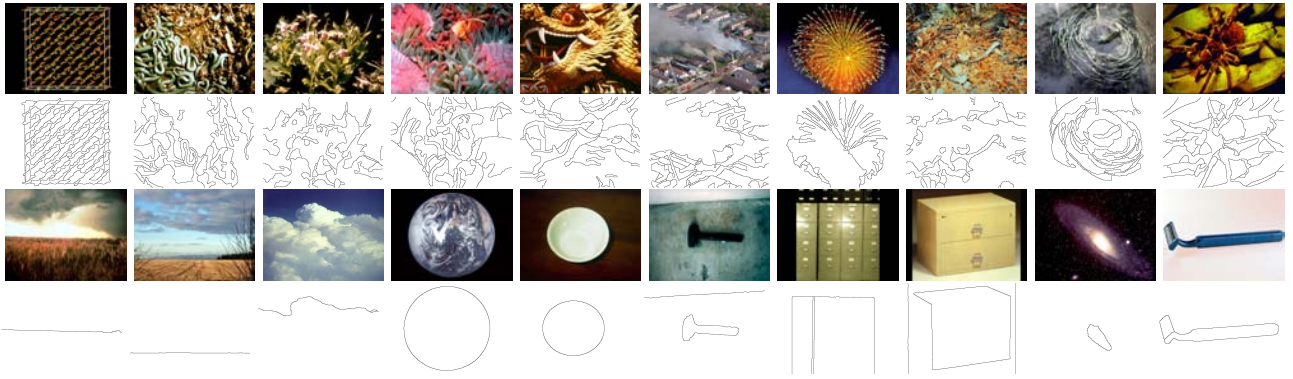


Figure 10: Images with highest and lowest number of angles.

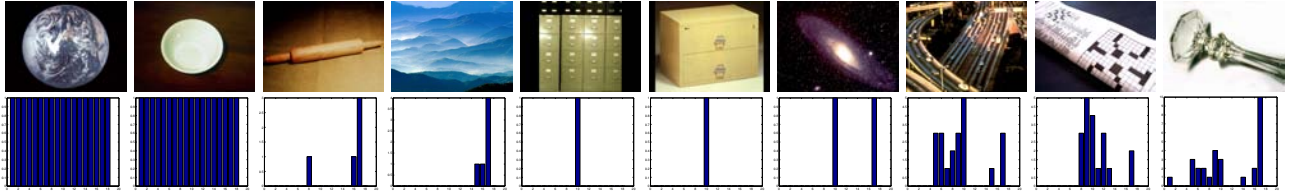


Figure 11: The distribution of angles in images.

have more curves with 60% – 100% fitness to ellipses and higher average curve count.

5. EXPERIMENTS

To demonstrate the relationship between proposed shape features and the felt emotions, the shape features were utilized in three tasks. First, we distinguished images with strong emotional content from emotionally neutral images. Second, we fit valence and arousal dimensions using regression methods. We then performed classification on discrete emotional categories. The proposed features were compared with the features discussed in Machajdik *et al.* [18], and overall accuracy was quantified by combining those features. Forward selection and Principal Component Analysis (PCA) strategies were employed for feature selection and to find the best combination of features.

5.1 Dataset

We used two subsets of the IAPS [15] dataset, which were developed by examining human affective responses to color photographs with varying degrees of emotional content. The IAPS dataset contains 1,182 images, wherein each image is associated with an empirically derived mean and standard deviation of valence, arousal, and dominance ratings.

Subset A of the IAPS dataset includes many images with faces and human bodies. Facial expressions and body language strongly affect emotions aroused by images, slight changes of which might lead to an opposite emotion. The proposed shape features are sensitive to faces hence we removed all images with faces and human bodies from the scope of this study. In experiments, we only considered the remaining 484 images, which we labeled as *Subset A*. To provide a better understanding of the ratings of the dataset, we analyzed the distribution of ratings within valence and arousal, as shown in Figure 14. We also calculated average variations of ratings in each rating unit (i.e., 1-2, 2-3, ..., 7-8). Valence ratings between 3 and 4, and 6 and 7, have

the maximum variance for single images. Similarly, arousal ratings between 4 and 5 varied the most.

Subset B are images with category labels (with discrete emotions), generated by Mikels [19]. *Subset B* includes eight categories namely, anger, disgust, fear, sadness, amusement, awe, contentment, and excitement, with 394 images in total. *Subset B* is a commonly used dataset, hence we used it to benchmark our classification accuracy with the results mentioned in Machajdik *et al.* [18].

5.2 Identifying Strong Emotional Content

Images with strong emotional content have very high or very low valence and arousal ratings. Images with values around the mean values of valence and arousal lack emotions and were used as samples for emotionally neutral images.

Based on dimensions of valence and arousal respectively, we generated two sample sets from *Subset A*. In Set 1, images with valence values higher than 6 or lower than 3.5 were considered images with strong emotional content and the rest to represent emotionally neutral images. This resulted in 247 emotional images and 237 neutral images. Similarly, images with arousal values higher than 5.5 or lower than 3.7 were defined as emotional images, and others as neutral images. With similar thresholds, we obtained 239 emotional images and 245 neutral images in Set 2.

We used the traditional Support Vector Machines (SVM) with radial basis function (RBF) kernel to perform the classification task. We trained SVM models using the proposed shape features, Machajdik’s features, and combined (Machajdik’s and shape) features. Training and testing were performed by dividing the dataset uniformly into training and testing sets. As we removed all images with faces and human bodies, we did not consider facial and skin features discussed in [18]. We used both forward selection and PCA methods to perform feature selection. In the forward selection method, we used the greedy strategy and accumulated one feature at a time to obtain the subset of features that maximized the classification accuracy. The seed features were also chosen at random over



Figure 12: Images with highest degree of curving.



Figure 13: Images with lowest degree of curving.

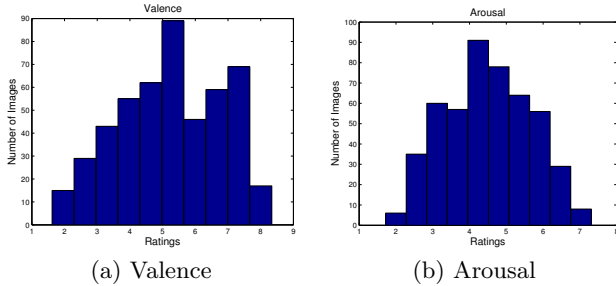


Figure 14: Distribution of ratings in IAPS.

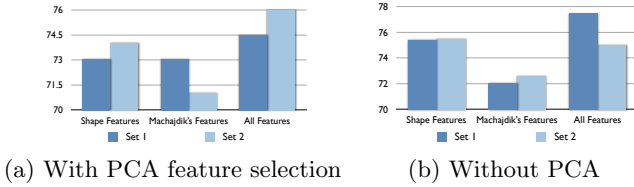


Figure 15: Classification accuracy (%) for emotional images and neutral images (Set 1 and Set 2 are defined in Section 5.2).

multiple iterations to obtain better results. Our analyses showed that the forward selection strategy achieved greater accuracy for Set 2, whereas PCA performed better for Set 1 (Figure 15). The feature comparison showed that the combined (Machajdik's and shape) features achieved the highest classification accuracy, whereas individually the shape features alone were much stronger than the features from [18] (Machajdik's features). This result is intuitive since emotions evoked by images cannot be well represented by shapes alone and can definitely be bolstered by other image features including their color composition and texture. By analyzing valence and arousal ratings of the correctly classified images, we observed that very complex/simple, round and angular images had strong emotional content and high valence values. Simple structured images with very low degrees of curving also tends to portray strong emotional content as well as to have high arousal values.

By analyzing the individual features for classification accuracy we found that line count, fitness, length span, degree of curving, and the number of horizontal lines achieved the best classification accuracy in Set 1. Fitness and line orientation were more dominant in Set 2.

We present a few example images, which were wrongly classified based on the proposed shape features in Figures 16 and 17. The misclassification can be explained as a shortcoming of the shape features in understanding the semantics. Some of the images generated extreme emotions based on image content irrespective of the low-level features. Besides the semantics, our performance was also limited by the performance of the contour extraction algorithm.

5.3 Fitting the Dimensionality of Emotion

Emotions can be represented by word pairs, as previously done in [23]. However, some emotions are difficult to label. Modeling basic emotional dimensions helps in alleviating this problem. We represented emotion as a tuple consisting of valence and arousal values. The values of valence and arousal were in the range of (1, 9). In order to predict the values of valence and arousal we proposed to learn a regression model for either dimension separately.

We used SVM regression with RBF kernel to model the valence and arousal values using shape, Machajdik's features, as well as the combination of features. The mean squared error (MSE) was computed for each of the individual features as well as combined for both valence and arousal values separately. The MSE values are shown in Figure 18(a). These figures show that the valence values were modeled more accurately by Machajdik's features than our shape features. Arousal was well modeled by shape features with a mean squared error of 0.9. However, the combined feature performance did not show any improvements. The results indicated that visual shapes provide a stronger cue in understanding the valence as opposed to the combination of color, texture, and composition in images.

We also computed the correlation between quantified individual shape features and valence-arousal ratings. The higher the correlation, the more relevant the features were. Through this process we found that angular count, fitness, circularity, and orientation of line segments showed higher correlations with valence, whereas angle count, angle metrics, straightness, length span, and orientation of curves had higher correlations with arousal.

5.4 Classifying Categorized Emotions

To evaluate the relationship between shape features and emotions on discrete emotions, we classified images into one of the eight categories, anger, disgust, fear, sadness,

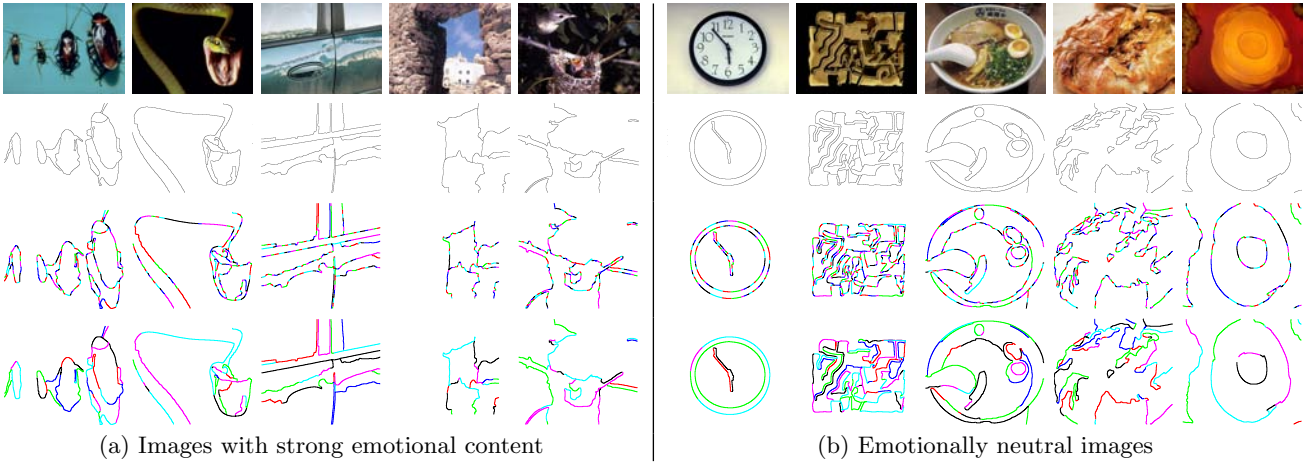


Figure 16: Examples of misclassification in Set 1. The four rows are original images, image contours, line segments, and continuous lines.

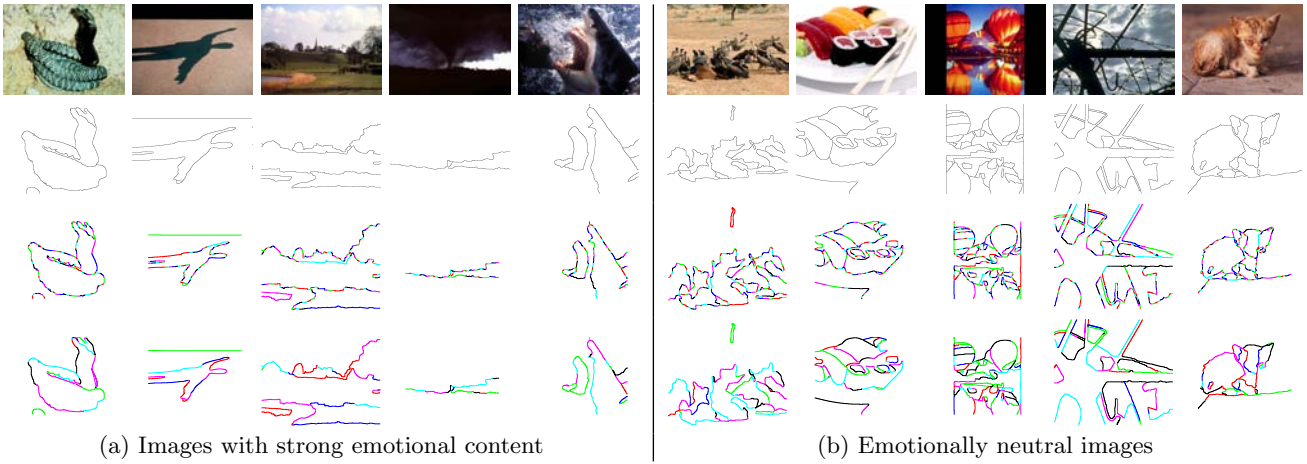


Figure 17: Examples of misclassification in Set 2. The four rows are original images, image contours, line segments, and continuous lines.

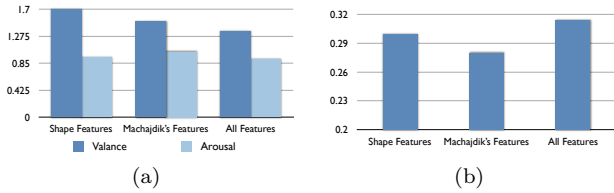


Figure 18: Experimental results. (a) Mean squared error for the dimensions of valence and arousal. (b) Accuracy for the classification task.

amusement, awe, contentment, and excitement. We followed Machajdik *et al.* [18] and performed one-vs-all classification to compare and benchmark our classification accuracy. The classification results are reported in Figure 18(b). We used SVM to assign the images to one of the eight classes. The highest accuracy was obtained by combining Machajdik's with shape features. We also observed a considerable increase in the classification accuracy by using the shape features alone, which proves that shape features indeed capture emotions in images more effectively.

In this experiment, we also built classifiers for each of the shape features. Each of the shape features listed in Table 4 achieved a classification accuracy of 30% or higher.

Table 4: Significant features to emotions.

Emotion	Features
Angry	Circularity
Disgust	Length of line segments
Fear	Orientation of line segments and angle count
Sadness	Fitness, mass of curves, circularity, and orientation of line segments
Amusement	Mass of curves and orientation of line segments
Awe	Orientation of line segments
Excitement	Orientation of line segments
Contentment	Mass of lines, angle count, and orientation of line segments

6. CONCLUSIONS

We investigated the computability of emotion through shape modeling. To achieve this goal, we first extracted contours from complex images, and then represented contours using lines and curves extracted from images. Statistical analyses were conducted on locally meaningful lines and curves to represent the concept of roundness, angularity, and simplicity, which have been postulated as playing a key role in evoked emotion for years. Leveraging the computational representation of these physical stimulus

properties, we evaluated the proposed shape features through three tasks: distinguishing emotional images from neutral images; classifying images according to categorized emotions; and fitting the dimensionality of emotion based on proposed shape features. We have achieved an improvement over the state-of-the-art solution [18]. We also attacked the problem of modeling the presence or absence of strong emotional content in images, which has long been overlooked. Separating images with strong emotional content from emotionally neutral ones can aid in many applications including improving the performance of keyword based image retrieval systems. We empirically verified that our proposed shape features indeed captured emotions in the images. The area of understanding emotions in images is still in its infancy and modeling emotions using low-level features is the first step toward solving this problem. We believe our contribution takes us closer to understanding emotions in images. In the future, we hope to expand our experimental dataset and provide stronger evidence of established relationships between shape features and emotions.

7. REFERENCES

- [1] P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik. Contour detection and hierarchical image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 33(5):898–916, 2011.
- [2] S. Arifin and P. Y. K. Cheung. A computation method for video segmentation utilizing the pleasure-arousal-dominance emotional information. In *ACM MM*, pages 68–77, 2007.
- [3] R. Arnheim. *Art and visual perception: A psychology of the creative eye*. 1974.
- [4] J. Aronoff. How we recognize angry and happy emotion in people, places, and things. *Cross-Cultural Research*, 40(1):83–105, 2006.
- [5] M. Bar and M. Neta. Humans prefer curved visual objects. *Psychological Science*, 17(8):645–648, 2006.
- [6] M. M. Bradley and P. J. Lang. The international affective picture system (IAPS) in the study of emotion and attention. In *Handbook of Emotion Elicitation and Assessment*, pages 29–46, 2007.
- [7] S. Brandt, J. Laaksonen, and E. Oja. Statistical shape features in content-based image retrieval. In *ICPR*, pages 1062–1065, 2000.
- [8] A. Chia, D. Rajan, M. Leung, and S. Rahardja. Object recognition by discriminative combinations of line segments, ellipses and appearance features. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 34(9):1758–1772, 2011.
- [9] G. Csurka, S. Skaff, L. Marchesotti, and C. Saunders. Building look & feel concept models from color combinations. *The Visual Computer*, 27(12):1039–1053, 2011.
- [10] R. Datta, D. Joshi, J. Li, and J. Z. Wang. Studying aesthetics in photographic images using a computational approach. In *ECCV*, pages 288–301, 2006.
- [11] R. Datta, J. Li, and J. Z. Wang. Algorithmic inferencing of aesthetics and emotion in natural image: An exposition. In *ICIP*, pages 105–108, 2008.
- [12] A. Hanjalic and L. Q. Xu. Affective video content representation and modeling. *IEEE Trans. on Multimedia*, 7(1):143–154, 2005.
- [13] D. Joshi, R. Datta, E. Fedorovskaya, Q. T. Luong, J. Z. Wang, J. Li, and J. Luo. Aesthetics and emotions in images. *IEEE Signal Processing Magazine*, 28(5):94–115, 2011.
- [14] P. J. Lang, M. M. Bradley, and B. N. Cuthbert. Emotion, motivation, and anxiety: Brain mechanisms and psychophysiology. *Biological Psychiatry*, 44(12):1248–1263, 1998.
- [15] P. J. Lang, M. M. Bradley, and B. N. Cuthbert. International affective picture system: Affective ratings of pictures and instruction manual. In *Technical Report A-8, University of Florida, Gainesville, FL*, 2008.
- [16] M. K. Leung and Y.-H. Yang. Dynamic two-strip algorithm in curve fitting. *Pattern Recognition*, 23(1-2):69–79, 1990.
- [17] K. A. Lindquist, T. D. Wager, H. Kober, E. Bliss-Moreau, and L. F. Barrett. The brain basis of emotion: A meta-analytic review. *Behavioral and Brain Sciences*, 173(4):1–86, 2011.
- [18] J. Machajdik and A. Hanbury. Affective image classification using features inspired by psychology and art theory. In *ACM MM*, pages 83–92, 2010.
- [19] J. Mikel, B. L. Fredrickson, G. R. Larkin, C. M. Lindberg, S. J. Maglio, and P. A. Reuter-Lorenz. Emotional category data on images from the international affective picture system. *Behavior Research Methods*, 37(4):626–630, 2005.
- [20] Y. Mingqiang, K. Kidiyo, and R. Joseph. A survey of shape feature extraction techniques. *Pattern Recognition*, pages 43–90, 2008.
- [21] R. Reber, N. Schwarz, and P. Winkielman. Processing fluency and aesthetic pleasure: Is beauty in the perceiver’s processing experience? *Personality and Social Psychology Review*, 8(4):364–382, 2004.
- [22] H. R. Schiffman. *Sense and Perception: An Integrated Approach*. 1990.
- [23] T. Shibata and T. Kato. Kansei image retrieval system for street landscape-discrimination and graphical parameters based on correlation of two image systems. In *International Conference on Systems, Man, and Cybernetics*, pages 274–252, 2006.
- [24] M. Solli and R. Lenz. Color based bags-of-emotions. *LNCS*, 5702:573–580, 2009.
- [25] H. L. Wang and L. F. Cheong. Affective understanding in film. *IEEE Trans. on Circuits and Systems for Video Technology*, 16(6):689–704, 2006.
- [26] V. Yanulevskaya, J. C. Van Gemert, K. Roth, A. K. Herbold, N. Sebe, and J. M. Geusebroek. Emotional valence categorization using holistic image features. In *ICIP*, pages 101–104, 2008.
- [27] H. Zhang, E. Augilius, T. Honkela, J. Laaksonen, H. Gamper, and H. Alene. Analyzing emotional semantics of abstract art using low-level image features. In *Advances in Intelligent Data Analysis*, pages 413–423, 2011.
- [28] S. L. Zhang, Q. Tian, Q. M. Huang, W. Gao, and S. P. Li. Utilizing affective analysis for efficient movie browsing. In *ICIP*, pages 1853–1856, 2009.