**Region-Based Color Transfer for Complex Content Images Using Intrinsic Component**

Wan-Chien CHIOU, Yi-Lei CHEN, and Chiou-Ting HSU*

*Department of Computer Science, National Tsing Hua University,
No. 101, Sec. 2, Kuang Fu Road, Hsinchu 30055, Taiwan

Received May 16, 2012; final version accepted August 6, 2012

This paper proposes an automatic color transfer method for complex content images. When given one or more high-quality reference images, our goal is to determine a set of best reference colors for transferring their color characteristics into the target image. Although several automatic color transfer methods have been proposed, there usually exists visible and unnatural artifact when processing images with complex content and lighting variation. In this paper, we represent each image in region level and propose to incorporate region attribute, region connectivity and intrinsic reflectivity to characterize the local organization within an image. We then determine the best-matched reference region for each target region using the proposed graph-theoretic region correspondence estimation. After determining the set of reference regions, we next conduct color transfer between the best-matched region pairs in a de-correlated color space. In order to reduce artifact across complex regions, we further propose a weighted color transfer in terms of intrinsic component. Both subjective and objective evaluation of our experiments demonstrates that the proposed method outperforms existing methods.

KEYWORDS: color transfer, region correspondence, intrinsic component

1. Introduction

The goal of color transfer is to transfer the color characteristics from one or more images into another so as to retain the original feel or aura of reference images in the target image. Applications of color transfer include colorization [1], color correction [2], non-photorealistic image rendering [3], and image quality improvement [4].

Existing color transfer methods can be divided into two categories: global and local methods. The global color transfer method [5] treats the whole image as a single entity and linearly transforms the color distribution from one reference image into the target image in the de-correlated $L_aB$ color space [6]. However, because the global method uses the same color transformation to process the whole image content, this method usually fails to process images with complex content or multiple color characteristics. On the other hand, local methods conduct color transfer for each region individually. For example, the method in [7] first performs probabilistic segmentation in the color space and then applies color transformation independently in each segmented region. Thus, the results obtained by local methods are usually of more colorfulness than by global methods. Nevertheless, the performance of both global and local methods heavily depends on the similarity between the reference and target images. Once the reference and target images contain dissimilar color distributions, these methods usually result in unnatural or over-saturated artifact.

Therefore, in [8], instead of using only one reference image, the authors proposed to select the best-matched reference region from multi-source images for each target region. However, the 3D GMM model adopted in [8] usually results in unsatisfactory segmentation result, which is a compromise between color and spatial connectivity. In addition, since the best-matched region pairs are determined independently in terms of only their corresponding color distributions, this method [8] usually results in unnatural color transition across regions.

From the above discussion, local color transfer from multiple reference images indeed improve the performance by considering local color characteristic and including more reference colors from multiple reference images. However, the performance could be further improved if spatially adjacent relationship is also included while determining the best-matched reference regions. Therefore, in this paper, we aim to propose a method to automatically transfer colors from multiple reference images by including spatial relationship between regions to determine the best-matched reference regions. To reduce the computational complexity, we first adopt content-based image retrieval technique to obtain a small number of high-quality images as the references. Next, we represent each image in region level as an attributed graph and adopt the graph-theoretic region correspondence estimation to determine the best-matched reference region for each target region. While conducting color transfer on images containing delicate lighting variation, we may still

* Corresponding author. E-mail: cthsu@cs.nthu.edu.tw
obtain poorly matched region pairs even with the graph-theoretic region correspondence estimation. Therefore, we propose to further incorporate the idea of intrinsic component [9] to better characterize the true region connectivity. Moreover, when processing images with complex content, we may see color-bleeding artifacts in the transferred result even with a good characterization of region connectivity. In order to eliminate false color during the color transfer procedure, we propose to use the intrinsic information [9] to determine a set of unreliable pixels and conduct the reflectance-weighted color transfer for these pixels according to their corresponding intrinsic reflectivity.

The rest of this paper is organized as follows. Section 2 presents the proposed method of automatic reflectance-weighted color transfer with graph-theoretic region correspondence estimation. Section 3 shows the experimental results and discussions. Finally, Section 4 gives the conclusion.

2. Proposed Method

Given a target image, we first employ content-based image retrieval technique [10] to select the top $K$ most relevant images from either a high-quality image database or internet as its reference images. Next, we segment each image into regions by the mean-shift based technique [11], which classifies pixels to its mode along the gradient direction in a joint feature space. Since we aim to incorporate spatially adjacent relationships to determine the best-matched region for each target region, we represent each image as an attributed and undirected graph so as to model both the region attributes and region connectivity.

To model both region attributes and region connectivity, we follow the method proposed in [10] and represent each image by a multivariate Gaussian in a certain color space. An image could therefore be modeled by a Gaussian mixture model (GMM), where each Gaussian model characterizes the color attributes of one corresponding region. However, this GMM representation captures mostly the region attributes but retains no region connectivity within the image. In order to model both region attributes and region connectivity, we follow the method proposed in [10] and represent each image as an attributed and undirected graph so as to model both the region attributes and region connectivity.

Let $G_T = (V_T, E_T, T)$ and $G_i^k = (V_i^k, E_i^k, S_i^k)$, $1 \leq i \leq K$ denote the target image and the $K$ references images, respectively; where $V$ is the set of nodes, $E$ is the set of edges, and $T_i^k, S_i^k$ are the corresponding adjacency matrices. Each node $x_a \in V$ indicates a region in the image, and the node attributes correspond to the Gaussian distribution of the color feature within the region. An edge $(x_a, x_b) \in E$ is built only between two spatially adjacent regions $x_a$ and $x_b$. The adjacency matrix $T$ with dimension $|V_T| \times |V_T|$ represents the region adjacency within the image $G_T$ and is defined by

$$ T_{ab} = \begin{cases} \gamma_{ab}, & \text{if } (x_a, x_b) \in E_T, \\ 0, & \text{otherwise} \end{cases} $$

(1)

where the edge weight $\gamma_{ab}$ is a function of similarity between nodes $x_a$ and $x_b$ and will be defined later. Similarly, the adjacency matrix $S_i^k$ of the reference image $G_i^k$ is with dimension $|V_i^k| \times |V_i^k|$ and defined by

$$ S_{mn}^i = \begin{cases} v_{mn}^i, & \text{if } (y_m, y_n) \in E_i^k, \\ 0, & \text{otherwise} \end{cases} $$

(2)

An example of our graph representation of the image in Fig. 1(a) is shown in Figs. 1(b)–1(c).

While designing the edge weights $\gamma_{ab}$ and $v_{mn}^i$ in Eqs. (1) and (2), an intuitive idea is to define the edge weight in terms of region attributes by [12]

$$ \gamma_{ab} \propto \exp[-d(x_a, x_b)], $$

(3)

and

$$ v_{mn}^i \propto \exp[-d(y_m, y_n)], $$

(4)

where $d(x_a, x_b)$ is the KL distance between the two region attributes:

$$ d(x_a, x_b) = D_{KL}(q_a || f_b) + D_{KL}(f_b || q_a). $$

(5)

In Eq. (5), $q_a$ and $f_b$ are the Gaussian color distributions in regions $x_a$ and $x_b$, respectively, and $D_{KL}(q_a || f_b)$ is KL divergence between $q_a$ and $f_b$. Figure 1(d) shows the adjacency matrix of Fig. 1(a) using the estimated edge weight by Eqs. (3) and (4).

However, when there exist lighting variations within an image, two adjacent regions which originally belong to the same object may have very different color distributions. Using the region attributes to define the edge weight may fail to characterize the true spatial organization between regions in complex content images. An example is shown in Fig. 2, where the four regions $X_1, \ldots, X_4$ in Fig. 2(b) all belong to the same foreground object. However, the estimated edge weights of $(X_1, X_2)$ and $(X_1, X_3)$ in Fig. 2(d) do not reflect their strong region connectivity.

2.1 Reflectance-weighted image representation

An image region is usually defined as a set of connected pixels with homogeneous color and thus could be modeled by a multivariate Gaussian in a certain color space. An image could therefore be modeled by a Gaussian mixture model (GMM), where each Gaussian model characterizes the color attributes of one corresponding region. However, this GMM representation captures mostly the region attributes but retains no region connectivity within the image. In order to model both region attributes and region connectivity, we follow the method proposed in [10] and represent each image as an attributed and undirected graph so as to model both the region attributes and region connectivity.

Let $G_T = (V_T, E_T, T)$ and $G_i^k = (V_i^k, E_i^k, S_i^k)$, $1 \leq i \leq K$ denote the target image and the $K$ references images, respectively; where $V$ is the set of nodes, $E$ is the set of edges, and $T_i^k, S_i^k$ are the corresponding adjacency matrices. Each node $x_a \in V$ indicates a region in the image, and the node attributes correspond to the Gaussian distribution of the color feature within the region. An edge $(x_a, x_b) \in E$ is built only between two spatially adjacent regions $x_a$ and $x_b$. The adjacency matrix $T$ with dimension $|V_T| \times |V_T|$ represents the region adjacency within the image $G_T$ and is defined by

$$ T_{ab} = \begin{cases} \gamma_{ab}, & \text{if } (x_a, x_b) \in E_T, \\ 0, & \text{otherwise} \end{cases} $$

(1)

where the edge weight $\gamma_{ab}$ is a function of similarity between nodes $x_a$ and $x_b$ and will be defined later. Similarly, the adjacency matrix $S_i^k$ of the reference image $G_i^k$ is with dimension $|V_i^k| \times |V_i^k|$ and defined by

$$ S_{mn}^i = \begin{cases} v_{mn}^i, & \text{if } (y_m, y_n) \in E_i^k, \\ 0, & \text{otherwise} \end{cases} $$

(2)

An example of our graph representation of the image in Fig. 1(a) is shown in Figs. 1(b)–1(c).

While designing the edge weights $\gamma_{ab}$ and $v_{mn}^i$ in Eqs. (1) and (2), an intuitive idea is to define the edge weight in terms of region attributes by [12]

$$ \gamma_{ab} \propto \exp[-d(x_a, x_b)], $$

(3)

and

$$ v_{mn}^i \propto \exp[-d(y_m, y_n)], $$

(4)

where $d(x_a, x_b)$ is the KL distance between the two region attributes:

$$ d(x_a, x_b) = D_{KL}(q_a || f_b) + D_{KL}(f_b || q_a). $$

(5)

In Eq. (5), $q_a$ and $f_b$ are the Gaussian color distributions in regions $x_a$ and $x_b$, respectively, and $D_{KL}(q_a || f_b)$ is KL divergence between $q_a$ and $f_b$. Figure 1(d) shows the adjacency matrix of Fig. 1(a) using the estimated edge weight by Eqs. (3) and (4).

However, when there exist lighting variations within an image, two adjacent regions which originally belong to the same object may have very different color distributions. Using the region attributes to define the edge weight may fail to characterize the true spatial organization between regions in complex content images. An example is shown in Fig. 2, where the four regions $X_1, \ldots, X_4$ in Fig. 2(b) all belong to the same foreground object. However, the estimated edge weights of $(X_1, X_2)$ and $(X_1, X_3)$ in Fig. 2(d) do not reflect their strong region connectivity.
To better reflect the region connectivity, we propose to incorporate the idea of intrinsic component to determine the edge weight. We follow the method in [9] to recover the intrinsic reflectivity image \( \rho \). First, we derive the illuminant invariant direction \( e^+ \) by entropy minimization. The axis \( e^+ \) is invariant to shading and illumination intensity. Next, we project all log band-ratio chromaticity points onto \( e^+ \) and derive the reflectivity image \( \rho \). As shown in Fig. 2(c), the pixels belonging to the same foreground object have similar reflectance values.

After we obtain the intrinsic reflectivity image \( \rho \), we then define the edge weight \( \gamma_{ab} \) and \( \nu_{mn} \) in Eqs. (1) and (2) according to their reflectance similarity by

\[
\gamma_{ab} \propto \exp\left\{-\frac{(\mu_{x_a}^+ - \mu_{x_b}^+)^2}{\sigma^2}\right\},
\]

and

\[
\nu_{mn} \propto \exp\left\{-\frac{(\mu_{y_m}^+ - \mu_{y_n}^+)^2}{\sigma^2}\right\},
\]
where $\mu_{\alpha}^{\mu}$ is the averaged reflectance value of all the pixels in the region $x_a$, and $\sigma_{\alpha}$ is the reflectance deviation of all the pixels in the image. Using the edge weight defined in Eqs. (6) and (7), the estimated adjacency matrix in Fig. 2(e) now better characterize the region organization in the original image.

### 2.2 Region correspondence estimation

Next, we need to define a region mapping function $f(\cdot)$ to measure the similarity between the target node $x_a$ and the reference node $y_m'$. We propose to incorporate both region attributes and spatially adjacent relationship between regions into the region mapping function:

$$f(x_a, y_m') \propto \{-(d(x_a, y_m') + w_{am})\}, \quad (8)$$

where $d(x_a, y_m')$ is the distance [as defined in Eq. (5)] between two region attributes, and $w_{am}$ is used to compensate the region distance $d(x_a, y_m')$ according to the matching degree between the neighboring region connectivity of $x_a$ and $y_m'$ by:

$$w_{am} = \sum_{b=1}^{\vert V_T \vert} \sum_{n=1}^{\vert V_T \vert} T_{abS_m}d(x_b, y_m'). \quad (9)$$

Equation (9) is a weighted summation of region distances between all the neighboring nodes of $x_a$ and $y_m'$. Thus, in Eq. (8), the weight $w_{am}$ penalizes the region distance $d(x_a, y_m')$ using the neighboring dissimilarity between $x_a$ and $y_m'$. The proposed region mapping function is therefore insensitive to different segmentation results. For example, assume we have a poor segmentation result where two spatially adjacent regions $x_a$ and $x_b$ belong to the same object. Then the weight $w_{am}$ of $x_a$ will highly depend on its neighboring node $x_b$ because of the higher edge weight in $T_{ab}$. If, otherwise, $x_a$ and $x_b$ are two nodes belonging to different objects, then $x_b$ would have less influence on $w_{am}$.

With the proposed region mapping function $f(\cdot)$, we next determine the best-matched reference region for each target region $x_a \in V_T$ by maximizing $f(x_a, y_m')$:

$$m^*, i^* \leftarrow \arg \max_{m, i} \{f(x_a, y_m')\} \mid 1 \leq i \leq K, \quad (10)$$

where $m^*$ denotes the index of the best-matched region in the $i^*$--th reference image. Figure 3 shows an example of the matched result for Fig. 1(a).

### 2.3 Automatic color transfer for images with complex content

After we determine the set of best-matched region pairs $\{(x_a, y_m')\}_{x_a \in V_T}$ for the target image $G_T$, we next have to transfer the color from the reference region to the corresponding target region. We first adopt the color transfer procedure [5, 8] and compute the transferred values in the chromatic $\alpha \beta$ channels for each pixel $z$ in the region $x_a$ by:

$$\bar{\alpha}_z = \frac{\sigma_{\alpha}^{y_m'}}{\sigma_{\alpha}} (\alpha_z \cdot \mu_{\alpha}^{y_m'}) + \mu_{\alpha}^{y_m'}, \quad (11)$$

and

$$\bar{\beta}_z = \frac{\sigma_{\beta}^{y_m'}}{\sigma_{\beta}} (\beta_z \cdot \mu_{\beta}^{y_m'}) + \mu_{\beta}^{y_m'}, \quad (12)$$

where $\mu_{\alpha}$ and $\sigma_{\alpha}$ are the mean and standard deviation of the pixels in the corresponding region, and the superscript $\alpha$ and $\beta$ denote the color statistics in $\alpha$ and $\beta$ channels, respectively. While transferring the luminance channel, in order to maintain the original contrast within the target image, we do not conduct region-level transfer. Instead, we conduct the global transfer to calculate the luminance value for each pixel $z$ by:

$$\bar{\ell}_z = \frac{\sigma_{\ell}^{y_m'}}{\sigma_{\ell}} (\ell_z \cdot \mu_{\ell}^{y_m'}) + \mu_{\ell}^{y_m'}, \quad (13)$$

*Fig. 3. The four best-matched regions from three reference images for the regions X1–X4 in Fig. 1(a).*
where $\mu_f$, $\sigma_f$, $\mu_s$, and $\sigma_s$ are the means and standard deviations of all the pixels in the target image and the $K$ reference images, respectively.

Figure 4 shows an example of the proposed color transfer for a high-quality target image. The transferred results by the global color transfer method [5] and the multi-source local method [8] are also shown for comparison. Figure 5 shows the histograms of the target image in Fig. 4(a) and its color transferred results [Figs. 4(b)–4(d)] in Blue–Yellow (the approximate $\alpha$) and Red–Green (the approximate $\beta$) color spaces. As shown in Fig. 5, global color transfer [5] tends to globally shift and scale the original distribution, multi-source local method [8] tends to locally and slightly shift and scale the original distribution, while the proposed method tends to obtain a new distribution. Figure 4 shows that all these methods result in good transferred results in this high-quality case.

However, using the color transfer procedure in Eqs. (11) and (12) may result in color bleeding artifact on images containing complex content. For example, in Figs. 6(a)–6(c), since we can only obtain a rough region segmentation
for the complex tree leaves area, the color transferred result contains visible artifact with false color on the tree leaves area. In order to solve this problem, we propose to fine tune some pixel values according to their corresponding intrinsic surface color.

We first map all the pixels within a region into the reflectance channel and define the pixels located away from the principal reflectance distribution as unreliable pixels. Next, we model each unreliable pixel $z$ as a linear combination of its neighboring pixels in the reflectance channel and derive their reflectance weight by minimizing the cost function:

$$
\epsilon(r) = \sum_z \rho_z - \sum_{x \in \eta(z)} r_x z^b \mu_{xz}^2,
$$

subject to the constraint $\sum_{x \in \eta(z)} r_x z^b = 1,$

where $\eta(\cdot)$ is the set of neighboring pixels determined in terms of Euclidean distance in the reflectance channel. Finally, we conduct the reflectance-weighted color transfer procedure to refine the chromatic values of the unreliable pixels by

$$
\tilde{a}_z = \sum_{x \in \eta(z)} r_x z^b \tilde{a}_x^b,
$$

and

$$
\tilde{b}_z = \sum_{x \in \eta(z)} r_x z^b \tilde{b}_x^b.
$$

The result in Fig. 6(d) shows that the proposed weighted color transfer procedure indeed successfully eliminates the color bleeding artifact and achieves visually pleasing chromaticity pixel values for the complex content area.

3. Experimental Results and Discussion

3.1 Color transfer for low quality images with complex content

In this experiment, we use 100 low-quality photos captured from 3 cell phone cameras (Nokia 6125, Sony Z800i, and Samsung E258) as our target images and the Corel photo gallery as our high-quality reference image database. To
reduce the computational complexity, we first employ content-based image retrieval technique [10] to retrieve the top 
$K = 3$ relevant images for each target image over 9500 Corel images as the reference images.

We compare the performance of color transfer with two related methods: “Global Color Transfer” [5] and “Multi-
Source Local Color Transfer” [8]. In order to have a fair comparison, we use the top 1 relevant image as the reference
image to implement the global method [5], and the top 3 relevant images as references to implement the multi-source
local method [8] and our proposed method. In the objective test, we adopt two criteria, colorfulness (C) [8] and entropy
(E) [12, 13], as our performance measurement. The colorfulness dissimilarity ($\Delta C$) adopted in [8] measures the
difference of colorfulness estimates between the original target image and its color transferred result. A lower
$\Delta C$ generally indicates a higher performance. On the other hand, entropy (E) measures the averaged color information of an
image. Higher entropy (E) indicates a higher diversity of transferred colors.

As already shown in Fig. 4, when processing a high-quality target image, all these methods result in good transferred
results. However, for lower-quality and noisy images, the two methods [5] and [8] usually result in poor quality. For
example, in Figs. 7(a) and 7(b), the global transfer method enhances the undesirable noises when scaling the blue color,
while the poor segmentation and independent region matching process in [8] deteriorate the performance. On the other
hand, because the region segmentation by mean-shift technique better maintains the spatial and color relationship, our
transferred result in Fig. 7(d) has better visual quality. In addition, our proposed region mapping function further
improves the region matching process by including both region attributes and spatially adjacent relationship into
consideration. Two more examples are given in Figs. 8 and 9, where our proposed method achieves better results in
terms of both visual quality and the two objective measurements.

Table 1 compares the quantitative measures of the averaged $\Delta C$ and E over the 100 target images. The results in
Table 1 show that the color transferred images processed with the proposed method achieve higher performance than
the multi-source local method [8] with lower colorfulness dissimilarity $\Delta C$ and higher entropy value E. Although the
proposed method has lower entropy value than the global method [5], the subjective performance of our method is
better than [5].

### 3.2 Application on black-and-white image colorization

We extend the proposed color transfer method to the application of black-and-white image colorization. This idea is
very similar to example-based colorization. Since the performance of example-based colorization highly depends on
the reference images, here we use 20 black-and-white images as our target images and select multiple images with
similar content as their reference images.

We use the region mapping criterion defined in Eq. (8) to determine the best-matched reference region for each
target region. Note that, because the black-and-white target images have no chromaticity information, the region
features involved in Eq. (8) include only luminance and Gabor texture energy maps. Finally, we perform the weighted
color transfer procedure defined in Eqs. (15) and (16) to transfer the chromatic channels from reference regions to each
black-and-white target region.
Fig. 8. (a) The original target image; (b) the result of [5] ($\Delta C = 10.1551$, $E = 8.7336$); (c) the result of [8] ($\Delta C = 1.7803$, $E = 9.9613$); and (d) the result of the proposed method ($\Delta C = 1.4451$, $E = 10.2547$).

Fig. 9. (a) The original target image; (b) the result of [5] ($\Delta C = 15.6377$, $E = 9.5056$); (c) the result of [8] ($\Delta C = 8.8854$, $E = 7.3661$); and (d) the result of the proposed method ($\Delta C = 6.5710$, $E = 9.0246$).

Table 1. Quantitative measures of color transfer methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Averaged $\Delta C$</th>
<th>Averaged $E$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multi-Reference Local Color Transfer [8]</td>
<td>10.08</td>
<td>9.07</td>
</tr>
<tr>
<td>Proposed method using intrinsic component</td>
<td>5.32</td>
<td>9.8854</td>
</tr>
</tbody>
</table>
We compare the performance of colorization with the global transfer method [5]. In the objective test, we also adopt colorfulness (C) [8] and entropy (E) [12, 13] as our performance measurement. As shown in Figs. 10 and 11, our proposed method achieves better colorization result and better objective measurements than the global-based method.

Table 2 compares the quantitative measures of the averaged $\Delta C$ and $E$ over the 20 target images and show that the proposed method achieves better performance than the global method [5] with lower colorfulness dissimilarity $\Delta C$ and higher entropy value $E$.

4. Conclusion

In this paper, we propose an automatic color transfer method for processing images with complex content based on intrinsic component. We use intrinsic information to better characterize the local organization within an image and to
eliminate the color-bleeding artifact across complex regions. We first determine the best-matched reference region for each target region using graph-theoretic region correspondence estimation combined with reflectance weighted scheme. Next, we conduct color transfer between the best-matched region pairs and perform weighted color transfer for unreliable pixels across complex regions in a de-correlated color space. We also extend the proposed method to colorization application for black-and-white images. Both subjective and objective performance evaluation of our experiments demonstrate that the proposed method outperforms the existing methods, especially for low-quality images with complex content.

REFERENCES


---

**Table 2. Quantitative measures of colorization.**

<table>
<thead>
<tr>
<th></th>
<th>Averaged ΔC</th>
<th>Averaged E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed method</td>
<td>8.317</td>
<td>9.2485</td>
</tr>
</tbody>
</table>