

Style-aware robust color transfer

Hristina Hristova¹, Olivier Le Meur¹, Rémi Cozot¹ and Kadi Bouatouch¹

¹University of Rennes 1, France

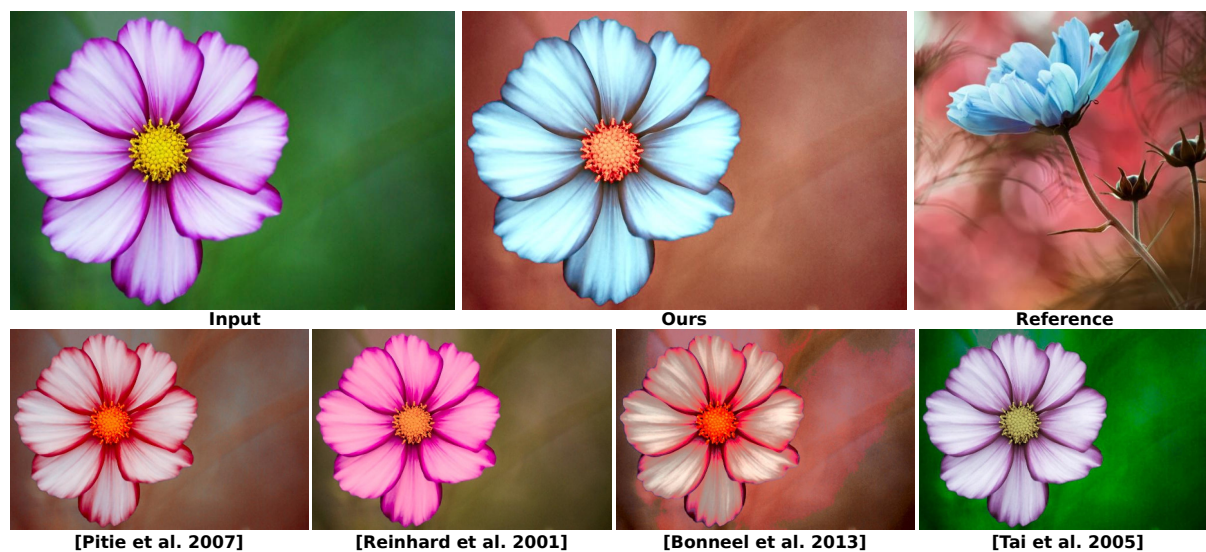


Figure 1: Comparison between our method and four state-of-the-art methods. The latter approaches aim at transferring the color proportion of the two main reference colors, i.e. red and blue. As shown, they do not respect the style of the reference image. On the other hand, our method succeeds in transferring the blue color of the reference flower to the input flower as well as the reddish reference background to the input background. This effect is achieved by color clustering and a proper mapping function between clusters.

Abstract

Transferring features, such as light and colors, between input and reference images is the main objective of color transfer methods. Current state-of-the-art methods focus mainly on the complete transfer of the light and color distributions. However, they do not successfully grasp specific light and color variations in image styles. In this paper, we propose a local method for carrying out a transfer of style between two images. Our method partitions both images to Gaussian distributed clusters by considering their main style features. These features are automatically determined by the classification step of our algorithm. Moreover, we present several novel policies for input/reference cluster mapping, which have not been tackled so far by previous methods. To complete the style transfer, for each pair of corresponding clusters, we apply a parametric color transfer method and a local chromatic adaptation transform. Results, subjective user evaluation as well as objective evaluation show that the proposed method obtains visually pleasing and artifact-free images, respecting the reference style.

1. Introduction

Color transfer aims at modifying the look of an original image considering the illumination and the color palette of a reference image. It can be employed for image and video enhancement by simulating the appearance of a given image or a video sequence [HLKK14, BSPP13]. It can also be applied to hallucinations of particular parts of the day [SPDF13].

The main objective of this paper is to propose an original way of carrying out a transfer of style between two images. In this paper, light and colors are the key features of a style. The style transfer strongly depends on these two characteristics. Although global methods produce results which respect the color and light distributions of the reference image, they do not grasp in a good manner some specific image feature variations in terms of light and colors. Indeed, in many cases, they produce results inconsistent with the style of the reference image and at odds with the concept of a balance between the different features in an image. For instance, such an imbalance may cause either over-saturated (under-saturated) images or images with a low visual quality.

Moreover, most global transformations rely on the assumption that the distributions of the input and reference images could be well-fitted by the multivariate Gaussian distribution. However, in general, the Gaussian assumption over the entire image distributions turns out to be too restrictive to ensure a good color transfer.

According to the paper of Reinhard et al. [RAGS01] and more recently, to the paper of Bonneel et al. [BSPP13], we can overcome these limitations with cluster-based techniques. In this case, color transfer is performed between clusters in images rather than between entire images. In this paper, we perform a local style-based approach in which Gaussian mixture models are used to fit the image distributions by partitioning the images into several clusters.

The main contributions of this paper are threefold:

- a novel automatic system that classifies an image as a light-based image or a colors-based image and determines the input/reference clusters;
- a cluster-based method consisting of several novel mapping policies between the input/reference clusters;
- a subjective user evaluation of the results as well as an objective evaluation consisting of two objective metrics.

The paper is organized as follows. Section 2 presents the related research work. Section 3 provides details on our classification system and introduces the three mapping policies which we have developed. Results, user evaluation and objective evaluation are presented in Section 4. Section 5 shows limitations and provides ideas for future research work. Finally, the last section concludes the paper.

2. Related work

Research works in the field of color transfer are classified by [FPC*14] into two main categories: geometry-based and statistical-based methods. Geometry-based approaches aim at finding content-based correspondences in pairs of images and ensure that these correspondences have the same colors.

We prefer statistical-based methods to geometry-based ones because the latter methods heavily depend on the image structure and therefore, it is not an easy task to apply them to image pairs of various contents. This paper focuses on statistical-based techniques in order to carry out transfer between images of various contents and styles. Generally, we could divide the statistical-based methods into two classes: non-parametric and parametric methods.

2.1. Non-parametric methods

The naive histogram matching is the first example of a non-parametric method. It attempts to borrow the thorough look of an image by matching two cumulative density functions. However, a full histogram transfer tends to be too harsh and results in artifacts. Therefore, recent works resolve this problem by matching histograms at different scales [PR10, PR11] and by performing gradient optimization [XM09]. Moreover, Pitié et al. [PKD05] match two 3D distributions through rotations and 1D histogram projections. Their method manages to entirely transform the input distribution to the reference distribution.

Recently, three papers have introduced new non-parametric approaches for colorizing an image [NKB14, FSDP14, HLKK14]. Hwang et al. [HLKK14] employ a non-linear interpolation using probabilistic moving least squares. Unlike Hwang et al. [HLKK14] who apply color transfer to images of the same scene, taken under different lighting conditions, Frigo et al. [FSDP14] designed their method for images with various contents. Similarly to [PKD07], Frigo's non-parametric approach is based on an optimal transportation problem and minimizes the overall displacement cost of the color mapping. However, due to smooth interpolation, Frigo's method could create colors which do not appear in the reference image. On the other hand, Nguyen et al. [NKB14] argue that a good color transfer is the one which does not produce out-of-gamut colors. Their illuminant-aware color transfer method is designed to respect the gamut of the reference image.

2.2. Parametric methods

Reinhard et al.'s [RAGS01] parametric color transfer method is specially well-designed for natural scenes. It assumes multivariate Gaussian distributions with diagonal covariance matrices for both the input and reference images. Another parametric method which adopts the multivariate Gaussian law was proposed by Pitié et al. [PKD07]. Their color mapping was presented as a closed form solution to the Monge-

Kantorovich optimization problem [Eva97]. Unlike Reinhard et al.’s approach [RAGS01], Pitié et al.’s [PKD07] method takes into account the dependencies between the channels of the employed color space.

Unlike the two latter methods which rely on a global transformation, local approaches cluster both images into several regions and build a local color map between each pair of corresponding regions. In the case of a mixture of Gaussian distributions, clustering can be carried out by the well-known technique Expectation Maximization (EM) [DLR*77, TJT05]. The paper of Bonneel et al. [BSPP13] introduces luminance-based clustering and employs Pitié’s color grading method [PKD07] to handle the color transfer. The mapping function between the clusters is based on the luminance component. Furthermore, Tai et al. [TJT05] apply a 3D EM algorithm to cluster the input and reference images. Like [BSPP13], they build a mapping function based on the mean luminance values of the input and reference images.

2.3. Limitations

Global color transfer methods rely on the assumption that the multivariate Gaussian distribution can fit the distributions of the entire input and reference images. This assumption turns out to restrict the color transfer, as it will be illustrated in Section 4.

On the other hand, local color transfer methods target only images with similar low-level characteristics. Moreover, they do not elaborate enough on the mapping functions linking the input and reference clusters. A good method should jointly consider light and color features.

Therefore, unlike the existing local methods in the field of color transfer, we take into account both characteristics in the clustering process and introduce novel style-based mapping policies which, to our knowledge, have not been considered so far.

3. Style-based cluster mapping

Our approach is designed to carry out a local color transfer on a given image by taking into account its main features as well as the main features of a reference image. We focus on light and colors as the two key features of image styles. These features are used in the process of clustering an input image I and a reference image J . As a result, we obtain a set of clusters for both images. Next, we propose several strategies to map the input and reference clusters. The mapping strategies and the number of clusters are automatically determined with regards to the main features of images in a pair. At the end, color transfer technique is applied to each pair of corresponding clusters to obtain the final result. The approach is illustrated in Figure 2 and the main notations used in this paper are given in Table 1.

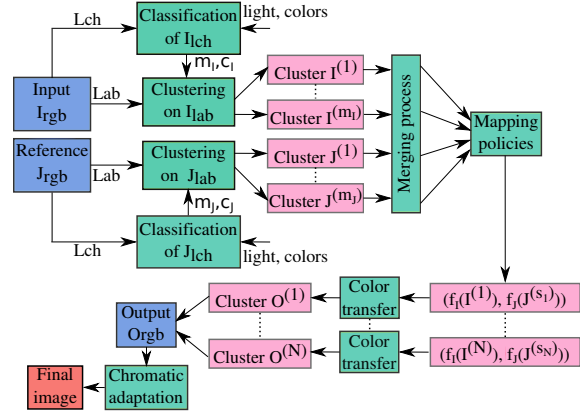


Figure 2: This framework illustrates the steps of the proposed approach. The blue boxes stand for images in the RGB color space and in violet are the image clusters. The green boxes illustrate the stages of the proposed approach.

3.1. Automatic light-based versus colors-based classification

The proposed automatic classification system determines which feature, among the two considered features, i.e. light and colors, carries more information about an image style. The modification of these two features influences the thorough look and perception of images [HLKK14]. Therefore, if we modify in a specific manner the light and the colors of an image, we will get close to a particular image style.

We propose to classify the images into two main categories: colors-based style images and light-based style images (refer to Figure 3 for illustration).

- Colors-based style images are images whose color information is sufficient enough to well-define at least two different and significant colors. An image of only one color is not considered as a colors-based style image. One color is not representative of the style of the image and there is a high probability that the light feature of that image will have a greater impact on its style.
- Images which are not classified as colors-based style images are classified as light-based style images because their light features are more meaningful than their color features.

The classification algorithm is performed in the Lch color space on both the input and reference images. It consists of three steps:

1. The image I_{rgb}^\dagger is first converted into the Lch color space

[†] In the description of the classification algorithm, I_{rgb} refers to either the input or reference images.

to obtain the image I_{lch} . The set G of gray points is extracted from the image I_{lch} as:

$$G : \{(l, c, h) \mid c < c_{min}, \forall (l, c, h) \in I_{lch}\} \quad (1)$$

where c_{min} is the threshold for the chroma component of the Lch color space.

- The hue histogram function $\phi : \Omega_\phi \subset \mathbb{R} \rightarrow \mathbb{R}$, where Ω_ϕ is the set of bin values in the Lch color space, is computed from the set of the remaining non-gray pixels denoted as $I_{lch} \setminus G$. We perform a linear search to obtain the local peaks of the hue histogram. Let P be the set of all the local peaks in the hue histogram and let δ define a small neighbourhood around a peak $p \in P$. Then, the set of significant peaks P_s is defined in the following way:

$$\forall p \in P : p \in P_s \Leftrightarrow \begin{cases} \sum_{\epsilon=-\delta}^{\delta} \phi(p+\epsilon) > s_{min}, \text{ and} \\ |p - p_s| > d_p, \forall p_s \in P_s \end{cases} \quad (2)$$

where s_{min} is the threshold for the minimum number of pixels defining a significant peak and d_p is the distance threshold between two significant peaks in the hue histogram.

- Finally, if $|P_s| > 1$, the original image is considered as a colors-based style image. Otherwise we classify the original image as a light-based style image.

We experimentally set the three parameters c_{min} , s_{min} and d_p to default values (refer to Section 3.5 for more details), although they can influence the final result as illustrated in Figure 3.

The advantages of our classification algorithm lie in its simplicity and effectiveness. We tried out the system on several image categories: natural scenes, city and street images, macro images, studio images and paintings. The first two rows of Figure 3 show images correctly classified by our classification system.

3.2. Clustering

Once the main features of the input and reference images are determined, a clustering is performed on both images. The optimal number of clusters is determined automatically. If an image is classified as a colors-based style image, the number of clusters is defined as the number of the significant peaks in its hue histogram, i.e. $|P_s|$. For a light-based style image, we adopt Bonneel’s idea [BSPP13] where three luminance clusters are considered, namely highlights, midtones and shadows. As a remark, there exist images for which one of these luminance clusters is not significant. In that case, we consider only two clusters as shown in Figure 4.

The clustering is performed by the EM algorithm in the Lch color space. Figure 4 illustrates the two types of image clustering. For light-based style images, we use their luminance histograms, because in the majority of the cases it is more meaningful than their hue histograms. On the other

Table 1: Notations. Ω_I and Ω_J denote the sets of spatial coordinates of the input and reference images respectively.

Notation	Definition
$CS \in \{rgb, lab, lch\}$	RGB, Lab or Lch
$I_{CS} : \Omega_I \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$	input image
$J_{CS} : \Omega_J \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$	reference image
$O_{CS} : \Omega_I \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$	output image
$R \in \{I, J\}$	image I or image J
m_R	number of clusters of R_{CS}
N	$\min(m_I, m_J)$
$R_{CS}^{(k)}$	k^{th} cluster of R_{CS}
$\zeta_{CS}^R = \{R_{CS}^{(1)}, \dots, R_{CS}^{(m_R)}\}$	set of clusters of R_{CS}
$\zeta_{CS}^R(v)$	v^{th} component of CS in ζ_{CS}^R
$f_{CS}^{(k)}$	pdf of $I_{CS}^{(k)}$
$g_{CS}^{(k)}$	pdf of $J_{CS}^{(k)}$
p_{ik}	probability of the i^{th} pixel to belong to $R_{CS}^{(k)}$
$O_{CS}^{(k)}$	k^{th} cluster of O_{CS}

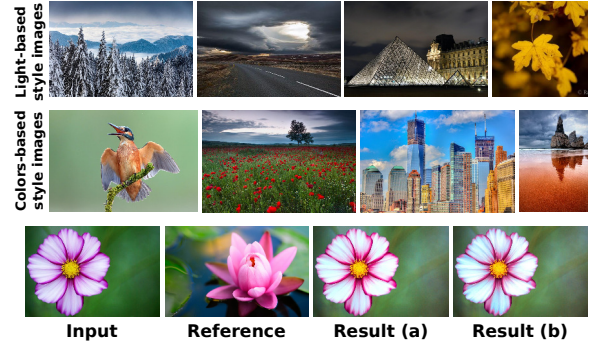


Figure 3: On the first two rows: images correctly classified as light-based style images and colors-based style images respectively. On the third row: result (a) is obtained with the default values of the three parameters of our classification system, whereas result (b) is obtained by setting d_p to half of its default value.

hand, the 2D Luminance-Hue distribution represents well the different color clusters for colors-based style images.

Note that the number of input and reference clusters can be different. In that case, if the number of input clusters is higher than the number of reference clusters, then the closest input clusters in terms of their means are merged, and vice versa. At the end of the clustering process, we obtain N clusters for both input and reference images.

3.3. Mapping policies

As a next step, a mapping function between the clusters has to be built. Hereafter, we introduce four mapping policies for different types of image pairs, namely Light to Colors, Colors to Light, Light to Light and Colors to Colors. Algo-

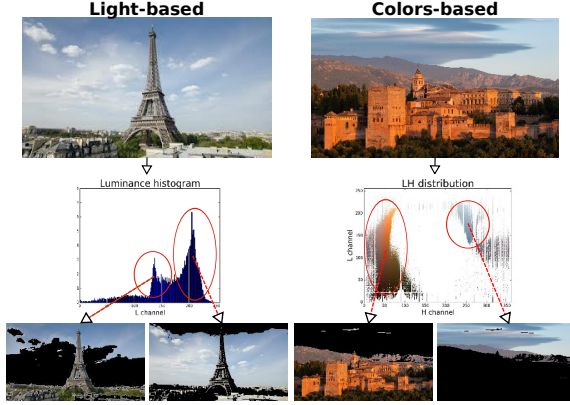


Figure 4: The luminance histogram of the light-based style image has two significant peaks and defines two luminance clusters. On the other hand, the 2D Luminance-Hue distribution of the colors-based style image clearly separates the two main colors of the image. Overlapping is included in the visualization of the clusters.

rithm 1 presents the global structure of our mapping strategies. In the following subsections, we present the mapping policies.

Algorithm 1 Mapping policies

```

1: for  $k = 1, \dots, N$  do
2:   if Light to Colors then
3:      $[s_I, f_{lab}^{(s_I)}] = \text{FindDarkestCluster}(\zeta_{lch}^I(l))$ 
4:      $[s_J, g_{lab}^{(s_J)}] = \text{FindColdestCluster}(\zeta_{lch}^J(h))$ 
5:   end if
6:   if Colors to Light then
7:      $[s_I, f_{lab}^{(s_I)}] = \text{FindColdestCluster}(\zeta_{lch}^I(h))$ 
8:      $[s_J, g_{lab}^{(s_J)}] = \text{FindDarkestCluster}(\zeta_{lch}^J(l))$ 
9:   end if
10:  if Light to Light then
11:     $[s_I, f_{lab}^{(s_I)}] = \text{FindDarkestCluster}(\zeta_{lch}^I(l))$ 
12:     $[s_J, g_{lab}^{(s_J)}] = \text{FindDarkestCluster}(\zeta_{lch}^J(l))$ 
13:  end if
14:  if Colors to Colors then
15:     $[s_I, s_J, f_{lab}^{(s_I)}, g_{lab}^{(s_J)}] = \text{FindMinDistPair}(\zeta_{lch}^I(h), \zeta_{lch}^J(h))$ 
16:  end if
17:   $[O_{lab}^{(k)}] = \text{PerformTranspOnAB}(f_{lab}^{(s_I)}, g_{lab}^{(s_J)})$ 
18:   $\zeta_{lch}^I \setminus J_{lch}^{(s_I)}$ 
19:   $\zeta_{lch}^J \setminus J_{lch}^{(s_J)}$ 
20: end for
21:  $[O_{rgb}^{final}] = \text{CATLocal}(O_{rgb}, J_{rgb})$ 

```

3.3.1. Light to Colors

Light to Colors policy is designed for images with distant styles. The input image is a light-based style image, whereas the reference image is a colors-based style image. As it will be presented in Section 4, such test cases are challenging

for the state-of-the-art methods. To deal with this issue, we developed a meaningful mapping function which links the light features from the input image to the color features from the reference image.

Usually people expect cold colors to be present in the shadows of an image, whereas warm colors are likely to appear as highlights. Additionally, the majority of photographers use the same approach as an artistic effect [Wis13]. They use cold colors to indicate shadows or background, and warm colors to highlight bright areas. Our Light to Colors mapping function is based upon these two arguments.

The policy begins with indicating which cluster among the set of input clusters ζ_{lch}^I has the minimum average luminance value. This problem is handled by the function $\text{FindDarkestCluster}(\zeta_{lch}^I(l))$, the output of which is the index s_I of the input cluster with the minimum average luminance value, and the pdf $f_{lab}^{(s_I)}$. Similarly, the function $\text{FindColdestCluster}(\zeta_{lch}^J(h))$ returns the index s_J of the reference cluster with the maximum average hue value, and the pdf $g_{lab}^{(s_J)}$. Our algorithm adopts standard hue wheel [Sto13]. The warmest color among a set of colors is defined as the one with the lowest hue value, whereas the coldest color as the one with the highest hue value. To this end, the function $\text{FindColdestCluster}(\zeta_{lch}^J(h))$ finds the reference cluster associated with the coldest color in the set of clusters ζ_{lch}^J .

Once computed, both probability functions $f_{lab}^{(s_I)}$ and $g_{lab}^{(s_J)}$, defined in the Lab color space, are passed to the function $\text{PerformTranspOnAB}()$. This function carries out the color transformation between the clusters $I_{lab}^{(s_I)}$ and $J_{lab}^{(s_J)}$. Finally, the latter clusters are removed respectively from the sets ζ_{lch}^I and ζ_{lch}^J . We repeat the procedure for each input cluster.

At the end of algorithm 1, our Light to Colors policy will have mapped the warmest colors of the reference image to the highlights of the input image and reversely, the coldest colors of the reference image to the shadows of the input image. Given this mapping function, we are able to carry out a transfer between a light-based style image and a colors-based style image. To the authors' best knowledge, there exists no other concrete explanation on how the mapping function between light features and color features could be held out. Therefore, our Light to Colors policy offers the first solution to that kind of mapping problem.

3.3.2. Colors to Light

Now, let us consider a colors-based style input image and a light-based style reference image. A policy that maps light features to color features has not been tackled by existing methods so far. We propose a strategy similar to Light to Colors strategy. Likewise, Colors to Light policy maps the highlight areas from the reference image to the warmest colors from the input image and vice versa, the darkest areas from the reference image correspond to the coldest colors from the input image.

3.3.3. Light to Light

The mapping between two light-based style images is handled by Bonneel’s [BSPP13] luminance-based mapping. Three luminance clusters are considered, namely shadows, midtones and highlights. There are images for which one of the clusters is insignificant. In that case, only two luminance clusters are considered. Light-to-Light policy maps shadows to shadows, midtones to midtones and highlights to highlights.

3.3.4. Colors to Colors

The last mapping policy has to map the clusters of two colors-based style images. There are several logical solutions to that kind of problem. Bonneel’s [BSPP13] approach applies luminance-based mapping for such kind of transfer. Nonetheless, for a style transfer between two colors-based style images, Bonneel’s method does not produce good results because it does not take into account the colors in the mapping process. As another solution, the clusters with the closest hue values could be mapped together. However, such an approach would not prevent the input clusters from being mapped only to one of the reference clusters. In like manner, we could also sort the input and reference clusters by their corresponding hue values and map them respectively. Conversely, we believe that users would expect similar colors to be mapped to similar colors which is not guaranteed by the latter mapping approach. We have experimented with these three ideas and finally, we came up with a more general solution described hereafter.

At each step of the algorithm, we map the two most similar input/reference clusters. They are those with the minimal Euclidean color distance between the centers. Put differently, the most similar cluster centers are nearest to each other in terms of hue values. $FindMinDistPair(\zeta_{lch}^I(h), \zeta_{lch}^J(h))$ is the function responsible for finding the most similar clusters $I_{lab}^{(sr)}$ and $J_{lab}^{(sj)}$ and their probability functions $f_{lab}^{(sr)}$ and $g_{lab}^{(sj)}$ at each iteration of Algorithm 1. The strategy ensures one-to-one mapping where the nearest colors are associated first. Like in the last two policies, the corresponding probability functions are passed to the function $PerformTranspOnAB()$, handling the color transformation[‡]. Finally, we exclude elements from the sets $\zeta_{lch}^I(h)$ and $\zeta_{lch}^J(h)$ at the end of each step. Therefore, the final stage of the algorithm will map the two most distant colors.

3.4. Color transfer method

Once we have mapped the input/reference clusters, a color transfer is performed between each pair of corresponding clusters. The color grading method consists

[‡] Whatever the used policy, the color transformation is always performed on the chroma axes (a and b) of the 3D distributions $f_{lab}^{(sr)}$ and $g_{lab}^{(sj)}$.

of a color transformation and a local chromatic adaptation. The color transformation is handled by the function $PerformTranspOnAB(f_{lab}^{(sr)}, g_{lab}^{(sj)})$ in Algorithm 1. Moreover, we carry out the color transformation in the CIE Lab color space. We separate the luminance channel from the chroma channels. Indeed, the human eye is much more sensitive to changes in the light conditions than to changes in the colors. Therefore, we apply the color transformation only on the chroma channels (whatever the determined policy). Finally, we use local chromatic adaptation transform (CAT), handled by the function $CATLocal(O_{rgb}, J_{rgb})$ in Algorithm 1, to reproduce the light of the reference image. These steps are discussed in the following subsections.

3.4.1. Color transformation on the chroma channels

The clustering step of our method performs a partitioning of the input and reference images into homogeneous clusters in the Lab color space. Their 2D ab distributions can be modeled by Gaussian distributions. Therefore, a parametric color transfer approach is used for carrying out the color transfer between the cluster ab distributions.

We adopted the parametric color grading technique proposed by Pitié et al. [PKD07]. Pitié’s method aims at building a mapping $t(I)$ between the image I and the image J . The mapping $t(I)$ transforms the input distribution $f(I)$ to the target distribution $g(J)$. To build the mapping $t(I)$, Pitié et al. assume that $f(I)$ and $g(J)$ follow a multivariate Gaussian law. For each corresponding pair of input/reference Gaussian clusters $I_{lab}^{(k)}$ and $J_{lab}^{(k)}$, we build a mapping $t_k(I_{lab}^{(k)})$ consistent with the proposed mapping by Pitié et al. [PKD07]. The mapping is derived as the closed form solution [OR93, DL82, PKD07] to an optimal transportation problem well-known as Monge-Kantorovich optimization problem [Eva97]. That way, for each pair of clusters, we build a unique mapping which minimizes the overall cost of the color transfer [PKD07].

3.4.2. Overlapping

When using a clustering technique, we need to take care of the strong color difference which may occur between the clusters. To achieve a smooth transition between the clusters and to lessen the visibility of eventual artifacts caused by the color grading, we let the input clusters overlap around their spatial boundaries. Associating pixels with more than one cluster is known as fuzzy (soft) clustering [Yan93, NN06]. Each pixel i is assigned a probability p_{ik} to belong to a cluster k . The formula is as follows:

$$p_{ik} = \alpha_{ik} / \sum_{j=1}^N \alpha_{ij} \quad (3)$$

where N is the number of clusters and $\sum_{j=1}^N p_{ij} = 1$. Additionally, α_{ik} is defined as $\alpha_{ik} = \exp(-D_M^2(x_i, f_{CS}^{(k)}))$, where

$D_M(x_i, f_{CS}^{(k)})$ is the Mahalanobis distance [DMJRM00] for 3D vector x_i of values in the Lab color space for the i^{th} pixel. Mahalanobis distance measures the distance of each overlapping pixel to one of Gaussian cluster distributions in our model. Finally, the values of the a and b channels for the output image O_{lab} are computed in the following manner:

$$O_i = \sum_{j=1}^N p_{ij} t_{ij} \quad (4)$$

where t_{ij} is the vector of chroma values for the i^{th} pixel, obtained from the transformation for the j^{th} cluster. O_i is the vector of new chroma values for the i^{th} pixel.

3.4.3. Local chromatic adaptation

Finally, we need to reproduce the light of the reference image. To this end, Bonneel et al. [BSPP13] apply naive histogram matching on the luminance channel. The naive histogram matching for the luminance channel tends to cause artifacts and high-saturated final results between images with very different lighting set-ups. Consequently, we consider that their approach would not be suitable for our purposes.

That is why, as a final stage of the color grading technique of our method, a local CAT algorithm is applied to the image O_{rgb} obtained with the color transformation. Local CAT aims at adapting pixel-wisely the colors of image O_{rgb} to the reference illuminant. This way, undesired color saturation is avoided and naturalism is preserved. Similarly to iCAM algorithm [KJF07], we apply CAT locally to the pixels of the input image by building a “white image” using Gaussian low-pass filter. Each input pixel is influenced by the chromatic transform and therefore, local luminance variations are captured effectively and reproduced in the result. Indeed, this approach enhances the contrast and prevents the image from becoming flat (refer to Example 3 and Example 5 in Figure 8).

Furthermore, Frigo et al. [FSDP14] were the first to apply the CAT algorithm iteratively. Instead of adapting the colors of an image to a well-known illuminant, they have used a global estimation of both the input and reference white points by assuming Gray World [HCWxW06]. Similarly, we estimate the reference illuminant as the average value of the non-gray pixels of the reference image [HCWxW06, FSDP14]. Unlike [FSDP14], we apply local CAT only once and not in an iterative manner, which results in a decrease of the computational time. As shown in Figure 5, local CAT adapts the colors of image O_{rgb} to the reference illuminant better than global CAT. Therefore, local CAT manages to better reproduce the colors of the reference image.

For all of the results in this paper, local CAT has been applied in the LMS color space. The adaptation factor in CAT was scaled by 0.3 [KJF07] and the value for the surround factor was set to 1.

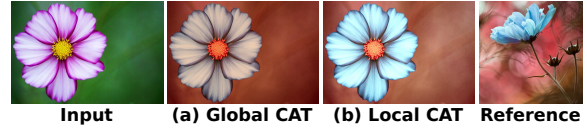


Figure 5: Comparison between global CAT and local CAT. Result (a) was obtained in iterative way, using global CAT [FSDP14] on image O_{rgb} .

3.5. Implementation details

The implementation of our algorithm begins with an image classification into two classes. Our feature detection system depends on three parameters whose default values are as follows: $c_{min} = 10$, $d_p = 30$, $s_{min} = 0.05 \times n$, where n is the number of pixels in the image. They are fixed throughout the computation. The default values were determined after several experiments.

Furthermore, the image classification and clustering are handled in the Lch color space. At the same time, we use the Lab color space to carry out the color transfer. We have implemented our algorithm in C++. The proposed algorithm has been performed on a laptop with an Intel Core i7 2.10GHz and 16Go RAM. For an image of 1000x1000 pixels, the average execution time is 25s (without optimization).

4. Results and evaluation

We compared results, obtained with the proposed method, to results, obtained with four other methods. On one hand, we chose the two state-of-the-art global transformations by Pitié [PKD07] and Reinhard [RAGS01]. On the other hand, we chose Bonneel’s and Tai’s local color transfer methods [BSPP13, TJT05] to show that a luminance-based mapping is not enough to ensure a good color transfer.

Figure 8 shows several results. As observed from the teaser example and Example 2 in Figure 8, the proposed method manages to transfer properly the colors of the foreground and the background without the need of segmentation [BSPP13] or saliency [FSDP14]. All of our results respect the semantic of the input image and the style of the reference image. As a result of the local CAT algorithm, applied at the end, over-saturated (under-saturated) images are unlikely to be produced by the proposed method.

4.1. User study evaluation

To compare the five methods in terms of style transfer and visual pleasingness of the results, we conducted a subjective evaluation study. We asked 15 users to evaluate 50 results obtained for 10 input and 10 reference images for the five methods. The input and reference images vary in content, semantics, lighting set-up, color features. The images used

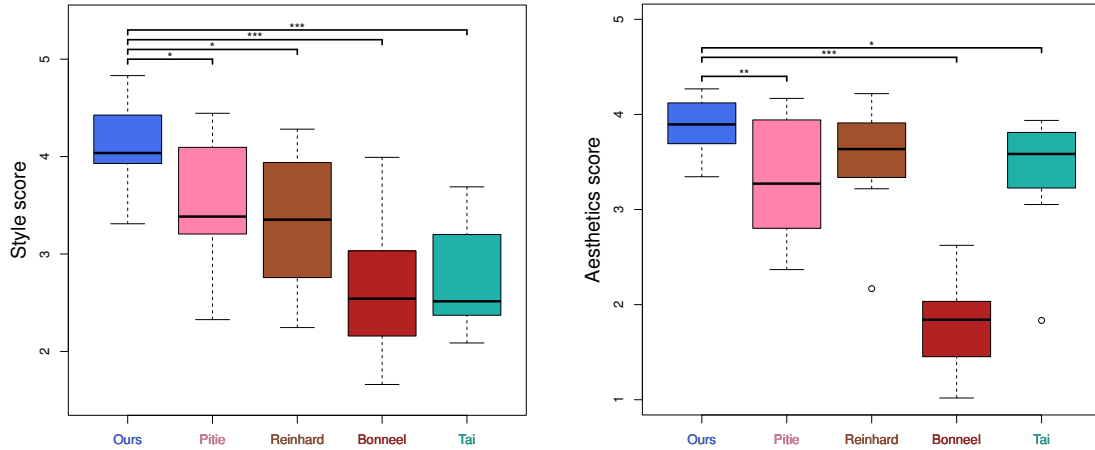


Figure 6: Box-and-Whiskers plots per method, displaying the distributions of the style and aesthetics scores. The connecting brackets indicate significant difference between our method and state-of-the-art methods (“*” stands for p-value ≤ 0.05 , “**” stands for p-value ≤ 0.01 and “***” stands for p-value ≤ 0.001).

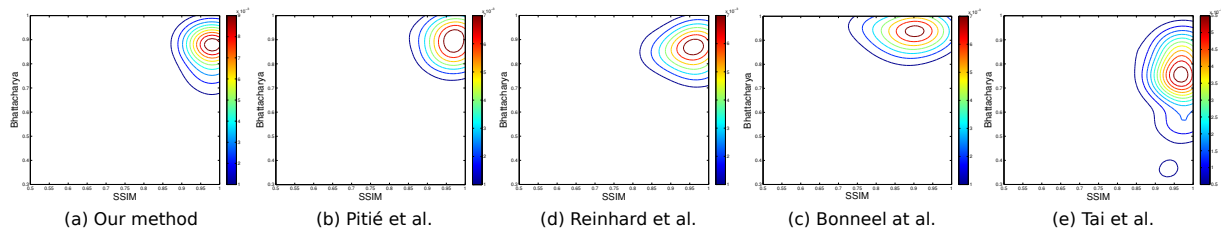


Figure 7: Contour plots for the set of metric pairs (SSIM, Bhattacharya) for each method. The majority of images has values which lie in the small areas displayed in red.

as test cases in this paper were selected from the various photographic collections of *500pixels.com* [§].

The participants were 23 to 53 years old. The majority of them had average image editing expertise. Five of them were wearing glasses and none of them was suffering from color vision deficiency. Each participant was presented with triples of images consisting of an input image, a reference image and a result, obtained from a color transfer between the input and reference images. Furthermore, we added a special image triple, referred to as baseline, to the set of the 50 image triples. The reference image in the baseline is different from the reference image used for obtaining the result in the baseline. That way we tested users’ judgments on the results.

First, the participants were asked to evaluate the match in styles (in terms of colors and light) between the result and the reference image with regards to their expectations. Second, the users were asked to evaluate how visually pleasing the results were. The evaluation of the visual pleasingness

was based on users’ perception of the aesthetics of the results. Five-point scale (5-excellent, 4-good, 3-acceptable, 2-poor, 1-bad) was used to evaluate the results for both tasks. Moreover, four repetitions per result were used to minimize any possible bias and to increase the robustness.

Each user evaluated the results individually. The indoor conditions, the display properties and the relative proximity of the user to the display were the same. In contrast, the order of displaying the image triples was random and different for each participant. Finally, a short training session took place before the real evaluation during which the users adapted to the tasks of the real test.

The scores of the four repetitions for each result were combined into a final score. Three different approaches were used to compute these final scores: average value, weighted value (giving different weights to each repetition) and median value. Paired t-tests for the final scores obtained with the three approaches have shown insignificant difference between them. Therefore, for all the analysis in this paper, average scoring was used. Finally, the average scores per user were normalized by the baseline scores to obtain the final scores per image.

[§] The input and reference images from Example 4 in Figure 8 were borrowed from the paper of Pitié et al. [PKD07].

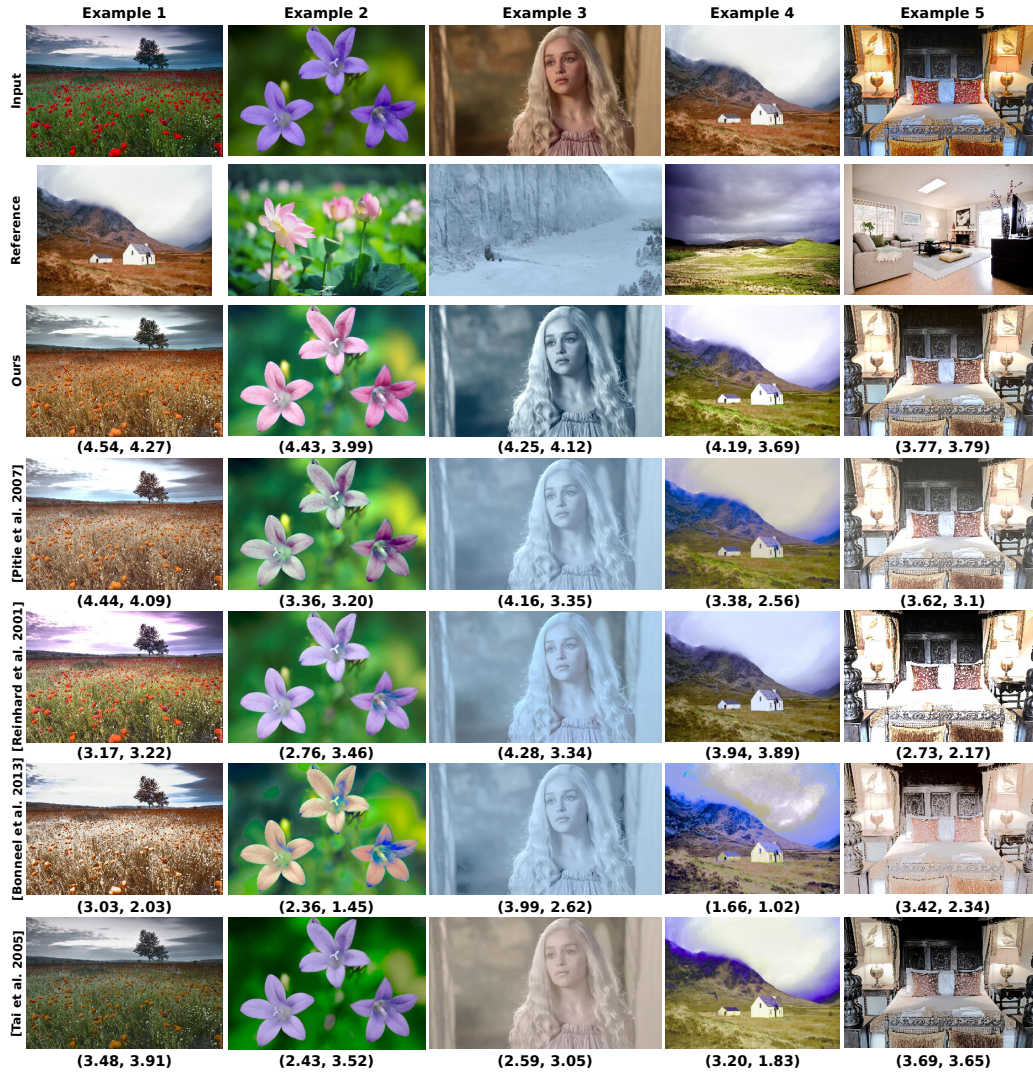


Figure 8: Results of our method and four state-of-the-art methods. The style and aesthetic scores are presented below each result (in the order: (style, aesthetics)). The optimal value for both scores is 5. Our method obtains the highest values regarding the visual pleasingness because it produces natural results with perceptually pleasing contrast. Likewise, the participants in the user study have given our results the highest style scores. An exception is Example 3 for which Reinhard et al.’s method [RAGS01] has obtained a slightly higher style score than the style score for our method.

We refer to the scores for the first and second questions respectively as style and aesthetic scores. Figure 6 displays the Box-and-Whiskers plots per method for these two types of scores. As observed, the mean values for both style and aesthetic scores are the highest for our method. Moreover, there is less deviation in both types of scores for our approach in comparison to the other methods.

Furthermore, we performed paired t-tests between the two types of scores for our method and for the other four methods. These tests have shown significant difference in the means of the style scores between our method and each

Table 2: Correlations between style and aesthetics scores.

Method	Ours	Pitie	Reinhard	Bonneel	Tai
Correlation	0.86	0.49	0.50	0.93	-0.05

of the other four approaches. Similarly, the tests have indicated that the aesthetic scores of our method differ significantly from the aesthetic scores of three of the state-of-the-art methods ([PKD05, BSPP13, TJT05]). Conversely, there is an insignificant difference between the aesthetic scores of

the proposed method and those of the method by Reinhard et al. [RAGS01]. Both methods obtain high aesthetic scores. To this end, both of them tend to produce visually pleasing images, as seen from the examples in the paper.

Additionally, as shown in Table 2, the style and aesthetic scores for our method are highly correlated, indicating that the proposed method produces consistently good results in terms of style transfer and visual pleasingness.

4.2. Objective metrics evaluation

In addition to the subjective user evaluation, an objective evaluation of the results was also carried out. A good color transfer has to ensure artifact-free images as well as to transfer properly the light and color distributions of the reference images. We believe that both criteria are equally important. Therefore, for the objective evaluation we use two objective metrics.

As depicted in [FSDP14, HLKK14], the metrics **SSIM** [WBSS04] can be used to measure the degree of artifacts in the final result. We apply this similarity metrics to the luminance channels of the input and final images. However, as the goal of a color transfer method is to transform the input luminance to the reference luminance, the input luminance differs from the luminance of the final result. That is why, unlike [FSDP14, HLKK14], we removed the luminance component of SSIM from the computation of the metrics and proceeded with only the structural and contrast components.

The **Bhattacharya coefficient** [Bha46] is used to measure the distance between two histograms [ATR98]. To evaluate how successful the color transfer is, we applied it to the color and luminance histograms of the result and the reference image in the Lab color space. The final metric value is obtained by averaging the scores for the luminance channel and the two chroma channels.

The pair of values (1, 1) is optimal for the pair (SSIM, Bhattacharya). It refers to results with the same visual quality as that of the input image and with exactly the same light and color distributions as those of the reference image.

For each method and for each image in a set of 40 image results per method (10 of which were used also in the user study), we plotted the pair of values for both SSIM and Bhattacharya coefficient. Then, contour plots were used to illustrate the density of the set of pairs for each method as shown in Figure 7. Several observations can be made.

First, we observe that our approach obtains the highest average value for SSIM axis as presented in Table 3. Moreover, paired t-tests have shown that SSIM values, obtained with our method, differ significantly from those obtained with the other four methods. Therefore, the proposed approach is the best one among the five methods when it comes to producing artifact-free final images.

Furthermore, the centers of the contour plots for our

Table 3: Mean values of SSIM and Bhattacharya coefficient, obtained for each method (a)-(e) (see Figure 7).

Metrics / Mean values	(a)	(b)	(c)	(d)	(e)
SSIM	0.98	0.97	0.95	0.89	0.96
Bhattacharya	0.86	0.88	0.85	0.92	0.73

method and Pitié’s approach are concentrated around the optimal value (SSIM, Bhattacharya) = (1, 1). Likewise, in [BSPP13, RAGS01], the authors transfer effectively the colors and the light. Although Bonneel et al. [BSPP13] obtain the highest mean value for Bhattacharya coefficient, the method is likely to cause significant number of artifacts to the final images. On the other hand, the two metrics show that Tai’s method is expected to produce less artifacts but it is less efficient in terms of color transfer. Finally, there is no significant difference between our method and both Pitié’s and Reinhard’s methods regarding the Bhattacharya coefficients.

To conclude, the proposed method succeeds in transferring the color and light distributions of the reference image with respect to the reference style while preserving the naturalism in the results and keeping the degree of artifacts low.

5. Limitations and future work

The results as well as the subjective and objective evaluations, presented in this paper, have pointed out that our approach outperformed results obtained with state-of-the-art techniques. Nevertheless, our method has also some limitations. Figure 9 shows one of them. According to the participants of our user study, the result in Figure 9 is far from their expectations due to the orange color of the buildings in the reference image which has not been transferred to the buildings in the input image. Therefore, this result is the lowest scored image, obtained with our method. Moreover, the state-of-the-art methods also fail to properly color the buildings in the input image. This challenging case could be solved by using additional constraints such as saliency [FSDP14] and content-based transfer [WDK*13].

There is even more room for future improvements. For instance, our automatic classification system depends on three parameters which, for now, are set to default values. We would like in the future to elaborate on the problem of finding the optimal values for each image as well as to enrich the image characteristics of the detection system. Finally, we are willing to go further by tackling the connection between the subjective judgment on the results and their objective evaluation.

6. Conclusion

Our work focused on developing a new way for style transfer for a wide class of image pairs. We introduced cluster-based style transfer method which outperforms state-of-the-art approaches. Furthermore, we developed an automatic way of

feature detection in images for the two image characteristics: light and colors. Our most important contribution lies in the development of three mapping policies which make it possible to carry out a local color transfer between pairs of images with various style features. Our method manages to solve some of the open questions in research and opens new venues for improvement.

Additional materials and examples can be found on the following web page: <http://people.irisa.fr/Hristina.Hristova>.

References

- [ATR98] AHERNE F. J., THACKER N. A., ROCKETT P. I.: The bhattacharyya metric as an absolute similarity measure for frequency coded data. *Kybernetika* 34, 4 (1998), 363–368. 10
- [Bha46] BHATTACHARYYA A.: On a measure of divergence between two multinomial populations. *Sankhyā: The Indian Journal of Statistics* (1946), 401–406. 10
- [BSPP13] BONNEEL N., SUNKAVALLI K., PARIS S., PFISTER H.: Example-based video color grading. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2013)* 32, 4 (2013), 2. 2, 3, 4, 6, 7, 9, 10
- [DL82] DOWSON D., LANDAU B.: The frechet distance between multivariate normal distributions. *Journal of multivariate analysis* 12, 3 (1982), 450–455. 6
- [DLR*77] DEMPSTER A. P., LAIRD N. M., RUBIN D. B., ET AL.: Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal statistical Society* 39, 1 (1977), 1–38. 3
- [DMJRM00] DE MAESSCHALCK R., JOUAN-RIMBAUD D., MASSART D. L.: The mahalanobis distance. *Chemometrics and intelligent laboratory systems* 50, 1 (2000), 1–18. 7
- [Eva97] EVANS L. C.: Partial differential equations and monge-kantorovich mass transfer. *Current developments in mathematics 1999* (1997), 65–126. 3, 6
- [FPC*14] FARIDUL H. S., POULI T., CHAMARET C., STAUDER J., TRÉMEAU A., REINHARD E., ET AL.: A survey of color mapping and its applications. In *Eurographics 2014-State of the Art Reports* (2014), The Eurographics Association, pp. 43–67. 2
- [FSDP14] FRIGO O., SABATER N., DEMOULIN V., PIERRE H.: Optimal transportation for example-guided color transfer. *12th Asian Conference on Computer Vision (ACCV)* (2014). 2, 7, 10
- [HCWxW06] HUO J., CHANG Y., WANG J., XIA WEI X.: Robust automatic white balance algorithm using gray color points in images. *IEEE Trans. Consumer Electronics* 52, 2 (2006), 541–546. URL: <http://dblp.uni-trier.de/db/journals/tce/tce52.html#HuoCWW06>. 7
- [HLKK14] HWANG Y., LEE J.-Y., KWEON I. S., KIM S. J.: Color transfer using probabilistic moving least squares. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on* (2014), IEEE, pp. 3342–3349. 2, 3, 10
- [KJF07] KUANG J., JOHNSON G. M., FAIRCHILD M. D.: icam06: A refined image appearance model for hdr image rendering. *Journal of Visual Communication and Image Representation* 18, 5 (2007), 406–414. 7
- [NKB14] NGUYEN R., KIM S., BROWN M.: Illuminant aware gamut-based color transfer. In *Computer Graphics Forum* (2014), vol. 33, Wiley Online Library, pp. 319–328. 2
- [NN06] NOCK R., NIELSEN F.: On weighting clustering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 28, 8 (2006), 1223–1235. 6
- [OR93] OLKIN I., RACHEV S.: Maximum submatrix traces for positive definite matrices. *SIAM journal on matrix analysis and applications* 14, 2 (1993), 390–397. 6
- [PKD05] PITIE F., KOKARAM A. C., DAHYOT R.: N-dimensional probability density function transfer and its application to color transfer. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on* (2005), vol. 2, IEEE, pp. 1434–1439. 2, 9
- [PKD07] PITIE F., KOKARAM A. C., DAHYOT R.: Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding* 107, 1 (2007), 123–137. 2, 3, 6, 7, 8
- [PR10] POULI T., REINHARD E.: Progressive histogram reshaping for creative color transfer and tone reproduction. In *Proceedings of the 8th International Symposium on Non-Photorealistic Animation and Rendering* (2010), ACM, pp. 81–90. 2
- [PR11] POULI T., REINHARD E.: Progressive color transfer for images of arbitrary dynamic range. *Computers & Graphics* 35, 1 (2011), 67–80. 2
- [RAGS01] REINHARD E., ASHIKHMEN M., GOOCH B., SHIRLEY P.: Color transfer between images. *Computer Graphics and Applications, IEEE* 21, 5 (2001), 34–41. 2, 3, 7, 9, 10
- [SPDF13] SHIH Y., PARIS S., DURAND F., FREEMAN W. T.: Data-driven hallucination of different times of day from a single outdoor photo. *ACM Transactions on Graphics (TOG)* 32, 6 (2013), 200. 2
- [Sto13] STONE M.: *A field guide to digital color*. CRC Press, 2013. 5
- [TJT05] TAI Y.-W., JIA J., TANG C.-K.: Local color transfer via probabilistic segmentation by expectation-maximization. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (2005), vol. 1, IEEE, pp. 747–754. 3, 7, 9
- [WBSS04] WANG Z., BOVIK A. C., SHEIKH H. R., SIMONCELLI E. P.: Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on* 13, 4 (2004), 600–612. 10
- [WDK*13] WU F., DONG W., KONG Y., MEI X., PAUL J.-C., ZHANG X.: Content-based colour transfer. In *Computer Graphics Forum* (2013), vol. 32, Wiley Online Library, pp. 190–203. 10
- [Wis13] WISSLAR V.: *Illuminated Pixels: The Why, What, and How of Digital Lighting*. Cengage Learning, 2013. 5
- [XM09] XIAO X., MA L.: Gradient-preserving color transfer. In *Computer Graphics Forum* (2009), vol. 28, Wiley Online Library, pp. 1879–1886. 2
- [Yan93] YANG M.-S.: A survey of fuzzy clustering. *Mathematical and Computer modelling* 18, 11 (1993), 1–16. 6



Figure 9: Failure case: the orange color of the buildings in the reference image has not been transferred to the buildings in the result, as expected by the majority of the participants in our user study.