

Saliency-Guided Consistent Color Harmonization

Yoann Baveye, Fabrice Urban, Christel Chamaret, Vincent Demoulin,
and Pierre Hellier

Technicolor Research and Innovation, Rennes, France
{baveyey,urbanf,chamaretc,demoulinv,hellierp}@technicolor.com
<https://research.technicolor.com/rennes/>

Abstract. The focus of this paper is automatic color harmonization, which amounts to re-coloring an image so that the obtained color palette is more harmonious for human observers. The proposed automatic algorithm builds on the pioneering works described in [3,12] where templates of harmonious colors are defined on the hue wheel. We bring three contributions in this paper: first, saliency [9] is used to predict the most attractive visual areas and estimate a consistent harmonious template. Second, an efficient color segmentation algorithm, adapted from [4], is proposed to perform consistent color mapping. Third, a new mapping function substitutes usual color shifting method. Results show that the method limits the visual artifacts of state-of-the-art methods and leads to a visually consistent harmonization.

Keywords: color harmonization, color segmentation, color mapping, saliency, visual attention.

1 Introduction

Harmony within a picture is somehow subjective and related to a specific field of investigation. While artists mention shape symmetry, photographers refer to image composition and movie directors perform rather at the color level. Harmony is a large term whose definition is context-dependent. In this paper, color harmony of picture will be referred to the way how to automatically change the color palette in order to enhance the global "look", rendering or feeling. The main motivation is to improve user experience by objectively creating more harmony when nature has associated different objects with each other, having global color disharmony, as illustrated in Figure 1.

During the last decades, increasing focus has been devoted to the area of color harmony, viewed as the color interaction within a spatial framework. Many color scientists have performed experimentations to design a definition and some objective criteria to color harmony characterization. Color combinations for defining comprehensively harmonious doublets or triplets have been widely discussed in the literature [13,15,8,18] especially associated to a range of moods and adjectives [20].

The color enhancement processing based on harmony criteria has been introduced recently by Cohen *et al.* [3] and named *color harmonization*. It is an automatic post-processing that may help color scientists or may completely guide novice for rendering image "look". It is a stand-alone application, but may also be a feature when learning and predicting aesthetic assessment [10] or emotion extraction [11]. The proposed algorithm builds on the pioneering work of Cohen-Or *et al.* [3] and aims at limiting its artifacts and limitations.

This paper proposes three main contributions. First, a new cost function is designed to choose harmonic colors that match original image. Saliency maps of a visual attention model [9] are used to select the most appropriate areas in the color harmonization process. It handles limitations of previous art regarding the full automation of such algorithm and more specifically the color mapping of critical areas such as skin, face or characters. Second, a new color segmentation algorithm is introduced and performances are discussed in the context of color mapping for color harmonization. Third, a color mapping function is proposed that better respects human perception.

The paper is organized as follows: in section 2, state-of-the-art techniques concerning color harmonization are presented and discussed. In section 3, the proposed method is described, including the template selection, the color segmentation technique and the color mapping functions. Finally, section 4 presents the results obtained and compares to state-of-the-art techniques.



Fig. 1. Example of color harmonization: (a) Original picture, (b) Harmonized picture using the technique described in this paper. In this example, the pink color of the hats that could be disruptive for some observers has been mapped to purple.

2 Related Work

In this section, previous work related to color harmony is presented and discussed. Generating harmonic colors has been an open problem among artists and scientists [5]. Modern color theory, which was developed at the beginning of the 20th century, deals mainly with representations of colors, but also discusses color harmony [7,1,14]. It is only in 1960 that scientists began to use a new kind

of color wheel with an emphasis on hue such as introduced by Itten [7]. Firstly, he took advantage of two complementary colors on the hue wheel to measure color harmony. Later, combinations of three and four-color harmony creating an equilateral triangle and square provided a reference for the assessment of harmony.

Pursuing the same idea, although expressed on HSV (hue H, saturation S and value V) color space, Matsuda [12] proposed height harmonic templates shown in Figure 2. The work presented in this paper relies on such template definition.

A predictive model of picture color harmony based on two-color harmonies of Ou *et al.* [15] have been developed by Solli *et al.* [19]. Psychophysical experiments confirmed the usefulness of their model as well as consistency in the human assessment of perceived multi-color harmony. Recently, new harmonic templates using unsupervised machine learning on the Adobe Kuler database have been proposed [16], consisting of hundreds of thousands color palettes used for creative purposes. However they are not precisely specified, since they have been obtained heuristically without strict rules on relationship between angles and angle ranges. For that reason, only templates from Figure 2 will be considered in the rest of the paper.

Color harmony has various applications in recoloring, recomposing, harmonizing, enhancing image or video, manual adjustment of final harmonious hue, background harmonization based on foreground data, example-based harmonization [17,21].

Harmonic templates defined on the color hue wheel [12] shown in Figure 2 are used in [3,6,21] to recolor or recompose color within a picture. The first step consists in determining the harmonious template type and its rotation angle that are the closest to the original picture by minimizing a cost function. Then, the colors are transformed so that the colors outside harmonious sectors are mapped

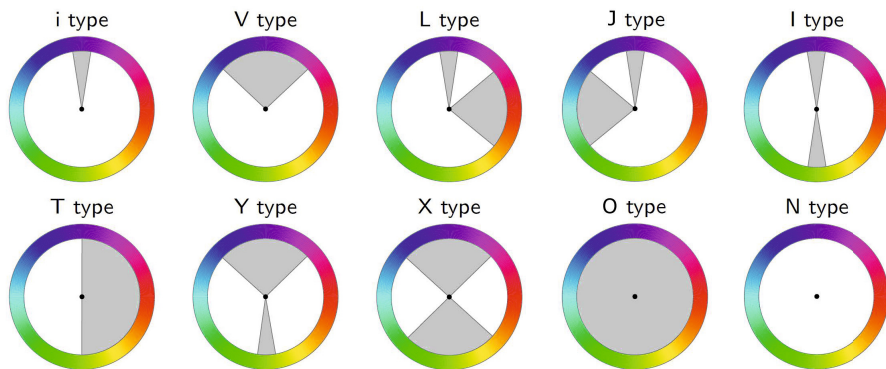


Fig. 2. Harmonious templates T_m on the hue wheel for a given angle [3]. The complete collection of harmonious templates is obtained by rotating all templates. For a given template at a given angle, grey areas enclose hues that form a harmonious set. In other words, an image where the histogram is strictly enclosed in the grey area is considered as harmonious.

inside a harmonious sector. Color segmentation has been identified as a crucial pre-processing before color mapping because visible artifacts can appear when two close colors, eventually associated to the same object, are mapped to two different sectors. Strategies differ in terms of cost function for template selection, color segmentation and color mapping.

A real-time color harmonization for video has been introduced by Sawant *et al.* [17] where a histogram splitting method is employed instead of a graph-cut approach to reduce computation cost. One dedicated template per group of frames is determined to guarantee temporal consistency. In [22], Tang *et al.* perform a foreground/background detection for their dedicated video implementation and apply the same template determination as [17] for a coherent group of frames.

3 Method

In this section, the harmonization method is described into details. An overview of the method is shown in Figure 3. We then describe the template determination, the pixel color mapping and the color segmentation technique.

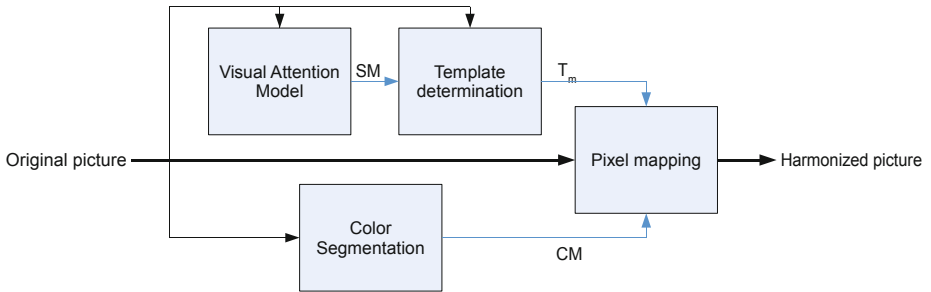


Fig. 3. Overview of the proposed algorithm. The first step (top part) consists in estimating the best harmonic template. Based on a measure of the salient areas, a template determination is performed so as to minimize a statistical distance between the histogram and a harmonious template. Prior to the color mapping, a color segmentation technique is performed to ensure that consistent colors are finally mapped into the same harmonious template sector.

3.1 Template Determination

This step consists in choosing the best template amongst one of the nine harmonious templates T_m ($m \in \{i, I, L, T, V, X, Y, J, O\}$) previously defined in [3] and depicted in Figure 2. Note that template O has been added since we want to leave unchanged (i.e., no harmonization) the pictures containing a large panel of hue values, like rainbow pictures for example.

The color histogram of the original image is computed in HSV space using L bins (typically, $L = 360$). However, to determine the template matching best the color distribution, a modified hue histogram M is weighted by the saturation and values as:

$$\forall i \in [1, L], M_i = \frac{1}{\sum_{(x,y)} S[x, y] \times V[x, y]} \times \sum_{(x,y) \in \{(u,v) \setminus H[u,v]=i\}} S[x, y] \times V[x, y]$$

Weighting the hue distribution by saturation and value allows to be closer to perceived amount of color compared to a simple hue histogram. This can be seen on the left part of Figure 6, where the pinkish and greenish parts have been emphasized.

Saliency Map. The saliency map SM is computed for the considered picture from the visual attention model of [9]. The saliency map provides a representation of the most visually attractive pixels (white pixels stand for the most attractive one.). The basic idea is that the estimation of the template should be driven mainly by visually representative areas of the image to obtain a visually consistent harmonization. As for the entire image, a weighted hue histogram $M_{saliient}$ is computed for the T most salient pixels. The threshold T has been empirically set to 1% of the total number of pixels in all experiments.

Cost Function. A cost function is required to determine the most appropriate template regarding M (and $M_{saliient}$). The appropriate template shape T_m and the associated orientation α that best fits the hue distribution M is chosen by minimizing the Kullback-Leibler divergence computed for each template and each orientation.

$$\min_{m, \alpha} \sum_i M_i \times \ln \left(\frac{M_i}{P_i(m, \alpha)} \right)$$

where $P_i(m, \alpha)$ is the normalized distribution of template T_m for the orientation α . Here $P_i(m, \alpha)$ typically represents a harmonized model or description and is composed of k sectors S_i^k . A distribution of a sector is defined as:

$$S_i^k(m, \alpha_k) = \begin{cases} \exp \left(-\frac{1}{1 - \left(\frac{2 \times \|i - \alpha_k\|}{w} \right)^{10}} \right) & \text{if } i \in]\alpha_k - \frac{w}{2}, \alpha_k + \frac{w}{2}[\\ 0 & \text{otherwise} \end{cases}$$

Therefore,

$$P_i(m, \alpha) = \begin{cases} \sum_{k=1}^n S_i^k(m, \alpha_k) & \text{if } \sum_{k=1}^n S_i^k(m, \alpha_k) \neq 0 \\ \frac{0.01}{L} \times \sum_i \left(\sum_{k=1}^n S_i^k(m, \alpha_k) \right) & \text{otherwise} \end{cases}$$

where $\| \cdot \|$ is the arc-length distance on the hue wheel (measured in degrees), α_k is the center and w the size of their associated sector and n is the number of sectors defined for template T_m .

The appropriate template T_m may be also identified by the couple $(m, \alpha)_{distribution}$. The cost function is applied independently on both M and $M_{salient}$ distributions, leading to the two sets of parameters $(m, \alpha)_M$ and $(m, \alpha)_{M_{salient}}$.

Merging the Two Templates. Both templates are then combined to form a new distribution or model P'_i . The combination consists in taking for each bin the maximum value in the histogram of the template computed on the whole image and in the template computed on salient pixels such as:

$$P'_i = \max(P_i(m, \alpha)_M, P_i(m, \alpha)_{M_{salient}}).$$

Once again, the most similar template to this new combined distribution, among the nine harmonious templates, is found by minimizing the Kullback-Leibler divergence between P'_i and the distribution computed for each template and each orientation.

Since minor hues in color histogram have to be taken into account due to their semantic or attentional relevance, two templates based on two different hue distributions (the entire image and only most salient pixels) are separately computed and then combined. The main advantage of using two hue distributions is that the method is parameter-free, compared to a weighting of salient pixels where the weights would have to be tuned carefully.

3.2 Pixel Color Mapping

Once an harmonious template has been estimated, colors that lie outside the template will be mapped into the template. This step is described here and the main purpose is to perform a larger modification for pixels whose color is outside the harmonious template. Some artifacts may be created during this step because two neighboring pixels that have similar colors can be mapped into two different sectors of a template. Figure 8 illustrates this effect and it be clearly seen that the girl's trouser is mapped to two different colors, what is visually inconsistent. To remove these artifacts, all pixels that are in the same segmented area are mapped in the same sector. To do so, a dedicated color segmentation will be used and is described in the next section of this paper.

A hue-mapping map is thus created where each pixel is assigned with the direction of mapping of its hue value:

- For each pixel p , its original hue quadrant between a sector angle α and a border is determined,
- each segmented area is sub-divided depending on the pixel hue where all pixels having their hues between a sector angle and the next are in the same sub-segment (this is because segmentation problems appear only at sector borders - each segment can be divided in two in example of fig. 4),
- for each sub-segment, the majority quadrant (between the 2 around a given border) is selected as the hue mapping direction for each pixel p , so that all pixels belonging to the same segment and having their hues around a sector border are harmonized in the same direction to avoid mapping artifacts.

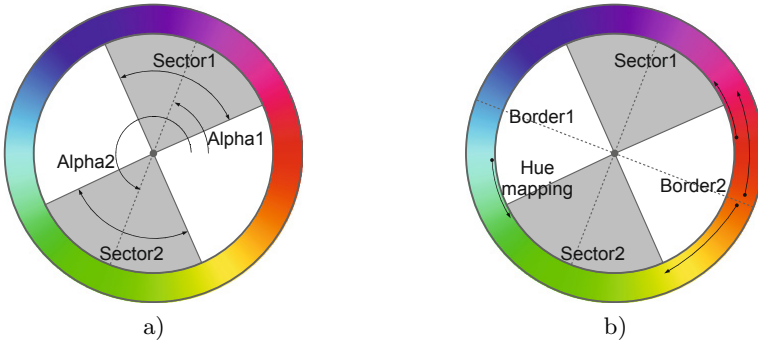


Fig. 4. Harmonic mapping of hue depending on the initial hue value. a) Chosen harmonic template with two sectors and their angles, b) hue mapping examples.

A sigmoid function is then used to transform the hue of each pixel p :

$$H'(p) = C(p) + \text{Sign}_{quadrant} \times \frac{w}{2} \times \tanh\left(\frac{2 \times \|H(p) - C(p)\|}{w}\right)$$

where $C(p)$ is the central hue of the sector associated with p , w is the arc-width of the template sector, $\|\cdot\|$ refers again to the arc-length distance on the hue wheel and $\text{Sign}_{quadrant}$ is the sign associated to the direction of mapping.

This sigmoid function has interesting properties for pixel color mapping: its asymptote in extreme values auto-clamps pixels in the template and its middle section (normal behavior) is nearly linear so at the center of a sector, hues are not changed. Original image and dynamics is less modified compared to [3] where the mapping function contracts hues even near the center of sectors. Since dominant colors are likely to be chosen near the middle section of harmonic sectors, they are less subject to hue change, thus the color semantics is conserved. Therefore automatic harmonization can be performed without any user interaction.

3.3 Color Segmentation

As mentioned previously, a color segmentation step is necessary to ensure the mapping consistency, in other words to maximize the probability that pixels belonging to the same semantic object are mapped to the same harmonious sector.

State-of-the-art techniques [3] and [6] apply conventional graph cut algorithm [2] for segmentation, while Tang *et al.* [22] perform a two-steps graph cut processing both at the region and pixel levels in order to handle issues regarding spatial inconsistency of graph-cut algorithm. Sawant *et al.* [17] do not use any image segmentation but implements histogram splitting on the hue histogram. Such a method does not account for spatial arrangement of the pixels and can be seen as color clusterization of the outlier pixels in two clusters.

For color harmonization, the spatial aspect of the color segmentation may not be compulsory and might even introduce artifacts as it can be seen on Figure 8. Therefore, we think that a histogram segmentation technique is adequate here, such as the popular K-means method. However, such a histogram segmentation should obey the following constraints:

- It should be unsupervised, meaning that the final number of color clusters should not be a parameter. As a matter of fact, the color harmonization would be very sensitive to a incorrect number of meaningful colors.
- The histogram segmentation technique should be capable of segmenting *small* modes of the histogram. In other words, small regions that could be seen as color outliers, such as the ladybug of Figure 7, should be detected as separate modes.

In order to meet these requirements, we propose here a color segmentation method that build on the work of Delon *et al.* [4] referred to as ACoPa (Automatic Color Palette). This color segmentation technique is based on a *contrario* analysis of the color histogram modes. A statistical estimation of meaningful histogram modes is performed. Instead of the hierarchical estimation of modes in the H , then S , then V space, we propose to perform a histogram decomposition of each component independently. The obtained modes are combined from all modes obtained, and segments with a very limited group of pixels are discarded. Finally, based on these histograms modes, a K-means post-processing is used to group the modes that are perceptually similar using a dictionary expressed in the Lab color space.

We have obtained a segmentation technique that is approximatively 10 times faster than the original version, and deals more efficiently with achromatic pixels.

4 Results and Discussion

In this section, all contributions of the paper, successively template determination, saliency, color segmentation and color mapping, are tested and compared to the state-of-the-art technique of Cohen-Or *et al.* [3]. Finally, results are shown on various images for psycho-visual assessment of the method.

4.1 Template Determination

We first here compare the determination of the template using our method compared to Cohen-Or *et al.* [3] and Tang *et al.* [21]. In [3] the cost function is the sum of the distance between each pixel hue and the template sectors, weighted by its saturation. In [21], the template is determined using the ratio between the number of pixels contained in the template and the maximum number of pixels that could be contained in the sectors.

We first computed harmonious templates using the three techniques on a database of 154 images. Results are shown in Figure 5. It can be seen that [3] tends to favour only the T and X templates, while the two other methods have a more balanced repartition of templates.

We also show the obtained template on a particular image on Figure 6. For the initial image (top left), the hue histogram and weighted hue histograms are shown on the hue wheel (bottom left). Templates estimated using the different methods are shown on the bottom row, and the harmonized images on the top row. While [3] and [21] have estimated the T template, our technique did fit a X template, leading to a visually more satisfactory result.

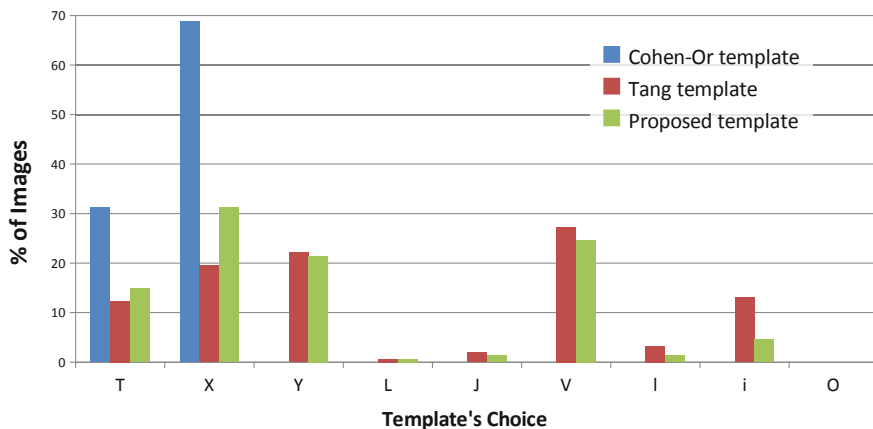


Fig. 5. Distribution of template selection for a database of 154 images regarding the Cohen-or [3], Tang [21] and our proposed energy computation. The Cohen-or technique tends to select the T and X templates.

4.2 Saliency

The loss of meaningfully minor colors in the global distribution is a challenge for automatic color harmonization. Modifying skin or sky colors in an unnatural way may have an annoying semantic impact. A manually defined mask is usually used to harmonize only selected pixel and thus preserve other areas [3]. Associated with a better selection of template, our approach fixes this limitation by taking into account saliency information in the refinement of the final template. The chosen harmonic template contains the colors of salient areas. The result is clearly illustrated

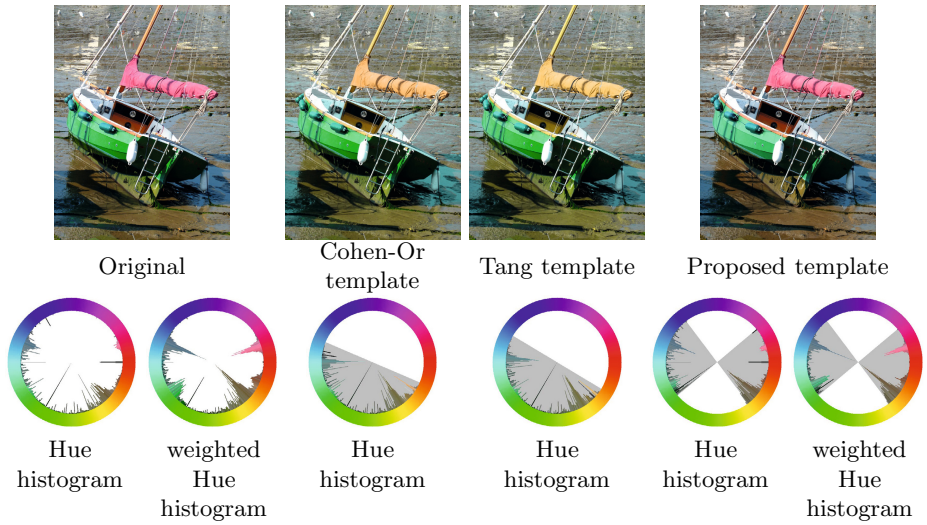


Fig. 6. Influence of distribution similarity computation regarding template choice using the proposed color mapping function

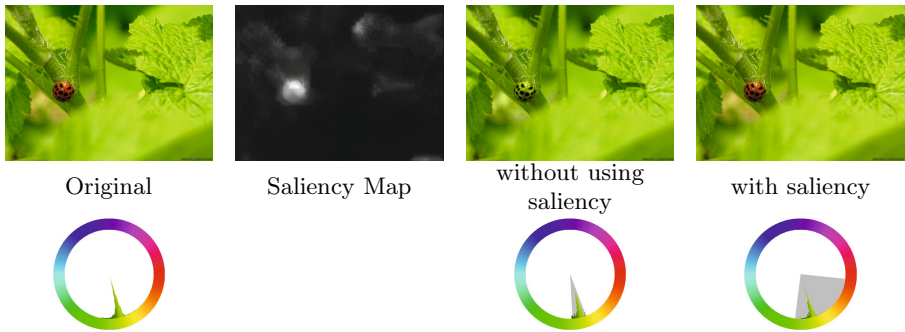


Fig. 7. Color harmonization and template choice when taking into account saliency map. A larger template allows preserving the semantically relevant red color.

in Figure 7, where a better choice to preserve contents for such a picture leads in little harmonization: the red color drives the template determination.

4.3 Color Segmentation

In previous section, we introduced the segmentation algorithm implemented as a preprocessing before the color mapping step. The need of such preprocessing has been discussed previously in [3] and the same strategy has been kept here. However, using a spatial segmentation might lead to split a disconnected object into two different components. Using a non spatial algorithm allows to treat all

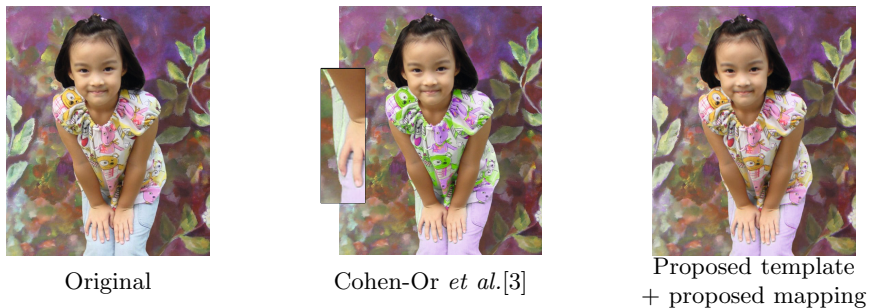


Fig. 8. Segmentation approach: Spatial segmentation dissociates two spatially disconnected regions of the same object. Our color segmentation technique leads to a more consistent hue-mapping.

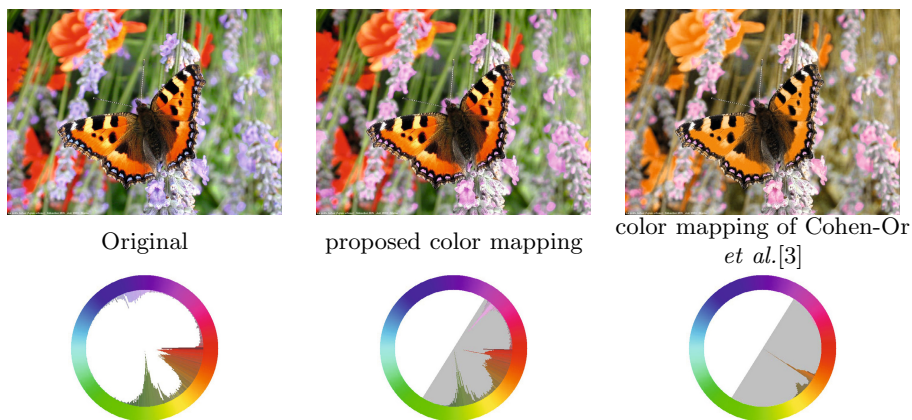


Fig. 9. Illustration of different mapping functions. Top row: original and harmonized images. Bottom row: original histogram and final histograms of harmonized images. For both methods, the template T has been selected. However, in [3], the proposed mapping function severely contracts the hues, while our proposed mapping functions better preserve original color dynamics.

pixels having the same colors without *a priori* on their position. Such a case is illustrated in Figure 8 where our approach maps regions of girl’s pants into the same harmonious sector.

4.4 Color Mapping

In state-of-the-art solutions, color mapping functions tend to highly contract hue even for colors that originally lie in the middle of harmonious sectors. Thus, the harmonization might over-modify original pictures. This effect may be sought

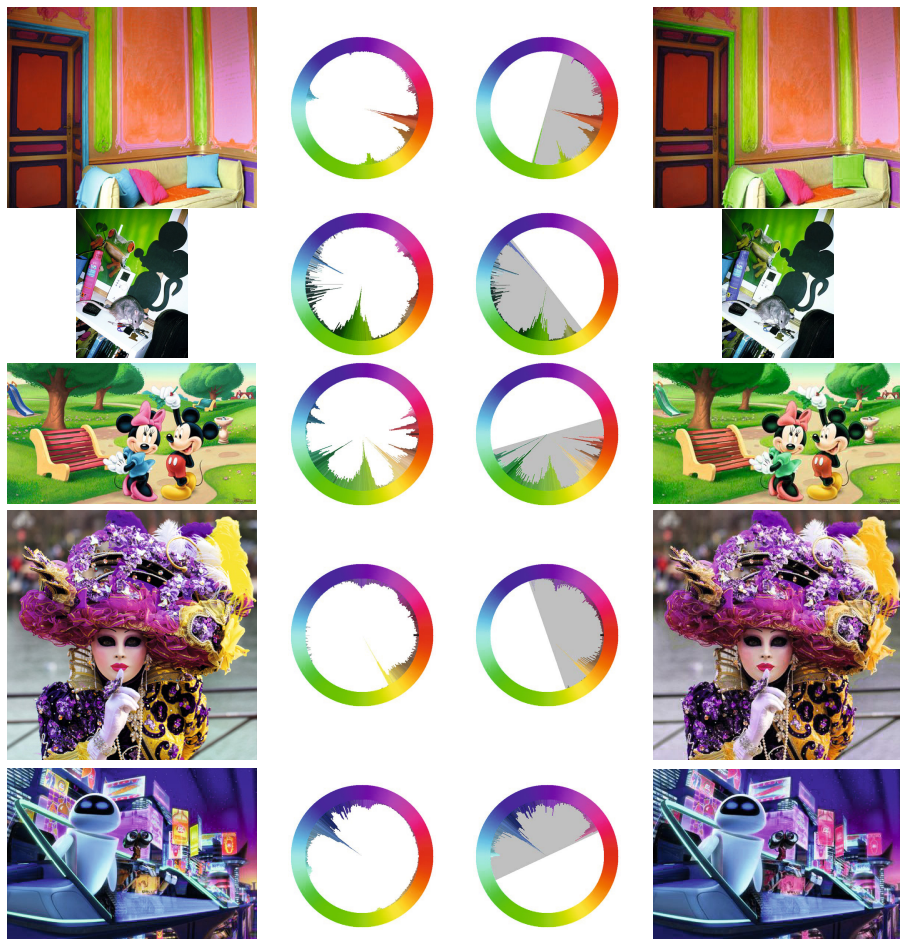


Fig. 10. Color harmonization and template choice on various images. For each image of the left column, the middle column shows the initial histogram and the estimated template. the final harmonized image is presented on the right. It must be stressed out that each image was obtained without any user interaction nor parameter tuning.

in some applications, but may be disruptive when the intent is to stick to the original image. However color dynamics are not preserved even for the pixels whose hue is near the middle of the template in the original image. The proposed mapping function (see section 3.2) guarantees original hue values at the center of the harmonious sectors and compresses more strongly hue values outside the template. Results of the color mapping and comparison with state of the art are illustrated in Figure 9. The harmonic colors are preserved, and only non-harmonic hues are modified.

4.5 Visual Results

In Figure 10, results are presented on various images, including natural images, stylistic images and cartoons. For every image of the left column, the harmonized result displayed on the right column was obtained without any user interaction nor parameter tuning. In addition, original histograms and estimated harmonious templates are presented in the middle column.

5 Conclusion

In this paper, a color harmonization method has been presented. Building on the work of Cohen-Or [3], the method proposes three main contributions to reduce possible artefacts and obtain a visually consistent result. First, a cost function measuring the statistical distance between a weighted hue histogram and the possible templates was used, jointly with the information of saliency. In Cohen-Or *et al.* [3] the cost penalizes only pixels that have a hue outside the template, and favors large templates. In Tang *et al.*[21], the cost penalizes large sectors when not necessary. Our approach considers hue histogram and templates as statistical distributions. The Kullback-Liebler divergence is used to quantify the distribution similarity and emphasizes histogram differences. In addition, the weighted hue histogram is closer to the perception of colors. Consequently, the selection of templates better fits the original image color distribution. Second, a new color mapping function was used. Results have shown that this mapping function contracts less the mapping within the desired template and leads to a more consistent mapping. Third, a dedicated color segmentation was used and was able to segment small histogram modes, while leading to a consistent result.

Further work should focus on the extension of such harmonization to videos. The temporal regularization of the harmonization should be enforced in order to avoid temporal artifacts.

References

1. Birren, F.: A grammar of color, a basic treatise on the color system of albert h. munsell (1969)
2. Boykov, Y.Y., Jolly, M.P.: Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In: Proceedings of the Eighth IEEE International Conference on Computer Vision, vol. 1, pp. 105–112. IEEE (2001)
3. Cohen-Or, D., Sorkine, O., Gal, R., Leyvand, T., Xu, Y.Q.: Color harmonization. ACM Transactions on Graphics (TOG) 25, 624–630 (2006)
4. Delon, J., Desolneux, A., Lisani, J.-L., Petro, A.-B.: A nonparametric approach for histogram segmentation. IEEE Transactions on Image Processing 16(1), 253–261 (2007)
5. Holtzschue, L.: Understanding color: an introduction for designers. Wiley, Hoboken (2011)
6. Huo, X., Tan, J.: An improved method for color harmonization. In: 2nd International Congress on Image and Signal Processing, pp. 1–4 (October 2009)

7. Itten, J.: The art of color: the subjective experience and objective rationale of color. Van Nostrand Reinhold, New York (1973)
8. Kawamoto, N.: Objective evaluation of color design. iii. *Color Research & Application* 20(1), 62–69 (1995)
9. Le Meur, O., Le Callet, P., Barba, D., Thoreau, D.: A coherent computational approach to model bottom-up visual attention. *IEEE Trans. Pattern Anal. Mach. Intell.* 28(5), 802–817 (2006)
10. Li, C., Chen, T.: Aesthetic visual quality assessment of paintings. *IEEE Journal of Selected Topics in Signal Processing* 3(2), 236–252 (2009)
11. Machajdik, J., Hanbury, A.: Affective image classification using features inspired by psychology and art theory. In: *Proceedings of the International Conference on Multimedia, MM 2010*, pp. 83–92. ACM, New York (2010)
12. Matsuda, Y.: *Color design*. Asakura Shoten (1995)
13. Mayer, T.: On the relationships between colors. *Color Research & Application* 25(1), 66–74 (2000)
14. Ostwald, W., Birren, F.: *The Color Primer: A Basic Treatise on the Color System of Wilhelm Ostwald*. Van Nostrand Reinhold Co. (1969)
15. Ou, L.-C., Ronnier Luo, M.: A colour harmony model for two-colour combinations. *Color Research & Application* 31(3), 191–204 (2006)
16. Przemyslaw, S., Michal, K.: Investigating human color harmony preferences using unsupervised machine learning. In: *European Conference on Colour in Graphics, Imaging, and Vision*, pp. 59–64 (2012)
17. Sawant, N., Mitra, N.J.: Color harmonization for videos. In: *Sixth Indian Conference on Computer Vision, Graphics & Image Processing*, pp. 576–582 (2008)
18. Sivik, L., Hrd, A.: Some reflections on studying colour combinations. *Color Research & Application* 19(4), 286–295 (1994)
19. Solli, M., Lenz, R.: Color harmony for image indexing. In: *2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops)*, pp. 1885–1892 (2009)
20. Sutton, T.: *The complete color harmony*. Rockport publishers, Massachusetts (2004)
21. Tang, Z., Miao, Z., Wan, Y.: Image composition with color harmonization. In: *2010 25th International Conference on Image and Vision Computing New Zealand (IVCNZ)*, pp. 1–8 (November 2010)
22. Tang, Z., Miao, Z., Wan, Y., Jesse, F.F.: Colour harmonization for images and videos via two-level graph cut. *Image Processing, IET* 5(7), 630–643 (2011)