Who's Afraid of Itten: Using the Art Theory of Color Combination to Analyze Emotions in Abstract Paintings

Andreza Sartori[1,2], Dubravko Culibrk[1,4], Yan Yan[1,3], Nicu Sebe[1]
[1]DISI, University of Trento, Italy;
[2]Telecom Italia - SKIL, Trento, Italy
[3]ADSC, Singapore
[4]FTS, University of Novi Sad, Serbia
{andreza.sartori, dubravko.culibrk, yan, sebe}@disi.unitn.it

ABSTRACT

Color plays an essential role in everyday life and is one of the most important visual cues in human perception. In abstract art, color is one of the essential means to convey the artist's intention and to affect the viewer emotionally. However, colors are rarely experienced in isolation, rather, they are usually presented together with other colors. In fact, the expressive properties of two-color combinations have been extensively studied by artists. It is intriguing to try to understand how color combinations in abstract paintings might affect the viewer emotionally, and to investigate if a computer algorithm can learn this mechanism.

In this work, we propose a novel computational approach able to analyze the color combinations in abstract paintings and use this information to infer whether a painting will evoke positive or negative emotions in an observer. We exploit art theory concepts to design our features and the learning algorithm. To make use of the color-group information, we propose inferring the emotions elicited by paintings based on the sparse group lasso approach. Our results show that a relative improvement of between 6% and 8% can be achieved in this way. Finally, as an application, we employ our method to generate Mondrian-like paintings and do a prospective user study to evaluate the ability of our method as an automatic tool for generating abstract paintings able to elicit positive and negative emotional responses in people.

Categories and Subject Descriptors

J.5 [**Computer Applications**]: Arts and Humanities - fine arts; I.2 [**Artificial Intelligence**]: Vision and Scene Understanding; H.3.1 [**Information Search and Retrieval**]: Content Analysis and Indexing

General Terms

Algorithms, Experimentation, Theory

MM'15, October 26–30, 2015, Brisbane, Australia.
© 2015 ACM. ISBN 978-1-4503-3459-4/15/10 ...$15.00.
DOI: http://dx.doi.org/10.1145/2733373.2806250.

Keywords

Visual Art, Abstract Paintings, Emotion Recognition, Color Combination

1. INTRODUCTION

Color has a strong influence on human perception and is commonly used to convey aesthetic information. It can affect people on an emotional level and determine the preferences, interpretation, memorization and reactions of people. Different amounts of one or more colors convey different messages – a mechanism which is closely connected with the mental and emotional states [6]. The relationship between emotion and color has been widely studied in many fields, such as psychology, art, design, marketing and other areas.

Artists such as Kandinsky, Itten, Albers and others, studied how the presence of colors, in different proportions and combinations, affects the viewer emotionally, and applied these studies to their artworks: "Discovery of relationships, mediated by the eye and brain, between color agents and color effects in man, is a major concern of the artist." [5].

Johannes Itten, a notable expressionist painter, theorist and professor of the Bauhaus School[1] studied and taught the aesthetic and expressive aspects of colors and their interpretation. In his studies on the relationship between the chromatic scale and the spiritual evolution, he provided a set of rules of colors and color combinations he called the "objective principles of colors" [5]. Itten stated that "The concept of color harmony should be removed from the realm of subjective attitude into that of objective principle." The fact that many paintings induce similar emotions in the viewer, suggests that there may be some consistency in the way people perceive colors in abstract paintings.

In this paper, we use state-of-the-art computer vision techniques to understand the contribution of colors and two-color combinations in evoking positive or negative emotions in viewers of abstract paintings. We focus on abstract paintings because they present low semantic content (e.g., color, shape and texture), thus color, in these paintings, is one of the most important elements used to convey emotion. Our study into the effect of color combinations is based on Itten's theory of color expression, which states that it is difficult to define the expressive properties of a color without relating it to other colors [5]. Guided by this principle, in his theory of

[1]The Bauhaus School was a German art school which has a "...strong influence of scientific studies of color on the development of abstraction..." [1]

color expression, Itten observed and defined the emotional aspects of two-color combinations, which we explored computationally in the present study.

The main contributions of this work are: (1) we study the role of color combinations on positive and negative emotions induced by abstract paintings; (2) we propose a novel approach which recognizes color combinations in abstract paintings and classifies them in positive and negative emotions, and tested our models in two different datasets of abstract artworks; (3) we are the first to infer color-group information based on sparse group lasso, increasing the relative classification accuracy by 6% to 8% over the approaches that do not use the group structure; (4) we apply our color visual vocabulary (Color Palette) to automatically generate Mondrian-like paintings and provide an empirical study which shows that our method can be used as a tool to generate paintings which evoke positive and negative emotions in people.

The rest of the paper is structured as follows. Section 2 presents the related work covering the color theory and the existing emotion recognition approaches using computer vision techniques. In Section 3 we describe the two datasets of Professional and Amateur abstract paintings, and describe the Relative Scaling method employed to annotate artworks as emotionally positive or emotionally negative. Section 4 presents the details of our proposed approach and Section 5 provides the respective evaluations and results. In Section 6 we describe the learning approach proposed which infers the emotions elicited by paintings based on the sparse group lasso. In Section 7 we present an empirical study to prove that our proposed approach is able to predict emotional response of people. Our last section is devoted to the final discussion and conclusions.

2. RELATED WORK

This section presents previous studies on color and art theory as well as the existing emotion recognition approaches using computer vision techniques and relate these approaches with our work.

2.1 Color Theory

Color is one of the most salient elements of an image and it is the subject of study in many areas, including physics, psychology, philosophy, neuroscience, art, computing and many others. These and other fields attempt to understand the nature of color, how it is constituted, how people perceive it, and how it influences people's life. Initially, Newton [13] observed the physical aspects behind the perception of color and discovered that if some wavelengths of light are reflected better than others, then the object appears colored: "When light falls on an object, some is absorbed. The light that isn't absorbed is reflected off the object's surface; this is the light we see." [10]. Opposed to Newton, Goethe proposed a color wheel based on physiological experiences (e.g., afterimage[2]) and psychological aspects (e.g., moral associations, symbolic and mystic use of colors) to describe complementary colors: "The chromatic circle... [is] arranged in a general way according to the natural order... for the colors diametrically opposed to each other in this diagram are those which reciprocally evoke each other in the eye. Thus, yellow demands purple; orange, blue; red, green; and vice versa: thus... all intermediate gradations reciprocally evoke each other; the simpler color demanding the compound, and vice versa." [21]. These and subsequent physiological and psychological theories of the effects of color, were used by famous painters and are still used today in order to improve the aesthetic aspects of paintings.

[2]Afterimage refers to an image continuing to appear in one's vision after the exposure to the actual image has ended.

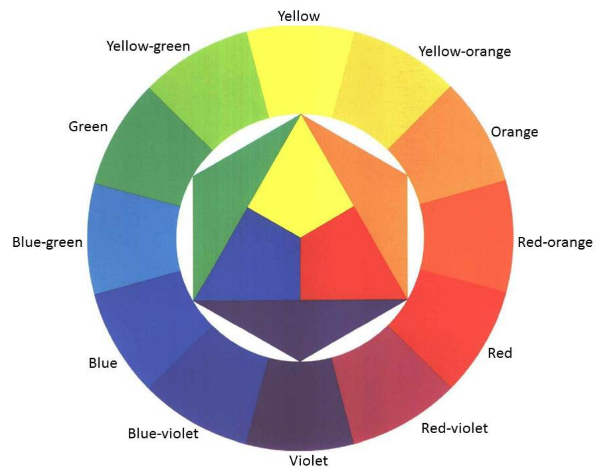

Figure 1: Itten's Color Wheel (taken from [5])

In abstract paintings, colors have an essential role. They are particularly studied and used by artists to convey the emotional message to the beholder. Kandinsky describes the emotional effects of individual colors, stating that it depends on two factors: a physical impression, "one of pleasure and contentment at the varied and beautiful colors" and a psychic effect, in which colors "produce a corresponding spiritual vibration, and it is only as a step towards this spiritual vibration that the elementary physical impression is of importance" [7].

More recently, Itten observed that it is difficult to define the expressive properties of a color without relating it to other colors [5]. Indeed, he formalized theories on the emotional effects of two-color combinations and their properties to generate harmonious artworks [5]. He describes the emotional values of the color combination hues defined on the 12-color circle (Figure 1). Itten's color wheel is widely used in art studies and is composed of 12 hues of primaries, secondary and tertiary colors which are respectively: yellow, red, blue, orange, violet, green, yellow-orange, red-orange, red-violet, blue-violet, blue-green and yellow-green. For Itten, the two diametrically opposed colors in the wheel are complementary, forming a harmonious dyad, for instance red/green, blue/orange, yellow/ violet, etc. These 12 hues are varied by five levels of luminance and three levels of saturation and are represented by Itten in a Color Sphere with a total of 180 colors. In the Color Sphere, the 12 hues are located in the equatorial zone, the luminance varies along the meridians, and the horizontal sections contain the degrees of saturation: "If I use the color sphere, I can get an indefinite number of harmonious dyads. The only requirement is that the two tones be symmetrical with respect to the center of the sphere. Thus if I take a tint of red, the corresponding green must be shaded in the same degree as the red is lightened" [5].

Using Itten's theory of color expression described in [5], we investigate how color combinations can be exploited to improve machine learning approaches to classify the emotions elicited by abstract paintings. In addition we discuss how the insight gained can, potentially, be used to revise and enhance the original theory.

2.2 Emotion Recognition

Recently there is a growing interest from different research communities in understanding the emotional response of the viewer while viewing colors in images, artworks, etc. A psychological study on the effects of colors on emotions based on Pleasure, Arousal and Dominance model concludes that brighter colors are more pleasant, less arousing, and induce less dominance than the darker colors [19]. Ou *et al.* investigate how eleven emotional scales are associated with three color-emotion factors (i.e., color activity, color weight and color heat) of single colors [14] and two-color combinations [15] determined by means of the factor analysis method. These studies show that there is consistency in the way people perceive colors.

From a computational perspective, color, along with shape and texture has been extensively used as a feature in the research focusing on artworks. Yanulevskaya et al. [23] proposed an emotion categorization system for masterpieces based on the assessment of local image statistics, applying supervised learning of emotion categories using Support Vector Machines. Machajdik and Hanbury [11] employed low-level features and combined them with concepts from psychology and art theory to categorize images and artworks emotionally. They obtained better accuracy in affective categorization of semantically rich images than abstract paintings. In [24] a Bag-of-Visual-Words model was trained to classify abstract paintings into those eliciting positive or negative emotions. With the backprojection technique, the authors determined which parts of the paintings evoke positive or negative emotion and concluded that, for instance, light colors and smooth shapes generate a positive emotion, whereas dark colors and sharp edges generate a negative emotion. Recently, sparse learning models have been used in different multimedia problems [22, 16, 9]. Sartori *et al.* [16] proposed to use both visual and text information in a joint learning model for abstract painting emotion recognition. Liu *et al.* [9] proposed a multi-task learning approach for painting style analysis. Zhao *et al.* [25] extracted emotion features of images based on the principles of art, including balance, emphasis, harmony, variety, gradation, and movement, which were shown to be effective for emotion classification of images. In our work, we analyse the emotional effects of color combinations present in abstract paintings.

3. DATASETS AND GROUND TRUTH COLLECTION

We conduct our analysis on two datasets of abstract paintings, which differ from each other in terms of the artists' background: one contains works by professional and the other by amateur artists[3]. Each dataset is composed of 500 abstract paintings. In the following subsections we detail the painting selection process, as well as the nature of each dataset.

[3]The datasets with their respective ground truths are publicly available at: http://disi.unitn.it/∼sartori/datasets/

3.1 MART Dataset: A Dataset of Professional Abstract Paintings

The MART dataset was collected in our previous work [24, 17] from the electronic archive of the Museum of Modern and Contemporary Art of Trento and Rovereto (MART). The selected paintings have been created by a total of 78 artists between 1913 and 2008. Most of these artists are Italians, however there are also European and American artists. Some of the artists present in this set, such as Wassily Kandinsky, Josef Albers, Paul Klee and Luigi Veronesi, are known for their studies on abstract art in terms of color, shapes and texture.

3.2 deviantArt Dataset: A Dataset of Amateur Abstract Paintings

The amateur collection of abstract paintings has been collected in our previous work [17] from the deviantArt (dA) website[4], an online social network devoted to user-generated art. DeviantArt is one of the largest online art communities with more than 280 million artworks and 30 million registered users. We selected the artworks that were under the category Traditional Art/Paintings/Abstract, and downloaded 8,000 artworks. We used the information reflecting how many times an artwork was favored as a parameter to downsize the collection from 8000 to 500 paintings[5]. We selected the paintings with the highest and least scores in terms of favourites, as well as added a random selection of paintings scoring in the middle range. Thus the collection has 500 paintings by 406 different authors.

3.3 User Study to Analyse Emotions Evoked by Artworks

To collect our ground truth for elicited emotions we used a relative score method from our previous work [17], in which we asked people to choose the more positive painting in a pair. The annotation was done online and we provided the following instructions to each annotator: "Which painting in the pair looks more positive to you? Let your instinct guide you and follow your first impression of the paintings."

We follow the method of TrueSkill ranking system described in [4, 12] to annotate and calculate the emotional scores from the paintings. The TrueSkill ranking system, developed by Microsoft Research for Xbox Live, recognizes and ranks the skills of the players in a game and matches players with similar skills for a new game. With this method, the annotation task is more manageable, as it gives a representative annotation with only 3,750 pairs of paintings, instead of 124,750 comparisons $(500 * (500 - 1) * 0.5)$, in case each painting is compared with all the remaining paintings in the dataset.

During the annotation process, we consider that all paintings initially have the same 'skill' and the painting which is chosen as more positive in a single trial wins a 'game'. Then, the rankings of the compared paintings are updated. Afterwards, the paintings with similar rankings are compared, until each painting is compared with at least 15 other paintings. We consider the results as emotional scores of the paintings, in which lower values correspond to negative feelings and the higher values to positive feelings.

[4]http://www.deviantart.com/

[5]We selected only 500 paintings to construct the deviantArt dataset in order to make an impartial comparison with the MART dataset.

In total, 25 subjects (11 females, 14 males) participated in the annotation of the MART dataset. Each person annotated from 19 to 356 pairs of paintings, 145 paintings on average. For the deviantArt dataset 60 people participated in the annotation process, 27 females and 33 males, respectively. Each participant annotated from 2 to 436 pairs of paintings, 63 paintings on average. The participants were anonymous and they participated in this annotation voluntarily, without getting any reward. There was no time-limit to the annotation procedure: each participant was free to annotate whenever he/she had time to do it.

Figure 2 shows a plot of the TrueSkill scores distribution for the MART and DeviantArt datasets. The mean score (μ) obtained for the MART dataset is slightly higher, while the standard deviation (σ) is slightly lower. The two distributions naturally differ, but not extensively so. The third distribution shown in the figure corresponds to synthetic images discussed in Section 7.

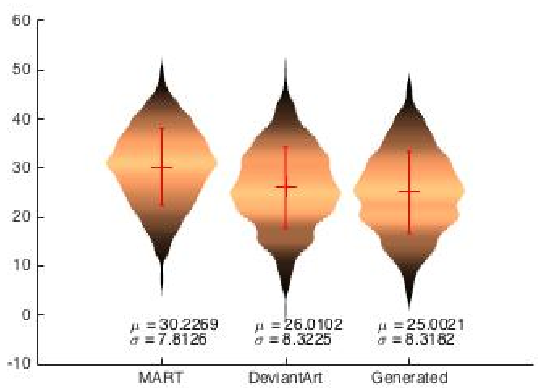

Figure 2: Distribution of TrueSkill scores for different datasets.

To compose the ground truth of both datasets, a threshold was defined to separate the paintings into those eliciting positive and and those eliciting negative emotions. As each painting was compared 15 times, we assume that the threshold should be related to one of the paintings which was chosen 8 times as more positive. Therefore, we consider the TrueSkill ranking values of these paintings, which are the ratings resulted from the comparisons of paintings by using the TrueSkill algorithm. The paintings with TrueSkill ranking equal or lower than the threshold are defined as negative and paintings with ranking higher than the threshold are defined as positive.

As a result, for the MART dataset, the painting with TrueSkill ranking value equal to 25.1 is used as threshold. We obtain a total of 131 paintings in the negative class and 369 in the positive class. To compose the ground truth for deviantArt, we considered as threshold the painting with TrueSkill ranking equal to 21.0. In total, 140 paintings were assigned to negative class and the 360 to the positive class.

4. PROPOSED APPROACH

In this section we provide details on how we use color combinations to automatically classify abstract paintings in positive and negative emotions. Specifically, we analyse the contribution of adjacent colors in a painting and use segmentation to generate the color combinations features to emotionally classify abstract paintings.

An overview of the proposed approach is shown in Figure 3. We first collect the database and the corresponding ground truth of positive and negative emotions evoked by abstract paintings. Afterwards, we use the method of van de Weijer *et al.* [20] to map the paintings into Color Name features and then cluster the entire dataset into 180 clusters (i.e., the largest number of colors in Itten's model) which were used as our *Color Palette*. Then, we calculate the closest value between the Color Palette and each pixel of the painting, and generate a new painting to segment. The segmentation yields blobs of uniform color that are further used to extract our features for image classification.

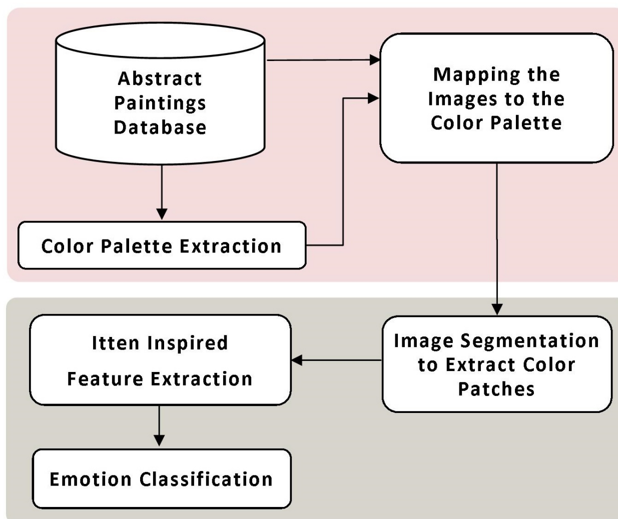

Figure 3: Overview of the proposed approach

4.1 Color Features

In this work, we propose a novel framework to analyze two-color combinations in abstract paintings and use this information to classify the images into those eliciting positive and negative emotions in the viewers. We use the Color Naming method of van de Weijer *et al.* [20] to create a visual vocabulary (Color Palette) able to model the human perception of colors.

van de Weijer *et al.* [20] describe colors through the confidence that a color can be described by 11 linguistic labels. These 11 linguistic labels are considered the ones which humans use to express the colors in the world (i.e., 'blue', 'red', 'yellow', etc.). The Color Name method shows to be more effective compared with photometric invariant descriptions, mainly with the description of achromatic colors such as white, black and grey which is harder to distinguish from a photometric invariance perspective [8, 18].

Algorithm 1 is used to extract the Color Palette. We use the method of [20] to extract initial color features, that should correspond to the way the colors are perceived by the viewers. Each painting in the RGB-color space was mapped to a corresponding image in the Color Names Feature (CNF) space. We used the default parameters of [20], calculating the mean of the patches with 21 pixels each, as the final value of the pixel in the CNF space.

Algorithm 1: Color palette extraction

Data: images
Result: color_palette
color_set ← {}
for $I \in images$ **do**
  $cni \leftarrow color\_name\_features(I)$; **for** $color \in cni$ **do**
    $color\_set \leftarrow color\_set \cup color$;
  end
end
$color\_palette \leftarrow kmeans(color\_set, 180)$;

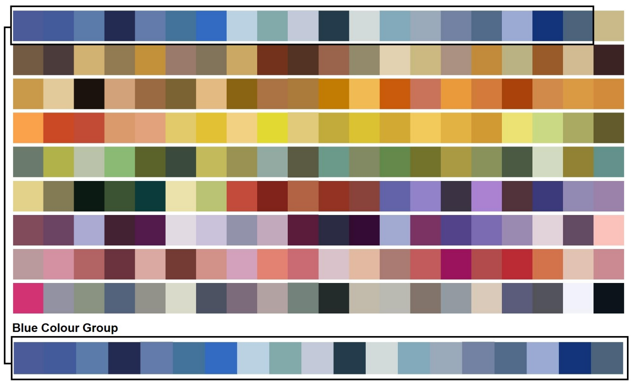

Figure 4: Color Palette extracted from MART dataset with 180 colors. Example of blue color group.

Once all the images have been mapped to the CNF space, we used the k-means algorithm to cluster the CNF values of the entire dataset of abstract paintings into 180 distinct CNF values (those closest to the cluster centroids) to form our CNF Color Palette (i.e., k is 180). The size of our palette matches the largest number of colors in Itten's color model [5].

Figures 4 and 5 display (unsorted) RGB versions of our Color Palettes with 180 colors created for MART and deviantArt datasets, which were subsequently used to represent all the colors in the images. The colors were manually arranged into color groups of blue, brown, orange, yellow, green, red, violet, pink, grey, white and black colors. Such grouping allows the proposed learning algorithm to use the group information to enhance classification performance.

Once our palette (i.e. color vocabulary) has been created for a dataset, the Euclidean distance is then used to to replace all the CNF values in the abstract paintings with the closest value in the palette. Finally the CNF values are mapped back to RGB to enable the use of a standard image segmentation algorithm to perform the segmentation into patches of same color.

Figure 6 shows a sample painting from the MART dataset and the corresponding RGB image after the mapping to our Color Palette.

4.2 Segmentation

To segment the paintings we used the graph-based image segmentation method of Felzenszwalb and Huttenlocher [3]. Their method effectively measures the evidence for a boundary between two regions. We used the publicly avail-

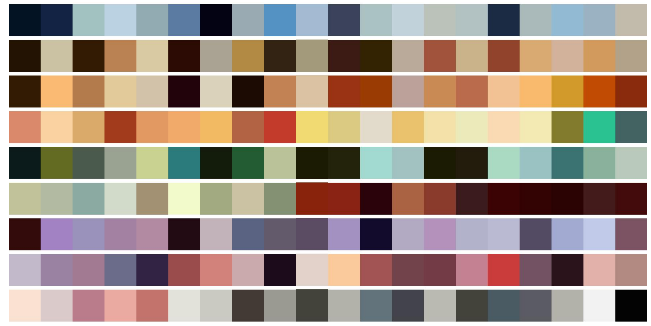

Figure 5: Color Palette extracted from deviantArt dataset with 180 colors.

Figure 6: MART paintings repainted with the colors in the derived Color Palette. The images on the left are the original paintings, and the images on the right RGB representations of the repainted paintings. (© *MART - Archivio fotografico e Mediateca*)

able implementation of the algorithm with the parameters $\sigma = 0.8$, the standard deviation of the Gaussian filter to smooth the image, and $k = 100$, which is used to control the threshold function. Figure 7 shows an example of the segmentation result.

4.3 Describing Color Combinations

In this section we detail how we use our computational approach to describe the color combinations in abstract paintings. Based on the art theory of Itten, we consider three aspects: the two-color combinations, the amount of colors and the position of colors. In the design of our features, we were guided by the principles identified by Itten, detailed in this section.

4.3.1 *The Color Pairs are Equally Important*

We consider that all color pairs are independent from each other and all color pairs are equally important. Indeed, Itten [5] affirms that "[...] the effect of a color is determined by its situation relative to accompanying colors. A color is always to be seen in relation to its surroundings."

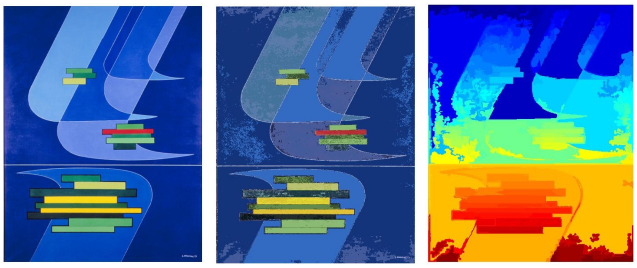

Figure 7: Example of our process of finding color combinations. From the left: (1) the original image (© *MART - Archivio fotografico e Mediateca*); (2) The image repainted with our color palette; (3) Visualization of all segments (patches) of the image. In (3) each color represents one patch.

4.3.2 The Amount of Colors Matters

We assume that the amount of colors affects the emotions elicited by abstract paintings. According to Itten, the harmonic areas of the paintings are generated by a balanced proportions of colors, which yield static, quiet effects. Itten states that the relative areas of two or more color patches, are neutralized when harmonious proportions of colors are used.

Itten gives as an example the work of Piet Mondrian. He points out that "Mondrian paintings employ two elementary resources, contrast of proportion and contrast of hue. [...] The quantitative proportions of the resulting areas assume a peculiarly independent life. Small configurations gain great significance by their placement in the field, while large forms recede and seem as if congealed. High sensitivity to proportion is required to organize all areas of a painting into a balanced whole."

4.3.3 The Position of Colors Matters

We consider that the emotions evoked by abstract paintings depend also on the position of colors. Itten points out that the distribution of colors is important to produce a distinct expression on the painting. In addition, he remarks that each direction in the area of a painting has its peculiar expression [5].

4.3.4 The Distance between the Colors Matters

We assume that the distances between colors might affect the viewer emotionally. In [5], Itten shows that the same color can appear to change in value, depending of the surrounding color.

- The weights of color areas act on either side of that axis. The left side is ordinarily the more passive, the right connoting activity. Right tends onward and upward, while left draws backward and downward.
- The horizontal direction denotes weight, distance, breadth.
- The vertical direction, conversely to the horizontal direction, denotes lightness, height and depth.
- The two directions together give an effect of area, a feeling of equilibrium, solidity and material hardness.
- A strong accent occurs where horizontal and vertical intersect.
- The diagonal directions generate movement and lead into the depth of the picture.

4.4 Feature Selection

We considered two types of features to represent the color combinations in an image:

1. Adjacent color co-occurrence features, which focus solely on the adjacent patches of color and are extracted on the level of the whole painting.
2. Patch-based color-combination features, taking into account the size, position and binary interactions between all the blobs of different color in a painting.

Since the segmentation yields segments (patches) that may not be completely homogeneous, each patch (p) is first assigned the most common value in it (mode), as its true color ($C(p)$).

To extract the adjacent color co-occurrence features, we consider a color combination to be present in the painting, if there are neighboring patches of the two colors in the combination. The adjacent color co-occurrence feature is then a histogram of all color combinations occurring in a painting.

To be able to model all the principles described in this section, we designed a more elaborate patch-level feature, which describes the size of each patch of color (amount of color in the patch), its position and its relationship to patches of other colors, present in the painting. To represent the amount of color contained in each patch, we calculated the relative size ($S(p)$) of each color patch, i.e. the size of each segment in pixels divided by the size of the image in pixels.

To allow for a scalable description of the position of each patch within a painting, the position is represented in relative terms by dividing the horizontal ($X(p)$) and vertical offset ($Y(p)$), from the top left corner of the image, of the center of the mass of the patch with image width (W) and height (H), respectively.

Once the position of all the patches has been determined, we calculate the relative distance of all other colors from a patch ($D(p,c)$) based on Eq. (1).

$$D(p,c) = \sum_{p_i \in P(c)} 1 - \frac{||L(p) - L(p_i)||_2 * S(p_i)}{\sqrt{W^2 + H^2}}, \qquad (1)$$

where $P(c) = \{p|C(p) = c\}$ is the set of all patches of color c in the painting and $L(p) = (X(p), Y(p))$ is the absolute location of the patch p within the image.

Our patch-based color-combination feature for a patch p is a concatenation $(X(p), Y(p), S(p), C(p), D(p, c_1), ..., D(p, c_N))$. N is the size of our palette. The painting is described by a set of features extracted for all the color patches in it.

Algorithm 2 outlines our elicited emotion classification approach. The images are first mapped to the corresponding palette, segmented and used to extract features, which are then input into the emotion classifier.

For the simple color co-occurrence features the classifier makes a decision based on a single feature vector extracted for a painting. For the color-combination features, the classifiers attempts to label each patch as positive or negative and the final classification for the painting is the mode of the labels obtained for each patch in the painting.

Algorithm 2: Elicited emotion classification

Data: image, color_palette
Result: emotion_label
cni ← $color\_name\_features(image)$;
for $pix\_val \in cni$ **do**
  | $pix\_val \leftarrow \min(dist(pix\_val, color\_palette))$;
end
rgb ← $color\_name\_to\_rgb(cni)$;
patches ← $color\_segmentation(rgb)$;
features ← $feature\_extraction(patches)$;
emotion_label ← $emotion\_classifier(features)$;

---

5. EVALUATION AND RESULTS

In this section we present the experimental results on the classification of abstract paintings into those eliciting positive and negative emotions.

We compare the results of the proposed approach to the state-of-the-art work of Yanulevskaya *et al.* [24], who applied a standard bag-of-words paradigm to extract color-based LAB visual words. They used a grid-based approach to create a color palette (visual vocabulary of colors) containing 343 LAB colors and represented the paintings as histograms of visual word occurrence. Their results show that colors are an important feature for the recognition of emotions elicited by abstract paintings, although they have not considered the effect of color-combinations.

To compare the proposed adjacent color co-occurrence features with those used in the work of Yanulevskaya *et al.* [24], we trained a Support Vector Machine with a linear kernel for supervised learning. The TrueSkill ratings are used as ground truth labels. We tested our model in a 2-fold cross-validation setup, where the images are assigned to folds randomly, and repeat the procedure 1,000 times.

Table 1 shows the classification accuracy results achieved using the adjacent color co-occurrence features approach and using LAB visual words approach of Yanulevskaya *et al.* [24].

Table 1: Linear SVM classification results.

Accuracy	LAB Visual Words	Color Combination
MART	0.673 ± 0.025	0.696 ± 0.028
deviantArt	0.678 ± 0.021	0.691 ± 0.022

Using the color co-occurrence framework, we obtained a 69.6% correct classification rate for the MART dataset. Our method of using color-combination related features improves the performance of the classification, with respect to the pure color-based methodology. Most of the paintings in the MART dataset contain uniform colors (i.e., 781.808 unique colors) which makes it simpler to extract our Color Palette.

The paintings in deviantArt dataset contain a larger amount of colors (i.e., over 4.4 million of unique colors), which made it more difficult to determine derive the Color Palette. To solve this problem, we split the set into smaller blocks of 100,000 colors and applied k-means on each block to get 100 centroids per block. As a result, we got 44.400 centroids from the blocks and then we clustered them into 180 k clusters. Although the deviantArt dataset presents a large amount of colors, the resulting Color Palette seems consistent with the colors represented in this dataset. Indeed, the classification accuracy is 69.1%, which is higher than LAB visual words that achieve 67.8% correct classification rate.

6. LEARNING APPROACH

The patch-level features described in Section 4.4 are specifically designed for inferring emotions elicited by abstract paintings from the color combinations present in the paintings. In order to exploit the structural relationships in the features, we propose inferring the emotions with a sparse group lasso.

6.1 Notation

The Lasso is a shrinkage and selection method for linear regression. It minimizes the usual sum of squared errors, with a bound on the sum of the absolute values of the coefficients. The traditional Lasso problem is formulated as:

$$\arg\min_{\beta} \|Y - X\beta\|_2^2 \qquad s.t. \ \|\beta\|_1 \leq \sigma \tag{2}$$

where $\sigma > 0$ is a constant which is used for bounding the optimal solution. $Y \in R^n$ is the emotion label vector for n training blob samples (1 for positive feeling and -1 for negative feeling), $X = [x_1, x_2, ..., x_n] \in R^{n \times d}$ is the feature matrix composed of d-dimensional blob feature vectors and $\beta \in R^d$ is the learned coefficient vector. Since the ℓ_1-norm is imposed on β, the learned β has sparse solutions.

6.2 Emotions from Color Combinations with Sparse Group Lasso

Since we have color groups information in our proposed features, we can formulate the problem as a sparse group lasso, where we enforce sparsity on both the group level (location group, color groups) and individual feature level. This is based on the observation that the color blobs can be very different from several groups of color, *e.g.*, red should be more similar to orange groups other than blue groups. To better exploit this group information, we formulate our problem as:

$$\arg\min_{\beta} \|Y - X\beta\|_2^2 + \lambda_1 \sum_{j=1}^{m} \left\|\beta_{G_j}\right\|_2 + \lambda_2 \|\beta\|_1 \tag{3}$$

where the training features are partitioned into m disjoint groups G_1, G_2, ... G_m, $\beta = [\beta_{G_1}, \beta_{G_2}, ... \beta_{G_m}]$ and β_{G_m} denotes the group of weights corresponding to group G_m. The first term in Eq.(3) is the reconstructed error. The second term in Eq.(3) is the combination of ℓ_1 and ℓ_2 norm.The ℓ_2 norm is used for the weights inside the same group, and the ℓ_1 norm is used to sum the results between groups. Using the group structure ensures more robust and accurate weights and still benefits from the sparsity. The third term in Eq.(3) is the ℓ_1 norm to enforce the sparsity on the coefficients in β, which benefits from the sparsity by selecting only a few groups with similar colors.

Testing: Since color blobs in each painting are used for training, the final emotion prediction for a new test painting is based on the majority vote from all blobs in the new test painting.

Table 2 shows the comparison with baseline methods on MART and DeviantArt dataset based on color group feature. In particular, we compared with SVM and Lasso which could not have structure of color groups. From Table 2, we observe that 8% and 6% relative increase in accuracy is achieved using the sparse group lasso approach for MART and DeviantArt dataset respectively. This shows the benefit of properly using color group information.

Table 2: Comparison with baseline methods on MART and DeviantArt dataset.

	MART	DeviantArt
SVM	0.676 ± 0.025	0.687 ± 0.031
Lasso (without group structure)	0.694 ± 0.021	0.701 ± 0.022
Proposed	0.751 ± 0.014	0.745 ± 0.019

7. APPLICATION

As an application, in this section we present an empirical prospective study done on the color combinations generated with our approach. Specifically, we generated Mondrian-like paintings from the Neo-plasticism style, using the colors obtained from our datasets and invited people to emotionally judge these simulated paintings.

Piet Mondrian was a noted Dutch painter and contributor to *De Stijl*, a Dutch artistic movement. The Neo-plasticism style, considered as "the new way of treating the form", is a geometrical style of abstract art which is essentially composed by horizontal and vertical straight black lines filled with primary colors [2]. Moreover, Mondrian's paintings employ the two theories of Itten, which are used in our study: the color combinations theory and the proportion of colors: "Mondrian confined himself to the fundamental colors of yellow, red, blue, white and black. Each of these colors has its unique character and special weight. The position of each color is very important, and so is its orientation, horizontal or vertical." [5]

We selected three Mondrian paintings and randomly changed the colors, replacing them with the colors from the color palette generated from the MART Dataset (Fig. 4). We selected paintings which highlight the aspects we want to analyse in the paintings: distance, amount and the positions of colors.

Figure 8 displays the original paintings selected for our study. We refer to the three paintings as simple, intermediate and busy, based on the quantity of colors and lines depicted in the painting.

The simple painting (Fig. 8 (a)), is composed of three filled small rectangles with primary colors. This painting is mainly composed of small blue, red, yellow and black areas and a large white area. The intermediate painting (Fig. 8 (b)) is composed of a big blue square divided in 3 parts for the black lines and small areas of red, black and yellow. It has also white areas between the colors. The busy painting (Fig. 8 (c)) is composed almost entirely of areas of colored rectangles, which are red, yellow, blue, gray, black and white.

To generate the Mondrian-like paintings we calculated the coordinates size of the horizontal and vertical lines according

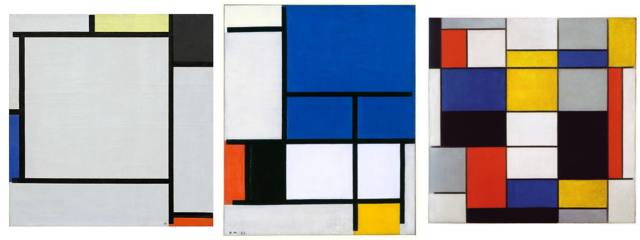

Figure 8: Piet Mondrian. From the left: (a) Tableau 2, 1922. Extracted from: Solomon R. Guggenheim Museum, New York; (b) Composition with Large Blue Plane, Red, Black, Yellow and Grey, 1921. Extracted from: The Dallas Museum of Art, Texas; (c) Grande composizione A con nero, rosso, grigio, giallo e blu, 1919 - 1920. Extracted from: Galleria Nazionale d'arte Moderna, Rome.

to the size of the original painting. Then, in the same position where the colors were used by Mondrian in the paintings (i.e., red, yellow, blue, gray and black) we automatically substituted the original colors with a random selection of colors from the color palette extracted from the MART dataset. In total, approximately 15.000 Mondrian-like paintings were generated for each selected Mondrian painting (i.e., simple, intermediate and busy). Figure 9 shows one example of Mondrian paintings generated by our algorithm.

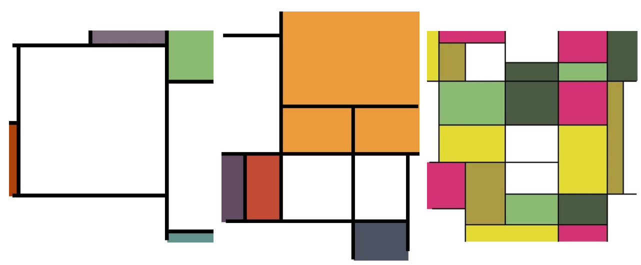

Figure 9: Example of Mondrian-like paintings generated with random colors from the MART Color Palette. From the left: (a) Mondrian Simple; (b) Mondrian Intermediate; (c) Mondrian Busy.

Afterwards, we applied the same segmentation method detailed in section 4.2 to generate the adjacent color co-occurrence features. We then used the linear Support Vector Machine classifier trained on the MART dataset to classify these images. We select the images with the highest classifier confidence to create a database of 500 Mondrian-like paintings.

7.1 User Study for Assessing Feelings Elicited by Mondrian-like Paintings

To annotate the paintings we use the same relative scale approach described in section 3.3. In total, 35 subjects participated in the annotation task, 18 were females and 17 males. Each subject annotated 104 paintings on average. The subjects participated voluntarily and were free to annotate at any time they wanted.

The distribution of TrueSkill scores obtained is shown in figure 2. As the plot shows, the generated images have almost the same mean and standard deviation of the scores, as the images in the DeviantArt dataset, and the general shape of the distribution matches both DeviantArt and MART datasets.

To define the threshold we use the same procedure of Section 3.3 and choose the TrueSkill score of the painting that was chosen as more positive 8 times. Therefore, the Mondrian-like painting with TrueSkill ranking value equal to 24.7 is used as threshold. In total, 241 paintings were assigned to the negative class and 259 paintings to the positive class. We then compared the TrueSkill ranking values of the Mondrian-like paintings with the results from the classification reported in the previous section. As a result, we have 60.6% correct classification rate using the Mondrian-like paintings with the MART color Palette.

We also matched the TrueSkill ratings obtained at the end of the user study to the individual annotations of the pairs, presented to all the annotators. The results show about 51.58% agreement, which can be considered a representative value of the mean human performance.

At last, Table 3 shows the comparison with baseline methods on Mondrian dataset based on color group feature. In particular, we compared with SVM and Lasso which could not have structure of color groups. From Table 3, we observe that 7% relative increase in accuracy is achieved using sparse group lasso approach on Mondrian dataset. This also shows the benefit of properly using color group information.

Table 3: Comparison with baseline methods on Mondrian dataset.

	Mondrian
SVM	0.576 ± 0.035
Lasso (without group structure)	0.625 ± 0.031
Proposed	0.671 ± 0.021

8. CONCLUSIONS

In this work, we propose novel features and computational approach, aiming to analyze the impact of color combinations on the positive or negative emotions an abstract painting will invoke in an observer.

We exploit art theory concepts to design our features and the learning algorithm. The features proposed allow a baseline classifier to infer the emotions better than the existing state-of-the-art features.

To use the color groups information better, we propose inferring the emotions elicited by paintings based on the sparse group lasso approach. The application of this method is able to increase the accuracy of the classifier even more.

Finally, as an application, we employ the colors extracted from our method to generate Mondrian-like paintings and do a prospective user study to evaluate the ability of our method as a tool for generating paintings able to elicit positive and negative emotional responses in people. The study shows that the proposed approach is able to predict the emotional response of people with accuracy higher than that of the mean human performance.

The limitation of our study is that the focus has been placed on abstract paintings and computational modelling of the art theory of emotions elicited by these paintings. It is unclear to which extent the proposed approach and the theory it is founded on can be applied to other types of paintings and natural images. Future work will focus evaluating possible extensions of the proposed approach to domains beyond abstract art.

Acknowledgments

We thank Victoria Yanulevskaya for the discussions in the initial phases of this work. This work has been partially supported by the MIUR Cluster project Active Ageing at Home, the EC project ACANTO and the A*STAR Singapore Human-Centered Cyber-physical Systems (HCCS) grant.

9. REFERENCES

[1] L. Dickerman, M. Affron, and M. of Modern Art. *Inventing Abstraction, 1910-1925: How a Radical Idea Changed Modern Art*. Museum of Modern Art (New York, N.Y.), 2012.

[2] G. Dorfles and A. Vettese. *Arti visive. Protagonisti e movimenti. VOL.3A: Il Novecento*. 2001.

[3] P. F. Felzenszwalb and D. P. Huttenlocher. Efficient graph-based image segmentation. *Int. J. Comput. Vision*, 59(2):167–181, 2004.

[4] R. Herbrich and T. Graepel. Trueskill(tm): A bayesian skill rating system. *no. MSR-TR-2006-80*, 2006.

[5] J. Itten. *The Art of Color: The Subjective Experience and Objective Rationale of Color*. Wiley, 1974.

[6] K. Ivanova, M. Dobreva, P. Stanchev, and G. Totkov, editors. *Access to Digital Cultural Heritage: Innovative Applications of Automated Metadata Generation.*, chapter APICAS - Content-Based Image Retrieval in Art Image Collections Utilizing Colour Semantics, pages 153–202. Univ. Publ. House Paisii Hilendarsk, 2012.

[7] W. Kandinsky. *Concerning the Spiritual in Art*. Dover Books on Art History Series. Dover, 1914.

[8] R. Khan, J. van de Weijer, F. Shahbaz Khan, D. Muselet, C. Ducottet, and C. Barat. Discriminative color descriptors. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2866–2873, June 2013.

[9] G. Liu, Y. Yan, E. Ricci, Y. Yang, Y. Han, S. Winkler, and N. Sebe. Inferring painting style with multi-task dictionary learning. In *International Joint Conferences on Artificial Intelligence (IJCAI)*, 2015.

[10] M. Livingstone. *Vision and art: the biology of seeing.* Harry N. Abrams, 2002.

[11] J. Machajdik and A. Hanbury. Affective image classification using features inspired by psychology and art theory. In *ACM Multimedia*, pages 83–92, 2010.

[12] J. Moser. True skill library. https://github.com/moserware/Skills/, 2010.

[13] I. Newton. *Opticks: Or, a Treatise of the Reflections, Refractions, Inflexions and Colours of Light. Also Two Treatises of the Species and Magnitude of Curvilinear Figures*. S. Smith, and B. Walford, 1704.

[14] L.-C. Ou, M. R. Luo, A. Woodcock, and A. Wright. A study of colour emotion and colour preference. part i: Colour emotions for single colours. *Color Research & Application*, 29(3):232–240, 2004.

[15] L.-C. Ou, M. R. Luo, A. Woodcock, and A. Wright. A study of colour emotion and colour preference. part ii: Colour emotions for two-colour combinations. *Color Research & Application*, 29(4):292–298, 2004.

[16] A. Sartori, Y. Yan, G. Ozbal, A. Salah, A. Salah, and N. Sebe. Looking at mondrian's victory boogie-woogie: What do i feel? In *International Joint Conferences on Artificial Intelligence (IJCAI)*, 2015.

[17] A. Sartori, V. Yanulevskaya, A. A. Salah, J. Uijlings, E. Bruni, and N. Sebe. Affective analysis of professional and amateur abstract paintings using statistical analysis and art theory. *ACM Transactions on Interactive Intelligent Systems*, 5(2):1–27, June 2015.

[18] F. Shahbaz Khan, R. Anwer, J. van de Weijer, A. Bagdanov, M. Vanrell, and A. Lopez. Color attributes for object detection. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 3306–3313, June 2012.

[19] P. Valdez and A. Mehrabian. Effects of color on emotions. *Journal of Experimental Psychology: General*, 123(4):394, 1994.

[20] J. van de Weijer, C. Schmid, J. Verbeek, and D. Larlus. Learning color names for real-world applications. *IEEE Transactions on Image Processing*, 18(7):1512–1523, 2009.

[21] J. von Goethe and C. Eastlake. *Goethe's Theory of Colours*. J. Murray, 1840.

[22] Y. Yan, E. Ricci, R. Subramanian, O. Lanz, and N. Sebe. No matter where you are: Flexible graph-guided multi-task learning for multi-view head pose classification under target motion. In *International Conference on Computer Vision (ICCV)*, 2013.

[23] V. Yanulevskaya, J. V. Gemert, K. Roth, A. Herbold, N. Sebe, and J. Geusebroek. Emotional valence categorization using holistic image features. In *IEEE ICIP*, pages 101–104, 2008.

[24] V. Yanulevskaya, J. Uijlings, E. Bruni, A. Sartori, E. Zamboni, F. Bacci, D. Melcher, and N. Sebe. In the eye of the beholder: employing statistical analysis and eye tracking for analyzing abstract paintings. In *ACM Multimedia*, 2012.

[25] S. Zhao, Y. Gao, X. Jiang, H. Yao, T. Chua, and X. Sun. Exploring principles-of-art features for image emotion recognition. In *ACM Multimedia*, pages 47–56, 2014.