Color Harmonization for Videos

Nikhil Sawant Niloy J. Mitra
Dept. of CSE, Indian Institute of Technology, Delhi
{nikhilus85,niloym}@gmail.com

Abstract

Color harmonization is an artistic technique to adjust the colors of a given image in order to enhance their visual harmony. In this paper, we present a method to automatically improve the color harmony of images. Harmonization is performed using a carefully designed optimization in the hue space, while keeping the saturation and intensity components unchanged. Finally, for videos, we pose the problem as an efficient joint optimization in space and time, thus minimizing flickering or visual artifacts in the harmonized output video. We report the performance of our algorithm on a variety of test images and video sequences.

1. Introduction

We are all aware that not all colors match or look nice with every other colors. Although the notion of harmony of color can depend on place, time, context, etc., there are some color arrangements specially those motivated by co-existing colors in nature, which are universally considered to be harmonious. *Color harmonization* refers to taking an input image, specially of a scene made of juxtaposition of various colored objects, and enhancing the harmony of the scene by minimally modifying or shifting the colors in the image. The situation is similar for videos where unharmonized scenes are common due to rapidly changing environment. The problem is compounded since in every frame the scene may be changing, along with fluctuations in illumination. Further, we may have indoor-outdoor scenes back to back which can be very disturbing if colors are not harmonized. We propose a novel method which harmonizes the given video by solving a joint space-time optimization. Our procedure tries to retain the original colors as far as possible, while minimally shifting the colors which are disturbing harmony. Our algorithm is simple and efficient, suitable for realtime use, and outputs a harmonized video while minimizing any flicker due to discontinuity in time and motion.

Study of color and light has been an active topic of interest since early days. As early as in the seventeenth century, Newton gave us the first color wheel. Subsequently, Maxwell came up with important contributions in the field of colors [3, 4]. Itten [2] was the first to introduce the color wheel based on hue information. He proposed a scheme of color harmony based on relative positions of colors on a color wheel. He introduced two (complementary colors), three (equilateral triangle), four (square), six (hexagon) color harmony schemes. Tokumaru [5] extended this work for harmony evaluation. He also introduced template-based harmonization schemes. Such templates attempt to quantify our perception and understanding of matching colors, allowing us to solve the color harmonization problem, whose goal is to improve the visual appeal of an image, in an optimization framework.

Our work is inspired by the recent work on image color harmonization by Cohen-Or and colleagues [1]. They used the templates proposed by Tokumaru et al. [5] to harmonize the images along with a graph cut method to ensure contiguous coloring. Such research efforts have led to the development of few professional tools for color harmonization for images. However, to our knowledge, no such system exists for color harmonizing video sequences.

In a conceivable approach for video color harmonization, one can take each individual frame and independently perform color harmonization. Such a technique is slow, im-

Figure 1: (Left) Input image. (Right) The color harmonized image is visually more pleasing. The output depends on the type of hue-template [5] used, template X in this case.

practical, and does not exploit the coherence between adjacent frames. More importantly, treating each frame independently results in artifacts in the forms of flickering as shown in Figure 6.

Producing harmonized videos is more involved compared to harmonized images for number of reasons like constant change in colors, light, scene, etc. Given a video sequence our method harmonizes the video so that viewer will have a pleasant experience. We also propose a novel method to address the segmentation problem, thus helping to maintain contiguous coloring. In order to maintain contiguous coloring in time, we automatically detect and make use of I-frames, commonly used in MPEG-styled video codecs.

The paper is organized as follows: In Section 2, we discuss image color harmonization and propose important improvements over existing methods. In Section 3, we introduce color harmonization for videos, and describe our method for video color harmonization addressing the associated challenges. We present various results in Section 4.

2. Color harmonization for images

Harmony is a pleasant combination of different things, encountered in various contexts — music, poetry, colors, etc. In absence of harmony, things seems to stand apart leaving an uncomfortable feeling. Lack of color harmony is often the reason behind boring or chaotic looking images. Historically color harmony often draws inspiration from nature which is mostly harmonized. In terms of a color wheel, analogous colors (adjacent colors) and complementary colors (opposite colors) leave a harmonized effect on the visual system. In our work, the color harmony is based on the templates defined by Takumaru et al. [5]. Figure 2 shows the various harmonic templates on the basis of hue information.

There are eight templates $\{i, V, L, I, T, Y, X, N\}$ defined in HSV space. We denote a template by t and its orientation by θ. The templates can have any orientation between $[0-2\pi]$, i.e., the shaded area within each template t can be rotated over the range $[0-2\pi]$. We call each intermediate

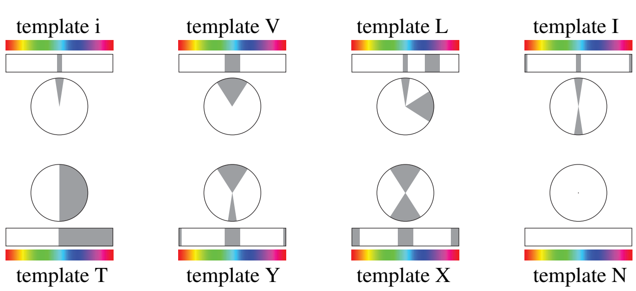

Figure 2: Color harmonization can be seen as fitting or approximating hue information using harmonizing templates, which were originally proposed by Tokumaru et al. [5].

position as an *orientation* of t and denote it by θ. The range of θ is $[0-2\pi]$ which has been discretized into 360 orientations. We did not observe significant benefits for finer discretization levels. Notice that in the color wheel, values are wrapped around going from 0 to 2π, i.e., the left and right end of the bars in Figure 2 are identified. An image is said to be harmonized under a particular template t only if the entire hue histogram of the image falls within the shaded region for a particular orientation θ. Combination of template t and orientation θ is denoted by θ_t, where $\theta \in [0-2\pi]$ and $t \in \{i, V, L, I, T, Y, X, N\}$. We leave out template N from further discussions since it is defined only for gray scale images.

To harmonize an image we adjust or shift the image hue histogram so that it falls under the gray region for some θ_t. However, colors may become undesirable to the viewer e.g. if sky is not blue or if grass is not green and so on. This can be avoided if minimum adjustment to the hue is performed during the process of color harmonization. The problem is to select t and its corresponding θ which requires minimum color adjustment — we denote such an optimal combination of θ and t by θ_H[1].

Given an image I, we define a potential function $P(\theta_t, I)$. The function measures the potential of the image I for an orientation θ under a template t, i.e., θ_t:

$$P(\theta_t, I) = [\#\text{ image pixels in the shaded region}].$$

This simply returns the number of pixels under θ_t. We do not take saturation [1] into consideration because it was found that results are not only comparable, but also this potential function is more computationally efficient, and hence is better suited for video processing.

Our first task is to find out an optimal orientation of θ_t for a fixed template t, which we call θ_{ot}[2]. The potential is calculated by considering all the values θ for the given template t. We select optimal θ, i.e., θ_{ot} such that:

$$\theta_{ot} = \max(P(\theta_t, I)), \quad \theta \in [0, 2\pi].$$

We maximize the value of potential function because it helps in minimizing the hue adjustment while harmonization. Likewise we calculate θ_{ot} for each template. To choose harmonizing template (θ_H) we choose θ_{ot} with maximum potential among all θ_{ot}:

$$\theta_H = \max(P(\theta_{ot}, I)), \quad T \in \{i, V, L, I, T, Y, X\}.$$

Therefore θ_H gives us the harmonizing template. $P(\theta_H, I)$ is the maximum possible value of the potential function for a given image I over all choices of t and θ.

[1] θ_H : Both θ and t are variable, we are not only required to find optimal θ but also optimal template t, which together forms the optimal solution.

[2] θ_{ot} : θ is variable but t is fixed i.e., θ corresponds to the fixed t, this gives the optimal solution for given template t.

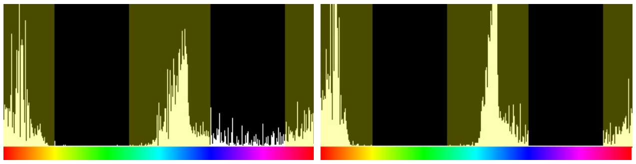

Figure 3: Representative hue histogram before (left) and after (right) color adjustment. Hue values are adjusted such that the non-zero parts of the hue-histogram lie inside the shaded (in green) region.

Since perception of colors additionally depends on many factor such as context, emotion, culture, likings, etc., we provide an additional option in our harmonization scheme to allow the user to choose a template t and a desired orientation θ.

Color adjustment. Once we find out the θ_H we know both template t and θ for which image potential is the highest. In order to make the given image optimally harmonized we need to adjust color such that entire hue histogram of the given image falls under θ_H. Simple linear contraction method has been used to adjust the hue values falling outside the current template to fall within the template as shown in Figure 3. We need to split the hue space in order to make image hue to fall under the shaded region. Splitting points are selected such that linear contraction needed to adjust the hue of the image is constant. After applying the contraction, the entire hue histogram lies within the θ_H producing a harmonized image. This simple method used for color adjustment helps in further speeding up the process which is essential when we work on videos.

Segmentation problem. The segmentation problem is due to splitting of colors present within the same region in an image. A continuous portion in the image can contain different shades of a color. These shades are marginally separated along the hue scale. During color adjustment it may happen that different shades of a color may get mapped to colors with larger distances on the hue scale. This makes dual colors to appear in the same region as shown in Figure 5 leading to visually objectionable artifacts.

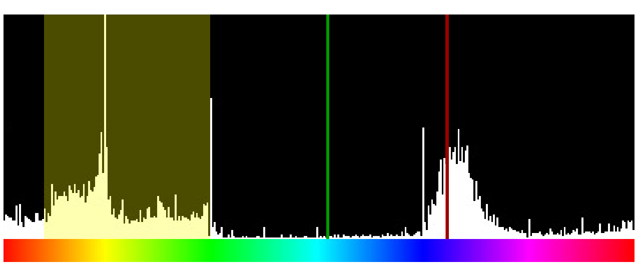

Figure 4: The red bar shows the splitting position according to a given template and orientation. The green bar is the position according to our histogram splitting method for reducing segmentation artifacts. Corresponding resultant images are shown in Figure 5.

In Figure 4, the red bar shows the original splitting position according to a selected template and orientation. If colors are adjusted according to the original splitting position, we might get a segmentation problem. We propose a simple technique to prevent such an artifact. A new splitting position is chosen as follows:

```
for i = 0 to I_H(S) do
  for j = 0 to δ do
    if I_H(S + j) = i then
      return (S + j)
    else
      if I_H(S − j) = i then
        return (S − j)
      end if
    end if
  end for
end for
return S
```

Here S is the original splitting position, $I_H(S)$ is the value of hue histogram of image I at S. We define a parameter δ which is maximum possible variation of splitting position from original splitting position S. This algorithm returns a new splitting position. The method considers δ locations from original splitting position in both direction and finds out the position with the minimum number of pixels present (see Figure 4). We expect a splitting position to be present in an empty region in the hue histogram, but

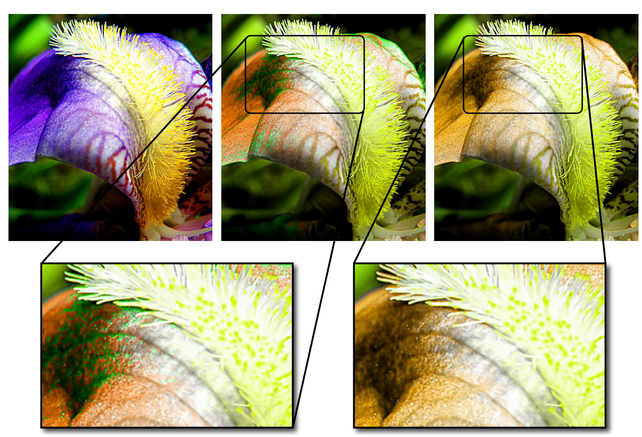

Figure 5: (Left) Original image, (middle) color harmonized result with segmentation artifacts, (right) final image using our splitting position refinement. Our method significantly reduces color bleeding.

Figure 6: A video sequence, when color harmonized by simply processing each frame individually, results in flickering.

this is not always possible. Figure 5 shows the improvement achieved when colors are adjusted according to the new splitting position.

The splitting position with zero value on hue histogram most of the times than not removes segmentation problem. But one can never guarantee that finding such a position will remove every occurrence of segmentation problem. The parameter δ should be chosen in such a way that it is sufficiently large to find a suitable position. Also care should be taken that this position does not come under shaded region for the given template and orientation.

Using this method it is possible to remove segmentation problem for most of the images. While graph cut based methods [1] give better results, they are usually slow. For real-time processing of video frames, we found that our method is more suitable without noticeable compromise in visual quality.

Two parameters variation. We further modify template X with two parameters variation. During color adjustment it is expected that image does not lose its original colors as far as possible and our method of two parameter variation extends this idea. We call this method two parameters variation because here we not only allow variation in θ but also allow variation in distance between shaded region also. We allow template to adjust itself so that it fits according to image hue histogram. Hence there is minimum color adjustment in the harmonized image. We limit the variation in template to keep the final image harmonized. Figure 7 shows an example of image harmonization using two parameter variation approach.

3. Color harmonization for videos

Often we encounter mismatched colors leading to unnatural appearances in videos. Such video sequences can be painful to the eyes. We extend the the basic scheme for image color harmonization to videos.

Videos are nothing but a sequence of image frames. Each frame differs slightly from the previous one to create the illusion of motion. Video is said to be harmonized if all the frames carry harmony not only individually, but also with respect to groups of adjacent frames. Such a video sequence is pleasant. A sequence of disturbing frames or even a single disturbing frame can affect the harmony of the entire video. We propose a technique which produces a video sequence which is harmonized and pleasant for the viewer.

Approach. We start with a simple approach, where we extract out the frames of a video and individually harmonize each frame. This is done by harmonizing all frames of a given video using the above stated method, i.e., we calculate θ_H. All the remaining steps like adjusting colors and handling segmentation problem are done similarly.

We follow this procedure for all the video frames one by one. The frames are stitched back to produce the output video.

Shortcomings. This simple approach makes each frame harmonized individually. But we observed that the obtained video contains some flickering artifacts. This flickering is due to noticeable change in the colors between the adjacent frames belonging to the same scene, as shown in Figure 6.

In the above approach each frame is harmonized irrespective of its adjacent frames. Since a frame shows a slight change from its previous frame, it is possible that this may cause a change in θ. If the change is small then difference between adjacent frames is not noticeable. However, if the

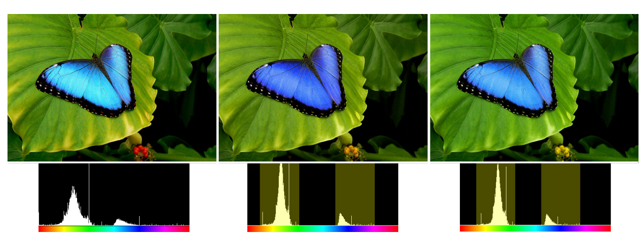

Figure 7: Original image (left), harmonized image using standard template X (middle), and the result after two parameter variation (right). The final result is not only color harmonized, but also exhibits less color variation as compared to the middle image.

Figure 8: Our joint space-time optimization approach results in a flicker-free color harmonized video. A naive approach on the same sequence results in artifacts as seen in Figure 6.

change is large, as in Figure 6, then it becomes noticeable. In order to eliminate this artifact, adjacent frames need to be correlated during the harmonization process. As a solution to this problem, we propose an improved method for video color harmonization.

Note that flickering artifacts are not always due to change in θ, it is also possible that the template t may change between adjacent frames. During the calculation of θ_H there is scope for varying θ as well as template t. Such changes in template never go unnoticed in video unless change of scene is present at the same junction. As a solution we select a template t initially, and subsequently only allow variations in θ, i.e., selecting θ_{ot} instead of θ_H. For the remaining discussion we fix the value of t and only consider change in θ. We have chosen template X as it harmonizes images with least change in color thus preserving the natural color content of images and producing more realistic images.

Improvement. As we discussed in the earlier sections, flickering artifacts are observed when each frame is harmonized individually. Hence we modify our approach to avoid such artifacts by coupling or correlating the adjacent frames.

Grouping. In order to remove the flickering artifacts we concentrate on processing of the frames in groups rather than individually. Certain number of subsequent frames are grouped together as shown in Figure 9. The mean of hue histogram of all the images belonging to a group is calculated. The orientation θ for a group is obtained by considering the mean hue values calculated for that group. All the frames belonging to that particular group are harmonized with this calculated θ. Here θ is the resultant of all the frames, hence it harmonizes image with acceptable (close to the minimum) color adjustment if not the minimum. Choice of number of frames per group depends on the video sequence and needs to be adjusted accordingly.

As a result no artifacts for the frames within a group are observed. But still with this approach there might be visible flickering at group boundaries. This indicates grouping removes intra-group flickering but can lead to inter-group artifacts. Such an inter-group flickers is a result of a large difference between mean hue values between two groups.

Overlapping. Overlapping method is slight improvement over grouping with a small variation. Instead of making disjoint groups of frames, we overlap frames between two adjacent groups as shown in Figure 9. This helps in reducing the inter-group flickering. The common frames between two groups help to change the mean hue values gradually rather than drastically. This method gives pleasing results with reduced artifacts at the cost of increased computing time.

We observed that inter-group flickering is due to change of scene. When the two adjacent groups contain the frames from same scene, flickering was often not noticeable. But when a group contains frames from two different scenes then the adjacent groups, both preceding and succeeding, show flickering. This suggests that the group containing frames from two different scenes causes the flickering. If we can avoid such groups then inter-group flickering can be completely removed. One method to remove such groups is by disallowing frames from two different scene in the same

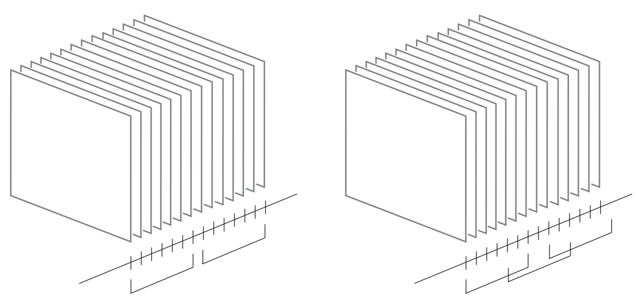

Figure 9: (Left) Non-overlapping grouping of frames for space-time color harmonization. Such a scheme can still result in minor flickers across group boundaries. A slightly more expensive alternative, which uses a sliding window (right) for space-time color harmonization, produces in smoother and more pleasing videos.

group. For this purpose we need to identify scene changes in the video sequence.

I-frame detection. We consider I-frame as the first frame of the scene, each scene always starts with I-frame. We detect I-frame so that we can avoid grouping the frames belonging to different scenes under the same group. Further sometimes there is an abrupt change between frames of the same scene. We would like to identify such frames and treat them as I-frames.

Our next job is to detect such frames. To detect I-frames we use the square of absolute difference for each pixel between adjacent frames. This error measure is commonly referred to as the mean square error (MSE). An appropriate threshold value is set: If the MSE is greater than the threshold value, then that particular frame is classified as an I-frame (see Figure 10).

Using this approach we find two consecutive I-frames. The frames between two I-frames are harmonized using overlapping approach with I-frame as a start of subsequence. Since between two I-frames there is little change in hue values of the frame no flickering artifact is seen, also change of θ is normally observed at I-frames which goes unnoticed because of change in scene. Hence this is a method to remove flickering artifacts, the result of this technique can be seen in Figure 8.

4. Results and applications

Color harmonization helps to harmonize the image, and improve the balance among the colors present in the image. Our software automatically calculates the best suited harmonic template and its orientation for a given image. The user can also change both the template and its orientation according to her preference. This allows the user to recolor the image according to her color preference. The technique can be used to build tools to help users choose harmonious color for webpages, presentations, wrappers for different products, posters, etc.

Figure 11: Input image (top left) and the color harmonized results using template T (top right), template V (bottom left) and template X (bottom right), respectively.

On an average, harmonizing images of resolution 640×480 took around 3.8 seconds. When a fixed template t is selected for harmonization the same set of images, the average time was around 0.67 seconds. This time remains comparable even if we chose different templates from the given set of templates. We worked on test set of 15 images. When both the template t and the θ were fixed for the same set of

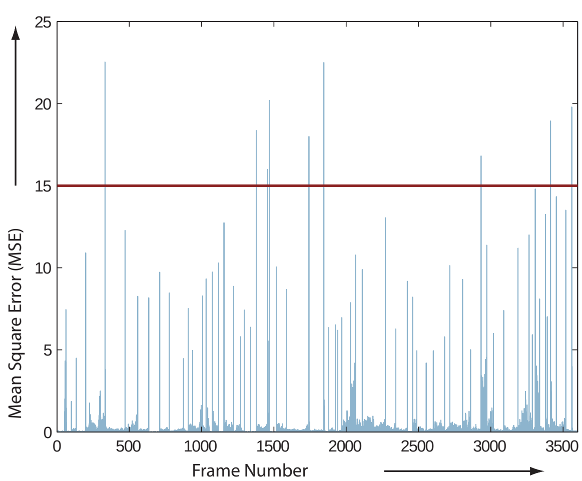

Figure 10: Mean square errors between two successive frames are thresholded using a suitable threshold, indicated in red, to identify I-frames.

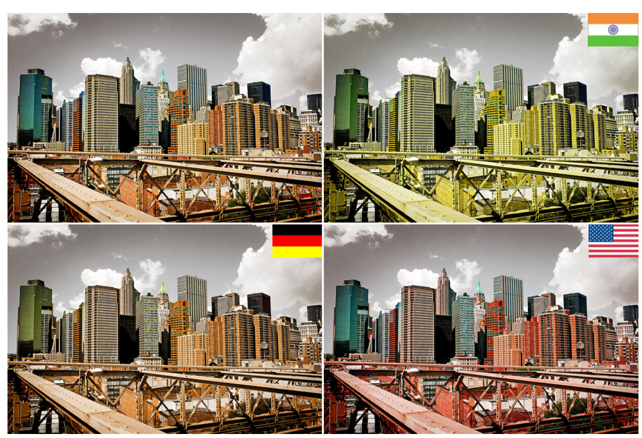

Figure 12: Adjusting the background color according to an input flag helps to maintain the harmony of the image even if the flag is added to the original image.

images, the average time went down to 0.14 seconds.

This technique can also used to improve the balance between background and foreground colors. As shown in the Figure 12, we adjust the color of a background image according to the foreground image, different flags in this example. One can also map colors between two images, as shown in Figure 13. We can transfer colors between two flowers, likewise we can use this technique for other objects too.

When it comes to videos, if we color harmonize each frame of the videos individually, we might introduce flickering (see Figure 6). This is caused by large differences in orientations or changes in the template across consecutive frames. Depending on the video sequence, such flickering may or may not be visible.

We found that our improved methods provide very good results for almost all videos. The resulting video frames are not only harmonized individually but consecutive frames are harmonized coherently so as to avoid flickering (Figure 8). It may happen that for some subset of sequences harmonization might create a bad result which can be corrected by manually selecting desired template and orientation for the given subset of frames. One can make an application for automatically harmonizing videos, while providing an intervention mode by which the user can manually tune parameters.

5. Conclusion

We presented a simple and fast technique for color harmonization of images building on state-of-the-art methods.

Figure 13: The flowers on the right derive its colors from the images on the left.

Figure 14: An original video frame (left), and the processed video frame after color harmonization (right).

For mid-sized images, our method gives close to real-time performance.

We also extended our technique to color harmonize videos. Video sequences can be more chaotic or disturbing since a number of image frames contribute to the impact of the video. Even a small subset of disturbing frames may create chaotic effects. Our technique is designed to reduce such effects. Using a joint space-time optimization, we harmonize each and every frame while maintaining coherency across successive frames.

In future work, we would like to give the user additional control to select subsections of videos in which she seeks harmonization. We also plan to incorporate user guidance for choosing a template and the corresponding θ-range for selected sections of the input video. Often the same object has slightly different colors in different scenes. If the user identifies such objects across frames, then we can also maintain color coherency in the object space.

References

[1] Daniel Cohen-Or, Olga Sorkine, Ran Gal, Tommer Leyvand, and Ying-Qing Xu. Color harmonization. *ACM Transactions on Graphics*, 25(3):624–630, 2006.

[2] Johannes Itten. *The Art of Color: the subjective experience and objective rationale of color*. Van Nostrand Reinhold, 1973.

[3] J. C. Maxwell. On the theory of compound colours. *Philosophical Transactions*, 150:57–84, 1860.

[4] J. C. Maxwell. Experiments on colour. *Transactions of the Royal Society Edinburgh*, 21:275–298, 1885.

[5] M. Tokumaru, N. Muranaka, and S. Imanishi. Color design support system considering color harmony. *Proc. of IEEE Fuzzy Systems*, 1:378–383, 2002.