Tone Correction with Dynamic Objects for Seamless Image Mosaic

Yong-Ho Shin[1], Min-Gyu Park[1], Young-Sun Jeon[2], Young-Su Moon[2], Shi-Hwa Lee[2], and Kuk-Jin Yoon[1]

[1] Gwangju Institute of Science and Technology
Cheomdan-gwagiro, Buk-gu, Gwangju, Republic of Korea
[2] Samsung Advanced Institute of Technology
Giheung-Gu, Yongin-Si, Gyeonggi-Do, Republic of Korea
{yongho,mpark,kjyoon}@gist.ac.kr,
{yong.jeon,mys66,hwa.lee}@samsung.com

Abstract. This paper presents a tone compensation method between images to make a seamless panoramic image. Different camera settings of input images, including white-balance, exposure time, and f-stops, affect the overall color tone of a resultant panoramic image. Although numerous methods have been proposed to deal with such color variations for seamless image stitching, most of them do not properly consider the dynamic scene in which different scene contents exist in input images. In this paper, we propose an efficient method that takes dynamic scene contents into account for compensating color tone difference. The proposed approach consists of three steps. First, we compensate the color tone difference by using the linear color transform with robust local features. Second, we filter out dynamic objects (i.e., dynamic scene contents) by measuring similarity between the linear transformed image and the reference image. Finally, we precisely correct the color variation with detected consistent regions only. The qualitative evaluation shows superior or competitive results compared to commercially available products.

Keywords: Panorama,Tone correction,Seamless image mosaic.

1 Introduction

Image stitching or panorama image mosaic has received substantial attention for decades. Since a resultant image provides wider field of view, it is well suited for some applications such as virtual tour guide and visual surveillance. Moreover, recent mobile phones and compact cameras have begun to embed panorama softwares to overcome their narrow field of view.

For the stitching, the registration of input images is essential that overlays multiple images of the same scene from the different view points at different time. The registration methods can be classified into two categories: direct methods and feature-based methods. Direct methods [1][2][3] use all the pixels in the overlapping area to register images. Although direct methods often require a

K.N. Kutulakos (Ed.): ECCV 2010 Workshops, Part II, LNCS 6554, pp. 104–117, 2012.

Fig. 1. Image stitching result using images with large photometric variation

user input to determine ordering of images, these show accurate results through minimization of the error function which is defined for entire pixels. On the other hand, feature-based methods [4][5][6] stitch each image by identifying local features such as blobs or corner points. Such features are robust under particular variations which cause degenerative results. Furthermore, distinctiveness and repetitiveness of features ensure the correct alignment of an unordered set of input images. For that reason, feature-based approaches are recently esteemed as a standard method in commercial products.

The feature-based image registration usually consists of feature extraction, matching, outlier filtering, and bundle adjustment steps. However, the registration of images does not guarantee a plausible stitching result since the difference of color tones and the presence of dynamic objects cause unnatural heterogeneity at stitched boundary. Therefore, consequent steps such as blending and optimal seam algorithms, performed after the image registration, focus on eliminating these artifacts to make seamless panorama images. Blending has been widely used for seamless image mosaic. Burt and Adelson [7] introduced the multi-band blending method that controls the degree of blending according to image contents. Perez el al. [8] suggested Poisson blending which is developed for sophisticate merging. Instead of blending visual seam along the image boundary, an optimal seam method finds a path that passes through the overlapping area with minimum difference. Davis [9] used Dijkstra's algorithm with difference of intensity. Uyttendaele et al. [10] used a vertex cover algorithm and Mills and Dudek [11] proposed compound difference of intensities and magnitude of gradients.

Since any blending techniques can be applied after the optimal seam selection, the combination of two methods can handle geometric miss alignment error as well as dynamic objects. It works fairly nice, unless color tone variation is significant. Fig. 1 illustrates the result of applying the combination of two methods under large color tone variation. From the geometrical point of view, input images are well aligned. However, we can see the unnatural seam owing to the photometrical variation.

(a) input images with dynamic objects

(b) image stitching result

Fig. 2. Degenerative result due to different scene contents when using the linear color transform [14]

Several methods have been proposed to compensate color tones while making different assumptions on the scene. Some methods [12][13] used a histogram matching technique and Tian el al. [14] proposed a correction method based on the linear color transform. However, they assumed static scenes and, therefore, dynamic objects or scenes seriously degrade the quality of a stitching result (see Fig. 2). To cope with dynamic objects, Mills and Dudek [11] used a linear transform with the random sample consensus (RANSAC) technique to filter out moving objects. However, the linear color transform used in that method cannot handle complex color variation sufficiently. Goldman and Chen [15] tried to estimate the camera response function (CRF) and vignette coefficients from given correspondences, and then compensated the color tone with the calculated CRF. However, the performance of CRF-based correction totally depends on the given correspondence and, therefore, it requires very accurate feature matching results across images. Furthermore, if some assumptions made on the camera parameters (such as constant focal lengths and f-stops) are not satisfied, the CRF-based methods do not ensure plausible color tone correction results.

To resolve the problems in variety cases, we propose a new color tone compensation method. The proposed method consists of a few steps. Initially, the linear

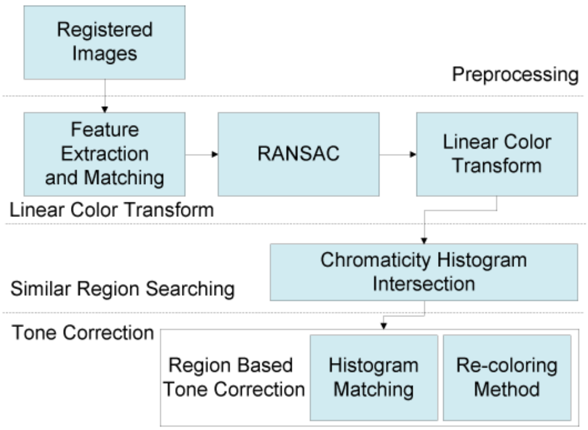


Fig. 3. Overall procedure of the proposed method

color transform is adopted to roughly compensate the different exposure and white balance under physically inconsistent situations. Then, geometrically consistent regions (corresponding to static scene contents) are computed by using the color histogram intersection method. Finally, globally adoptable tone correction methods such as histogram specification and re-coloring methods generate photometrically consistent panoramic images.

2 Proposed Color Tone Correction Method

As mentioned, the proposed algorithm consists of three main parts. The first part is the linear color transformation with robust local features. The second part is the color histogram based intersection to detect geometrically consistent regions. Finally, region based tone correction methods compensate the color variations of input images. Fig. 3 shows the overall procedure of the proposed method. Here, we assume that the input images are already well aligned geometrically in the preprocessing stage.

2.1 Linear Color Transform with Robust Features

The objective of this step is finding a linear color transform matrix as in [14] which compensates different camera settings such as white balances and exposures. To acquire a reliable color transform matrix, measurement data points should be physically consistent. Since the scene is not always static in practice, we need a methodology of finding reliable matches which correspond to the same physical scene points. For this, we use scale invariant feature transform (SIFT) [16] which is commonly used in many computer vision applications. SIFT features are invariant to rotation and scale changes, and it is also invariant to illumination variances. Because SIFT generates feature descriptors from the image gradients and normalizes the descriptor vectors, photometric variations merely affect the feature descriptors. Therefore, SIFT-based correspondence searching ensures physically consistent matches and it is suitable for calculating color transformation matrix as well.

Here, the linear color transformation approach assumes that optical sensors work fairly linear so that pixel color or luminance values are proportional to the amount of incoming light. Base on this assumption, a diagonal model [17] might be sufficient for the color tone correction. However, more accurate transform is required to deal with more complex color transition having non-linearity which is common in practice. For example, the auto white balance compensates different color temperatures non-linearly and the analog-digital converter for digital cameras does not guarantee the linearity, neither. Therefore, to solve these problems approximately as possible, we define the transformation matrix having twelve degrees of freedom as below,

$$\mathbf{M} = \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \end{pmatrix}, \mathbf{I_t} = \begin{pmatrix} R_t \\ G_t \\ B_t \\ 1 \end{pmatrix}, \mathbf{I_r} \begin{pmatrix} R_r \\ G_r \\ B_r \end{pmatrix}, \tag{1}$$

where the subscript r and t indicate the reference image and the target image respectively. The reference is automatically set among overlapping images, which has the largest contrast of brightness. After that, the matrix representing linear transformation between the sets of pixels, can be calculated by

$$\mathbf{MI_t} = \mathbf{I_r}. \tag{2}$$

The matrix $\mathbf{M}$ can be computed in the least square sense. The linearly transformed image is shown in Fig. 4. It works well when corresponding features are obtained accurately.

2.2 Region Searching with Chromaticity Histogram Intersection

To apply region-based tone correction techniques, we first find geometrically consistent regions. Since the previous step is a feature-based technique, the transform uses insufficient color information from the overlapping area. The main concern of this step is that, although the feature-based linear tone correction has its own strengths, it does not guarantee plausible results if there are just a few correspondences. To deal with these problems, we introduce a consistent region searching method based on the chromaticity histogram, which examines geometrically consistent regions inside the overlapping area. Instead of using the direct comparison of RGB values, we transform the color space from RGB to chromaticity as

$$c_r = \frac{R}{(R+G+B)}, c_b = \frac{B}{(R+G+B)}, \tag{3}$$

(a) input images with dynamic objects

(b) image stitching result

Fig. 4. Image stitching result with the linearly transformed image

where R, G, B represent each RGB value and (c_r, c_b) represents the chromaticity. Since this conversion reduces the dimension of the color space, it reduces computational complexity while increasing the reliability of color matching under luminance changes.

Here, instead of using individual pixel values, we define the histogram and region based similarity measure to avoid the effect of geometrical misalignment as

$$\frac{\sum_{c_r}\sum_{c_b} min\left(h(c_r, c_b), g(c_r, c_b)\right)}{\sum_{c_r}\sum_{c_b} h(c_r, c_b)}, \tag{4}$$

where h and g represent the histogram of each image in the overlapping area. The denominator is a normalizing constant which counts the number of pixels in the overlapping area, and the numerator indicates summation of the intersected histogram. As the similarity ratio increases, the intersected histogram preserves an original shape of histogram. Therefore, it measures the region based similarity based on the color histogram which is not or merely affected by the geometrical misalignment errors.

Based on the color histogram intersection method, the proposed method finds reliable regions with the following procedure. First, we initialize a binary mask which represents the overlapping area and outside of the overlapping area (see Fig. 5). Each pixel serves as a label where white indicates inliers and black denotes outliers. Second, we compute the similarity ratio and compare it with the

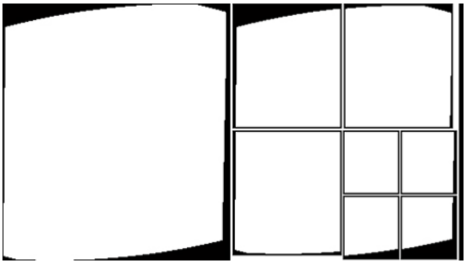


Fig. 5. A mask and the illustrative procedure of splitting

predefined threshold (we use 95 percent). If the similarity is less than the threshold, the region split into quad sub blocks as shown in Fig. 5. For each sub block, we repeat the same procedure until the size of a split block becomes smaller than the predefined size. The final result of the algorithm is the mask representing inliers (geometrically consistent scene contents) and outliers (geometrically inconsistent scene contents) as shown in Fig. 6. Based on this mask, region based tone compensation is performed in the next step.

Fig. 6. Two input images and the result of the consistent region examination

2.3 Region Based Tone Correction

Since the inlier mask represents geometrically consistent regions, the last stage of the proposed method is to create photometrically consistent results through region based tone correction methods such as the histogram specification and the re-coloring [18].

Histogram Based Tone Correction. The histogram based correction method performs histogram specification for each RGB channel. It finds an mapping function from the cumulative histogram H_2 of the target image to the reference histogram H_1 as follow,

$$z = H_1^{-1}(H_2(r)), \tag{5}$$

where r and z are continuous random variables representing input and output images, respectively. Although histogram specification is simple and quite conventional, it shows reasonable results. However, this method has a problem in cope with the tone correction when the cumulative histogram changes abruptly. For example, if one particular color is dominant and the distribution is narrow, numerous pixels are mapped to the dominant color.

Color Transfer Based Tone Correction. In this approach, we use a color transfer method instead of histogram specification. Color transfer or re-coloring [18] technique initially transforms the color space from RGB into $\ell\alpha\beta$, developed by Ruderman at al. [19], and corrects the color tone using mean and standard deviation of each image.

Since the $\ell\alpha\beta$ space has been developed to reduce correlation among channels assuming that the human visual system is ideally suitable for natural scenes. This color space shows the least correlation between each plane especially for natural scenes, thus, we do not need to change the value of pixel in a coherent way. Finally, by using characteristics of the each images in the $\ell\alpha\beta$ space, the re-colored image is obtained as below,

$$\begin{aligned} \ell' &= \frac{\sigma_r^\ell}{\sigma_t^\ell}(\ell - \mu_t^\ell) + \mu_r^\ell \\ \alpha' &= \frac{\sigma_r^\alpha}{\sigma_t^\alpha}(\alpha - \mu_t^\alpha) + \mu_r^\alpha \\ \beta' &= \frac{\sigma_r^\beta}{\sigma_t^\beta}(\beta - \mu_t^\beta) + \mu_r^\beta, \end{aligned} \tag{6}$$

where σ and μ indicate standard deviation and mean, and r and t denote the reference and the target images. Fig. 7 shows the results by using the histogram based and re-coloring based methods, respectively.

3 Experiment

The experimental images are taken by differentiating white-balance, exposure time, ISO, and f-stops using a Cannon 1Ds camera with a manual focus Canon 24-70mm lens. Table 1 and 2 provide the camera settings for input images in Fig. 8 and Fig. 9. The experiment includes the results of the linear color transformation, the histogram matching, and the proposed method. In addition, three commercial products, PTGui [20], Hugin [21], and Autostitch [22], are compared for the performance evaluation.

Fig. 8 shows the result of each approach for the static scene and the same scene with a dynamic object. For the comparison, the second and the third rows illustrate the results from the static scene, (a) and (b), whereas the fifth and the sixth rows describe the results from the scene with the dynamic object, (i) and (j). First of all, the linear transformation [14] results, (c) and (k), are hazy because it cannot fully handle non-linear photometric variations. Moreover, (k) is also influenced by the dynamic object. Next, histogram matching deals with non-linearity and generates a reliable result, (d), for the static scene. However,

(a) histogram matching result

(b) $\ell\alpha\beta$ re-coloring result

Fig. 7. Results of histogram matching and $\ell\alpha\beta$ re-coloring

since histogram matching does not consider dynamic objects, it is dominated by the yellow artificial object as shown in (l). The proposed methods are shown to be robust in both cases, (e) and (m), regardless of inconsistence scene contents.

Table 1. Different camera settings for input images in Fig. 8

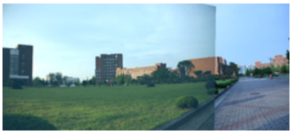
-	Input image 1	Input image 2
White-balance	Day light	Tungsten light
Exposure time	1/500 sec	1/320 sec
ISO	ISO-100	ISO-200
F-stops	f/3.5	f/4.5

Table 2. Different camera settings for input images in Fig. 9

-	Input image 1	Input image 2	Input image 3	Input image 4	Input image 5
White-balance	Day light	Cloudy	Day light	Tungsten light	Day light
Exposure time	1/200 sec	1/200 sec	1/100 sec	1/160 sec	1/320 sec
ISO	ISO-200	ISO-160	ISO-500	ISO-200	ISO-100
F-stops	f/8	f/6.3	f/8	f/11	f/5

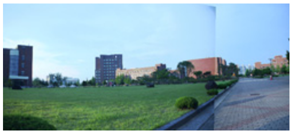
(a) input image 1-1

(b) input image 1-2



(c) linear transformation 1

(d) histogram matching 1



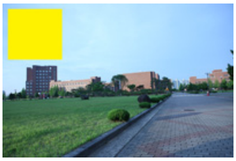
(e) proposed method 1

(f) PTGui 1 [20]

(g) Hugin 1 [21]

(h) Autostitch 1 [22]

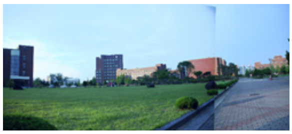
(i) input image 2-1



(j) input image 2-2

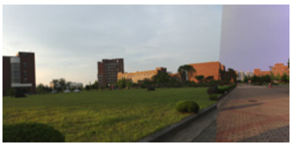
(k) linear transformation 2

(l) histogram matching 2

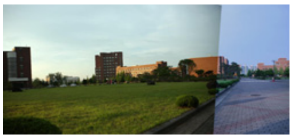


(m) proposed method 2

(n) PTGui 2 [20]

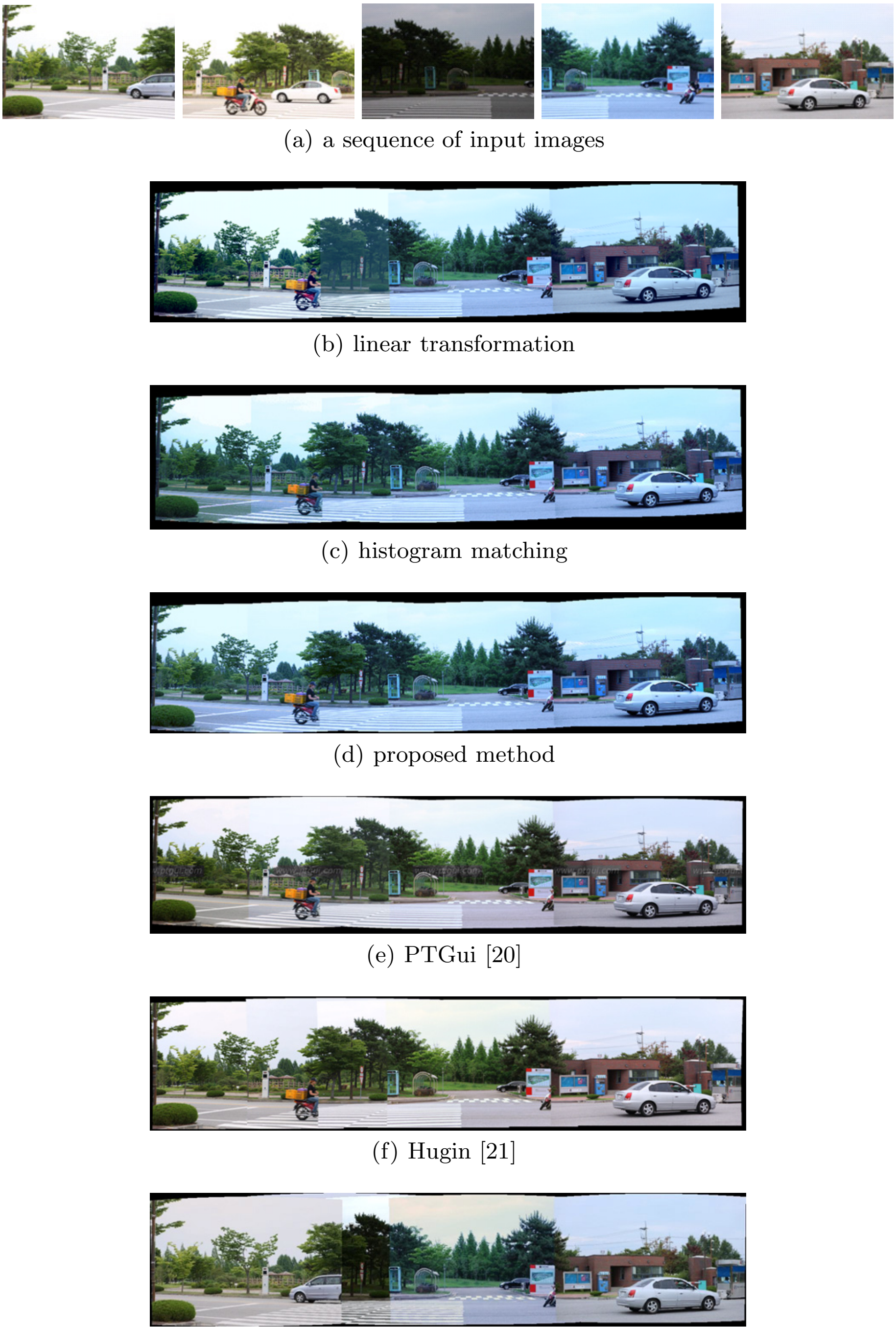


(o) Hugin 2 [21]



(p) Autostitch 2 [22]

Fig. 8. Comparison of cropped results with other methods



(a) a sequence of input images

(b) linear transformation

(c) histogram matching

(d) proposed method

(e) PTGui [20]

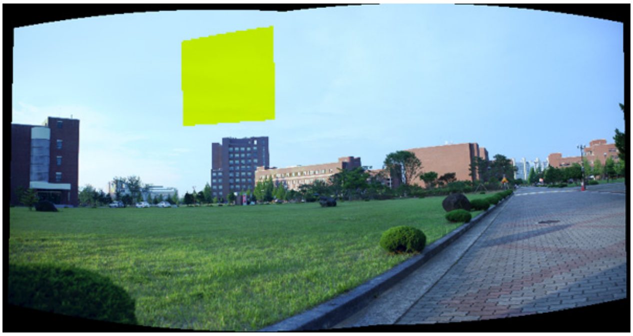
(f) Hugin [21]

(g) Autostitch [22]

Fig. 9. Comparison of results with other methods for dynamic scene contents

(a) image stitching result for input images in Fig. 2

(b) image stitching result for input images in Fig. 4



(c) image stitching result for input images in Fig. 8

(d) image stitching result for input images in Fig. 9

Fig. 10. Results of the proposed algorithm combined with the optimal seam selection and the multi-band blending

In addition, commercial products are evaluated for the comparison. CRF-based approaches, PTGui and Hugin, compute CRF based on local features, thus, they are not influenced by dynamic objects. However, variation of un-modeled camera parameters such as ISO and f-stops severely degrade quality of the results, (f),(g),(n),and (o). Autostitch which uses gain compensation method , adjusts global intensity for each channel. Thus, (h) and (p) looks better than the original images but it also has a problem with complex color variation, plus, (p) is dominated by the yellow artificial object.

Fig. 9 illustrates the results from a sequence of images having several dynamic objects and different camera settings. Similar to Fig. 8, dynamic objects degrade the results of (b), (c), and (g). Next, variation of un-modeled camera parameters spoil the results, (e) and (f). Lastly, complex color variation causes unsatisfactory results in (b) and (g). The proposed method can be combined the optimal seam and multi-band blending techniques to produce more natural and realistic results as shown in Fig. 10.

4 Conclusion

We have presented a new method which solves color tone correction problem for seamless image stitching, especially when the scene is not static and input images are taken with different settings. The proposed method redresses the photometric inconsistency through the linear color transform, chromaticity histogram intersection, and region based compensation. The feature-based linear transform is robust under different camera settings and the existence of moving objects. However, it is usually inadequate to solve entire color tone correction problem. To make up this limitation, chromaticity histogram intersection is adopted to find geometrically consistent regions. Afterwards, region based approaches compensate color tone difference more accurately with the consistent regions only. The overall procedure is designed to be used in the real environment where dynamic scene contents exist. The experimental results show that the proposed method is superior to the previous methods.

References

1. Szeliski, R.: Image mosaicing for tele-reality applications. In: IEEE Workshop on Applications of Computer Vision, pp. 44–53 (1994)
2. Szeliski, R., Corporation, M.: Video mosaics for virtual environments. IEEE Computer Graphics and Applications 16, 22–30 (1996)
3. Shum, H.Y., Szeliski, R.: Construction of panoramic mosaics with global and local alignment. International Journal of Computer Vision 36, 101–130 (2000)
4. Zoghlami, I., Faugeras, O., Deriche, R., Antipolis Cedex, F.S.: Using geometric corners to build a 2d mosaic from a set of images. In: CVPR 1997, pp. 420–425 (1997)
5. Mclauchlan, P.F., Jaenicke, A., Xh, G.G.: Image mosaicing using sequential bundle adjustment. In: Proc. BMVC, pp. 751–759 (2000)

6. Brown, M., Lowe, D.: Automatic panoramic image stitching using invariant features. International Journal of Computer Vision 74, 59–73 (2007)
7. Burt, J., Adelson, H.: A multiresolution spline with application to image mosaics. ACM Transactions on Graphics 2, 217–236 (1983)
8. Pérez, P., Gangnet, M., Blake, A.: Poisson image editing. ACM Trans. Graph. 22, 313–318 (2003)
9. Davis, J.: Mosaics of scenes with moving objects. In: CVPR 1998: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, p. 354. IEEE Computer Society, Washington, DC (1998)
10. Uyttendaele, M., Eden, A., Szeliski, R.: Eliminating ghosting and exposure artifacts in image mosaics. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, p. 509 (2001)
11. Mills, A., Dudek, G.: Image stitching with dynamic elements. Image and Vision Computing 27, 1593–1602 (2009)
12. Zheng, L., Zhang, J., Luo, Y.: Color matching in colour remote sensing image. In: IMSCCS 2006: Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences, vol. 1, pp. 303–306. IEEE Computer Society, Washington, DC (2006)
13. Azzari, P., Bevilacqua, A.: Joint Spatial and Tonal Mosaic Alignment for Motion Detection with PTZ Camera. In: Campilho, A., Kamel, M.S. (eds.) ICIAR 2006. LNCS, vol. 4142, pp. 764–775. Springer, Heidelberg (2006)
14. Tian, G.Y., Gledhill, D., Taylor, D., Clarke, D.: Colour correction for panoramic imaging. In: Proceedings of the Sixth International Conference on Information Visualisation, vol. 0, pp. 483–488. IEEE Computer Society, Los Alamitos (2002)
15. Goldman, D.B.: hung Chen, J.: Vignette and exposure calibration and compensation. In: ICCV 2005: Proceedings of the Tenth IEEE International Conference on Computer Vision, vol. 1, pp. 899–906. IEEE Computer Society (2005)
16. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60, 91 (2004)
17. Finlayson, G.D., Drew, M.S., Funt, B.V.: Color constancy: Generalized diagonal transforms suffice. J. Opt. Soc. Am. A 11, 3011–3020 (1994)
18. Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. IEEE Computer Graphics and Applications 21, 34–41 (2001)
19. Ruderman, L., Cronin, W., Chiao, C.C.: Statistics of cone responses to natural images: Implications for visual coding. Journal of the Optical Society of America A 15, 2036–2045 (1998)
20. PTGui, `http://www.ptgui.com`
21. Hugin, `http://www.hugin.sourceforge.net`
22. Autostitch, `http://cvlab.epfl.ch/~brown/autostitch/autostitch.html`