IMAGE WARMNESS — A NEW PERCEPTUAL FEATURE FOR IMAGES AND VIDEOS

Michail Dimopoulos, Thomas Winkler

Fraunhofer IAIS
Schloss Birlinghoven
Sankt Augustin, Germany

ABSTRACT

Many basic, but very useful features, for characterizing an image or calculating the similarity between two images are based on color information. Psychological studies show that beyond the tone of color, different colors are also associated with different emotions. Thus, two colors that trigger the same impression are most likely considered to be more similar than two colors which trigger the opposite impression. We introduce a new feature called *image warmness*. It is based on the *cold* or *warm* impression that a single color triggers in the brain of the beholder. Image warmness provides a measure about how cold or warm an entire image is perceived by humans based on colors it contains. In a survey and evaluation with 90 images and 101 participants we show that the values of image warmness calculated by the proposed formula are close to the average rating of the survey participants.

***Index Terms*—** image warmness, feature extraction, image similarity, human perception

1. INTRODUCTION

Image, as well as video similarity estimation, needs various features suited to characterize aspects of similarity adequately. Major attempts have been made in the field of research of image and video similarity during the last years towards combining visual, audio, and textual features into one integrated framework for video similarity which, for example, can be used for video indexing and retrieval. According to [1] and [2] features used in video indexing are mainly color, texture and motion. Typical examples of such features are histograms, color moments, contrast, orientation histograms, edge histograms and motion vector fields. In [3] the authors present an online video advertising system called VideoSense. Similarity matching between content and advertisements is mostly based on dominant color HSV histograms and motion intensity.

Aforementioned approaches use established visual content features that are already proven to work well. Such features though do not have a direct focus in higher level human perception. Researchers have focused their interest in the perceptual direction only in the last few years. There have been some proposals on perceptual color features, for example in [4, 5], but mostly with a small scale evaluation. In this paper, we introduce a perceptual visual feature which we call **image warmness**. It enables the measurement of perceived cold or warm impression triggered by colors of an image and we believe it could be interesting as an additional feature for video and image similarity estimation. In particular, it should be interesting for applications of the creative industries, as well as online advertising, for example automatically placing an advertisement in a cold or warm context to stress important characteristics of a product like a cold and refreshing soft drink.

2. MOTIVATION

Several studies in the field of psychology show that human perception is not limited to distinguishing between different colors, but also that these colors provoke certain emotions in people. For example, red color is associated with danger, failure or (sexual) passion depending on the context [6]. In [7] a more detailed model of color and psychological functioning is proposed. The authors argue that color information is paired to impressions from everyday life and that humans are often unaware of the influence of color on them.

Thus, color influences the way that humans perceive visual similarity in several aspects, as also shown in [8]. Effects of cold and warm colors, as an aspect of human color perception, have been mostly studied by psychologists. They split visible colors into two groups, one of warm and one of cold colors respectively based on their impact on people. In this paper we attempt to capture this psychological aspect in a concrete feature that we call image warmness.

The color circle is divided into two parts, warm colors (red and yellow regions) and cold colors (blue and green regions). The exact angles separating warm and cold colors vary slightly in literature. They are found between yellow and green (around $75°$ in HSV color space) and blue and red (around $285°$). A cold or warm feeling is considered to be a *binary* characteristic triggered by a single color. In case of an entire image additional steps are required for calculating the influence of each pixel or color bin on the overall perceived warmness. In addition to *hue*, which determines if a color is warm or cold, *saturation* and *brightness* of each pixel must be considered as they influence the impact each pixel of an image has on the viewer.

3. ALGORITHMIC APPROACH

Calculation of image warmness is based on *quantized colors* of an image, determined in the preprocessing step. As a first step, *color warmness* of each of the quantized colors is calculated in HSV color space by a binary assignment of cold or warm and a weighting by saturation and value. Image warmness is then defined as the weighted average over color warmness values of all color bins. The algorithm is described in detail in this section.

3.1. Preprocessing

We assume, that human perception does not differentiate between all possible color tones, but focuses on the dominant colors instead. Thus, we reduce the large number of colors to N dominant ones by means of color quantization in the first preprocessing step. We use the median cut algorithm for color quantization [9]. Next, we calculate the histogram $\mathcal{H}$ and look up table (*LUT*) of the quantized image. For image warmness calculation we finally transform RGB primary values of colors in the LUT into their corresponding HSV values.

3.2. Binary Assignment

Based on the hue, H_n, of color bin n we define a binary assignment function $T_n(H_n)$, which reflects if the color bin is considered to be warm or cold according to color theory:

$$T_n(H_n) = \begin{cases} -1 & \text{, if } 75^\circ < H_n < 285^\circ \\ +1 & \text{, if } 0^\circ \leq H_n \leq 75^\circ \text{ or } 285^\circ \leq H_n \leq 360^\circ \end{cases} \tag{1}$$

We assign a negative weight $T_n(H_n) = -1$ for cold colors ($75^\circ < H_n < 285^\circ$) and a positive one $T_n(H_n) = +1$ for warm colors ($0^\circ \leq H_n \leq 75^\circ$ or $285^\circ \leq H_n \leq 360^\circ$).

3.3. Weighting by Saturation and Value

Furthermore, we define a weighting function $w_n(S_n, V_n)$ which quantifies the level of impact of each color bin based on saturation and value. Three different weighting functions $w_n^{(1)}$, $w_n^{(2)}$ and $w_n^{(3)}$ are introduced and evaluated.

Weight $w_n^{(1)}$ is calculated straight forward as the **product** of saturation, S_n, and value (brightness), V_n:

$$w_n^{(1)}(S_n, V_n) = S_n\, V_n, \forall S_n, V_n \in [0, 1] \tag{2}$$

With both saturation and value between 0 and 1, weight $w_n^{(1)}$ is also in the range of $[0, 1]$. A very bright and very saturated color gives a strong impression of warmness or coldness, weight $w_n^{(1)}$ is approaching 1. When the color is dark or less saturated there is little to no warm/cold impression, $w_n^{(1)}$ is approaching neutrality, 0.

The second weighting function $w_n^{(2)}$ considers a weighting related to **Euclidean distance**. With a maximum weighting of 1 for $S_n = 1$ and $V_n = 1$ and euclidean distance $r_n(S_n, V_n) = \sqrt{(1 - S_n)^2 + (1 - V_n)^2}$ for any S_n, V_n to this maximum, the weight is calculated in the following way:

$$w_n^{(2)}(S_n, V_n) = \begin{cases} 1 - r_n(S_n, V_n) & \text{, if } r_n(S_n, V_n) \leq 1 \\ 0 & \text{, if } r_n(S_n, V_n) > 1 \end{cases} \tag{3}$$

The point of maximum perceptual impact at $S_n = 1$, $V_n = 1$ has a weight of $w_n^{(2)}(1, 1) = 1$, all points in the SV plane with a euclidean distance of 1 or more from this point have a weight of 0 and are considered neutral. The cutoff area with euclidean distance of 1 or more is a rough approximation of a region with no major apparent influence on perceived warmness.

The third weighting function, $w_n^{(3)}$ is based on the **harmonic mean** of saturation and value:

$$w_n^{(3)}(S_n, V_n) = \begin{cases} 2\, S_n\, V_n / (S_n + V_n) & \text{, if } S_n \neq 0 \text{ or } V_n \neq 0 \\ 0 & \text{, if } S_n = 0 \text{ and } V_n = 0 \end{cases} \tag{4}$$

The weight, following the harmonic mean, is always between both values S_n and V_n but tends to give more importance to the lower of the two. Weight $w_n^{(3)}$, compared to weight $w_n^{(1)}$ which uses the product of both values, provides higher weights for low saturation or value.

3.4. Calculating Color Warmness

Color warmness, θ_n, is then calculated using one of the weighting functions (Equations 2–4) to quantify the binary warm/cold impression T_n determined by the hue, H_n, of bin n:

$$\theta_n = T_n(H_n)\, w_n(S_n, V_n) \tag{5}$$

Color warmness θ_n provides values in the range from $[-1, 0]$ for cold colors and $[0, 1]$ for warm colors.

3.5. Calculating Image Warmness

Image warmness, Θ, is calculated from the color warmness values θ_n of all color bins of the color-quantized image. Based on a histogram $\mathcal{H}_N$ of all N quantized colors we derive the relative frequency $f_n = \frac{c_n}{I}$, where c_n is the number of pixels with color n and I is the number of all pixels in the image. The image warmness Θ is then calculated from the color warmness values θ_n weighted by the relative frequency f_n of the bin:

$$\Theta = \sum_{n=1}^{N} f_n \theta_n \tag{6}$$

Image warmness Θ is a single value in the range $[-1, 1]$ with a negative value for cold colored images and a positive value for warm colored images as well as neutral warmness (neither cold nor warm) at 0. As such, we can easily express the warmness of an image and compare the similarity of two images.

4. USER SURVEY

We carefully prepared and conducted a questionnaire survey to gather ground truth data of perceived image warmness from human participants.

4.1. Preparation

Image warmness as introduced in the previous section has values in the range $[-1, 1]$. For a simple and fast rating during our survey we introduce a modified discrete scale with 7 possible values for image warmness from -3 (cold) to $+3$ (warm) in steps of 1 and with neutrality at 0. This scale can easily be mapped to the calculated range of $[-1, 1]$ during evaluation. In total, *90 images* were selected from the MIR-FLICKR '08 dataset [10]. The selection process was random.

We added 7 images (each roughly representing one of the 7 discrete warmness values) before the 90 images of the survey. These extra images were provided to the participants to tune their perception to the available range of warmness impressions and are not used in the calculation of the ground truth. The participants were not informed about the tuning process and were only asked to rate all 97 images. The only restrictive criterion explicitly mentioned was that participants should choose their answer according to the "feeling" of warmness or coldness that the colors in the images introduce to them, independent of the presented object or scenery each image depicts.

101 participants took part in the survey on a voluntary basis. Most of the participants are undergraduate or postgraduate students or researchers in the age group of 18–35 years, mainly with a technical studies background. A few have a background in human sciences and very few are in the 35–55 age group from various scientific backgrounds. Even though the group of people was not selected strictly as representative of the general population, the size of the survey with 101 participants and 90 images should allow us to draw sufficiently reliable conclusions concerning the proposed image warmness feature and weighting functions.

4.2. Analysis

Judging from some random samples of the survey data several aspects beyond color itself often seemed to influence the rating of the participants. Such influences are difficult to measure, but two aspects seemed to be particularly important: First, the object, scenery etc. depicted in an image may cause a particular cold or warm impression independent of the actual color. Second, the center of attention seems to be relevant; warm objects or elements in the foreground of an image, for example, seem to have a stronger impact on the felt warmness than a cold background making up most of the pixels of the image. Since warmness is based on the perception of each individual, it is impossible to eliminate such underlying subjective factors completely. For this reason a rather high variance in the participants ratings is expected, but we think that the average values of all participants' ratings provide a solid basis for evaluating the influence of the colors on the perceived warmness of an image.

The discrete scale from $[-3, 3]$ used for the survey is linearly scaled to $[-1, 1]$ to match the range of the calculated image warmness and the average value for all 101 participants is calculated for each of the images. The lowest average warmness value is -0.5578 and the highest 0.6304. The mean value μ is slightly positive but close to neutral at almost 0.05. The average variance per image is 0.1970 and varies between 0.1037 and 0.3778 indicating rather diverse impressions of warmness caused by some of the images and more similar impressions caused by others. In Figure 1 we can see two example images of the survey. On the left hand side the image with the highest agreement amongst the participants (lowest variance $\sigma^2 = 0.1037$ with $\mu = 0.3597$) is shown with its distribution of votes. It shows a Gaussian-like distribution of all participants' votes. On the right hand side we see the image with the highest deviations from the mean value of votes (highest variance $\sigma^2 = 0.3778$ with $\mu = 0$). Even though the mean value of all votes is almost exactly 0 hardly anyone rated the image to be neutral. We can see two distinct rating behaviors: The first group of people rated the image to be rather cold (mainly -1 and -2). Rescaled to the range of $[-1, 1]$ this average of these values approximately matches the calculated value of -0.3950 ($w_i^{(1)}$, $q = 256$) for

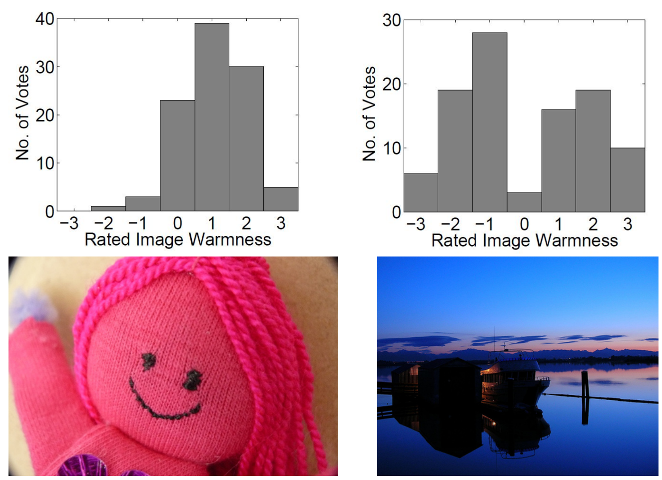

Fig. 1. Example images with their rating histograms from the survey. Left: image with lowest σ^2 (highest agreement); right: image with highest σ^2 (highest disagreement).

this image. The second group of people rated the image as rather warm (around 2). This second group probably considered other factors than the pure color information while rating the image. Such factors might be the calm and nice ("warm") scenery or a stronger focus on smaller areas of the image with, for example, warm colors of the sunset. While colors introduce a "feeling" of warm or cold, this example shows that other aspects of an image also influence our impression of warmness.

5. EVALUATION

The survey results averaged over all ratings for each image and linearly scaled to $[-1, 1]$ provide the ground truth for the evaluation of our image warmness feature and the different weighting options.

5.1. Image Warmness

As we try to match the human perception of warmness based on the color information in an image we compare the average rating of image warmness Θ_{sur} of the survey participants with the calculated values for image warmness $\Theta^{(1)}$, $\Theta^{(2)}$, and $\Theta^{(3)}$ (using the respective weighting approaches $w_n^{(1)}$ to $w_n^{(3)}$). We calculate the root mean squared error (RMSE) of the calculated image warmness to the ground truth of the survey. The RMSE for the calculated image warmness for some of the evaluated quantization levels and the different weighting functions are presented in Table 1. The lowest RMSE for all weighting functions is achieved for a quantization into $q = 256$ color bins. The first weighting function $w^{(1)}$ (weighting by the product of S and V) produces the lowest RMSE with 0.2263 at a quantization of $q = 256$. The RMSE for the second function is slightly higher with 0.2399 while the third function produces a significantly higher RMSE of 0.3055. Hence, we suggest to calculate image warmness with a quantization of $q = 256$ and the first weighting function $w^{(1)}$, which yields the best results in our evaluations. All following results are based on this setup.

Table 1. RMSE for different weighting functions and different levels of quantization q and without quantization $(-)$. Best results are marked.

q	$w^{(1)}$	$w^{(2)}$	$w^{(3)}$
8	0.2345	0.2500	0.3276
30	0.2307	0.2448	0.3147
100	0.2275	0.2413	0.3077
256	**0.2263**	**0.2399**	**0.3055**
$-$	0.2630	0.2749	0.3303

Disagreement amongst the participants also leads to a larger gap between calculated image warmness and average image warmness from the survey. In case of the 45 images with the highest agreement amongst all participants (lowest variance), the RMSE falls to 0.1776. For the 45 images with the highest variance it rises to 0.2662.

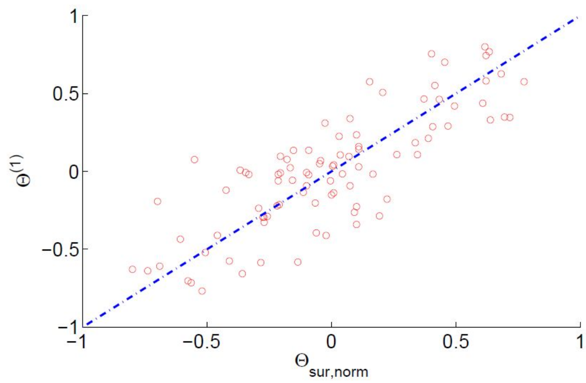

Fig. 2. Scatter plot for $\Theta^{(1)}$ dependent on the normalized average survey ratings $\Theta_{sur,norm}$.

In Figure 2 the calculated image warmness is plotted against the normalized average survey ratings. For normalization the average values of the survey are shifted by the difference in mean ($\Delta\mu = 0.0425$) and multiplied by the gain (the quotient of maximum absolute value of calculated and the shifted rated image warmness, $g = 0.7996/0.6037$). The normalization is applied to compensate for a reduced range and gradient of the survey results due to the averaging process over all participants. Thus, the plot mainly emphasizes the correlation between calculated values and survey results. The trend of the scatter plot almost exactly follows the ideal line. The calculated image warmness is generally able to predict the human rating. The variance is not much higher than the variance amongst human ratings.

5.2. Discussion of Outliers

Figure 3 shows some examples where calculated warmness and rated warmness differ significantly. The first image (upper left) contains mainly cold colored pixels but is rated slightly warm by the participants. It shows flowers with neutral to warm colors in the foreground which might influence the participants to vote towards a slightly warm impression. The image at the bottom left shows a jellyfish with warm colors in an environment (background) with mainly cold colors. The jellyfish might drag the attention of the participants, so that the comparatively small amount of warm colored pixels might be given a higher relevance by the viewer. The images on the right hand side picture objects with a mainly negative connotation that might introduce a colder impression to the viewer than the pure color information indicates. The presented images support the assumption, that the warmness impression of the image on a human being seems to be in-

Fig. 3. Some examples of images with large absolute distance in warmness between ground truth and calculated image warmness.

fluenced by additional factors like semantic information or center of attention in addition to the pure color information.

6. CONCLUSIONS

In this paper we introduced image warmness as a new feature for measuring perceived warmness of images. In a survey with 101 participants we derived the ground truth of the perceived warmness impression of 90 different images. In our evaluation we could show that our new feature has a high correlation with the average impression of the participants and is suitable to estimate the perceived image warmness. This could be useful, for example, to automatically place an advertisement of a cold refreshment in a contrasting warm or in a matching cold environment to trigger a particular emotional reaction.

Major differences between calculated image warmness and average perceived image warmness seem to occur mainly in two cases: First, images with foreground objects of a warm color in an image with mainly cold colors were often rated warmer than calculated. Second, associations on a semantic level related to the image scenery or object seem to influence the felt warmness in addition to the pure color information. Thus, a stronger consideration of the center of interest might improve the calculated values for image warmness.

7. ACKNOWLEDGMENTS

The authors acknowledge the great support of all people who participated in the image warmness survey.

REFERENCES

[1] Weiming Hu, Nianhua Xie, Li Li, Xianglin Zeng, and Stephen J. Maybank, "A survey on visual content-based video indexing and retrieval," *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, vol. 41, no. 6, pp. 797–819, November 2011.

[2] Darshak G. Thakore and A. E. Trivedi, "Content based image retrieval techniques - issues, analysis and the state of the art," in *Proceedings of International Symposium on Computer Engineering and Technology 2010*, 2010.

[3] Tao Mei, Xian-Sheng Hua, and Shipeng Li, "Videosense: A contextual in-video advertising system," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 12, pp. 1866–1879, December 2009.

[4] Marcal Rusinol, Farshad Noorbakhsh, Dimosthenis Karatzas, Ernest Valveny, and Josep Llados, "Perceptual image retrieval by adding color information to the shape context descriptor," in *Proceedings of the 2010 20th International Conference on Pattern Recognition*, Washington, DC, USA, 2010, ICPR '10, pp. 1594–1597, IEEE Computer Society.

[5] Serkan Kiranyaz, Murat Birinci, and Moncef Gabbouj, "Perceptual color descriptor based on spatial distribution: A top-down approach," *Image Vision Comput.*, vol. 28, no. 8, pp. 1309–1326, Aug. 2010.

[6] Arlen C. Moller, Andrew J. Elliot, and Markus A. Maier, "Basic hue-meaning associations," *Emotion*, vol. 9, pp. 898–902, December 2009.

[7] A.J. Elliot and M.A. Maier, "Color and psychological functioning," *Current Directions in Psychological Science*, vol. 16, pp. 250–254, 2007.

[8] Bernice E. Rogowitz, Thomas Frese, John R. Smith, Charles A. Bouman, and Edward Kalin, "Perceptual image similarity experiments," in *Human Vision and Electronic Imaging III, Proceedings of the SPIE, 3299*, 1998.

[9] Paul Heckbert, "Color image quantization for frame buffer display," in *Proceedings of the 9th annual conference on Computer graphics and interactive techniques*, New York, NY, USA, 1982, SIGGRAPH '82, pp. 297–307, ACM.

[10] Mark J. Huiskes and Michael S. Lew, "The mir flickr retrieval evaluation," in *MIR '08: Proceedings of the 2008 ACM International Conference on Multimedia Information Retrieval*, New York, NY, USA, 2008, ACM.