Color-Mood Analysis of Films Based on Syntactic and Psychological Models

Cheng-Yu Wei
Electrical Engineering Dept.
Columbia University
nelson@ee.columbia.edu

Nevenka Dimitrova
Philips Research
Briarcliff Manor, NY 10510
nevenka.dimitrova@philips.com

Shih-Fu Chang
Electrical Engineering Dept.
Columbia University
sfchang@ee.columbia.edu

Abstract

The emergence of peer-to-peer networking and the increase of home PC storage capacity are necessitating efficient scaleable methods for video clustering, recommending and browsing. Based on film theories and psychological models, color-mood is an important factor affecting user emotional preferences. We propose a compact set of features for color-mood analysis and subgenre discrimination. We introduce two color representations for scenes and full films in order to extract the essential moods from the films: a global measure for the color palette and a discriminative measure for the transitions of the moods in the movie. We captured the dominant color ratio and the pace of the movie. Despite the simplicity and efficiency of the features, the classification accuracy was surprisingly good, about 80%, possibly thanks to the prevalence of the color-mood association in feature films.

1. Introduction

Film entertainment is one of many forms of storytelling. The art of filmmaking is, for example, to turn a news paragraph about a murder story into a 90-minute thriller. One important aspect in the production process has been the color composition in the mise en scene. To maintain mood consistency in a scene, filmmakers consider the colors in each scene, ranging from the colors of the costumes to the background colors (hues, saturation, and brightness). The Film Theory [9] advocates that color perception and lighting are important contributors to the mood tones the director wants to bring forward in any scene in addition to dialogues, storyline, actors, etc. Color can also be used to create particular surprising or artistic effects (e.g. Kieslowski's "Blue," "White" and "Red"). This is also supported in the psychology literature [2]. Our goal is to develop methods that can be used to approximate the moods of films using computable content features to facilitate high-level tools for organizing films in personal storage. Specifically, we propose an automated, color content-based system to analyze the color distribution and related feelings brought to viewers. We use the notion of family histogram to represent each shot. We derive a movie palette histogram and a mood dynamics histogram to represent the dominant colors and mood transitions in the movie from all family histograms.

We studied two methods for using this information for mood and genre classification. The first method is to deterministically map the final histograms of movie palette and mood transitions into a mood type and film genres corresponding to the global dominant colors. The second method is to use a sliding window over groups of shots in order to produce features and statistically determine the correct movie classification. In this process we classify observed low-level feature sequences into high-level mood tones. And then, c-SVC Support Vector Machine [8] is adopted to classify the multidimensional feature vectors and to detect mood types. The experiment is conducted in two directions: (i) discriminatively map the color-related features, representing the whole film, to a fine-genre and a primary mood typology. (ii) Use the c-SVC SVM to classify the feature vector of windows accompanying the labels of emotional tones.

Existing work in video retrieval has explored affective content analysis [10]. For example, Motion Activity, Sound Energy, Density of Shot Change, as well as the Color features are used for affective content analysis [7]. The experimental results in [4] reported that quantized color, motion and shot cut rate provide high classification rate for affective events such as fear, sadness, joy and normal. To the best of our knowledge, this is the first system that explores formal, extensive psychophysical models of color-mood association and applies them to automatic film genre classification at a finer level.

2. Mood Analysis

In the case of films, the rich representations of film plot and characters' feelings unfolded on-screen all contribute to the affective "landscape" that is presented to the viewer, and they all help the viewer to interpret the mood tone of a scene. In psychological and cinematographic aspects of Color-Mood analysis, we have found different studies about the association of color and mood [2][3]. Mood and emotion are different concepts (although, of course related). Emotions are aroused by specific objects or events (and usually only for a relative short period of time), whereas mood has no such specific referent (and usually lasts for a relatively longer period of time). The American psychologist Mahnke studied the associations of emotion terms with colors. He compiled the results from multiple studies [3]. Color associations are not necessarily made with only a single emotion term. For example, red is not the only color associated with love, but the term is associated with red-violet. We should note here, that in the presented studies there is a varying level of agreement among the subjects. Based on the information in above literature, we have established our association of

corresponding colors to mood tones (emotions), which is summarized in table 1.

Colors	Associated Mood Tones (Emotion Terms)
Black	Hatred, Mourning, Sorrow, Indefinite
White	Mourning, Grief, Depression
Red	Love, Hatred, Life, Noble
Orange	Jovial, Happy
Yellow	Happy, Luminous, Jovial
Green	Tranquility, Peace, Life
Blue	Peace, Tranquility, Noble
Purple	Love, Noble, Authoritative

Table 1. Color-Mood Tones Associations

3. System Overview

Figure 1 is a system overview for color-mood analysis. First, we decode each MPEG-1 input into a series of image matrices in a specifically chosen color space, CIELUV [5]. CIELUV has the advantage of Perceptual Uniformity, i.e. the same distance between two different points makes equal perceived color difference. Utilizing the color histograms, we segment these videos into shots, select key-frames and compute "Family Histogram" from key-frames [1].
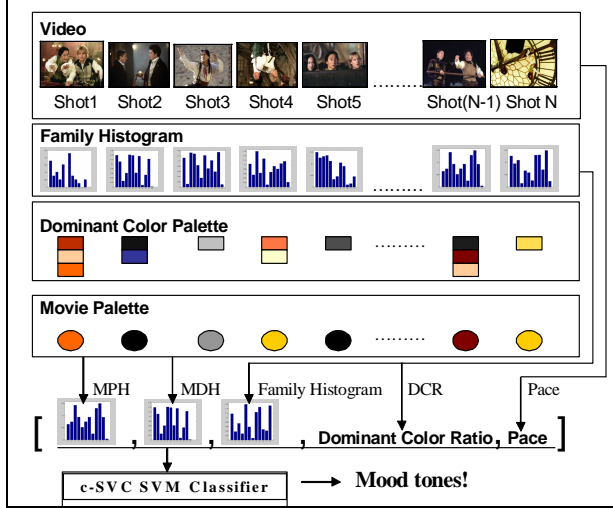
Figure 1. Color Transformation into Mood

For each Family Histogram, we extract the most dominant 1-3 color(s) corresponding to the largest 1-3 bin value(s), and the weighted sum of the dominant color(s) represents one "Representative Dominant Color (RDC)" for each shot. For a whole movie, we compute the "Movie Palette (MP)" by comparing the colors in "RDC Sequence" against the colors in our pre-established database and identifying the closest matching colors. Using the MP statistics we derive the "Movie Palette Histogram (MPH)" as we will explain in section 4.1. From the statistics of the color-transitions between consecutive shots in MP, we derive the "Mood Dynamics Histogram (MDH)" as we will explain in section 4.2. In addition, we extract the dominant color ratio, pace, and family histogram as we will explain in section 4.3. These five features are used for Support Vector Machine classification [8] of shots into moods as explained in section 4.4.

4. Color Features and Representation

Here we present a novel color representation and the chosen features for efficient color-mood clustering. We use this representation to achieve the following:

(i) Derive a global representation, MPH, in order to find the mapping to major mood tones. In addition we use another global representation, MDH, to gauge the dynamics of the color change in the movie. These measures are purely based on color and we investigate how much they can influence the association of mood tones and genre to a particular movie. We employed the approach from "Color Psychology" in [3], and extracted dominant bins from computed MPH and MDH for the Color-Mood mappings.

(ii) Mood classification according to the color characterization and pace features using SVM. The observation vector contains 422 values from the features:
- *Movie Palette Histogram (MPH)*, consisting of the closest colors in database - 12 values
- *Mood Dynamics Histogram (MDH)* - 144 values
- *Dominant Color Ratio (DCR)* - 1 value
- *Pace,* # of shots over total # of frames – 1 value
- *Family Histogram* - 264 values

We established a number of mood types (as shown in table 2), and each video belongs to one of these. We map the whole movie to a certain mood type and also classify the primary emotions for scenes. The c-SVC SVM module with 3 different types of kernels is established for training the hyper-planes and classifying the associated moods.

4.1. Movie Palette Histogram

We gather color statistics in MP for all shots, to obtain Movie Palette Histogram (MPH). We start by computing a Family Histogram $H_k(i)$ of shot k:

$$H_k(i) = \sum_{x,y} \begin{cases} 1, if\ Pix\,(x,y) = i \\ 0, if\ Pix\,(x,y) \neq i \end{cases}$$, where $Pix(x,y)$ is the pixel

value of position (x,y); i is a bin number in CIELUV color space, i=1~264; shot number k=1~M, where M is the number of shots within a movie. Next, we compute Dominant Color Palette: $DCP(k) = [P_1^k, P_2^k, P_3^k]$, for k=1~M, where P_1^k, P_2^k, P_3^k are the 3 bin numbers of corresponding top 3 most dominant colors in histogram H_k.

$$P_1^k = \arg\ \{ \max_{i=1\sim264} [H_k(i)]\}$$

$$P_2^k = \begin{cases} P_2^k, if\ H_k(P_2^k) > [H_k(P_1^k) \times Th\,] \\ 0, otherwise \end{cases}$$

$$P_3^k = \begin{cases} P_3^k, if\ H_k(P_3^k) > [H_k(P_1^k) \times Th\,] \\ 0, otherwise \end{cases}$$

the dominant color bin counts are as follows:

$$D_1^k = H_k(P_1^k)$$ where Th : 0~1,

$$D_2^k = \begin{cases} H_k(P_2^k), if\ H_k(P_2^k) > [H_k(P_1^k) \times Th\,] \\ 0, otherwise \end{cases}$$

$$D_3^k = \begin{cases} H_k(P_3^k), if\ H_k(P_3^k) > [H_k(P_1^k) \times Th\,] \\ 0, otherwise \end{cases}$$

a fixed threshold value. Based on the dominant colors we compute Representative Dominant Color Sequence *RDCS(k)* :

$$RDCS(k) = \sum_{l=1}^{3} P_l^k \times W_l^k \,,\ W_l^k = \frac{D_l^k}{D_1^k + D_2^k + D_3^k},$$

for k=1~M, and l=1~3.

Next, Movie Palette (MP), consists of a sequence of the reference colors closest to those in RDC, for all shots. We have established a database consisting of reference colors R(m), m=1~12, which uniformly divide the CIELUV color space. Movie Palette *MP(k)* is computed as:

$$MP(k) = \arg_{R(m)}\{\min_{m=1\sim 12}[distance(RDCS(k), R(m))]\}, \text{ for } k=1\sim M.$$

We can understand the frequency of appearance and the distribution of those MP colors by introducing Movie Palette Histogram *MPH(m)* as follows:

$$MPH(m) = \sum_{k=1}^{M} \begin{cases} 1, & if\ MP(k) = R(m) \\ 0, & if\ MP(k) \neq R(m) \end{cases}, \text{ for } m=1\sim 12$$

The MP may reveal several moods that the filmmakers want to bring to audiences, while the MPH can display the major one. The bins of MPH correspond to the pre-determined number of reference colors in the database (see figure 2). The largest few bins correspond to the major mood of that film. A large number of reference colors may lead to too "sparse" MPH, which would lose their representative characteristics for the main moods inside. We have tried using the numbers of 70, 50, 12, and 8. We found that with 12 bins we achieve higher accuracy in mood type classification.

4.2. Mood Dynamics Histogram

Filmmakers compose not only color-distribution in a scene, but also the transitions between scenes, to affect human feelings. The transitions of colors between shots in the MP may lead to mood dynamics [3]. We acquire the Mood Dynamics Histogram (MDH) from the statistics of color transitions in MP. *MDH((m1-1)*M+m2) =*

$$\sum_{k=2}^{M} \begin{cases} \frac{1}{M}, & if\ [MP(k-1)=R(m_1)]\ \&\ [MP(k)=R(m_2)]\ \&\ [m_1=m_2] \\ 1, & if\ [MP(k-1)=R(m_1)]\ \&\ [MP(k)=R(m_2)]\ \&\ [m_1 \neq m_2] \\ 0, & otherwise \end{cases},$$

for m1=1~M, m2=1~M, where M is the number of shots. Using above equations, we put more emphasis on the transition from reference color R(m1) to R(m2), based on the assumption of that transition from a certain color to the same color itself, is not really a transition. Examples of MPH and MDH are given in Figure 3.

4.3. Additional Color and Pace Features

In addition to MPH and MDH, other features such as the pace, dominant color ratio, family histogram, can help determine the finer category of mood tones, as well as their corresponding genres. They are defined as:

- Dominant Color Ratio: *DCR = |P_d| / |P|*, where P_d is the set of dominant color pixels, and P is the set of all pixels in a frame [4]. Shots that have fewer dominant colors are regarded as having higher DCR. We set 3 possible categories for DCR – high(3), medium(2), and low(1).
- Pace: Pace is defined as the ratio of the number of shots over the total number of frames. Shown in Figure 2.
- Family Histogram: merged histogram of shot's keyframes.

Film	RLR	RP	MR
Pace/DCR	4.9P/	6.0P/	5.6P/

Figure 2. Features: Pace and DCR (P: 10⁻⁴)

4.4. Classification

Emotions are to a great extent unconscious processes which cannot be classified to show a logical plan that links various parts together. However, psychologist Plutchik has identifies eight behavioral patterns as the foundation of eight primary mood types, shown in Table 2. Many emotion terms are associated with those mood types [3][6], which are the primary moods used in our ground truth labeling. Everyday we generally experience mixed emotion terms, and not always pure primary mood types in their various intensities. It is interesting that he also devised an emotional wheel in [3] in the form of an eight-part color wheel. He let people pick one or more colors from the eight predetermined colors to some emotions without thinking about specific objects. The experiment is all about the association people make with the concept or idea of the terms shown in Table 1.

Primary Mood Types	Example Emotion Terms
(1)Anger	Rage, wrath, annoyance, irritation, violence
(2)Fear	Terror, apprehension, timidity, nerve, tense
(3)Joy	Ecstasy, rapture, fun, happiness, pleasure, sex, euphoria, excitement, pride, optimism, hope
(4)Sorrow	Anguish, distress, grief, dejection, pensiveness, sadness, somber, loneliness, desolation
(5)Acceptance	Love, affection, tolerance, adoration, sympathy
(6)Rejection	Dislike, loathing, tiresomeness, aversion, boredom, remorse, confusion
(7)Surprise	Amazement, shock, astonishment, strangeness
(8)Expectancy	Anticipation, interest, attentiveness, calm

Table 2. Selected Emotion Tokens for 8 Mood Types

For a fixed-length sliding window (e.g. 6 shots, moved by 1 shot), MPH, MDH, and Pace, are extracted from the shots within the window, while other shot-level features, i.e. Family Histogram and DCR, are represented by the mean of those values in the window. According to Table 2, we mapped the emotion terms to associated established mood types. Next, the mood type of each window is determined based on the types of shots within the window. In training, the Sample Matrix of size MxN (M is the feature vector dimension, N is the number of training windows) is fed into SVM. In testing, the input matrices contain the feature vectors of the rest windows and their class labels to generate predicted class labels and the accuracy of classification.

5. Experiments

For evaluation of our proposed color-mood approach we analyzed 15 full length feature films of different genres. We extracted all the needed features from sliding windows and whole films, as described in Sec. 4.1-4.3. Eight mood types are also established for all emotion terms shown in Table 2, forming the class labels.

5.1. Film Level: Similarity Comparison

We looked up the genres of movies from the Internet Movie Database (http://us.imdb.com/). The associated mood types of each movie are mapped from MP/MPH according to the associations in Table 1. In Figure 3, we list titles, genres, MPH, and MDH after our experiments (due to lack of space, we only show 5 films out of 15.) We noticed that movies of the same genre have similar MPHs, e.g. *FF* and *SK*, *RLR* and *BN*, while those of different genres have much dissimilar

ones, e.g. *HS* and *RLR*, *HS* and *FF*. We observed dominant mood transitions happen at the largest values in MDH. Although *FF* and *SK* have similar MPHs, but much different MDHs, for *FF* is a thrilled/action movie while *SK* is an action/adventure comedy. However, *BN* and *RLR* have both similar MPHs and similar MDHs, because they are of more similar genres: drama/Avant-garde, shown in Table 3. So the combination of MPH and MDH can help in distinguishing subgenres for movies with common genre.
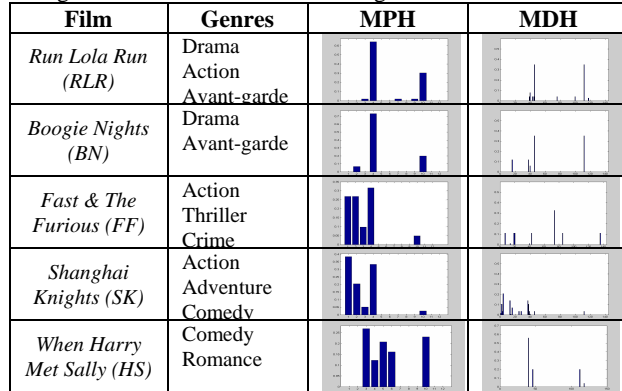
Film	Genres	MPH	MDH
Run Lola Run (RLR)	Drama Action Avant-garde		
Boogie Nights (BN)	Drama Avant-garde		
Fast & The Furious (FF)	Action Thriller Crime		
Shanghai Knights (SK)	Action Adventure Comedy		
When Harry Met Sally (HS)	Comedy Romance		

Figure 3. MPHs and MDHs of Five Films

Similarity=1–Dissimilarity, where *Dissimilarity* is the L-2 difference between 2 histograms of the same type.

Films	Common Genres	Similarity of MPHs	Similarity of MDHs
RLR & BN	Drama, Avant-garde	85.1%	83.8%
RLR & HS	N/A	62.4%	39.1%
FF & SK	Action	74.4%	47.1%
FF & HS	N/A	62.5%	35.9%

Table 3. Similarity Comparison between Genres

5.2. Window Level: Content Classification

To establish the ground truth, the emotions of each shot are manually labeled by a person who had no idea about the project goal. There was a list of 16 emotion terms to describe each shot. He can also include an additional emotion description. Next, each window consisting of W shots (W=6 in our experiments) is mapped into a mood type according to Table 2. We designate a mood type as the representative of the window if it exhibits prolonged, continuous shots of that type. If the window contains emotions of W/2 (3 in our experiments) or more mood types, we designate this window as "mood type transition." We adopted c-SVC SVM for mood classification and determination [8]. We have experimented with shrinking heuristics and kernel functions of "linear," "polynomial," and "RBF." We adopted "polynomial" kernel for its leading to best classification accuracy. We performed two experiments: (i) Training and testing windows are from a single film, and (ii) Cross-Training using Leave-One-Out strategy.

(i) For a certain film, we use the odd/even number windows as training samples, while the even/odd number (or all) windows of the same film as test samples. We found the accuracy always higher than that with the complementary half windows tested only. The accuracy for films "Minority Report" (*MR*), "Wall Street" (*WS*), "Run Lola Run" (*RLR*), and "The Lord of the Rings" (*LR1*) are shown in Table 4.

Train/Test Film	Even/Odd Accuracy	Even/All Accuracy	Odd/Even Accuracy	Odd/All Accuracy
MR	73.2%	78.2%	69.3%	75.5%
WS	68.7%	71.2%	74.3%	77.8%
RLR	75.8%	82.4%	77.1%	80.3%
LR1	73.7%	77.3%	69.8%	76.5%

Table 4. Single Movie Mood Classification Accuracy

(ii) Cross-Training (Leave-one-out): we took windows of all films as the training samples excluding the film whose windows are test samples. The result is shown in Table 5. In this case, the training set contains more diverse categories, higher number of training data, and therefore leads to higher classification accuracy than single movie classification, except for *RLR*. We can speculate that for an Avant-garde movie like *RLR*, directors compose colors in a more novel way. Different color-combination may be used, with special associations between color features and mood types. Thus, its accuracy is lower than that of previous experiment.

Test Film	MR	WS	RLR	LR1
Classification Accuracy	81.5%	84.6%	78.8%	80.0%

Table 5. Mood Classification Accuracy w/ Cross-Training

6. Conclusions

In this paper we investigated color characterization of movies at global (full) and local (a group of shots) level for mood analysis. We introduced two global measures reflecting color palette and mood dynamics. In our extensive experiments on 15 full-length films, we saw that movie dynamics is a very discriminative measure for characterizing mood tones. We also captured the dominant color ratio and the pace of the movie. With these measures we applied a statistical classification approach to classify the windows of shots into mood types reflecting the nuances in the features. Our experiments with color mood association at a window (group of shots) level show about 80% accuracy for mood tone classification.

In this study we focused on color as a primary source for color-mood association. In our future work we plan to include aural elements for mood classification.

7. References

[1] N. Dimitrova, J. Martino, H. Elenbaas. L. Agnihotri, "Color SuperHistograms for Video Representation," *IEEE ICIP, 1999*.

[2] L. B. Wexner, "The degree to which colors (hues) are associated with mood-tones," *J. Applied Psychology, 38, 1954*. pp. 432-435.

[3] F.H. Mahnke, "Color, Environmental and Human Response" Ch.3, 4, *Van Nostrand Reinhold*, NY, 1996

[4] H.B. Kang, "Affective Content Detection using HMMs," *Proc. of the 11th ACM Multimedia 2003*, Berkeley, CA, pp.259-262

[5] D. Park, J. Park, J.H. Han, "Image Indexing Using Color Histogram in the CIELUV Color Space," *Proc. 5th Japan-Korea Wksp. Computer Vision, 1999*, Korea, pp.126-132

[6] A. Salway and M. Graham, "Extracting Information about Emotions in Films," *ACM Multimedia 2003*, CA

[7] A. Hanjalic, L.-Q. Xu, "User-oriented Affective Video Content Analysis," *Proc. IEEE CBAIBL'01*, Kauai, HI, pp.50-57, 2001.

[8] Christopher J. C Burges, "A Tutorial on Support Vector Machines for Pattern Recognition," *Journal of Data Mining and Knowledge Discovery*, 1998, pp. 121-167

[9] H. Zettl. "Sight Sound Motion: Applied Media Aesthetics." *Wadsworth Publishing, 3rd edition, 1999.*

[10] Truong, B.T., Venkatesh, S., Dorai,C., "Application of Computational Media Aesthetics Methodology to Extracting Color Semantics in Film," *ACM Multimedia, France Dec, 2002*.