COLOR CORRECTION THROUGH REGION MATCHING LEVERAGED BY POINT CORRESPONDENCES

Dieu-Sang Ly, Serge Beucher and Michel Bilodeau

MINES ParisTech, PSL - Research University, CMM - Centre for mathematical morphology
35 rue St Honoré 77300 Fontainebleau, France

ABSTRACT

This paper addresses the problem of region-based color correction. Our solution begins with the image segmentation by marker-controlled watershed transformation, which is faster and produces more uniform regions with better adherence to object boundaries than the segmentation in previous works. Next, regions between two images are matched using point feature correspondences which are invariant to geometric transformation and illumination changes. Finally, the color-corrected image is generated from the color transfer functions of corresponding regions. We demonstrate the results of this approach using different sets of images.

***Index Terms*—** Color correction, color transfer, region segmentation, region matching

1. INTRODUCTION

Color correction refers to modifying the color of a given image (denoted as input image) so that it is similar to the color of another image (reference image). Methods of color correction can be classified into two categories: one based on pixel correspondences [1] and the other based on color statistics [2, 3]. The first one relies on correspondence accuracy whereas the second one does not require exact correspondences. There are two main groups of statistical approaches: histogram matching and color distribution modification. The basic histogram matching method [4] computes a mapping that aligns the histograms of two images. It assumes that both images are captured from the same viewpoint and under the same illumination. In order to deal with other cases when these strong assumptions are violated, [5] used the histogram matching as the building block and proposed several solutions based on the computation of consistent color mappings. The second group of statistical approaches utilizes the color transfer function to scale and shift the color distribution of the input image towards the reference one. The global color correction [2] computes the parameters of the color transfer from the entire images and hence can produce a correct result uniquely when two cameras observe the same regions under the same illumination. On the other hand, the local color correction [6, 7] applies a different color transfer function to each region. Most of the local (or region-based) approaches consist of three steps: region segmentation, region matching to compute the color mappings and color transfer.

In this work, we present an approach of local color correction following the general framework mentioned above. First, we propose the region partition using the marker-controlled watershed transformation, which is much faster and produces more uniform regions with better adherence to region boundaries than the segmentation in previous color correction methods. Next, the region matching is leveraged by point feature correspondences and can handle images captured under different conditions. Lastly, the input image is corrected using the color transfer functions of all region matches.

The rest of this paper is organised as follows. Section 2 reviews related work on local color correction. Section 3 describes our proposed approach. The evaluation results are given in section 4. Section 5 concludes the paper.

2. RELATED WORK

According to a survey of color correction algorithms in [8], the region-based approach in [6] produced the best results among several compared approaches. It is composed of three steps: (i) segment both images by Expectation-Maximization (EM) algorithm and characterize each image by a Gaussian Mixture Model, where each region is associated with a Gaussian component, (ii) match the regions between two images based on their Gaussian mean value or their overlapping rate and (iii) produce the color-corrected image by combining the color transfer functions of all regions.

Concerning the region segmentation, [7] proposed an improvement of [6] by using the Mean-Shift algorithm which is less time-consuming than the EM algorithm and does not require the predefined number of regions as the EM algorithm does. In this work, we propose to decompose the images into regions using watershed transformation, which is also unsupervised but much less time-consuming than the Mean-Shift segmentation. For instance, given a 960x540 image as illustrated in figure 1 - left, it takes 55 msec. by watershed trans-

This work has been performed in the project PANORAMA, co-funded by grants from Belgium, Italy, France, the Netherlands, and the United Kingdom, and the ENIAC Joint Undertaking.

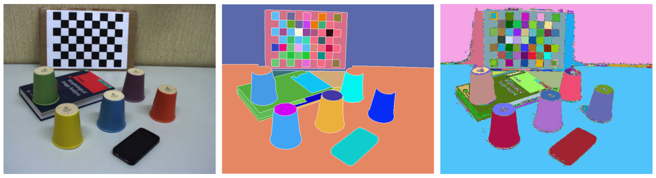

Fig. 1. Segmentation of the left image by the watershed transformation (middle) and the Mean-Shift algorithm (right)

Fig. 2. Region matching. **Left:** reference (upper) and input (lower) images with region segmentation. **Middle:** matching by region color and location. Matched regions are displayed with the same color. **Right:** matching by region mapping from the input to the reference images. Each segmented region of the input image is illustrated by a color. The mapping is uniquely correct around the building.

form (plus 1 sec. by region fusion) and 27 sec. by Mean-Shift. It can be seen that the Mean-Shift algorithm does not return uniform regions with a good adherence to region boundaries as the watershed transform does. In order to suppress the redundant details and sharpen the region contours in Mean-Shift, one must increase the spatial and color range parameters, and consequently augment the computation time dramatically.

Regarding the region matching, the first solution is based on region similarity, i.e. pairing regions using their mean color [6], location, area or other characteristics. On one hand, using color as matching criterion requires that the color of two regions be sufficiently similar, therefore is irrelevant in color correction application. On the other hand, location and area criteria can fail in case of complex transformation between two images, for example important zooming or translation. Figure 2 - middle shows the region matching between two segmented images using their color and location. First, the region color is calculated from the average of the color of all region pixels in LAB space. The region location is assumed to be at the region centroid. Next, given a threshold of region location, we search, in the neighbourhood of each region of the input image, the region of the reference image having the closest color. As the depth of the scene is important in these two images, the regions distant from the cameras have a significant translation between two images whereas the regions close to the cameras do not. Consequently, the correspondences are erroneous when we use a single threshold of region location. In other words, searching region correspondences based on location criterion is not robust to image transformation. Moreover, the luminance of these two images is so different that using the color in region matching produces incorrect result. The second solution assumes known image registration [7]. Each region in the segmented input image is mapped to the unsegmented reference image using the image transformation and the overlaid region is considered as its match. The first drawback of this method is that it requires coarsely registered images. In particular, [7] assumes that the input image is entirely included in the reference image to facilitate the image registration. This strong assumption rarely holds when two cameras have random field of view. In case of complex image transformations, it is impossible to map all regions from one image to the other due to the extrapolation problem as a transformation can correctly map the image region straddled by points used to compute that transformation and is less accurate with distance from this region [9]. Figure 2 - right presents the mapping of regions from the input image to the reference one. Even if point correspondences are well distributed over the images, the transformation still can handle only regions straddled by the most accurate matched points. The second drawback of this method is that the region projection will result in wrong matches in case of occultation: if a region of the input image is mapped to an occult region of the reference image, the computation of the color of this occult region is incorrect. As a consequence, the parameters of the color transfer computed from this region are wrong. In this paper, we present a region matching process based on point feature correspondences, which is independent of image registration and able to handle images under different acquisition conditions.

3. REGION-BASED COLOR CORRECTION

The proposed approach is composed of three tasks: segment both input and reference images into regions, search for region correspondences between these images and apply the color transfer to the input image.

3.1. Region segmentation

The original image is segmented into regions using watershed transformation [10]. The idea is to consider a gray-scale image as a topographic relief and to flood this relief from different sources until they start to merge. This results in watershed lines separating different catchment basins. In addition, predefined markers can be used as flooding sources to control the segmentation, e.g. to avoid over segmentation. The marker-controlled watershed segmentation can be described as follows

1. **Computation of segmentation criterion and markers:** In order to partition the image into homogeneous regions,

we can use the image gradient as the segmentation criterion (or the topographic relief mentioned above) since the gradient value is low within a homogeneous region and high at its boundary. The markers should locate inside the regions, hence can be computed from the local minima of the gradient image or by applying a threshold to the gradient image. Note that we compute the gradient by the maximum of the gradients of all color channels in order to preserves region boundaries better than the gradient from gray-level image.

2. **Marker-controlled watershed segmentation:** The image gradient and markers are provided to watershed segmentation. If the resulting regions are more numerous than expected, we can run an additional region fusion: if the color difference between two adjacent regions is inferior to a given threshold, we eliminate their inner boundary and keep their outer boundaries with other regions in order to avoid incorrect boundary elimination and region fusion.

3.2. Region matching

As the challenge is to search region correspondences between two images under severe geometric transformation and illumination changes, we seek for an approach of region pairing using a criterion robust to these variations. In this work, we propose a method of region matching leveraged by image point features as follows

1. Segment the input and reference images.
2. Compute point correspondences between these images.
3. Match two regions if they are straddled by matched points. In addition, merge regions in case of one-to-multiple matching, which may happen when a region in one image corresponds to several adjacent regions in the other image.

Point features are detected and described using SIFT [11], which is well known to be invariant to image translation, rotation, scaling and illumination changes. Next, point correspondences are estimated using the Brute-Force matcher and refined by RANSAC [12] with the fitting model being the homography between two images. Note that we use RANSAC to obtain inliers but discard the output model. If we terminate the point matching here, the correspondences are the best matches fitted to the best estimated homography, therefore it is not guaranteed that they are spatially distributed all over the image. In order to overcome this limited distribution, we implemented an incremental tiling approach: after one set of point correspondences is found, we mask the image part straddled by these points in both images and search for point correspondences within the unmasked image part. This mask can be computed by a rotated rectangle or a convex hull bounding a set of point. The point matching with mask is repeated until the final mask covers most of the image, 80% in our case.

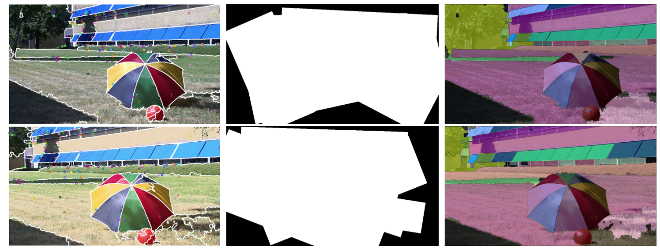

Fig. 3. Region matching leveraged by point correspondences. **Left:** point correspondences by incremental tiling. **Middle:** the final masks bounding these points. **Right:** region correspondences in the same color.

Figure 3 illustrates the result of point matching by the incremental tiling technique and the region matching leveraged by these points.

3.3. Color transfer

The general color transfer function scales and offsets the color distribution of an input image towards a reference image.

$$\mathbf{C}_{\mathrm{o}} = \mu_{\mathrm{r}} + \frac{\sigma_{\mathrm{r}}}{\sigma_{\mathrm{i}}}(\mathbf{C}_{\mathrm{i}} - \mu_{\mathrm{i}}) \tag{1}$$

where $\mathbf{C}_{\mathrm{i}}$ is the input image, $\mathbf{C}_{\mathrm{o}}$ the output of color transfer, $(\mu_{\mathrm{i}}, \sigma_{\mathrm{i}})$ and $(\mu_{\mathrm{r}}, \sigma_{\mathrm{r}})$ the (mean, standard deviation) of the input and reference images.

For each pair of regions k found in the region matching step, we compute their color distribution μ_{i}^k, σ_{i}^k, μ_{r}^k and σ_{r}^k, which are the parameters of the color transfer between these two regions. Given N region matches between the input and reference images, the color correction is a combination of N local color transfers. In addition, in order to ensure a smooth color shading across the color-corrected image, each local color transfer is weighted by an influence mask $\mathbf{IM}$, which measures the similarity of each pixel of the input image and the mean color of the region in consideration.

$$\mathbf{IM} = e^{a\mathbf{p}^b} \text{ with } \mathbf{p} = 1 - \frac{\mathbf{d}}{\max(\mathbf{d})} \tag{2}$$

where $\mathbf{d}$ is the Euclidean distance between every pixel of the input image $\mathbf{C}_{\mathrm{i}}$ and the mean color of region k in LAB space $\mathbf{d} = ||\mathbf{C}_{\mathrm{i}} - \mu_{\mathrm{i}}^k||$. We introduce the distance $\mathbf{p}$ to normalize the maximum of $\mathbf{d}$ which varies from different regions. An element of $\mathbf{p}$ approaches 1 when the color of the corresponding pixel in $\mathbf{C}_{\mathrm{i}}$ is close to μ_{i}^k and 0 vice versa. In our experiments, $a = 10$ and $b = 2$. Example of influence masks is shown in figure 4.

The region-based color correction combines the color transfer functions of N region correspondences weighted by N influence masks

$$\mathbf{C}_{\mathrm{o}} = \frac{\sum_{k=1}^{N}(\mu_{\mathrm{r}}^k + \frac{\sigma_{\mathrm{r}}^k}{\sigma_{\mathrm{i}}^k}(\mathbf{C}_{\mathrm{i}} - \mu_{\mathrm{i}}^k)) \times \mathbf{IM}^k}{\sum_{k=1}^{N}\mathbf{IM}^k} \tag{3}$$

4. EXPERIMENTAL RESULTS

Our first result of color correction is presented in figure 5 - left. The reference and input images are captured from two cameras at different viewpoints and under the same illumination. The color values between them vary due to the different photometric parameters of two cameras. It can be seen that the global approach is erroneous. For example, the blue color (of the umbrella and the outdoor blinds) and the yellow color (of the grass, the umbrella and the building wall) are not well corrected. Since these images includes non-corresponding regions, applying the global color transfer to the entire image produces incorrect result. On the contrary, the local technique provides good correction output.

In order to provide quantitative evaluation, it is necessary to use the ground-truth of the color correction. The ground-truth image should have the same region content as the input image and the color similar to the reference image. Therefore, beside the image set without the ground-truth image (figure 5 - left), we use two other sets of images, ”Football” and ”Desk” (figure 5 - middle and right), each one acquired by the same camera and including a reference image, an input image captured at a different viewpoint and with different photometric parameters from the reference image together with a ground-truth image captured at the same viewpoint as the input image, and with the same photometric parameters as the reference image.

We use the metric based on the color similarity proposed in [8]. The color similarity (CS) between two images $\mathbf{A}$ and $\mathbf{B}$ is defined as their peak-signal-to-noise-ratio ($PSNR$). A higher $PSNR$ generally indicates that the color between two images are closer.

$$CS(\mathbf{A}, \mathbf{B}) = PSNR(\hat{\mathbf{A}}, \hat{\mathbf{B}}) = 20 \log_{10} \frac{\max \mathbf{I}}{RMSE(\hat{\mathbf{A}}, \hat{\mathbf{B}})} \quad (4)$$

where:

- $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ are the overlapped areas between $\mathbf{A}$ and $\mathbf{B}$.
- $\max \mathbf{I}$ is the highest possible pixel value of the image. Since each pixel is represented by 8 bits, $\max \mathbf{I} = 255$.
- $RMSE$ stands for the root mean square error. In this case, it is computed from all pixels of $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ in RGB channels.

The evaluation result of the color correction with the ”Football” and ”Desk” sets is given in figure 5 and table 1. We measure the color similarity between the ground-truth image $\mathbf{C}_{gt}$ and (i) the input image $\mathbf{C}_i$, (ii) the image after the global color correction $\mathbf{C}_{global}$ and (iii) the image after the region-based color correction $\mathbf{C}_{local}$. It can be seen that both approaches of color correction improve the color of the input image, and the local method performs better than the global one. The reason why the performance of global approach is quite good for these two image sets is that the reference and input images contain very similar regions.

5. CONCLUSIONS

We proposed an approach of region-based color correction. First, images are segmented to regions by rapid watershed transform. Then, regions are paired using point correspondences which are invariant to geometric transformation and illumination variations. Finally, the weighted color transfer is applied to modify the image color distribution. This method can be used in television broadcasting to correct the color of video streams from various cameras, which may have different internal settings and external illumination conditions.

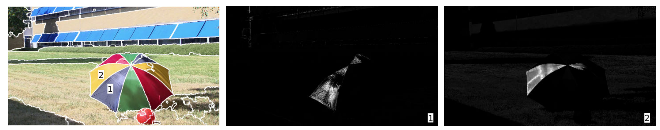

Fig. 4. Example of color influence masks

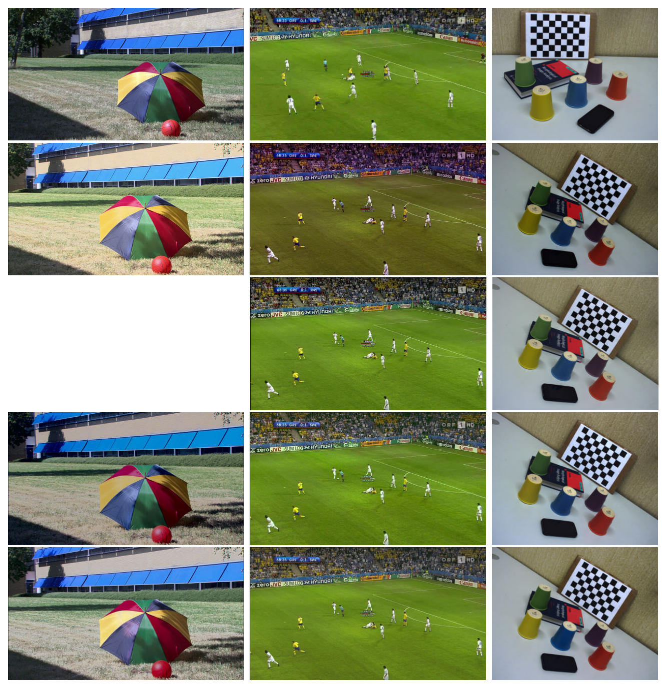

Fig. 5. Color correction: Row 1: reference image. Row 2: input image. Row 3: ground-truth image. Row 4: global color correction. Row 5: local color correction

	”Football”	”Desk”
$CS(\mathbf{C}_{gt}, \mathbf{C}_i)$	19.125	21.173
$CS(\mathbf{C}_{gt}, \mathbf{C}_{global})$	22.840	30.392
$CS(\mathbf{C}_{gt}, \mathbf{C}_{local})$	23.328	32.872

Table 1. Color similarity

6. REFERENCES

[1] Frank M. Candocia and D. A. Mandarino, "A semi-parametric model for accurate camera response function modeling and exposure estimation from comparametric data.," *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1138–1150, 2005.

[2] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley, "Color transfer between images," *IEEE Comput. Graph. Appl.*, vol. 21, no. 5, pp. 34–41, Sept. 2001.

[3] Francois Pitie, Anil C. Kokaram, and Rozenn Dahyot, "N-dimensional probablility density function transfer and its application to colour transfer," in *Proceedings of the Tenth IEEE International Conference on Computer Vision - Volume 2*, Washington, DC, USA, 2005, ICCV '05, pp. 1434–1439, IEEE Computer Society.

[4] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing (3rd Edition)*, Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 2006.

[5] Sefy Kagarlitsky, Yael Moses, and Yacov Hel-Or, "Piecewise-consistent color mappings of images acquired under various conditions.," in *ICCV*. 2009, pp. 2311–2318, IEEE.

[6] Yu-Wing Tai, Jiaya Jia, and Chi-Keung Tang, "Local color transfer via probabilistic segmentation by expectation-maximization.," in *CVPR (1)*. 2005, pp. 747–754, IEEE Computer Society.

[7] Miguel Oliveira, Angel Domingo Sappa, and Vítor Santos, "Unsupervised local color correction for coarsely registered images," in *CVPR*, 2011, pp. 201–208.

[8] Wei Xu and Jane Mulligan, "Performance evaluation of color correction approaches for automatic multi-view image and video stitching," in *CVPR*, 2010, pp. 263–270.

[9] Richard Hartley and Andrew Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, New York, NY, USA, 2nd edition, 2003.

[10] Serge Beucher and Christian Lantuejoul, "Use of watersheds in contour detection," in *International Workshop on Image Processing*, 1979.

[11] David G. Lowe, "Object recognition from local scale-invariant features," in *ICCV*, 1999, pp. 1150–1157.

[12] Martin A. Fischler and Robert C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, June 1981.