Adaptive As-Natural-As-Possible Image Stitching

Chung-Ching Lin, Sharathchandra U. Pankanti,
Karthikeyan Natesan Ramamurthy, and Aleksandr Y. Aravkin
IBM Thomas J. Watson Research Center, 1101 Kitchawan Road, Yorktown Heights, NY 10598
{cclin,sharath,knatesa,saravkin}@us.ibm.com

Abstract

The goal of image stitching is to create natural-looking mosaics free of artifacts that may occur due to relative camera motion, illumination changes, and optical aberrations. In this paper, we propose a novel stitching method, that uses a smooth stitching field over the entire target image, while accounting for all the local transformation variations. Computing the warp is fully automated and uses a combination of local homography and global similarity transformations, both of which are estimated with respect to the target. We mitigate the perspective distortion in the non-overlapping regions by linearizing the homography and slowly changing it to the global similarity. The proposed method is easily generalized to multiple images, and allows one to automatically obtain the best perspective in the panorama. It is also more robust to parameter selection, and hence more automated compared with state-of-the-art methods. The benefits of the proposed approach are demonstrated using a variety of challenging cases.

1. Introduction

Algorithms for aligning and stitching images into seamless photo-mosaics are among the oldest and most widely used in computer vision [11]. The holy grail of image stitching is to seamlessly blend overlapping images, even in the presence of parallax, lens distortion, and scene illumination, to provide a mosaic without any artifacts that looks as natural as possible. Evidently, there is some subjectivity in interpreting how natural a panorama or a mosaic looks. Furthermore, the stitching techniques must be able to extrapolate well to the regions of the panorama where there is information only from a single image.

Early methods focused on obtaining global 2D transformations to align one image with the other [11]. However, assuming a single global transformation such as a homography, will be incorrect except under special conditions, and this will lead to misalignments and ghosting effects. Arguably, most of the problems in 2D image stitching happen because it is impossible to estimate the stitching field accurately due to the complex interaction between the 3D scene and the camera parameters, both of which are unavailable. However, several assumptions can be posed on the stitching field during image alignment [12, 9, 4, 2] and tolerance to parallax can also be imposed [13].

We propose a new method that incorporates several techniques to make the panorama look more natural. To mitigate perspective distortion that occurs in As-Projective-As-Possible (APAP) [12] stitching, we linearize the homography in the regions that do not overlap with any other image. We then automatically estimate a global similarity transform using a subset of corresponding points in the overlapping regions. Finally, we interpolate smoothly between the homography and the global similarity in the overlapping regions, and similarly extrapolate using the linearized homography (affine) and the global similarity transform in the non-overlapping regions. The smooth combination of two stitching fields (homography/linearized homography and global similarity) help us achieve: (a) a fully continuous and smooth stitching field with no bending artifacts, (b) improved perspective in the non-overlapping regions using a global similarity transform, (c) full benefits of the state-of-the-art alignment accuracy offered by APAP.

2. Related Work

A description of fundamental concepts in image stitching and the many associated transformations is available in [11]. Some special cases where cylindrical and spherical image stitching algorithms can be used are also discussed. For example, the cylindrical model can be used when the camera is known to be level and rotating around its vertical axis. Parallax error could be minimized with this assumption, but ghosting occurs when it is violated. A simple extension for computing a single global homography is introduced in [7] by separating a single scene into a distant plane and a ground plane. A weight map is used to smoothly combine two homographic transformations over the target image, but this technique is limited to scenes without local perspective variation.

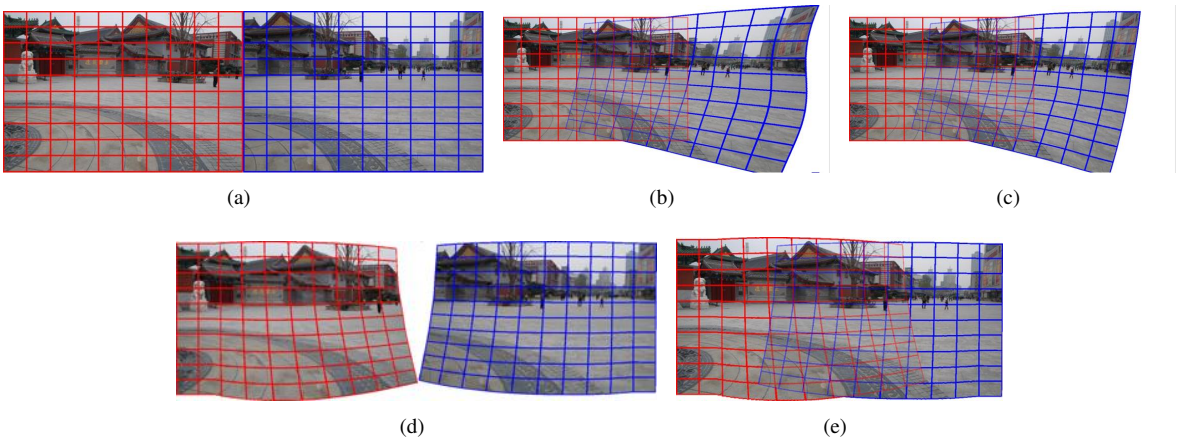

Figure 1: Illustration of proposed algorithm. (a) Original images, (b) Warp after applying moving DLT with Gaussian weighting. (c) Extrapolation of non-overlapping areas using homography linearization and Student's t-weighting, (d) Proposed final warps after integrating global similarity transformation, and (e) Final stitched image.

One of the first approaches that estimates a smooth stitching field is the smoothly varying affine (SVA) stitching proposed in Lin *et al.* [9]. A global affine transform is estimated, which is then relaxed to form a smoothly varying affine stitching field, using an EM style formulation. It is flexible enough to handle parallax while retaining the extrapolation and occlusion handling properties of parametric transforms. Although it can handle local variations well, it fails to impose global projectivity. This drawback is alleviated by the APAP approach proposed in [12], which estimates a smoothly varying projective stitching field and hence provides excellent alignment accuracy. A simple moving Direct Linear Transformation (DLT) method is used to estimate the local parameters, by providing higher weights to closer feature points and lower weights to the farther ones.

Since APAP extrapolates the projective transform in the non-overlapping regions, it introduces severe perspective distortions in regions far from the boundary. The authors in [4], propose the shape-preserving half-projective (SPHP) warp to preserve shapes in the non-overlapping areas. They analyze the projective transform along a rotated co-ordinate axis [5] and propose an approach to gradually change the warp from projective to similarity, as we move from the overlapping to the non-overlapping regions. The stitching provides for shape preservation, but does not guarantee against parallax. Although the combination of SPHP and APAP can be claimed to the state-of-the-art approach, it is sensitive to parameter selection. Furthermore, if the overlapping areas have multiple distinct planes, deriving a single global similarity transformation from the global homography may not be sufficient. This may lead to undesirable and unnatural visual effects in the mosaic.

Carroll *et al.* [3] proposed a novel method to manipulate the perspective of a single image employing the user's annotations of planar regions, the straight lines, and associated vanishing points. This method can synthesize a new image from images with different perspectives. Kopf *et al.* [8] proposed a method to obtain an image with more plausible visual effects by post-processing a panorama availing user's annotations on unnatural regions. However, both the above methods cannot perform perspective preserving image stitching automatically.

3. Proposed Algorithm

We will provide a detailed presentation of the proposed algorithm. We first describe the moving DLT method to estimate the local homography, and proceed to propose an approach to linearize it in the non-overlapping regions. Then, we explain the computation of a global similarity transformation between the reference and the target images. Since many similarity transformations are possible, we automatically choose the one with the lowest rotation angle as the best candidate. Finally, the details of the proposed warp, which is constructed by combining the homography or its linearized version across the whole image with the global similarity, are presented.

3.1. Local Homography Model

Let the target and the reference images be denoted by I and image I'. Given a pair of matching points $\mathbf{p} = [x\ y]^T$ and $\mathbf{p}' = [x'\ y']^T$, between I and I', the homographic trans-

formation $\mathbf{p}' = \mathbf{h}(\mathbf{p})$ can be represented as

$$\mathbf{h}_x(\mathbf{p}) = \frac{h_1 x + h_2 y + h_3}{h_7 x + h_8 y + h_9}, \quad (1)$$

$$\mathbf{h}_y(\mathbf{p}) = \frac{h_4 x + h_5 y + h_6}{h_7 x + h_8 y + h_9}. \quad (2)$$

In homogeneous coordinates $\hat{\mathbf{p}} = [x\ y\ 1]^T$, and $\hat{\mathbf{p}}' = [x'\ y'\ 1]^T$, it can be represented up to a scaling using the homography matrix $\mathbf{H} \in \mathbb{R}^{3\times 3}$ as

$$\hat{\mathbf{p}}' \sim \mathbf{H}\hat{\mathbf{p}}. \quad (3)$$

The columns of $\mathbf{H}$ are given by $\mathbf{h}_1 = [h_1\ h_4\ h_7]^T$, $\mathbf{h}_2 = [h_2\ h_5\ h_8]^T$, and $\mathbf{h}_3 = [h_3\ h_6\ h_9]^T$. Taking a cross product on both sides of (3), we obtain

$$\mathbf{0}_{3\times 1} = \hat{\mathbf{p}}' \times \mathbf{H}\hat{\mathbf{p}} \quad (4)$$

which can be re-written as follows

$$\mathbf{0}_{3\times 1} = \begin{bmatrix} \mathbf{0}_{3\times 1} & -\hat{\mathbf{p}}^T & y'\hat{\mathbf{p}}^T \\ \hat{\mathbf{p}}^T & \mathbf{0}_{3\times 1} & -x\hat{\mathbf{p}}^T \\ -y\hat{\mathbf{p}}^T & x'\hat{\mathbf{p}}^T & \mathbf{0}_{3\times 1} \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \end{bmatrix}. \quad (5)$$

We will denote the 9×1 vector in (5) as $\mathbf{h}$. Since only two rows of the 3×9 matrix in (5) are linearly independent, for a set of N matching points $\{\hat{\mathbf{p}}_\mathbf{i}\}_{i=1}^N$, and $\{\hat{\mathbf{p}}'_\mathbf{i}\}_{i=1}^N$, we can estimate $\mathbf{h}$ using

$$\mathbf{h} = \underset{\mathbf{h}}{\operatorname{argmin}} \sum_{i=1}^{N} \left\| \begin{bmatrix} \mathbf{a}_{i,1} \\ \mathbf{a}_{i,2} \end{bmatrix} \mathbf{h} \right\|^2 = \underset{\mathbf{h}}{\operatorname{argmin}} \|\mathbf{A}\mathbf{h}\|^2, \quad (6)$$

where $\mathbf{a}_{i,1}$ and $\mathbf{a}_{i,2}$ correspond to the two rows of the matrix in (5). We will also incorporate the constraint $\|\mathbf{h}\|^2 = 1$ since the homographic transformation has only 8 degrees of freedom.

In [12], authors introduced moving DLT framework to estimate local homography by including locality-enforcing weights in the objective of (6). The local homography at the location $\mathbf{p}_j$ is estimated as

$$\mathbf{h}_j = \underset{\mathbf{h}_j}{\operatorname{argmin}} \sum_{i=1}^{N} \omega_{i,j} \left\| \begin{bmatrix} \mathbf{a}_{i,1} \\ \mathbf{a}_{i,2} \end{bmatrix} \mathbf{h} \right\|^2 \quad (7)$$

which can be writtenas $\mathbf{h}_j = \underset{\mathbf{h}}{\operatorname{argmin}} \|\mathbf{W}_j\mathbf{A}\mathbf{h}\|^2$, where $\mathbf{W}_j = \text{diag}([\omega_{1,j}\ \omega_{1,j}\ \ldots\ \omega_{N,j}\ \omega_{N,j}])$. In [12], the weights are generated using the offsetted Gaussian which assumes high value for pixels in the neighborhood of $\mathbf{p}_j$ and equal values for those that are very far, i.e., $\omega_{i,j} = \max(\exp(-\|\mathbf{p}_i - \mathbf{p}_j\|^2/\sigma^2),\ \gamma)$. The parameter $\gamma \in [0\ 1]$ is the offset used to prevent numerical issues. Note that the local homography can be computed only in the regions of the target image that overlap with the reference image.

For each pixel in the non-overlapping regions, the transformation is computed as a weighted linear combination of the local homographies in the overlapping regions. Here it becomes important to choose a proper offset to avoid extrapolation artifacts. This is demonstrated in Figure 1(b), where setting $\gamma = 0$ leads to "wavy" effects due to the isotropic nature of Gaussian weighting, whereas choosing a proper offset leads to a good result. Even in this case, the perspective distortion in the non-overlapping area is apparent with APAP, as noted in [4] as well.

In our proposed method, we use the moving DLT without offset in overlapping area to estimate the local homography, and extrapolate to the non-overlapping area using homography linearization, as described in the following section. This reduces the perspective distortion, and our proposed weighting scheme for extrapolation is less dependent on the choice of the parameters.

3.2. Homography Linearization

The extrapolation of homographic transformation in the non-overlapping areas produces extreme and unnatural scaling effects, as seen in Figure 1(b). The reason for this effect can be understood by considering the 1-D perspective transform, $x' = \frac{ax+b}{cx+d}$. If we estimate the parameters $\{a, b, c, d\}$ using a set of corresponding points, outside the range of the available corresponding points as well, the relationship between x and x' will be non-linear. This translates to severe perspective distortion in 2-D. However, this distortion can be minimized by linearizing the transformation.

With images, the linearization of homography at any point $\mathbf{q}$ in the neighborhood of the anchor point $\mathbf{p}$ can be understood by considering the Taylor series of the homographic transformation $\mathbf{h}(\mathbf{q})$, where $\mathbf{h} : \mathbb{R}^2 \to \mathbb{R}^2$

$$\mathbf{h}(\mathbf{q}) = \mathbf{h}(\mathbf{p}) + \mathbf{J}_\mathbf{h}(\mathbf{p})(\mathbf{q} - \mathbf{p}) + o(\|\mathbf{q} - \mathbf{p}\|), \quad (8)$$

where $\mathbf{J}_\mathbf{h}(\mathbf{p})$ is the Jacobian of the homography $\mathbf{h}$ at the point $\mathbf{p}$. The first two terms in (8) provide the best linearization for $\mathbf{h}(\mathbf{q})$. Since, if $\mathbf{h}$ is differentiable at $\mathbf{p}$, $\mathbf{J}_\mathbf{h}(\mathbf{p})$ is invertible, the linearization of homography is an affine transformation. However, it is not straightforward to compute linearization at an arbitrary point $\mathbf{q}$ in the non-overlapping region as in the case of $1-$D data, since the boundary between the overlapping and the non-overlapping regions could contain multiple points and we would not know where the Jacobian has to be computed. Therefore, we consider anchor points in the boundary for linearization and compute a weighted average of the transformations.

For a set of R anchor points $\{\mathbf{p}_i\}_{i=1}^R$ at the boundary with possibly different local homographies, the weighted combination of linearizations is given as

$$\mathbf{h}^L(\mathbf{q}) = \sum_{i=1}^{R} \alpha_i \left(\mathbf{h}(\mathbf{p}_i) + \mathbf{J}_\mathbf{h}(\mathbf{p}_i)(\mathbf{q} - \mathbf{p}_i)\right). \quad (9)$$

We assume α_i to be a function of $\|\mathbf{q} - \mathbf{p}_i\|$, and in particular we consider the Gaussian weighting where $\alpha_i = \exp(-\|\mathbf{q} - \mathbf{p}_i\|^2/\sigma^2)$, or the Student's t-weighting where $\alpha_i = \left(1 + \frac{\|\mathbf{q}-\mathbf{p}_i\|^2}{\nu}\right)^{\frac{-(\nu+1)}{2}}$. Student's t-weighting is more robust since that tail of the distribution decays slowly compared to Gaussian and hence when $\mathbf{q}$ is far from anchor points, all the anchor points are given similar weighting. However, if Gaussian weighting is chosen, the tail should be made flat at the offset parameter γ to avoid "wavy" effects.

The stitching result using our extrapolation method is shown in Figure 1(c). Using the linearized homography to extrapolate the non-overlapping area has less perspective distortions than the result using APAP. The result is similar to the stitching result using dual-homography warping [7]. However, with our method, there is no need to estimate two homographies of distant plane and ground plane. Our method can adapt to the more complicated scenes and is a generalized method in comparison of dual-homography method. It does not need the parameter γ that should be determined on a case-by-case basis [12], hence being less dependent on parameter choices.

3.3. Global Similarity Transformation

In previous section, we introduced a method to linearize the homography and hence reduce the perspective distortion in the overlapping areas. In the following sections, we will propose approaches to further reduce the distortions and hence make the panorama look natural. The idea is to use a similarity transformation in the non-overlapping areas in the target image, since it will not introduce any perspective distortions.

If the global similarity transformation approximates the camera motion between the the target and the reference images, the estimated similarity transform can be used to compensate for the camera motion. However, finding a global similarity transformation using all point matches may result in non-optimal solution particularly when the overlapping areas contain distinct image planes. The problem is apparent in Figure 2, which shows the stitching result of SPHP. Note that SPHP uses the global homography transformation to uniquely determine the global similarity, which may not approximate the camera motion well.

Given that there are multiple planes in the scene and an image projection plane at the focal length of the cameras, every plane in the scene has an intersection angle with the image projection plane, and each plane in the scene corresponds to a local homography transformation. The homography transformation corresponding to the plane that is most parallel (smallest intersection angle) to image projection plane can be used to derive the optimal similarity transformation that represents the camera motion.

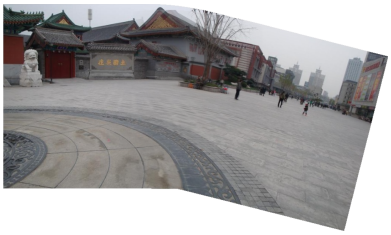

Figure 2: Stitching result for *Temple* image dataset using SPHP [4].

We propose an approach to compute an optimal similarity transformation between the reference and the target images by segmenting the corresponding points in the following manner. After obtaining the feature point matches, we first remove the outliers using RANSAC [6] with threshold ε_g. Then, we use RANSAC with a threshold ε_l to find a homography of the plane with largest inliers, where $\varepsilon_l < \varepsilon_g$, and we remove the inliers. This is repeated until the number of inliers is small than η. Each group of matched inlier points is used to calculate an individual similarity transformation. Then, the rotation angles corresponding to the transformations are examined and the one with the smallest rotation angle is chosen.

Figure 3 shows an example of the grouping results. The green and yellow circles on the figure belong to two different groups of point correspondences. The red circles do not belong to any group. In this example, the group with yellow points generate the optimal global similarity transformation with the least rotation angle.

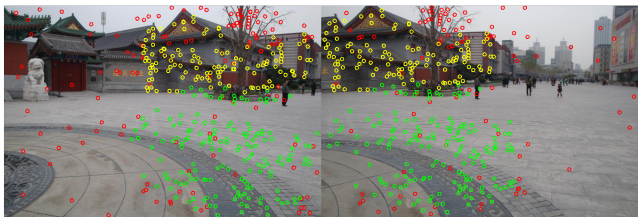

Figure 3: Grouping feature points for computing the optimal global similarity transform.

3.4. Integration of Global Similarity Transformation

After the global similarity transformation is calculated, it is used to adjust warps of target image in order to mitigate the perspective distortions in the overall panorama. If we only adjust the transformations on the non-overlapping area, the stitching result may have an unnatural visual effect. Hence, we gradually update the local transformations

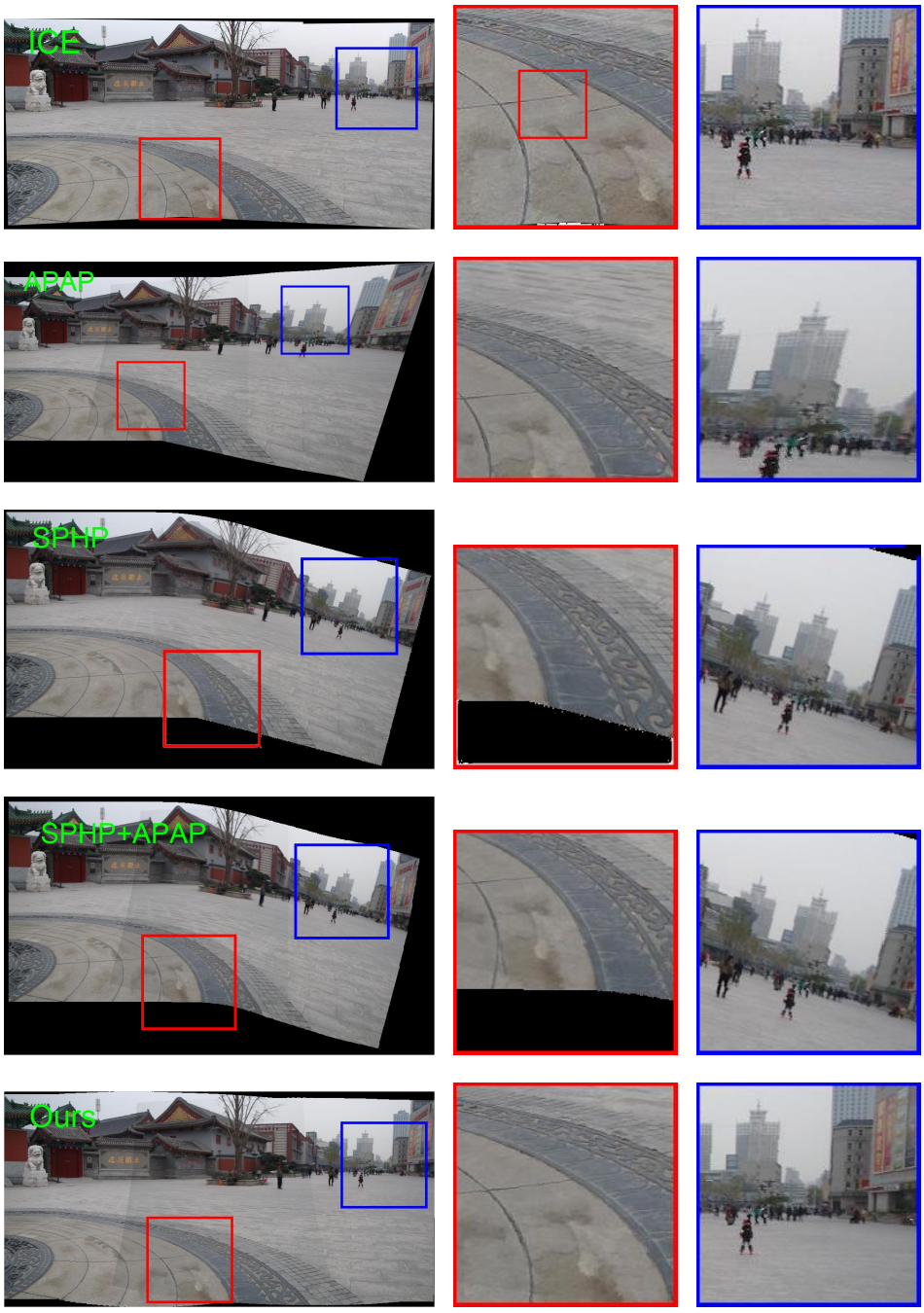

Figure 4: Comparisons with state-of-the-art image stitching techniques on the *Temple* image dataset.

of entire target image to the global similarity transformation using the following equation:

$$\hat{\mathbf{H}}_i^{(t)} = \mu_h \mathbf{H}_i^{(t)} + \mu_s \mathbf{S}. \tag{10}$$

Here, $\mathbf{H}_i^{(t)}$ is i^{th} local homography, $\hat{\mathbf{H}}_i^{(t)}$ is updated local transformation, $\mathbf{S}$ is the global similarity transformation. μ_h and μ_s are weighting coefficients. The superscript (t) refers to the target image and the superscript (r) denotes the reference image. We also constrain $\mu_h + \mu_s = 1$, where μ_h and μ_s are between 0 and 1. They are computed as,

$$\mu_h(i) = \langle \overrightarrow{\kappa_m p(i)}, \overrightarrow{\kappa_m \kappa_M} \rangle / \left| \overrightarrow{\kappa_m \kappa_M} \right|, \tag{11}$$

$$\mu_s(i) = 1 - \mu_h(i), \tag{12}$$

where κ is the projected point of warped target image on the $\overrightarrow{o_r o_t}$ direction. o_r and o_t are the center points of the reference image and the warped target image. κ_m and κ_M are the points with smallest and largest value of $\langle \overrightarrow{o_r p(i)}, \overrightarrow{o_r o_t} \rangle$ respectively. Here, $p(i)$ is the location of the i^{th} location in the final panorama.

Updating the warps of target image with global similarity transformation causes misalignment of overlapping areas between reference image and target image that were previously aligned. Therefore, we need to compensate the changes by appropriately propagating the changes from the target image to the reference image. The local transformation of the reference image can be now obtained as

$$\hat{\mathbf{H}}_i^{(r)} = \hat{\mathbf{H}}_i^{(t)} (\mathbf{H}_i^{(t)})^{-1} \tag{13}$$

Fig. 1(d) show the final warping results of reference image and target image. The final stitching result shown in Fig. 1(e) clearly resembles a natural-looking panorama.

4. Experiments

We have conducted comparative experiments of proposed algorithm on a variety of existing datasets made available by [12]. The compared methods include Microsoft Image Composite Editor (ICE) [1], APAP [12], SPHP with global homography [4], SPHP with local homographies computed with APAP (SPHP+APAP). In our experiments, we use the same set of parameters suggested in the respective papers. We use the code provided by the authors of the papers to obtain the results for comparison. The corresponding points are detected using SIFT [10]. For the moving DLT, we set σ as 12.5, for Student's t, we set ν as 5, the threshold, ε_g, for global RANSAC error function is set as 0.1, the threshold, ε_l, for the local RANSAC error function is fixed at 0.001, and the threshold for inlier number, κ, is fixed at 50. The proposed method typically takes from 20 to 30 seconds with a 2.7GHz CPU and 16GB RAM to stitch two images with 800×600 resolution. To keep the paper concise, we show provide results only for *Temple* and *Railtracks* datasets, but more results are available in the supplementary material.

The results for the *Temple* dataset are provided in Figure 4. Each row is a result of different methods. The results are in the following order: ICE, APAP, SPHP, SPHP+APAP, and our method. Two areas of each results have been highlighted. Red boxes show parallax error in overlapping areas, and blue boxes show the perspective distortion in non-overlapping areas. The result of ICE look good visually. The perspective is preserved but there is some misalignment on the ground region. The APAP results on the second row, as discussed in the previous section, show good alignment on the overlapping areas, but the perspective distortion on non-overlapping area is non-negligible, for the reasons discussed before. The third row shows the results of SPHP method. As described in SPHP paper, it pays more attention to mitigating the perspective distortion but not the alignment accuracy. The result shows the shape is preserved but parallax errors exist. To alleviate the parallax errors, authors of SPHP suggest combining SPHP with APAP. The results in the next row show the parallax errors are improved. However in both SPHP and SPHP+APAP, the buildings on the right side are not parallel to temples. This is because the similarity transformation is derived from the global homography and hence may not be optimal. This is particularly true, if there are multiple distinct planes in the overlapping areas, just like in the image. This can be corrected only if the rotation angle of the camera is known. The results in the last row show that our method mitigates the perspective distortion and can also successfully handle the parallax issue.

Figure 5 compares the results for the *Railtracks* dataset. We can still see parallax error in ICE/SPHP and perspective issues in APAP. Without manually correcting for the rotation angle, the results from SPHP and SPHP+APAP do not look natural. The proposed method maintains alignment accuracy and shows robustness in this challenging example. The panorama examples that follow demonstrate the performance of our proposed method with multiple images. The image dataset in Figure 6 consist of a truck, a roundabout, and an arced ground surface. The images in Figure 7 includes skylines, buildings, trees, and a swimming pool. Our method works well in both datasets, maintaining the integrity of image contents, and providing a natural look to the panorama. There are no visible parallax errors and perspective distortions.

5. Conclusion

In this work, we have presented a novel stitching method that uses a smooth stitching field derived from local homography or its linearized version and a global similarity transformation. Results show that our approach provides a more natural panorama with no visible parallax in the overlapping regions and mitigates the perspective distortion issue in the non-overlapping regions. Furthermore, it is less dependent on the choice of the parameters and computes the appropriate global similarity transform automatically. Experimental comparisons to existing methods show that the proposed method yields the best stitch compared to the state-of-the-art methods. Future research developments will include compensating for parallax when large motion exists, which can be performed by integrating seam-cut methods into this framework. This will make our approach a one-stop solution that addresses all major problems in image stitching.

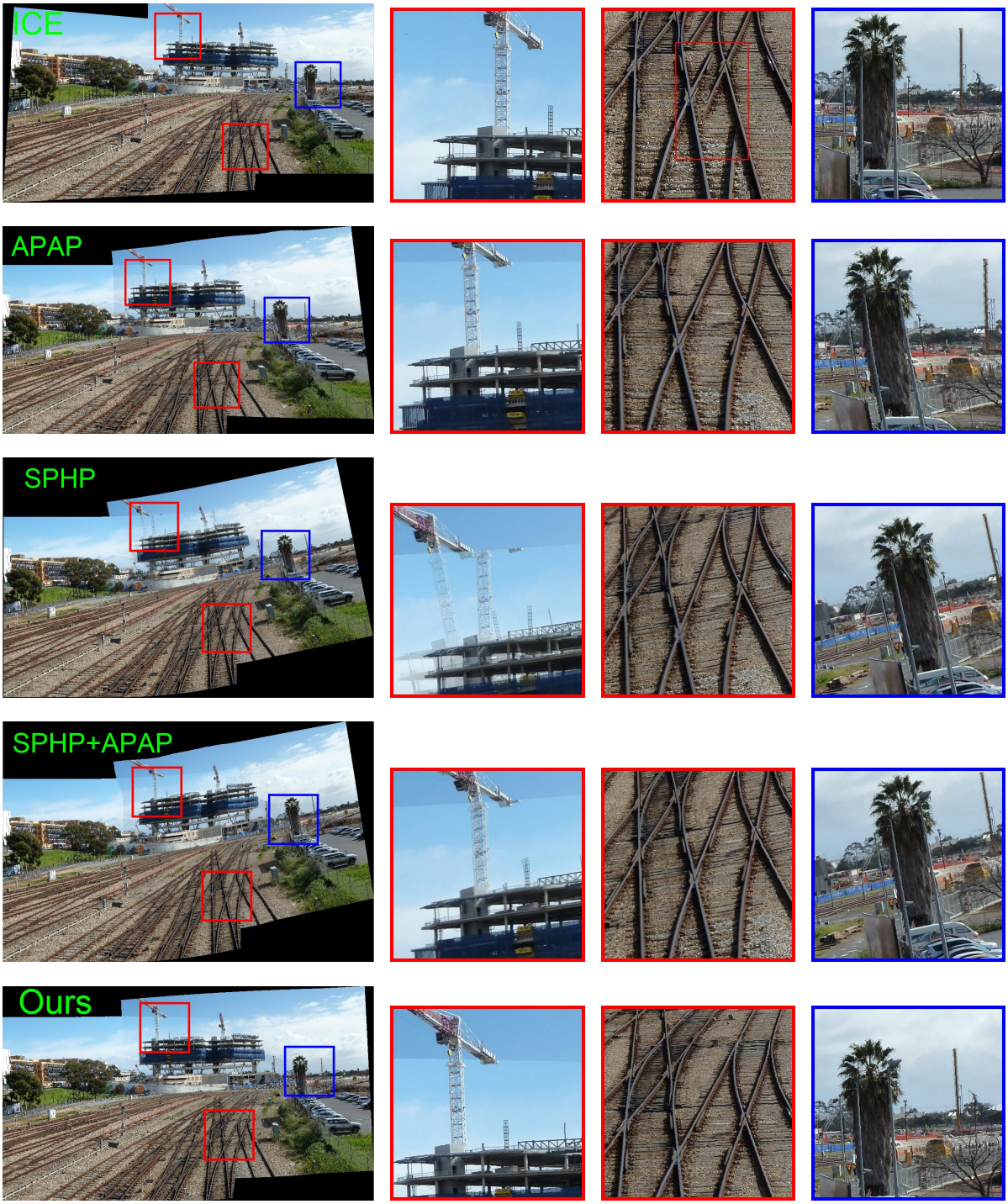

Figure 5: Comparisons with state-of-the-art image stitching techniques on the *Railtracks* image dataset.

Figure 6: Panorama of roundabout images.

Figure 7: Panorama of skyline images.

Acknowledgments

This work is sponsored in part by Defense Advanced Research Projects Agency, Microsystems Technology Office (MTO), under contract no. HR0011-13-C-0022. The views expressed are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government. This document is: Approved for Public Release, Distribution Unlimited.

References

[1] Microsoft research image composite editor. `http://research.microsoft.com/en-us/um/redmond/groups/ivm/ice/`.

[2] M. Brown and D. G. Lowe. Automatic panoramic image stitching using invariant features. *International journal of computer vision*, 74(1):59–73, 2007.

[3] R. Carroll, A. Agarwala, and M. Agrawala. Image warps for artistic perspective manipulation. In *ACM Transactions on Graphics (TOG)*, volume 29, page 127. ACM, 2010.

[4] C.-H. Chang, Y. Sato, and Y.-Y. Chuang. Shape-preserving half-projective warps for image stitching. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, pages 3254–3261. IEEE, 2014.

[5] O. Chum, T. Pajdla, and P. Sturm. The geometric error for homographies. *Computer Vision and Image Understanding*, 97(1):86–102, 2005.

[6] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and aarticleutomated cartography. 1981.

[7] J. Gao, S. J. Kim, and M. S. Brown. Constructing image panoramas using dual-homography warping. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 49–56. IEEE, 2011.

[8] J. Kopf, D. Lischinski, O. Deussen, D. Cohen-Or, and M. Cohen. Locally adapted projections to reduce panorama distortions. In *Computer Graphics Forum*, volume 28, pages 1083–1089. Wiley Online Library, 2009.

[9] W.-Y. Lin, S. Liu, Y. Matsushita, T.-T. Ng, and L.-F. Cheong. Smoothly varying affine stitching. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 345–352. IEEE, 2011.

[10] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[11] R. Szeliski. Image alignment and stitching: A tutorial. *Foundations and Trends® in Computer Graphics and Vision*, 2(1):1–104, 2006.

[12] J. Zaragoza, T.-J. Chin, M. S. Brown, and D. Suter. As-projective-as-possible image stitching with moving DLT. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2339–2346. IEEE, 2013.

[13] F. Zhang and F. Liu. Parallax-tolerant image stitching. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014.