Learning moods and emotions from color combinations

Gabriela Csurka*
XRCE-Xerox Research Centre Europe
6, Chemin de Maupertuis
Meylan, France
gabriela.csurka@xerox.com

Sandra Skaff
XRCE-Xerox Research Centre Europe
6, Chemin de Maupertuis
Meylan, France
sandra.skaff@xerox.com

Luca Marchesotti
XRCE-Xerox Research Centre Europe
6, Chemin de Maupertuis
Meylan, France
luca.marchesotti@xerox.com

Craig Saunders
XRCE-Xerox Research Centre Europe
6, Chemin de Maupertuis
Meylan, France
craig.saunders@xerox.com

ABSTRACT

In this paper, we tackle the problem of associating combinations of colors to abstract categories (e.g. capricious, classic, cool, delicate, etc.). It is evident that such concepts would be difficult to distinguish using single colors, therefore we consider combinations of colors or *color palettes*. We leverage two novel databases for color palettes and we learn categorization models using low and high level descriptors. Preliminary results show that Fisher representation based on GMMs is the most rewarding strategy in terms of classification performance over a baseline model. We also suggest a process for cleaning weakly annotated data, whilst preserving the visual coherence of categories. Finally, we demonstrate how learning abstract categories on color palettes can be used in the application of color transfer, personalization and image re-ranking.

1. INTRODUCTION

Color is one of the most instantaneous methods for conveying media messages [6]. Color can not only symbolize thoughts, but it can often contribute significantly to aesthetic, emotional and stylistic impressions of what is being visualized. The ubiquitous presence of color in media such as documents and images naturally raises the question of whether it is possible to learn such impressions from colors, thus enabling a new range of tools which would allow retrieval and organisation of digital assets using abstract concepts. For example, such tools would be advantageous in many applications related to graphic design for personalizing a specific asset according to the context of use, user preference or profile.

*Corresponding author

ICVGIP '10, December 12-15, 2010, Chennai, India

Research in this domain however is hindered by the shortage of data annotated with abstract categories arising from emotional, mood, stylistic or aesthetic labels. Furthermore, due to the nature of these categories, such data tends to be 'noisy' in the label space: user preferences differ, the concepts themselves are difficult to define in a consistent manner, and any data actually labelled in this way is often incomplete (for example an asset may be tagged as 'sunny' and 'warm', but not tagged as 'summer' or 'peaceful' even though the latter two labels may be equally applicable). Despite these difficulties, research has been conducted into linking color to abstract categories. Most methods, however, have so far been aimed at associating color name tags to single colors. This was performed either through user tests in constrained environments [2, 4, 10, 1] or by training statistical models using data collected on the web [17]. To our knowledge, only very few studies [12][18] explored the relationship between perceived colors and the emotion induced, but the results again involved a limited set of individual colors and categories.

In this paper, we go beyond color name tags and directly tackle the problem of associating combinations of colors to abstract categories. We create a vocabulary of natural language tags (typically adjectives) which describe 15 categories (e.g. capricious, classic, cool, delicate, etc., see Table 1 for the complete list) that are indicative of some of the moods which people may associate to an asset when visualizing combinations of colors in images or documents. It is evident that such concepts would be difficult to distinguish using single colors, therefore we consider combinations of colors or *color palettes*.

Our work has several technical contributions: 1. We analyse two novel databases for color palettes, one that is *strongly annotated* by skilled designers, and a second larger database created by amateurs which is *weakly annotated* and contains a great deal of label uncertainty. 2. We suggest and compare different feature representations for color palettes. The main result is that using a Fisher representation based on GMMs, we can reliably model strongly annotated data and achieve good classification performance over a baseline model. 3. We suggest a process for cleaning weakly annotated data, whilst preserving the visual coherence of categories. We give both quantative analysis based on a retrieval task, and qualitive results on palette ranking to show the effectiveness of such cleaning. 4. We also demonstrate how learning such abstract categories on color palettes can be used in the application of color transfer and personalisation. Furthermore, we show qualitative results on how such models can be used in reranking image collections.

The paper is organized as follows: in section 2 we describe the databases we created and used for the experimentation, and section 3 introduces various palettes representations. In section 4 we provide the results of the tests performed around palette categorization, and also introduce and discuss the effectiveness of our cleaning method (section 5). Sections 6, 7 and 8 demonstrate the use of such tools for image retrieval and color transfer. This is followed be section 9 which outlines the conclusions and the future work.

2. THE DATABASES

We created two novel databases of color palettes. One of them contains palettes generated by color consultants and experts in the field of graphic design. The second was derived from a popular social network of amateur designers that put their palettes online as well as associated metadata such as textual annotations, comments, bookmarks, and so on. This leads to two very distinct data sets with two key challenges: the first set has a minimal number of examples per class; the second is richer, but more noisy in the textual annotations.

2.1 Communicating with Colors Database

The first database we present is derived from [6], in which hundreds of color combinations were created and organized according to semantic categories. The taxonomy of these categories is flat and it describes the concepts such as joyful, classic, romantic, cool, etc. (see Table 1). These concepts are based on color combinations which are suitable for creating marketing collaterals such as advertisements, flyers, and brochures. Among the categories in the book, we selected the 15 categories with palettes formed by at least 3 colors. Since palettes in this dataset were carefully annotated by experienced color consultants, the swatches, which constitute the individual colors of a palette, were selected from the colors of the Pantone Matching System (PMS), a proprietary color space considered as a “de-facto” standard in a variety of industries, primarily printing, colored paint, fabric, and plastics [6]. In the database, there are 134 unique Pantone colors combined into 360 distinct palettes. We denote this dataset by CC.

2.2 ColourLovers Database

We downloaded approximately 22000 color palettes from a popular social network of graphic design enthusiasts [8] using the keywords associated to each class of CC and derived from [6]. Palettes are loosely tagged by users with natural language words. This leads to a more challenging data set if one wishes to perform classification or retrieval tasks: concepts are less well defined, ambiguities and incoherences are introduced, labels are often missing and so on. Furthermore, in most part of the cases, the palettes consist of five color swatches, where colors are chosen among the 16^3 combinations available in the RGB space and not among the 1116 colors forming the Pantone Color System. For these reasons, we shall refer to this dataset as “weakly annotated”. In figure 1, we plotted the distribution of colors in the R,G,B cube for two reference categories (Romantic and Playful) in CC and CL. As expected, the distributions in CC are more compact and clusters of colors appear in the distribution. However, at least for certain categories such as “Romantic”, the distributions in CC and CL share overlapping clusters.

Table 1: 15 style and emotional palettes categories along with associated keywords. These keywords disambiguate the sense of the category. They are used for the annotation of the “Communicating witih Colors“ database (CC) and for the “Colour Lovers“ database (CL). The brakedown of the number of palettes per class is indicated on the columns in the right for CC and CL database.

Id	Concepts	Keywords	$\|CC\|$	$\|CL\|$
1	**capricious**	eclectic, intricate, diverse, unexpected, surprise, eureka	24	205
2	**classic**	simplicity, timeless, neutral, basic, essential, uncomplicated, minimal, beige, stone, sand, pebble, rock, cavern, cave, ancient	24	1468
3	**cool**	purifying, wet, soothing, refreshing, fresh, crisp, revive, rejuvenate, refresh, water, diving, splash, breath, mint, crystalline, sparkling, air	24	3443
4	**delicate**	infantile, cherished, wispy, fragile, fine, mild, gentle, lightest, pale, vulnerable, baby	24	767
6	**earthy**	rustic, sturdy, sheltering, woodsy, wholesome, autumnal	24	563
5	**elegant**	distinguished, refined, sophisticated, polished, cultivated, luxurious, accomplished, dark, amethist, aubergines, peridot, emerald, malachite, sapphire, ruby, opulence, gold, bronze, silver, copper	24	288
7	**luscious**	sweet, appealing, scrumptious, delectable, sugary, delicious, ice cream, gelatos, mocha, fudge, cherry, raspberry	24	2618
8	**playful**	outgoing, comical, comic, fun, fun-loving, joyful, joy, child, children, child-like, kid, noisy, play, active, spontaneous	24	2769
9	**robust**	rich, vintage, resonant, lush, tasty, savory, hearty	24	2504
10	**romantic**	loving, intimate, charming, soft, affectionate, beloved, candlelight, mid-tone, rose, pale, diffuse	24	2527
11	**sensual**	fiery, tempestuous, alluring, exotic, daring, captivating, passionate, erotic, sexy, provocative	24	345
12	**serene**	calming, tranquil, peaceful, quiet, reassuring, uplifting, peace, placid, sky, clean, clouds, lavender, snow	24	801
13	**spicy**	tangy, piquant, sizzling, bold, gusto, acidic, punguent, zesty	24	965
14	**spiritual**	meditative, enchanting, sublime, mysterious, mystical, inspirational, lofty , transcendent, enigmatic, ghost, ghostly, surreal	24	316
15	**warm**	sunbaked, temperate, mild, comfy, balmy, toasted, dry, sunshine, sun, creamy, golden, non-invasive, easy-going, tangerine, sundance	24	2826

3. COLOR PALETTE REPRESENTATIONS

A color palette is a sequence of colors, which are referred to as *swatches*. Each swatch is represented by a 3D vector in a given color space, such as RGB, HSV, CMY, etc. A palette of c color swatches is represented by a $c * 3$ dimensional

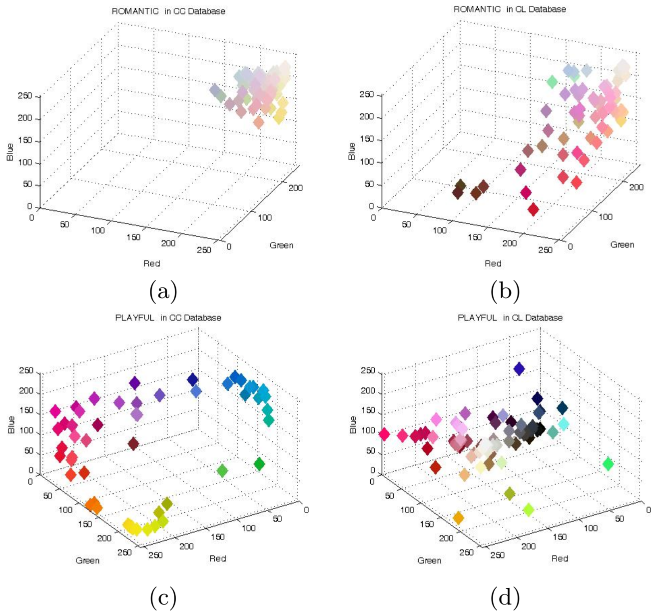


Figure 1: Distribution of colors in the R,G,B cube for the two categories "Romantic" and "Playful" plotted using samples drawn from the CC (a)-(c) and 1% of colors randomly drawn from CL (b)-(d).

vector as such: $[x_1, .., x_k, .., x_c]$, where x_k is a 3D vector denoting a single color swatch. In the case of the palette databases used, the *CC* and *CL* palettes have three and five swatches respectively. In this paper, we take RGB to be the reference color space and show detailed comparisons to representations of the swatches in other color spaces, both perceptually coherent and non coherent ones. Four different feature extraction schemes were considered:

- *Raw*: The individual color swatches are concatenated into a feature vector X_t.

- *Mean*: The means of swatches are computed for each channel, such as each of R, G and B for the RGB color space representations, and concatenated into a single 3D vector X_t.

- *Bag of 'Colors' (BOC)*: A color vocabulary is created by sampling at random the colors in the chosen database [5]. We derive a probability density function (pdf), denoted by p, which models the process of generation of colors in the database. The color vocabulary is modeled with a Gaussian mixture model (GMM), where each Gaussian corresponds to a color [7, 13]. Let $\lambda = \{w_i, \mu_i, \Sigma_i, i = 1...N\}$ be the set of parameters of p where w_i, μ_i and Σ_i denote respectively the weight, mean vector and covariance matrix of Gaussian i and where N denotes the number of Gaussians. Let p_i be the distribution of Gaussian i such that we have $p(x) = \sum_{i=1}^{N} w_i p_i(x)$, where x denotes an observation of a color in this case. Let $\gamma_i(x_k)$ denote the occupancy probability that the color x_k is assigned to Gaussian i. This quantity can be computed with Bayes' formula and then normalized using:
$$\gamma_i(x_k) = \frac{w_i p_i(x_k)}{\sum_{j=1}^{N} w_j p_j(x_k)}. \quad (1)$$
The color swatch x_k is then transformed into the high-level N-dimensional descriptor:
$$\gamma(x_k) = [\gamma_1(x_k), \gamma_2(x_k), \ldots, \gamma_N(x_k)].$$
The palette feature vector X_t is then obtained by summing $\gamma(x_k)$ over the c color swatches :
$$X_t = [\sum_{k=1}^{c} \gamma_1(x_k), \sum_{k=1}^{c} \gamma_2(x_k), \ldots, \sum_{k=1}^{c} \gamma_N(x_k)].$$

- *Fisher Vector (FV)*: The palettes are represented using the Fisher kernel framework [9], which is an alternative to the BOC. The FV of a palette describes in which direction the parameters of the model p (as given in the previous section) should be modified to best fit this sample. We assume diagonal covariance matrices and we consider the gradients of $\log p(x_k|\lambda)$ with respect to only the mean and standard deviation parameters; as the gradient with respect to the weight parameters provides negligible additional information [13]:
$$f_{\mu_i^d}(x_k) = \frac{\partial \log p(x_k|\lambda)}{\partial \mu_i^d} = \gamma_t(i)\left[\frac{x_k^d - \mu_i^d}{(\sigma_i^d)^2}\right],$$
$$f_{\sigma_i^d}(x_k) = \frac{\partial \log p(x_k|\lambda)}{\partial \sigma_i^d} = \gamma_t(i)\left[\frac{(x_k^d - \mu_i^d)^2}{(\sigma_i^d)^3} - \frac{1}{\sigma_i^d}\right].$$
where the superscript $d = 1 \ldots M$ denotes the d-th dimension of the vector in the M-dimensional feature space. The Fisher Vector $f_\lambda(x_k)$ of the color swatch x_k is the concatenation of all these partial derivatives leading to a $2 * M * N$ dimensional vector. Finally, to obtain the palette representation X_t we take the sum over the Fisher Vectors from all the swatches:
$$X_t = \sum_{k=1}^{c} f_\lambda(x_k).$$

Both Fisher Vector and BOC representations are advantageous over the other features described above because they are independent from the number of color swatches and their order. More importantly, since they are high-level feature representation techniques, they are more richer descriptors leading to more accurate results.

Any discriminative classifier may be used in order to classify palettes. For our experiments we use a Sparse Logistic Regression (SLR) classifier, with an element wise square root normalization of the feature vector (known as the Bhattacharrya kernel).

4. PALETTE CATEGORIZATION

In this section, we compare categorization performance using the various palette representations outlined above. All results were obtained using 5 fold cross-validation (80% training, 20% test). and averaging the classification accuracy (ACC) over the folds. The BOC and the FV were computed using 16 Gaussian[1] "color vocabulary" (GMM)

[1]We made several tests using a bigger sized vocabulary (e.g 32 Gaussians) but improvements on the categorization performance were in general small.

which was trained on colors randomly selected from the CL database. We used a subset of colors randomly selected from a much larger number of color swatches (approximately 0.5 million) to ensure good generalization of the vocabulary and to avoid bias of the features. The Fisher Vectors were also computed using the same color vocabulary (GMM).

Table 2: The first three rows show average classification accuracy (in %) using different palette features in the RGB, HSV, and Lab color spaces. The fourth line shows ACC when late fusion (LF) is used on all features, and the fifth line also includes MCM features.

ACC	*CC*				*CL*			
	Raw	Mean	BOC	FV	Raw	Mean	BOC	FV
RGB	51.1	26.4	69.2	70	14.4	12.1	16.5	17.1
HSV	56.9	31.4	64.7	70.6	13.8	12.3	16.9	18.2
Lab	54.7	26.1	65.6	71.1	14.7	12.4	17.9	18.2
LF	58.3	31.9	75.3	73.9	15.6	13.1	18.6	18.9
LF+MCM	66.9	57.8	76.9	74.2	16.6	15.5	18.8	19.2

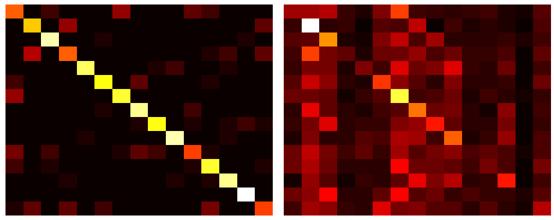
In Table 2 we show overall class accuracies averaged over the 15 classes for different features and different color spaces. For CC, we are able to achieve high classification accuracies which is not surprising as it is a reliably annotated database with a limited number of samples. The classification results on the CL dataset are generally poor for most categories due to the fact that the tags of these palettes are very noisy (either due to mislabelling or and incomplete set of tags, see example of fig. 1).

Nevertheless, we see a similar trend on both datasets for the different features. The simple mean of color swatches (a 3 dimensional feature vector) is the worst performing feature. The Raw features are able to capture some of the concepts in the CC database, however they fail completely in the case of the CL dataset. The higher level features (BOC and FV) outperform the low level features for both datasets. This gives additional evidence to the hypothesis that the problem for CL does not lie at a feature design level, but it concerned with the overwelming amount of noise in the annotations. BOC and FV perform more or less the same if evaluated in color spaces other than RGB. We also experimented with color spaces such as XYZ or YCbCr, leading to very similar performance to RGB as one might expect. Omitting luminance channel from YCbCr and focusing only on chrominance yielded significantly worse results. Therefore we conclude that luminance is an important descriptor for this kind of analysis. Finally, we also extracted a set of features specifically designed to capture emotional content [11, 15], but these also obtained much poorer classification accuracies.

Table 3: Classification accuracies (in %) obtained using early fusion.

ACC	Raw	Mean	BOC	FV
CC	60.8	68.9	74.7	71.9
CL	16.8	17.3	18.5	17

We also experimented with the low level features, BOC and FV combined in a late and early fusion strategy. In Table 2, the 4th row shows the results of late fusion of the features where we average the classification scores from each feature representation in the 3 color spaces. Table 3 gives the results of an early fusion approach which consists on concatenating the different features into a unique vector before classification. Figure 2 shows an example confusion matrix for the two datasets, illustrating the difficulties of label noise in the CL data set. The confusion matrices were evaluated using late fusion, however all representations show a similar trend.



(a) (b)

Figure 2: Confusion matrices using the late fusion of the FV representations of palette in RGB, HSV and LAB spaces for (a) CC, (b) CL.

We can see that the fusion always improves the individual performances, but in many cases the improvement is relatively small. The less significant increase in performance can be observed for the Fisher Vector. This is motivated by inner richness of the FV representation that does not leave margin for further improvement.

The BOC and Raw representations are somehow poorer and hence benefit more of the redundant information coming from different color spaces. The gain on early fusion leads to results not far from those obtained with more complex high level features (BOC and FV). This new 9 dimensional feature vector (the concatenation of RGB, HSV and Lab means) will be referred in the rest of the paper as MCM (Mean Color Means). We also combined with a late fusion strategy the the MCM results with the ones obtained with the other individual features (see fifth row of Table 2) and we obtained a further improvement of the classification accuracy.

Performance of raw features could potentially be boosted by considering all possible permutations of swatches and by making the features order independent. For some applications however, where palettes have been carefully designed, the order may provide relevant information. Furthermore, representations of raw colour features are dependant on the size of the palettes[2], whereas BOC, FV and MCM are indepedent on the number of swatches. We intend to investigate in future work the effect of using different size palettes and considering in more detail how order and insertions/deletions affect performance in various application scenarios.

5. DATABASE CLEANING

When dealing such diverse datasets such as CC and CL we are faced with a challenge: we would like to enrich CC

[2]In this work we focus on 5-color swatches palettes from ColorLovers, however palettes with different size exist in the database.

with the larger number of palettes available from CL; however we know that the latter suffers from severe label noise. The presence of noise is not surprising since the concepts we are learning are highly subjective. Therefore we implemented a strategy to clean the database by finding a consensus in the annotations and by removing outliers. Another option is to rely on extra human annotation, but it would be unpractical given the large number of examples forming our database.

We find relevant and compact clusters for each category where several annotators agreed to label similar palettes with the same tag. Finding coherent clusters and removing samples not validated by the majority of the annotators is part of the wider research on outlier removal. Developing such techniques is not aim of this paper, however we propose here a simple and efficient method that can be used to clean a weakly annotated databases.

Given a tag and a related palette, we measure the density of the similar palettes labeled with the same tag:

$$dens(p_i) = \frac{1}{|\mathcal{KNN}_i^c|} \sum_{p_j \in \mathcal{KNN}_i^c} dist(p_i, p_j) \tag{2}$$

where, $\mathcal{KNN}_i^c$ is the set of K nearest neighbor palettes for palette p_i labeled with the concept c, and $dist(p_i, p_j)$ is the distance between the palettes. The cleaning process then seeks to keep those palettes that are below a given density threshold TH.

As we have seen from previous experiments, the order of color swatches was not really meaningful in the case of the CL dataset. We therefore defined the following order independent distance between palettes:

$$dist(p_i, p_j) = dist(p_i, \tau^*(p_j)) \tag{3}$$

where $\tau^*(p_j)$ is a reordered version of the palette p_j with permutation τ. We choose the τ^* from all possible permutations of $(1, 2, ..N)$, N being the number of swatches in the palette, as the one that minimizes:

$$\tau^* = argmin_{\tau \in \mathcal{P}_N} \sum_{n=1}^{N} \|s_i^n - s_j^{\tau(n)})\| \tag{4}$$

where s_i^n is the n^{th} swatch (color) of the palette p_i (here in the RGB space).

Figure 3 demonstrates the effect of this methods when varying the threshold TH and the number of neighbours considered K.

5.1 Cleaning effect on categorization

In this section we show how classification results are impacted by the cleaning operation. To this end, we retrained our classifier with a late fusion strategy combining RGB , HSV, LAB Fisher Vectors and the MCM features (FV+MCM, cell in the last raw and last column of Table 2). Figure 4 shows the different average class accuracies (ACC) obtained by varying the threshold TH and the number of neighbors K taken into account.

For $TH = 150$ we retain nearly all categories and we observe an improvement in classification. As expected, the denser the clusters, the smaller is the number of categories we obtain. Figure 5 shows the individual class performances for $TH = 100$ varying K and Figure 6 shows the effect of varying TH with K fixed ($K = 50$). For an easier interpretation of the individual class performances we use Mean

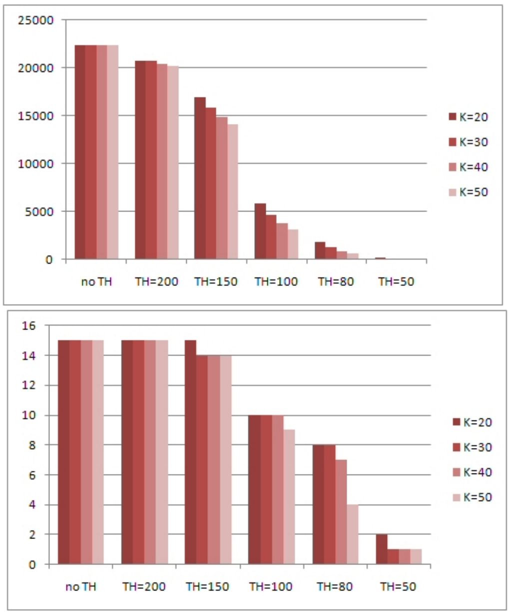


Figure 3: The number of palettes (up) and number of classes (down) in the dataset after the visual cleaning process varying TH and K.

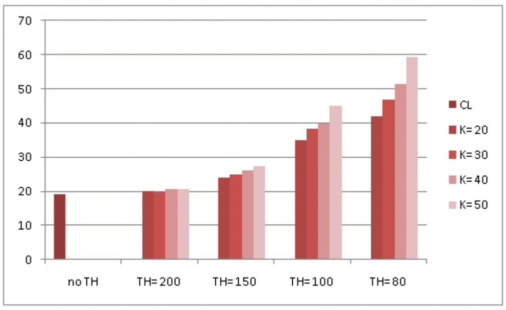


Figure 4: The average class accuracy after the visual cleaning process varying the threshold TH and neighborhood size K.

Average Precision (MAP) instead of the ACC. In fact, the MAP measures how well the relevant palettes are ranked given a classifier (independently from the other classifiers), whilst ACC depends on scores of all classifiers[3].

We can observe that increasing K generally increases the accuracy. Similarly, decreasing TH it also increases the performances. This latter parameter is a penalization term and it aggressively remove outliers. We can see that when $TH = 100$ or below, several classes have no palettes verifying these constraints, and those classes are accordingly removed (see Figure 3).

For further experiments in the paper, we set $TH = 100$ and $K = 50$, as this provides a compromise between accuracy (45%) and number of abstract concepts considered (9).

6. PALETTE RETRIEVAL

[3]A palette receives the label with the highest classification score among all the classifiers.

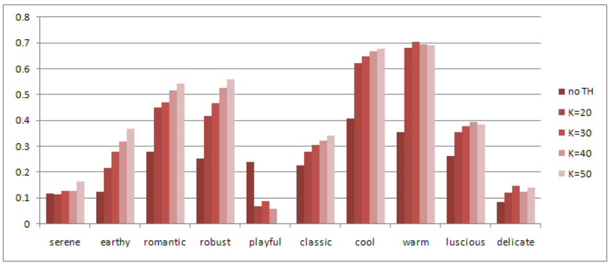


Figure 5: Class MAPs after the visual cleaning process varying K with the fixed threshold $TH = 100$.

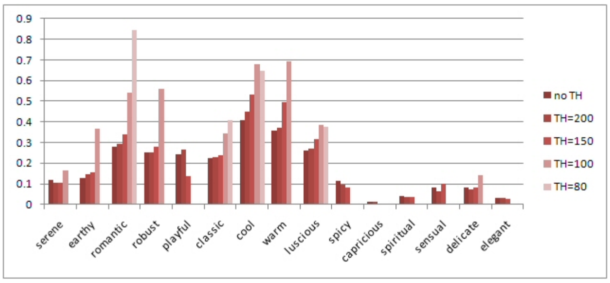


Figure 6: Class MAPs after the visual cleaning process varying the threshold TH with $K = 50$.

Retrieval experiment were carried out using the following databases: CC, CL and CL after the cleaning (cCL). They were used independently or merged in a 5-fold cross validation setup: 4 folds were used to train each classifier whereas the palettes of remaining folder were ranked using the classification scores. In fig. 7 we show palette retrieval results for classification models created using different training sets. We employed several measure such as mean average precision (MAP), break even point (BEP = 1-EER) and area under curve (AUC). Since similar behaviour was observed with these different measures, here we show only the MAP values. Here we only provide the results for the FV+MCM features combination. However, similar conclusions can be deduced using other feature combinations.

As it can be seen, all the classifiers perform rather poorly on CL. Instead, the fact that all classifiers perform better on cCL in comparison to CL it confirms the advantages of the cleaning operation. Once again the best performances are obtained on CC where the distribution of color palettes is more compact and perceptually coherent. Combining all training samples of CC to CL improves the retrieval performances on CC with respect to models trained on CL (or cCL) and it generates similar performances if we retrieve on CL.

For a better interpretation of these results, we also examined the palettes that received top classification scores in the retrieval experiment performed on cCL. In particular, in Figure 7 we show palettes ranked with decreasing score (from left to right) for the *serene* category for three classifiers trained on CC, cCL and $CC + cCL$. As expected the CC model retrieves very coherent palettes sharing similar colors whereas cCL finds more diverse palettes. We judge that the combination $CC + cCL$ provides a tradeoff which might be interesting for users willing to explore colors with some degree of variation over the dominant hues of CC and CL.

Table 4: Mean average precisions using FV+MCM. The algorithm is trained and tested on different combinations of CC, CL,cCL (cleaned CL).

	MAP	FV + MCM	
		CC	CL
Train	CC	81	15.8
	CL	51.7	20.2
	$CC + CL$	61.7	20.2
	cCL	51.5	18.9
	$CC + cCL$	65.9	19

7. IMAGE RETRIEVAL

We can apply the classification models used for palettes retrieval and classification to retrieve images. To verify this possibility, we proceeded as follows: we downloaded 220 images from Google Images for each of the 15 categories listed in Section 2. The queries used for each category were the category name concatenated with the tag "colors" (e.g "earthy colors"). We extracted five-swatch palettes from each image as described in step 2 of Section 8. The features used to represent the palettes were LF+MCM, and the palettes were classified using the classifer trained on $CC + cCL$ or $CC + CL$ (where the class was not in cCL such as *playful*). Example results of top retrieved images for the *delicate* and *playful* concept can be seen in Figure 8 and Figure 9 respectively. Additional results are provided in the supplementary material.

8. IMAGE COLOR TRANSFER

We present in this section another potential application where we can use the color palettes annotated by our classification models. In particular we focus on image color transfer based on a target concept. In the past, image color transfer was applied to harmonize the colors of the foregrounds and backgrounds in images [3], to color correct an undesirable color cast or to perform gamma correction [14]. In these cases, a target image is needed along with the input image to perform the transfer.

We propose to use a target concept among the ones studied in this paper, to perform the transfer and thereby eliminating the need of selecting a desired target image. In particular, given an image I_{in} to transfer, the application performs the following steps:

1. The user selects a target concept from the ones available.

2. An *input palette* pal_{in} is extracted from I_{in} following an approach similar to that of Tai *et al.* [16]. First, we convert the image into the Lab color space. We then fit a GMM using the image pixels as observations, denoted by x, while taking into account only the channels a and b. The parameters of the GMM μ_i, σ_i are estimated by maximizing the log of the likelihood function $p(x|\lambda)$ using the Expectation Maximization (EM) algorithm. We set the number of Gaussians to five as the pal_{in} should have the same number of swatches

Figure 7: Top 8 retrieved palettes from the CL dataset for the category *serene*, when the classifier is trained on the CC, cCL, and $CC + cCL$ datasets.

Figure 8: Top retrieved images from the downloaded GoogleImages with *delicate colors* and top images of our *delicate* categorizer from the 15*220 downloaded images.

as the target palette. The latter is obtained from the CL database which contains palettes of five swatches each. For each pixel x, we compute the occupancy probabilities $\gamma_i(x)$ which indicate the likelihood that the specific pixel was generated by Gaussian i.

3. A *target palette* pal_t, modeled by a GMM, is obtained by learning its parameters using palettes which received high scores for the specific target concept. The learning is performed using the method above.

4. The transformation between pal_{in} and the target palette pal_t is computed as in Reinhard *et al.*'s method for color transfer [14]. The colors of pal_{in} are ordered by decreasing weight of the corresponding Gaussian functions. Before performing the color transfer, we need to find a linear transform that matches the statistics of pal_{in} to those of pal_t. Since our transfer method takes into account only the means of the GMM's, we can write: $b_i = pal_t(i) - \mu_i$, where i denotes the color swatch in the target palette p_t as well as the index of the corresponding Gaussian in the GMM model. We then compute the color transfer map $b(x)$ as such: $b(x) = \sum_{i=1}^{5} \gamma_i(x) b_i$. Finally the output image is computed as follows: $I_{out}(x) = I_{in}(x) + b(x)$.

In figure 10 we show color transfer performed on two images, with "sensual" as the target concept. Other examples are provided in the additional material.

9. CONCLUSIONS

In this paper, we showed that it is possible to classify color palettes using challenging emotion and mood categories, provided that reliably annotated data is available (e.g. CC dataset). Since this is not the case, especially with datasets collected on the web, we also investigated methods to clean "weakly annotated" datasets. Several palettes representations were tested in different frameworks ("late fusion", dataset cross-evaluation, etc.) and we deduced that the most performing features are based on the Fisher Vector

Figure 10: From left to right and top to bottom: two input images, target palette ("sensual"), and color transfer results.

Figure 9: Top retrieved images from the downloaded GoogleImages with *playful colors* and top images of our *playful* categorizer from the 15*220 downloaded images.

and on the Bag of Colors approaches. Finally, we showed how the palettes classifiers can be used to retrieve palettes, images or for color transfer. The development of new methods for reducing the annotation noise in the current datasets as well as the design of new palettes features will be further investigated in the future.

10. ACKNOWLEDGMENTS

This work was supported by the Agence Nationale de la Recherche through the OMNIA project (ANR-07-MDCO-009-02).

11. REFERENCES

[1] R. Benavente, M. Vanrell, and R. Baldrich. Parametric fuzzy sets for automatic color naming. *Journal of the Optical Society of America A*, 25(10):2582–2593, October 2008.

[2] B. Berlin and P. Kay. *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press, University of California, 1969.

[3] D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, and X. Ying-Qing. Color harmonization. In *Proceedings of ACM SIGGRAPH*, volume 25, pages 624–630, 2006.

[4] D. Conway. An experimental comparison of three natural language colour naming models. In *East-West International Conference on Human-Computer Interactions*, pages 328–339, 1992.

[5] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray. Visual categorization with bags of keypoints. In *ECCV Workshop on Statistical Learning for Computer Vision*, 2004.

[6] L. Eiseman. *Pantone Guide to Communicating with Color*. Graffix Press, Ltd., 2000.

[7] J. Farquhar, S. Szedmak, H. Meng, and J. Shawe-Taylor. Improving "bag-of-keypoints" image categorisation. Technical report, University of Southampton, 2005.

[8] http://www.colourlovers.com/.

[9] T. Jaakkola and D. Haussler. Exploiting generative models in discriminative classifiers. In *Advances in Neural Information Processing Systems 11*, 1999.

[10] J. M. G. Lammens. *A Computational Model of Color Perception and Color Naming*. PhD thesis, University of Buffalo, 1994.

[11] L. Ou, M. Luo, A.Woodcock, and A. Wright. A study of colour emotion and colour preference. part iii: Colour preference modeling. *COLOR research and application*, 29:381Ű–389, 2004.

[12] 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 And Application*, 29(3):232–240, June 2004.

[13] F. Perronnin, C. Dance, G. Csurka, and M. Bressan. Adapted vocabularies for generic visual categorization. In *ECCV*, 2006.

[14] E. Reinhard, M. Ashikhmin, B. Gooch, and P. Shirley. Color transfer between images. *IEEE Computer Graphics Applications*, 21(5):34–41, September/October 2001.

[15] M. Solli and R. Lenz. Color emotions for image classification and retrieval. In *CGIV*, 2008.

[16] Y. Tai, J. Jia, and C. Tang. Local color transfer via probabilistic segmentation by expectation-maximization. In *IEEE International Conference on Computer Vision and Pattern Recognition*, volume 1, pages 747–754, 2005.

[17] 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, July 2009.

[18] C.-K. Yang and L.-K. Peng. Automatic mood-transferring between color images. *IEEE Computer Graphics and Applications*, 28(2):52 – 61, March-April 2008.