Unsupervised Local Color Correction for Coarsely Registered Images

Miguel Oliveira*
mriem@ua.pt

Angel D. Sappa‡
asappa@cvc.uab.es

Vitor Santos*
vitor@ua.pt

*Department of Mechanical Engineering
University of Aveiro, Campus U. Santiago
3800 Aveiro, Portugal

‡Computer Vision Center
Edifici O, Campus UAB
08193 Bellaterra, Barcelona, Spain

Abstract

The current paper proposes a new parametric local color correction technique. Initially, several color transfer functions are computed from the output of the mean shift color segmentation algorithm. Secondly, color influence maps are calculated. Finally, the contribution of every color transfer function is merged using the weights from the color influence maps. The proposed approach is compared with both global and local color correction approaches. Results show that our method outperforms the technique ranked first in a recent performance evaluation on this topic. Moreover, the proposed approach is computed in about one tenth of the time.

1. Introduction

Recent years have proven the importance of image mosaicing. This area of research and other similar variations such as image compositing and stitching have found a vast field of applications ranging from satellite or aerial imagery [1] to medical imaging [2], street view maps [3], city 3D modeling [4], super-resolution [5] or texture synthesis [6], to name a few.

In general, whenever merging two or more images of the same scene is required for comparison or integration purposes, the correspondence problem should be faced. The final result of correspondence between images, known as the *mosaic*, is a couple of images that are as similar as possible, both geometrically and photometrically [7].

The *geometric* correspondence is usually referred to as image registration: "*the process of overlaying two or more images of the same scene taken at different times, from different viewpoints, and/or by different sensors. It geometrically aligns two images—the reference and sensed images*" [8]. This problem has been extensively studied and is out of the scope of the current paper. In other words we assume that the given images are coarsely registered. On the other hand, the *photometrical* correspondence between a pair of images is still an open problem, which has attracted less attention from the research community. In general, image mosaicing uses sets of images taken with the same lighting conditions and with a single camera, or similar ones. In this way, the colors present on both images are very similar and the problem of photometric correspondence is overlooked.

Just to show one of the motivations of the current work we can mention that recently Google launched the Street View feature of Google Maps, which provides an interface that can display street-level images in a natural way that enables convenient navigation between images without losing the map context [3]. The street maps where captured using a six camera spherical vision system onboard the company's vehicles. Registration provides the final mosaic we see in street view. What if one could use publicly available pictures of the streets on the web, instead of having to go onsite to photograph? Automatic image registration methods already provide the means of aligning images. In fact, publicly available images have already been coarsely registered on internet sites [10]. The missing piece is photometric correspondence: how to balance the color of one picture to match the color of another. This operation is referred to as *color correction*.

In the current paper we propose a new technique to perform unsupervised color correction between two coarsely registered images. It is completely automatic and obtains better results than current state of the art color correction techniques. It is robust to coarsely registered images, which are usually available in public domains (e.g., Google [9], Flickr [10]).

The remainder of this paper is organized as follows. Section 2 describes previous work on color correction. Section 3 presents the proposed technique. Qualitative and quantitative results are provided in Section 4. Finally, conclusions and future work are given in Section 5.

2. Related work

Several color correction methods have been proposed in literature. They can be divided into model-based parametric approaches [11][12][13] and modeless non parametric approaches [15]. Usually, parametric approaches outperform their non-parametric counterparts [16]. Parametric approaches are based on [14], where a simple

Figure 1. *(left)* Model image. *(right)* Target image.

statistical framework to perform color correction of a given image (denoted as *target image*) using the colors of another image (the *model image*) was proposed. The color correction is performed in the *lαβ* color space. Cross channel correlation is present on most color spaces, like RGB. The *lαβ* color space is used because it minimizes correlation between channels for many natural scenes, which enables the application of different operations in different color channels with some confidence that undesirable cross channel artifacts won't occur [14]. In [14], the authors use the mean and standard deviation of both the model and target image to recover the value of the color correction. The original colors are corrected by scaling and offsetting according to the mean and standard deviation of the target image [15]. The correction is performed on each channel independently, by the color transfer function defined bellow:

$$c_t^{new}(i,j) = \mu_m + \frac{\sigma_m}{\sigma_t}\left(c_t(i,j) - \mu_t\right), \qquad (1)$$

where $c_t^{new}(i,j)$ and $c_t(i,j)$ are the resulting and original values of the three color channels from the target image pixel (*i, j*), respectively. The statistical measures μ and σ represent the mean and standard deviation of both the target and model images for all three channels. Figure 1 shows the model and target image pair used as a case study in the current paper. The objective is to correct the color components of the target image so that it would result in a more similar image to the model.

Figure 2 *(left)* shows the color corrected target image using the global color correction defined in (1), notice that after color correction white portions of the image are bluish since blue is the dominant color.

There is a major limitation from the technique presented in [14], since it uses the entire image to collect statistical information—i.e., assumes a constant color correction function. In complex scenes this assumption does not hold due to differing optics, sensor characteristics, and hardware processing employed by video cameras [15]. Although there are several color correction methods to deal with this problem, most involve strong assumptions, such as constant illumination, which are in general, difficult to fulfill in complex environments. Hence, some authors propose to use learning approaches or non-linear techniques (e.g., neural networks) to find an appropriate color transfer function for color correction. However, learning approaches have the limitation of requiring a specific training for different setups. Offline learned techniques won't be enough to solve the general problem.

Figure 2. *(left)* Target image with global color correction [14]; after color correction white portions of the image are bluish since blue is the dominant color. *(right)* Result from local EM based color correction [17].

A recent survey of color correction algorithms has compared nine color correction methods [16]. The aproach presented in [17] outperformed all others and is reccomended as the first option to try for a general image and video stitching aplication in pratice [16]. It consists of segmenting both the target and model images into several regions using an expectation maximization algorithm (*local EM based color correction*). Then, regions from the target image are matched to the model image. This is done by projecting each target image region onto the model image in order to assess the highest overlaping region in the model image. The match of regions from the target and the model image provides the statistical parameters required in (1). In addition to the segmented regions, the local EM based color correction algorithm also computes a weight mask for each region. These weight values indicate the probability that a given pixel belongs to that region. The final pixel color is obtained by adding up the contributions of each region's color transfer function weighted by its corresponding weight.

Although it was the best performing algorithm in the evaluation done in [16], we believe the local EM based color correction algorithm can be improved in two major aspects. First, the expectation maximization segmentation stage is computationally demanding: authors state that this step takes four minutes to converge while segmenting a 512x512 image. Since the local EM based color correction must segment both the target and model image, the segmentation can take about 8 minutes. Second, the expectation maximization stage requires a parameter to define the desired number of regions: this is not interesting if unsupervised color correction is required. Figure 2 *(right)* shows the color corrected target image using this local EM based approach. Our proposed color correction approach addresses both aspects allowing unsupervised applications and lower computation times. Up to our knowledge, [17] is best algorithm published in the

Figure 3. *(left)* Result from meanshift color segmentation (notice that one of the windows and the white bands of the pink building have been included onto other regions). Four regions are highlighted and numbered for future references. *(right)* Result from the proposed color correction technique.

literature, hence we have used it to quantitatively and qualitatively evaluate the performance of our proposal (a Matlab implementation available in [16] has been used).

3. Proposed Approach

The proposed approach consists of four different stages applied consecutively. Initially, the target image is segmented into a set of regions according to their color information. In the current version the well known mean shift algorithm has been used [18]. Secondly, a color transfer function is computed for every region. It is obtained using color information from the given region and its corresponding pixels in the model image. The correspondence between regions and model image pixels results from the given coarse registration, which is out of the scope of current work. Thirdly, a color influence map is computed for every region [21] in order to mix the previously computed local color transfer functions and generate a smoother result. Finally, the color from the current image is corrected using the local transfer functions weighted by the color influence maps. Each one of these stages is detailed bellow.

3.1. Image segmentation

In order to segment the target image into a set of similar colored regions the mean shift algorithm [19] is used—in particular the implementation provided by [20]. The mean shift algorithm automatically segments the target image into a set of regions. Some care must be taken while tuning the mean shift parameters: very large regions may include a large set of different colors and could result in very similar results for local color correction to its global counterpart; on the other hand, very small regions are sensitive to the lack of accuracy in image registration. All in all, the mean shift algorithm's output should reasonably represent the different colors present in the image. Nonetheless, we have employed the same mean shift parameters for all the images tested with good results (except for the minimum region size parameter which is a function of the image's size). This shows the robustness to the color segmentation stage of the proposed local color correction techniques. Figure 3 *(left)* shows the result of the mean shift algorithm applied to the target image presented in Fig. 1 *(right)*. It should be mentioned that in [17] the number of colors (regions) in the image is an input parameter, while meanshift automatically estimates this value. Therefore, it seems feasible to assume that in meanshift the input parameters are less sensitive than the ones required by [17]. A more detailed analysis is certainly required, but is out of the scope of this paper.

3.2. Local Color transfer Functions

The outcome of the previous stage is a set of regions where each one of them represent a different color; the objective now is to define a local color transfer function for each one of these regions. Let *k* be a given region from the target image; a pixel from that region will be denoted as $(i,j)_t^k$. The whole set of pixels defining that region $(\mathbf{i},\mathbf{j})_t^k$ is used for calculating the target mean and standard deviation measures, μ_t^k and σ_t^k respectively. The corresponding region in the model image is obtained by projecting the pixels from the region of the target image onto the model image.

$$(i,j)_m^k = \mathbf{T}(i,j)_t^k, \tag{2}$$

where $\mathbf{T}$ is the target to model transformation matrix obtained after the registration procedure. Using the projected pixels the model's mean and standard deviation for region *k* (μ_m^k and σ_m^k respectivelly) are computed, Thus, a local color correction function can be easily formulated by adapting (1) to the local case:

$$c_t^{new}(i,j) = \mu_m^k + \frac{\sigma_m^k}{\sigma_t^k}\left(c_t(i,j) - \mu_t^k\right). \tag{3}$$

The color transfer function, be it global or local, can be expressed generically as a function of several parameters:

$$c_t^{new}(i,j) = \mathbf{f}\left(\mu_t, \mu_m, \frac{\sigma_m}{\sigma_t}, c_t(i,j)\right). \tag{4}$$

A global color transfer function will have the first three parameters constant for the whole image, while the local approach maintains those parameters constant merely for every color segmented region. The initial hypothesis of this paper was that due to different surface reflective properties and non uniform illumination, the global image color statistics would generate only a rough approximation of all the color transfer functions.

Figure 4 plots the global and local function parameters for three channels of the *lαβ* target and model images corresponding to Fig. 1. Dashed lines represent the global color transfer function parameters for each channel of the *lαβ* space. Dots represent the parameters of local color

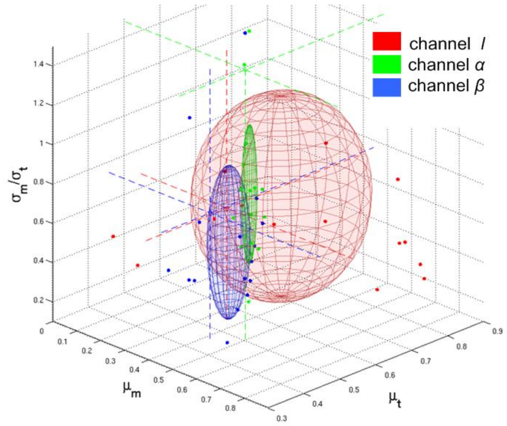

Figure 4. Dashed lines represent the global color transfer function parameters for each channel. Dots represent the parameters of local color transfer functions. Ellipsoids are centered at the average of the mean color of all regions; their size corresponds to one σ in every direction.

transfer functions computed from Fig. 3 (*left*). As expected, the parameters from the local color correction functions (dots in Fig. 4) are different for every region. It is also possible to observe that local color transfer functions sometime have differences to the global parameters of over 0.2 (20%, since the offset values are normalized from 0 to1 on all image channels). Hence, it is possible to conclude that using a single global color transfer function as done in [14] is merely a rough approximation of the real color transfer functions.

Another conclusion can be drawn by noting the larger size of the red ellipsoid when compared to the other two: luminance has higher standard deviations than the *αβ* channels. This indicates that luminance is the channel with highest variability which reinforces the initial hypothesis that assuming a constant illuminant is not feasible for complex environments

3.3. Color Influence Map

Although the importance of using multiple local color transfer functions has been established, the application of the color transfer functions to each pixel must be addressed in order to achieve natural color transition across regions. Hence, in this section we propose a methodology for combining the different local color transfer functions. It is based on the use of Color Influence Maps (CIM), which are computed for every region. The CIM [21] is a weight mask that measures the similarity between each color pixel and the mean color of that particular region. Since in the *lαβ* color space the different channels are uncorrelated, the color similarity can be computed as an Euclidian distance.

$$\mathbf{CIM}^k(i,j) = \mathrm{f}_{\mathbf{CIM}}\left(\left\| c_t(i,j) - \mu_t^k \right\|\right), \qquad (5)$$

Figure 5. Color Influence Maps for the regions in Fig. 3 (*left*).

where $\mathrm{f}_{\mathrm{CIM}}$ is an arbitrary response function; in the current work the function proposed in [21] has been used:

$$\mathrm{f}_{\mathbf{CIM}}(x) = e^{-3x^2}. \qquad (6)$$

Figure 5 shows the CIMs computed for the four regions highlighted in Fig. 3 *(left)*. In these illustrations pixels with a color similar to the mean color value of the considered region are represented with a high value (i.e., white), even though they may not belong to the same region.

3.4. Weighted Color Correction

Finally, in order to merge the different CIMs and get a single color correction value for every pixel we propose a weighted color correction scheme. The final color for a given pixel is obtained by adding the contributions of every color transfer function, weighted by the corresponding CIM:

$$c_t^{new}(i,j) = \frac{\sum_{k=1}^{N}\left(\left(\mu_s^k + \frac{\sigma_s^k}{\sigma_t^k}\left(c_t(i,j) - \mu_t^k\right)\right) \times \mathbf{CIM}^k(i,j)\right)}{\sum_{k=1}^{N} \mathbf{CIM}^k(i,j)}, \qquad (7)$$

where N is the total number of segmented regions of the target image. The final color corrected target image is shown on Fig. 3 *(right)*.

Figure 6 shows the mosaics obtained by merging the model and target images using the following corrections: *(top-left)* non-corrected image; *(top-right)* global color correction; *(bottom-left)* local EM based color correction strategy; *(bottom-right)* proposed local color correction.

4. Results

The proposed approach has been applied to a set of images and compared with both a global approach [14] and the best local color correction approach [17] in the literature [16]. In all the examples the proposed approach

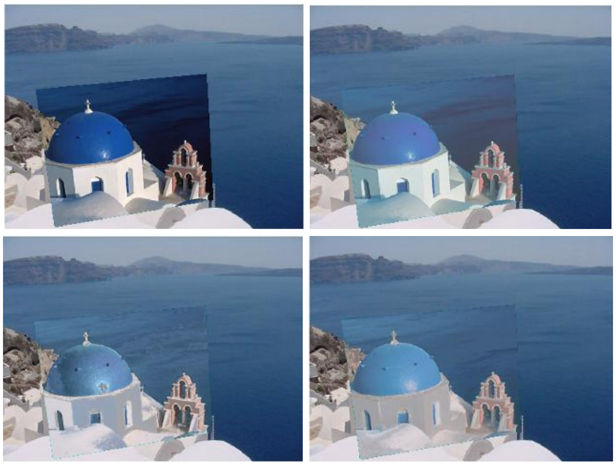

Figure 6. Model image with the target image using the following corrections: *(top-left)* non-corrected image; *(top-right)* global color correction; *(bottom-left)* local EM based color correction; *(bottom-right)* proposed local color correction.

obtains the best results. The color correction ground truth for the target image is computed by assuming a perfect target to model image registration. Hence, the ground truth image borrows the color from the corresponding pixel in the model image. Then, the ground truth image is built using the pixel projection from (2):

$$c_{gt}\left(i,j\right)_t = c\left(i,j\right)_m \ , \qquad (8)$$

where c_{gt} is the ground truth image, $(i, j)_t$ is the target image pixel and $(i, j)_m$ the corresponding pixel in the model image. The ground truth image, although obtained using a sub optimal registration, has many differences to the original target image. In fact, if the ground truth image would be equal to the target image but with its colors replaced by the model's colors, the problem of color correction for registered cameras would be a simple matter of replacing the color of each pixel in the target image by the color of the corresponding pixel in the model image. This is not the case since registration is never optimal. Nonetheless, we use this ground truth image to evaluate quantitatively the results from the proposed approach; comparisons with [14] and [17] are also provided.

In order to evaluate the results of color correction techniques [16] proposes to measure both structure and color similarity. In the current paper, structure similarity cannot be measured due to the inaccurate registrations that exist for the tested image pairs. Therefore, we propose a color similarity criterion inspired in the one presented in [16]. The proposed Color Similarity (*CS*) criterion is defined as the three channel Euclidian distance between the color corrected image and the ground truth image.

$$CS = \left\| c_{gt}\left(i,j\right) - c_t^{new}\left(i,j\right) \right\| . \qquad (9)$$

Then, the color similarity between the original target image and the ground truth image denoted as CS_{base} is computed. This value represents the initial situation that any color correction algorithm should improve. We define color correction improvement ratio (**CC$_{g}$**: for global; **CC$_{EM}$**: for local EM based approach; and **CC$_{PA}$**: for the proposed approach) as the improvement in *CS* obtained by a color correction algorithm over the CS_{base}:

$$CC_{method} = \frac{CS_{base} - CS_{method}}{CS_{base}} \times 100 \ . \qquad (10)$$

The metric presented in (10) has been used for evaluating the performance using different images. Table 1 shows comparisons of the color correction improvement ratio both for local and global methods. It shows that both local methods obtain better results than the global approach. It also shows that our method outperforms the local EM based color correction method, which is the best reported approach in the literature [16]. Two reasons can be mentioned to explain this improvement. First the mean shift algorithm performs better at segmenting the regions. Second, the local EM based is more sensitive to registration inaccuracies, because it matches the blobs based on maximum region overlap, while in our approach we do not segment the model image but collect statistics directly from the projection of every region in the target image to the model image.

Figures 7 and 8 show the images corresponding to the results presented in Table 1. Mosaics are shown without color correction, with global color correction, with local EM based color correction and with the approach proposed in the current paper. It is visible that local approaches obtain better results, especially in images where there is a great variety of colors. The Astro Clock (Fig. 7, 3rd col) is a special case because the clock is in a different position; it was included to try to assert how a small area that has an evident miss registration is handled by the algorithms. In [17], the blue region is not moved while in the proposed approach the blue is somewhat transferred to the left portion of the clock, where it should

Table 1. Comparisons on the improvement of Color Similarity (CS). The improvements over the original target image are shown for global color correction (**CC$_{g}$**) [14], local EM based color correction (**CC$_{EM}$**) [17], and the proposed approach (**CC$_{PA}$**). **CC$_{PA}$** always gets the best result.

Image Pair	CC$_{g}$	CC$_{EM}$	CC$_{PA}$
Santorini, Greece (Figure 1)	48.9	60.0	66.0
Westminster Abbey, London (Figure 7)	8.6	13.7	26.2
Golden Gate, San Francisco (Figure 7)	52.1	55.9	69.9
Astronomical Clock, Prague (Figure 7)	1.9	4.9	6.4
Big Ben, London (Figure 7)	44.6	65.9	69.0
Sagrada Familia, Barcelona (Figure 8)	71.4	73.3	73.3
Ponte Vecchio, Firenze (Figure 8)	14.6	21.6	30.2
Times Square, New York (Figure 8)	25.8	31.2	31.4

be. Regarding Ponte Vecchio (Fig. 8, 2nd col.), the proposed approach was able to correct the yellow houses on both sides of the bridge. In [17], both houses are painted white, which is a clear color correction failure. In Times Square (Fig. 8, 3rd col.) although the sky seems better using [17], the reflection of the yellow beer bottle on the left is supposed to disappear on the corrected image. In this detail, the proposed approach clearly does a better job than [17]. In Sagrada Familia (Fig. 8, 1st col.), a single global color correction obtains the same results as the local techniques, because the target image has a low variety of colors. This can be also appreciated in the quantitative evaluation presented in Table 1. In Big Ben (Fig. 7, 4th col.), the daylight from the target to night in the model is very hard to handle by the global approach.

Finally, although out of the scope of the comparisons, we can mention that the proposed approach requires on average 60 sec. to obtain the result, while the local EM based algorithm requires, on average, over 10 min.

5. Conclusions

This paper presents a new parametric method for local color correction of two coarsely registered images. The method uses the well known mean shift algorithm and builds weight masks using color influence maps. Results show that the proposed approach overtakes the current state of the art. Furthermore, it is computed in about one tenth of the time of the local EM based algorithm.

Future work will include the problem of how to infer color transfer functions when the target image is not entirely contained by the model image. In these cases, the overlapping region must provide information that must be extrapolated to the rest of the image.

Acknowledgements

This work was supported by the Portuguese Foundation for Science and Technology under grant SFRH/43203/2008 and the Spanish Government under Projects TRA2007-62526/AUT and TRA2010-21371-C03-01; Research Program Consolider Ingenio 2010: MIPRCV (CSD2007-00018).

References

[1] Y. Lin and G. Medioni. Map-enhanced UAV image sequence registration and synchronization of multiple image sequences. In *IEEE Int. Conference on Computer Vision and Pattern Recognition,* Minneapolis, USA, 2007.

[2] K. Loewke, D. Camarillo, K. Salisbury, and S. Thrun. Deformable image mosaicing for optical biopsy. In *IEEE Int. Conference on Computer Vision,* Rio de Janeiro, Brazil, October 2007.

[3] L. Vincent. Taking online maps down to street level. *Computer*, 40(12): 118-120, 2007.

[4] B. Micusik and J. Kosecka. Piecewise planar city 3D modeling from street view panoramic sequences. In *IEEE Int. Conference on Computer Vision and Pattern Recognition,* Miami, Florida, USA, June 2009.

[5] D. Capel and A. Zisserman. Automated mosaicing with super-resolution zoom. In *IEEE Int. Conference on Computer Vision and Pattern Recognition,* Santa Barbara, USA, 1998.

[6] V. Lempitsky and D. Ivanov. Seamless mosaicing of image-based texture maps. In *IEEE Int. Conference on Computer Vision and Pattern Recognition,* Minneapolis, USA, June 2007.

[7] A. Levin, A. Zomet, S. Peleg, and Y. Weiss. Seamless image stitching in the gradient domain. In 8th *European Conference on Computer Vision*, Prague, Czech Republic, May 2004.

[8] B. Zitová and J. Flusser. Image registration methods: a survey. *Image and Vision Computing*, 21(11): 977-1000, October 2003.

[9] "Google maps with street view." [Online]. Available: http://www.google.com/intl/en_us/help/maps/streetview/. [Accessed: 10-Nov-2010].

[10] "Welcome to flickr—photo sharing." [Online]. Available: http://www.flickr.com/. [Accessed: 10-Nov-2010].

[11] Y. Xiang, B. Zou, and H. Li. Selective color transfer with multi-source images. *Pattern Recognition Letters*, 30(7): 682-689, May 2009.

[12] F. Pitié, A. C. Kokaram, and R. Dahyot. Automated colour grading using colour distribution transfer. *Computer Vision and Image Understanding*, 107(1-2): 123-137, July 2007.

[13] H. Siddiqui and C. Bouman. Hierarchical color correction for camera cell phone images. *IEEE Trans. on Image Processing,* 17(11): 2138-2155, 2008.

[14] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley. Color transfer between images. *IEEE Computer Graphics and Applications,* 21(5):. 34-41, 2001.

[15] J. Yin and J. R. Cooperstock. Color correction methods with applications to digital projection environments. *Journal of WSCG*, vol. 12: 1-3, 2004.

[16] W. Xu and J. Mulligan. Performance evaluation of color correction approaches for automatic multi-view image and video stitching. *In IEEE Int. Conference on Computer Vision and Pattern Recognition*, San Francisco, USA, 2010.

[17] Y.-W. Tai, J. Jia, and C.-K. Tang. Local color transfer via probabilistic segmentation by expectation-maximization. *In IEEE Int. Conference on Computer Vision and Pattern Recognition*, San Diego, USA, June 2005.

[18] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(5): 603-619, 2002.

[19] D. Comaniciu, P. Meer: Mean shift analysis and applications. *In IEEE Int. Conference on Computer Vision,* Kerkyra, Greece, September 1999.

[20] S. Bagon, Matlab mean shift implementation, Available: http://www.wisdom.weizmann.ac.il/~bagon/, last visited October 2010.

[21] A. Maslennikova and V. Vezhnevets. Abstract interactive local color transfer between images. *In GraphiCon*, Moscow, Russia, June 2007.

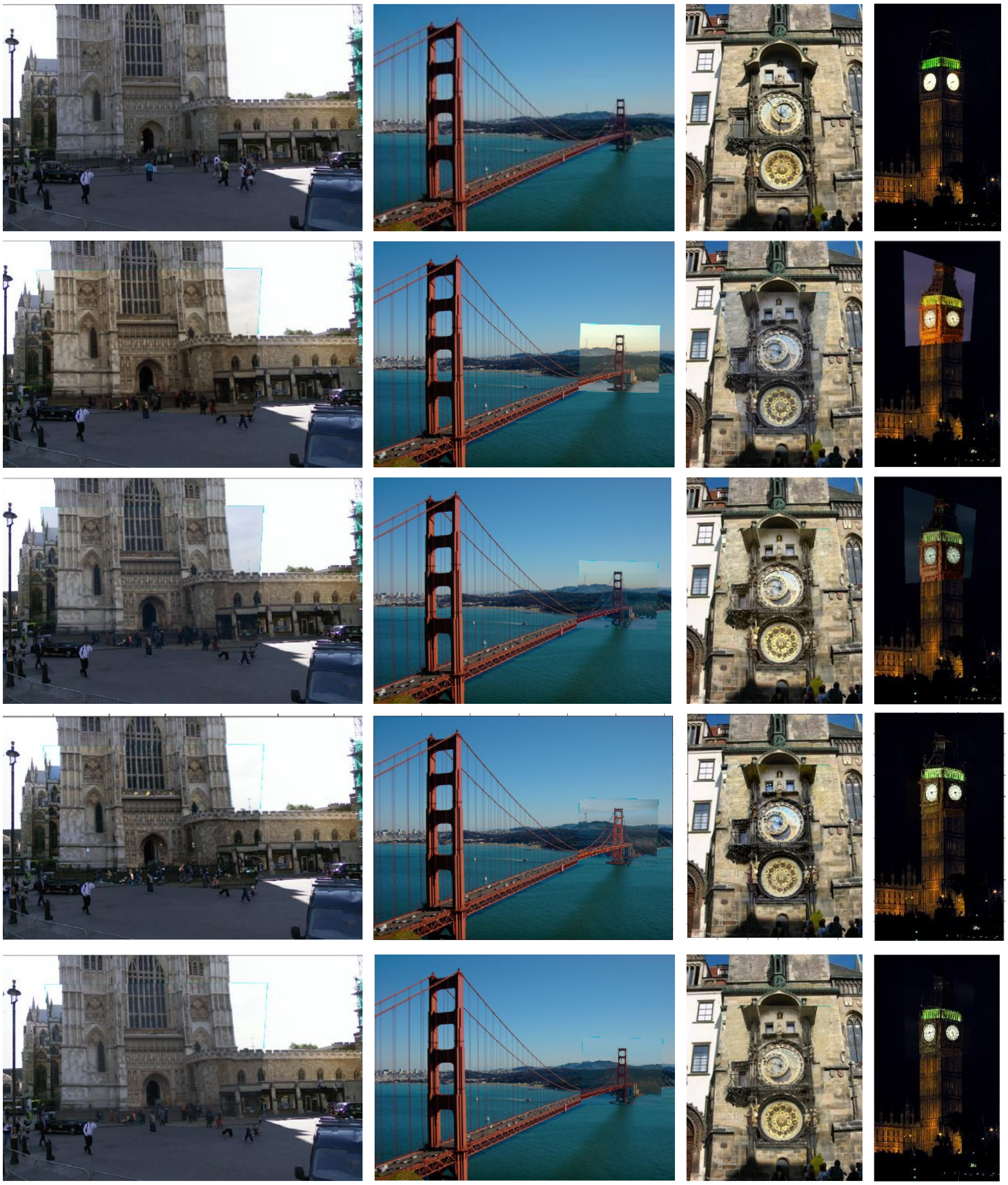

Figure 7. Mosaics for some of the test images mentioned in Table 1. (*from left to right*) Westminster Abbey, London; Golden Gate, San Francisco; Astronomical Clock, Prague; Big Ben, London. (*from top to bottom*): **1st row**, original model image (the objective is to have an image very similar to this one after mosaicing with the target image); **2nd row**, mosaic with the original non color corrected target image; **3rd row**, mosaic with the global color correction [14]; **4th row**, mosaic with the local EM based color correction [17]; **5th row**, mosaic with the proposed local color correction technique.

Figure 8. Mosaics for some of the test images mentioned in Table 1. (*from left to right*) Sagrada Familia, Barcelona; Ponte Vecchio, Firenze; Times Square, New York. (*from top to bottom*): **1st row**, original model image (the objective is to have an image very similar to this one after mosaicing with the target image); **2nd row**, mosaic with the original non color corrected target image; **3rd row**, mosaic with the global color correction [14]; **4th row**, mosaic with the local EM based color correction [17]; **5th row**, mosaic with the proposed local color correction technique.