Assessment approach for image mosaicing algorithms

Li-hui Zou
Jie Chen
Juan Zhang

Assessment approach for image mosaicing algorithms

Li-hui Zou, Jie Chen, and Juan Zhang
Beijing Institute of Technology University
School of Automation, Beijing, China, 100081, and Ministry of Education, Key Laboratory of Complex System Intelligent Control and Decision (Beijing Institute of Technology)
Beijing, China, 100081
E-mail: zoulihui@bit.edu.cn

Abstract. In this paper, a novel quantitative assessment approach for image mosaicing algorithms is proposed by evaluating the quality of an output mosaic image and comparing with the input images. Five evaluation indices—entropy, clarity, registration error, peak signal-to-noise ratio, and structural similarity—are calculated and combined with the analytic hierarchy process for comprehensive assessment. This approach can overcome the lack of panoramic image and is easy to implement. Experimental results show the objectivity and the validity of the proposed approach. © *2011 Society of Photo-Optical Instrumentation Engineers (SPIE).* [DOI: 10.1117/1.3646746]

1 Introduction

Image mosaicing has impacted a wide variety of application areas.[1] It is to combine multiple images with overlapping areas into a wide field or high-resolution image.[2] The process is divided into two steps. The first step is image registration whose purpose is to find the coordinate transformation between a pair of images and align them. The second is image blending which aims at smoothing the transition areas and minimizing visual flaws, such as ghost, exposure difference, or misalignments. The performance of the algorithm is, therefore, both related to these two steps.

To evaluate the performance of the image mosaicing algorithm quantitatively is indispensable for algorithm comparisons and improvements. However, this aspect of research is still in an early stage of development due to the difficulties in obtaining the reference panoramas. The general solutions usually require manual interventions. In Ref. 3, it rates the mosaicing algorithm based on the comparison between the output yielded on standard data sets and the ground truth stitched manually. In Ref. 4, the input data sets are collected by virtual camera, a software module simulating the imaging process of an actual device. The shooting parameters are fully customizable so that it is possible to compose the ground truth. In Ref. 5, a reference image is specified as the real world scenery in a test environment. Similarly, the inputs are simulated by applying a group of common imaging distortions on narrow views of the reference image.

In this paper, an assessment approach for image mosaicing algorithms without reference panoramas is proposed. Five evaluation indices are calculated and combined with the analytic hierarchy process (AHP) for comprehensive assessment. It can reduce manual interventions and make the evaluation much more convenient.

2 Assessment Scheme for Image Mosaicing

A novel assessment scheme for image mosaicing is established as seen in Fig. 1. The most notable advantage of the approach is that it does not need a panoramic image but the original input narrow-angle image instead for comparison. It is deemed that the gathered image for stitching is the most realistic view from that angle in that projection. It is fair enough for partial comparison. The indices extracted in this aspect belonging to full-reference assessment are calculated by comparing the similarity between the original input images and their neighboring images, or their inverse warped partial mosaics. As for the whole mosaic image, it is up to its own performance in perception. The related indices belonging to no-reference assessment are calculated by measuring the output mosaics globally.

3 Evaluation Criteria and Comprehensive Assessment

3.1 No-Reference Criteria

The no-reference criteria include two indices. The first one is entropy for characterizing the texture and information amount of the output mosaics I_{mos}, denoted as C_1. The higher the value of C_1, the richer the image information. The second one is clarity and defined as follows.

$$C_2 = \mathrm{Clarity}_{\mathrm{Sobel8}}(I_{\mathrm{mos}}) = \sum_{x=1}^{M}\sum_{y=1}^{N}|H(x,\,y)|^2,\, H(x,\,y) > T_N,$$

$$H(x,\,y) = \sqrt{\sum(I_{\mathrm{mos}}(x,\,y)*S_j)^2},\qquad (1)$$
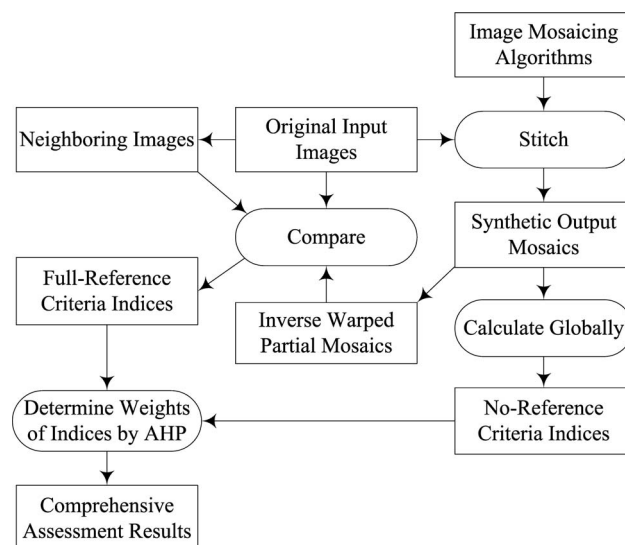
Fig. 1 The framework of the proposed approach.
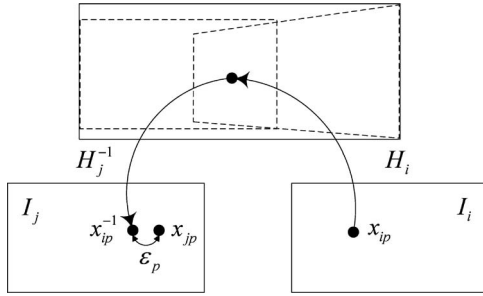
Fig. 2 The computation process of registration error.

Fig. 3 The hierarchy tree of evaluation indices for visual quality.

where S_i are Sobel masks in eight directions (0 deg, 45 deg, . . . , 315 deg),[6] T_N is the standard deviation of I_{mos}, and $M \times N$ is the size of I_{mos}. The higher the value of C_2, the less the blur of the mosaic image.

3.2 *Full-Reference Criteria*

The full-reference criteria include registration error, peak signal-to-noise ratio (PSNR) and structural similarity (SSIM), which are defined as C_3, C_4, and C_5.

$$C_3 = \frac{1}{N-1} \sum_{n=1}^{N-1} \varepsilon_n$$

$$= \frac{1}{N-1} \sum_{n=1}^{N-1} \left(\frac{1}{P} \sum_{p=1}^{P} \| x_{jp} - H_j^{-1} H_i x_{ip} \| \right) \quad (2)$$

$$C_4 = \frac{1}{N} \sum_{n=1}^{N} \mathrm{PSNR}(\hat{r}_n, s_n) = \frac{1}{N} \sum_{n=1}^{N} 20 \lg \frac{255}{\sqrt{\mathrm{MSE}(\hat{r}_n, s_n)}}, \quad (3)$$

$$C_5 = \frac{1}{N} \sum_{n=1}^{N} \mathrm{SSIM}(\hat{r}_n, s_n)$$
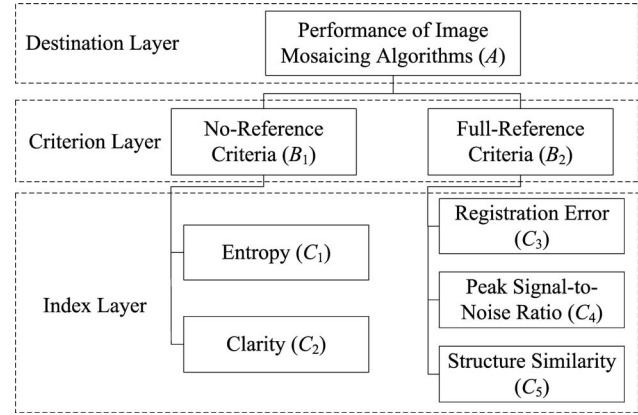
$$= \frac{1}{N} \sum_{n=1}^{N} \frac{1}{K} \sum_{k=1}^{K} [l_n(x_k, y_k)]^{\alpha} [c_n(x_k, y_k)]^{\beta} [s_n(x_k, y_k)]^{\gamma}, \quad (4)$$

where P is the number of feature pairs in neighboring images, and the transform relations are shown in Fig. 2, N is the number of input images or interested regions, $\hat{r}_n$ is the inverse warped partial mosaic image, and s_n is the input image. Refer to Ref. 8 for more parameter details of SSIM. The lower the value of C_3, the smaller the registration error. The higher the value of C_4, the more continuous in blending intensity. The higher the value of C_5, the less structure difference between the input sequence and the synthetic mosaics.

3.3 *Comprehensive Assessment Based on AHP*

For different applications, the demands for image mosaicing are usually different. The above criteria play roles of varying importance. To determine the weight coefficients, the problem is modeled as a hierarchy, shown in Fig. 3, and pairwise comparisons of the AHP are performed. The comprehensive assessment result is obtained by three steps.

The first step is constructing a set of pairwise comparison matrices U using a scale of values ranging from 1 (equal importance) to 9 (absolute importance).

$$U = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & & u_{2n} \\ \dots & & & \\ u_{n1} & u_{n2} & & u_{nn} \end{bmatrix}, \quad (5)$$

where u_{ij} is the relative importance value of the i'th factor compared to the j'th, therefore, $u_{ij} = 1/u_{ji}$, $u_{ii} = 1$. The elements of U can be determined by statistical data, intensity of requirement on issues, or experts' knowledge. Taking the application in photograph editing for instance, they can be defined as Table 1. For the criterion layer, the local full-reference criteria are deemed to be slightly important over the global no-reference criteria, since photograph editing requires high accuracy and the evaluation usually pays close attention to local details of the mosaics. Other matrices can be also determined in this way, seen in Table 1.

The second step is computing a vector of priorities in each layer. It equals the normalized eigenvector $[w_1, w_2, \dots, w_n]^T$ corresponding to the maximum eigenvalue of U, where w_i is calculated by

$$w_i = \sqrt[n]{\prod_{j=1}^{n} u_{ij}} \Bigg/ \sum_{i=1}^{n} \sqrt[n]{\prod_{j=1}^{n} u_{ij}}. \quad (6)$$

In addition, the judgments must pass through a consistency check.[9] The computation results of the above case are also listed in Table 1.

The third step is synthesizing priority by aggregating the relative weights up the hierarchy. The comprehensive

Table 1 An example on comparison matrices and weights in the hierarchy.

Criterion layer				Index layer B_1				Index layer B_2				
A	B_1	B_2	w_i	B_1	C_1	C_2	$w_1{}^i$	B_2	C_3	C_4	C_5	$w_2{}^i$
								C_3	1	1	1/2	0.240
B_1	1	1/2	0.333	C_1	1	1/3	0.250	C_4	1	1	1/3	0.210
B_2	2	1	0.667	C_2	3	1	0.750	C_5	2	3	1	0.550

Table 2 Values of the evaluation indices.

		C_1	C_2	C_3	C_4	C_5
sc. 1	#1	6.8308	22187.4	0.3278	24.0025	0.9991
	#2	6.8020	45455.2	0.4309	25.6590	0.9991
sc. 2	#1	7.2475	16138.5	0.2098	27.8449	0.9998
	#2	7.2059	28174.2	0.2166	26.7563	0.9997
sc. 3	#1	7.2618	12553.5	0.3816	26.9503	0.9998
	#2	7.2386	23478.5	0.3699	26.2622	0.9998
sc. 4	#1	6.7956	16256.3	0.2025	26.4237	0.9997
	#2	6.5323	30427.2	0.2739	26.1030	0.9996
sc. 5	#1	7.0709	15771.4	0.2677	25.1656	0.9996
	#2	7.0754	32270.6	0.3801	24.1162	0.9995

assessment result is expressed as

$$A = \sum_{i=1}^{m} \left(w_i \sum_{j=1}^{n} w_j^i C_i \right), \tag{7}$$

where w_i is weight coefficient of evaluation factor in criterion layer, and w_j^i is weight coefficient of evaluation index of the i'th criterion factor in the j'th index layer, C_i is the normalized index data. In the example, the computed overall weight vector becomes $W = [0.083, 0.250, 0.160, 0.140, 0.367]$.

4 Experiments and Conclusions

Two algorithms for photograph editing are compared here, denoted by #1 and #2. They are both based on a projective model, but extracting different features for registration. The blending methods are also different, #1 using center-weighting but #2 using optimal seam stitching. In order to test the validity of the proposed method, a series of mosaics stitched by these algorithms were evaluated by 15 volunteers. They observed each pair of mosaics under the same viewing

Table 3 The comparison of assessment results.

		Proposed method				Compared method	
		No-ref.	Full-ref.	Overall		rms error	
sc. 1	#1	0.5061	0.6269	0.5867	√	0.2556	×
	#2	0.8504	0.5911	0.6775		0.2623	
sc. 2	#1	0.5501	0.6132	0.5922	√	0.1258	√
	#2	0.8271	0.6018	0.6769		0.1126	
sc. 3	#1	0.5307	0.6068	0.5814	√	0.1031	√
	#2	0.8379	0.6083	0.6848		0.0905	
sc. 4	#1	0.5337	0.6355	0.6015	√	0.1151	√
	#2	0.8348	0.5833	0.6671		0.1068	
sc. 5	#1	0.5060	0.6422	0.5968	√	0.1041	×
	#2	0.8507	0.5778	0.6687		0.1127	
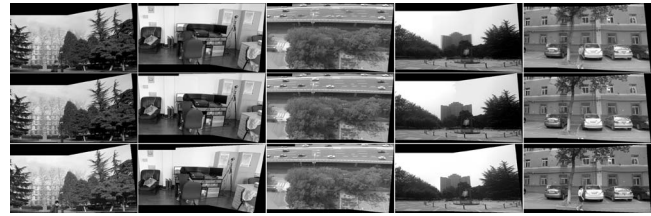
Fig. 4 Assessment experiments: the first row is the mosaicing results of #1; the second row is the mosaicing results of #2; the third row is manual stitched panoramas.

conditions and voted for the algorithm. The algorithm with more votes was deemed as the better one. For all the test scenes, #2 was superior to #1 in our subjective experiments. Due to space limitations, we present five typical scenes with different complexities and textures for illustration. We also compared with another evaluation method on the basis of rms intensity error between the final mosaics and the reference panoramas stitched by hand. The mosaicing results and the reference panoramas are shown in Fig. 4. The computed index data in the proposed scheme are shown in Table 2. The relative parameters are set as recommendations in references. Their assessment results are shown in Table 3. They are all consistent with subjective judgments. Although some local moving objects might be lost by #2 as seen in Fig. 4, which results in lower full-reference evaluation, its global impression significantly defeats the local details by avoiding the blending ghost. However, some evaluation results of the compared method are contrary to subjective judgments as seen in Table 3, marked by "×," mainly caused by moving objects which increase the rms intensity error if taking a large portion in the input image and lost in the final mosaics.

The above experiments show the objectivity and the stability of the proposed method. Owing to multicriteria and AHP weights selection, the assessment is robust even if there are variances and disturbances in different scenes. Indices extension and parameter analysis will be the focus in future research.

References

1. L. H. Zou, J. Chen, J. Zhang, and J. Lu, "Real-time wide-field monitoring system based on image mosaicing technique and its application," *Proceedings of the 18th IFAC World Congress*, Milano, Italy, pp. 14928–14933 (2011).
2. H. Lin, J. Si, and G. P. Abousleman, "Dynamic point selection in image mosaicking," *Opt. Eng.* **45**, 030501 (2006).
3. M. Brown, "Multi-image matching using invariant features," Ph.D. thesis, University of British Columbia, Vancouver, BC, pp. 74–87 (2005).
4. P. Azzari, L. D. Stefano, and S. Mattoccia, "An evaluation methodology for image mosaicing algorithms," *Proceedings ACIVS 2008*, Juan-les-Pins, France, pp. 89–100 (2008).
5. J. Boutellier, O. Silvén, M. Tico, and L. Korhonen, "Objective evaluation of image mosaics," *Commun. Comp. Info. Sci.* **21**, 107–117 (2009).
6. T. Jiang, Y. Tan, and Q. Liu, "Research of a clarity-evaluation function of image based on Sobel," *Comp. Dig. Eng.* **36**(8), 129–131 (2008).
7. W. Xu and J. Mulligan, "Performance evaluation of color correction approaches for automatic multi-view image and video stitching," *IEEE Conference on CVPR*, San Francisco, USA, pp. 263–270 (2010).
8. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**(4), 1–14 (2004).
9. E. Kwon and W. I. Ko, "Evaluation method of nuclear nonproliferation credibility," *Ann. Nucl. Energy* **36**, 910–916 (2009).