

# Spatially-Varying Image Warps for Scene Alignment

Che-Han Chang  
Graduate Institute of  
Networking and Multimedia  
National Taiwan University  
Taipei, Taiwan 106  
Email: frank@cmlab.csie.ntu.edu.tw

Chiu-Ju Chen  
Graduate Institute of  
Networking and Multimedia  
National Taiwan University  
Taipei, Taiwan 106  
Email: dogip@cmlab.csie.ntu.edu.tw

Yung-Yu Chuang  
Department of Computer Science  
and Information Engineering  
National Taiwan University  
Taipei, Taiwan 106  
Email: cyy@csie.ntu.edu.tw

**Abstract**—This paper proposes a method to align a set of images captured from multiple view points. Traditional methods using image warps parameterized by global transformations suffer from the problem of misalignment due to parallax effects induced by camera motions between images and depth variations of the scene. Our method parameterizes warps using mesh deformation and achieves spatially-varying transformations to alleviate the misalignment problem. The proposed method has two stages: a hybrid image alignment stage which combines direct-based methods and feature-based methods, followed by a shape-preserving aggregation stage which further refines the result. Experiments show that our method achieves better alignment and provides visually pleasing image summaries for scenes.

## I. INTRODUCTION

Limited by the camera’s capability, a single view often can not fully capture a scene as we perceive it. To better depict the scene, it has become common to take multiple pictures of the same scene and compose them into a single seamless image of the scene, called a scene collage. It is a more informative representation than a single viewpoint can provide. Artists often incorporate multiple view points into a single painting to produce more informative representations than a single viewpoint can [1]. A panorama can be taken as a special case of scene collage, often taken from a single viewpoint. Unfortunately, limited by alignment methods, photographers do not have the freedom to change the view point smoothly and to select which aspects of the scene will be represented.

This paper proposes a spatially-varying-warp-based method to align multiple images captured from multiple view points so that users can enjoy more freedom to select views and produce compelling and informative scene collages. Other than few exceptions [2], [3], [4], most image alignment methods assume a global transform for the motion between images [5]. Thus, they often produce misalignment when aligning images from multiple view points as in our application. A previous method of scene collage [1], [6] uses global warps and the alignment is not always perfect. Our method overcomes the problem by using spatially-varying warps driven by mesh deformation and provides more flexibility for generating scene collages. In addition, users have controls on the final composite by setting optional constraints. Thus, users can create more visually pleasing and seamless scene collages with our method. In addition to the proposed application, by providing more accurate alignment between images, our method can be used in various applications such as finding correspondences, annotation transfer and so on.

Although a few recent methods also allow local deformation for better alignment [2], [4], our method bears some differences with them. First, for areas without alignment constraints, we use conformal energy for regularization instead of imposing globally affine/projective assumptions. Second, our method provides more ease on controlling results by adding constraints.

## A. Related work

Existing image alignment methods can be mainly categorized into two types: direct (pixel-based) methods and feature-based methods. Direct methods utilize all pixels for alignment. It can be very accurate, but has a limited range of convergence [5]. Feature-based methods have success in matching images with large differences in scales, orientations, lighting [7]. A well-known application is panorama. However, most panorama construction methods assume a global transformation to model the motion between images. Few explored spatially-varying models. Lin et al. [2] proposed a smoothly varying affine stitching field, which is globally affine while allowing local deformations, for image alignment. The core idea is to pose the problem as nonrigid point-set registration. Zaragoza et al. [4] proposed the as-projective-as-possible warp which is globally projective while allowing local deviations for better alignment. Gao et al. [3] proposed dual-homography warping to construct panoramas. Since they focus on the panorama application, they assume that most panoramic scenes are mainly composed by two dominant planes and find two homographies for each image. The goal of our method is related to multi-viewpoint panorama generation [8], [9], [10]. For reducing shape/area distortion while stitching multiple images, Chang et al. [11] proposed the shape-preserving half-projective warp, a spatial combination of a projective transformation and a similarity transformation. It provides good alignment accuracy as projective warps while preserving the perspective of individual image as similarity warps.

## B. Overview

In the following sections, we first develop the proposed method using spatially-varying warps (Section II). Our method combines feature-based and direct methods with advantages of both: a feature-based method provides us a roughly correct alignment and a direct method improves upon it for more accurate results. We then present results using our technique and comparisons with an existing method (Section III). Finally, we conclude with a summary and ideas for future research (Section IV).

## II. THE PROPOSED METHOD

Given a set of  $m$  images  $I_1, I_2, \dots, I_m$ , the goal is to jointly align these images into a single one. Adopting group-wise image alignment is a natural choice. However, the number of variables are huge and the optimization is time-consuming. Since image-to-image alignment techniques is more mature and stable than groupwise alignment, we instead opt to adopt pairwise alignment techniques first.

Our method consists of two stages: pairwise alignment and aggregation. The first stage finds pairs to match and use pairwise alignment to align image pairs. The aggregation stage integrates the pairwise results altogether using content-preserving warps. In addition, our framework provides users with tools for improving the aggregated result. After alignment, any blending technique can be used to obtain the final composite. Here we simply use layering to better demonstrate quality of the alignment.

We start by defining notations. Let  $W$  be a parametric warping function parameterized by its parameters  $\mathbf{p}$ . A point  $\mathbf{x}$  is mapped to the position  $W(\mathbf{x}; \mathbf{p})$  by the warping function. We use a uniform grid mesh to guide spatially-varying image warping. A uniform grid mesh for the  $i$ -th image  $I_i$  is denoted as  $\{\mathbf{V}_i, \mathbf{E}_i, \mathbf{Q}_i\}$ .  $\mathbf{V} = [\mathbf{v}_1^T, \mathbf{v}_2^T, \dots, \mathbf{v}_n^T]^T$  is the vector by concatenating all vertex positions of the mesh.  $\mathbf{E}$  and  $\mathbf{Q}$  represent edge set and quad set. The warping parameters are defined as the vertex position differences between the original mesh and the deformed mesh,  $\mathbf{p} = \tilde{\mathbf{V}} - \mathbf{V}$ . Using this definition makes sure that  $W(\mathbf{x}; \mathbf{0})$  is the identity warp, which is natural and is a necessary requirement for the optimization below. The warping function becomes

$$\tilde{\mathbf{x}} = W(\mathbf{x}; \mathbf{p}) = \Psi(\mathbf{x})^T \tilde{\mathbf{V}} = \Psi(\mathbf{x})^T (\mathbf{V} + \mathbf{p}), \quad (1)$$

where  $\Psi(\mathbf{x})$  is a vector containing  $\mathbf{x}$ 's bilinear interpolation coefficients in terms of  $\mathbf{V}$ .

### A. Pairwise alignment

In this step, we consider the problem of aligning two images,  $I_i$  and  $I_j$ , at a time. The goal is to find their respective warping functions  $W_i$  and  $W_j$  determined by parameters  $\mathbf{p}_i$  and  $\mathbf{p}_j$ , so that (1) the two images align well with each other and (2) the distortions induced by warping are small. Our approach combines feature-based alignment and direct-based alignment for achieving good alignment.

First, we adopt SIFT feature matching to find image pairs and their feature matches [7]. Feature matches are verified by RANSAC with the Homography-based motion model. We denote the set of resulting feature pairs between  $I_i$  and  $I_j$  as  $\mathbf{F}_{ij}$ . Given  $\mathbf{F}_{ij}$ , we attempt to align  $I_i$  and  $I_j$  by minimizing an energy function comprised of three terms: *feature alignment*, *pixel alignment* and *shape distortion*.

**Feature alignment**  $E_f$ . Assume the  $k$ -th feature pair contains features  $(\mathbf{f}_i^k, \mathbf{f}_j^k)$ . To guarantee good alignment, the positions of matched features after warping should coincide. Thus, we define the term by measuring the distance between matched features after warping:

$$E_f(\mathbf{p}_i, \mathbf{p}_j) = \sum_{k=1}^{|\mathbf{F}_{ij}|} \left\| \tilde{\mathbf{f}}_i^k - \tilde{\mathbf{f}}_j^k \right\|^2, \quad (2)$$

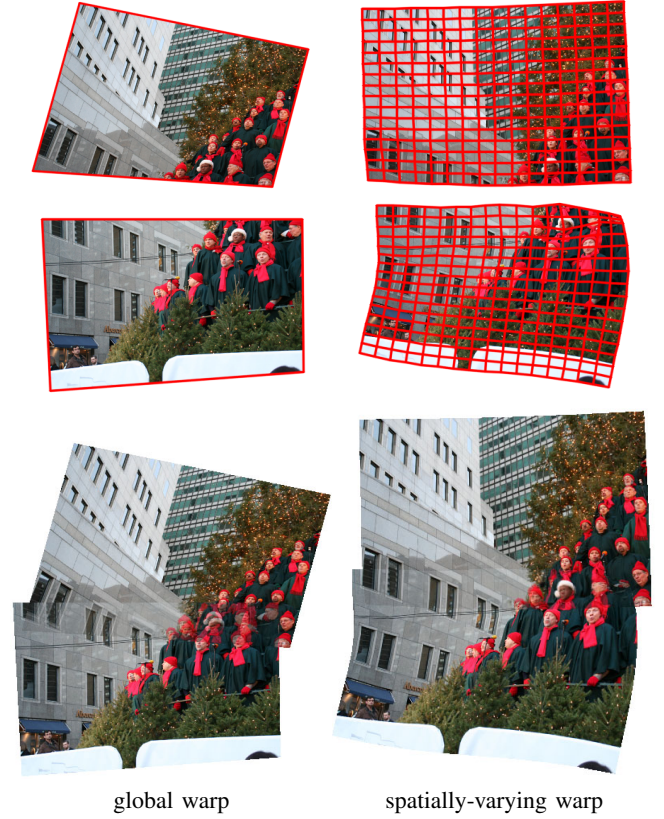


Fig. 1. Pairwise alignment using global and spatially-varying warps. On the left, two aligned images using global warps and their composition are shown from top to bottom. The right displays results using spatially-varying warps. In the overlapped region, misalignment is much reduced with our method using spatially-varying warps.

where  $\tilde{\mathbf{f}}_i^k = W_i(\mathbf{f}_i^k; \mathbf{p}_i)$  and  $\tilde{\mathbf{f}}_j^k = W_j(\mathbf{f}_j^k; \mathbf{p}_j)$  are positions of features after warping.

**Pixel alignment**  $E_a$ . This term further requires that the visual appearance of overlap regions of both images after warping should be similar. Let  $I(\mathbf{x})$  denote the image intensity at position  $\mathbf{x}$ . According to the warping function  $W$  with parameter  $\mathbf{p}$ , by using inverse mapping, the intensity of the warped image  $\tilde{I}$  at  $\mathbf{x}$  should be

$$\tilde{I}(\mathbf{x}) = I(W^{-1}(\mathbf{x}; \mathbf{p})). \quad (3)$$

This term is thus defined by measuring the intensity differences between the overlap regions of  $\tilde{I}_i$  and  $\tilde{I}_j$ ,

$$E_a(\mathbf{p}_i, \mathbf{p}_j) = \sum_{\mathbf{x} \in \mathbf{X}} \left( \tilde{I}_i(\mathbf{x}) - \tilde{I}_j(\mathbf{x}) \right)^2, \quad (4)$$

where  $\mathbf{X}$  is the set of pixel locations where  $\tilde{I}_i$  and  $\tilde{I}_j$  overlaps.

**Shape distortion**  $E_d$ . This term attempts to maintain the shapes of the quad meshes, preventing the warped image from much distortion. We use the conformal energy [12] for measuring distortion. It prefers that each quad undergoes a similarity transformation by minimizing the deviation of the deformed quad from a similarity transform. Since  $E_f$  and  $E_a$  usually only cover a subset of all vertices, including  $E_d$  not only minimizes distortions but also ensures that the energy function has a unique solution.

**Optimization.** The total energy can be written as a linear combination of these three terms as follows:

$$E(\mathbf{p}_i, \mathbf{p}_j) = E_a(\mathbf{p}_i, \mathbf{p}_j) + \lambda_f E_f(\mathbf{p}_i, \mathbf{p}_j) + \lambda_d (E_d(\mathbf{p}_i) + E_d(\mathbf{p}_j)). \quad (5)$$

We used  $\lambda_f = 5000$  and  $\lambda_d = 1$  in our experiments. Since  $E_a$  is nonlinear, we used iterative optimization for minimization. Assume that an initial guess of  $\mathbf{p}_i$  and  $\mathbf{p}_j$  is given, the optimization alternatively solves for  $\mathbf{p}_i$  and  $\mathbf{p}_j$ . Assuming that  $\mathbf{p}_j$  is constant, Equation 5 becomes

$$E(\mathbf{p}_i) = E_a(\mathbf{p}_i) + \lambda_f E_f(\mathbf{p}_i) + \lambda_d E_d(\mathbf{p}_i). \quad (6)$$

Since  $E_a$  is the term that complicates the optimization, we take a closer look at  $E_a(\mathbf{p}_i)$ . By denoting  $\mathbf{y} = W_i^{-1}(\mathbf{x}; \mathbf{p}_i)$ ,  $E_a(\mathbf{p}_i)$  can be reformulated as

$$E_a(\mathbf{p}_i) = \sum_{\mathbf{y} \in \mathbf{Y}} \left( I_i(\mathbf{y}) - \tilde{I}_j(W_i(\mathbf{y}; \mathbf{p}_i)) \right)^2, \quad (7)$$

where  $\mathbf{Y}$  is the set of pixel locations where  $I_i(\mathbf{y})$  and  $\tilde{I}_j(W_i(\mathbf{y}; \mathbf{p}_i))$  overlaps. Equation 7 becomes the formulation of traditional image alignment problem [13]. The well-known Lucas-Kanade framework can be used for solving Equation 6. We use the inverse compositional Lucas-Kanade algorithm [14] to minimize it. In the view of Lucas-Kanade algorithm,  $I_i$  and  $\tilde{I}_j$  play the roles of *template image* and *input image* respectively.  $E_f(\mathbf{p}_i)$  and  $E_d(\mathbf{p}_i)$  serve as regularization terms. Please refer to [14] for more details.

For the initial guesses of  $\mathbf{p}_i$  and  $\mathbf{p}_j$ , we use the optimal homography computed from matched features. Let  $H_{j \rightarrow i}$  be the homography transformation from  $I_j$  to  $I_i$ , and  $\mathbf{V}_{j \rightarrow i}$  be the warped vertex positions  $\mathbf{V}_j$  by  $H_{j \rightarrow i}$ . We set  $W_i$  to be the identity warp and  $W_j$  to be  $H_{j \rightarrow i}$  as the initial guess. That is,  $\mathbf{p}_i = \mathbf{0}$  and  $\mathbf{p}_j = \mathbf{V}_{j \rightarrow i} - \mathbf{V}_j$ . To converge faster and avoid local minimum, we employ a coarse-to-fine strategy in terms of the number of quads in meshes. At the coarsest level, the mesh has only  $1 \times 1$  quad. When converged, to advance to the next level, a quad is divided into four quads. Five levels were used in our experiments. Thus, at the finest level, the mesh consists of  $16 \times 16$  quads. Figure 1 shows a sample result for pairwise alignment and compares global warps and spatially-varying warps.

### B. Aggregation

The goal of aggregation is to align  $I_1, I_2, \dots, I_m$  altogether to obtain a group-aligned image by combining pairwise alignment results from the previous stage. Since an image is usually involved with more than one image pair, it has multiple deformations for the pairs it is involved. We adopt an energy minimization approach to jointly estimate the deformed vertices  $\tilde{\mathbf{V}}_1, \tilde{\mathbf{V}}_2, \dots, \mathbf{V}_m$  for all images for combining the deformations derived from pairwise alignment. The energy encourages the following three properties: (1) Aligned image pairs should remain well aligned after aggregation. (2) The distortions induced by aggregation should be small. (3) User can refine the shape of the group-aligned image by imposing constraints. The energy has three terms, each corresponding to a property.

**Correspondences  $E_c$ .** This term attempts to preserve the correspondences induced from pairwise alignment. It plays the



Fig. 2. The smoothness energy term ensures that the shape of mesh is deformed smoothly.

same role with  $E_f$  in pairwise alignment. Correspondences are extracted as follows. Let  $\mathbf{A}$  denotes the set of image pairs from the previous stage. Consider an image pair  $(i, j) \in \mathbf{A}$ . The warping functions  $W_i$  and  $W_j$  estimated from pairwise alignment describe how  $I_i$  and  $I_j$  should be warped to align to each other. We uniformly extract point correspondences as constraints in aggregation. Specifically, for  $(i, j)$ , we construct a set of correspondences  $\mathbf{C}_{ij}$  by corresponding vertices to their positions after warping, that is

$$\mathbf{C}_{ij} = \left\{ (\mathbf{v}_i^k, W_j^{-1}(\tilde{\mathbf{v}}_j^k; \mathbf{p}_j)) \mid k = 1, 2, \dots, n \right\} \cup \left\{ (W_i^{-1}(\tilde{\mathbf{v}}_j^k; \mathbf{p}_i), \mathbf{v}_j^k) \mid k = 1, 2, \dots, n \right\}. \quad (8)$$

We would like to maintain the positions of corresponding points to be the same after aggregation. The correspondence energy  $E_c$  is thus defined as

$$\sum_{(i,j) \in \mathbf{A}} \sum_{k=1}^{|\mathbf{C}_{ij}|} \|\tilde{\mathbf{c}}_i^k - \tilde{\mathbf{c}}_j^k\|^2, \quad (9)$$

where  $\tilde{\mathbf{c}}_i^k = \Psi_i(\mathbf{c}_i^k)^T \tilde{\mathbf{V}}_i$  and  $\tilde{\mathbf{c}}_j^k = \Psi_j(\mathbf{c}_j^k)^T \tilde{\mathbf{V}}_j$ .

**Shape preservation  $E_d$  and  $E_s$ .** Like pairwise alignment, we want to minimize distortions of the deformed meshes using the same term  $E_d$  defined in the previous stage. We also requires that the deformations are smooth. We use the smoothness energy  $E_s$  proposed by Carroll et al. [15], which attempts to minimize the second derivatives of deformed vertices. Figure 2 compares results with and without the smoothness energy term. We associate each image with both energy terms. Minimizing the above energies yields solutions up to a global similarity transformation. To make the minimization problem well-posed, we adopt point constraints to resolve the ill-posedness.

**Point constraints  $E_p$ .** This term constrains the location of a pixel to be fixed in a place specified by the user. To ensure that the minimization has a unique solution, we apply  $E_p$  at two positions to make it well posed. This term was used by others for a similar purpose [15].

**User refinement.** We provide three options for users to improve the shape of the stitched image. The first is to add more point constraints. By doing so, users are able to change the overall shape of the stitched image by changing positions of control points. Figure 5 gives an example.

The second option is to constrain orientations for some images for creating more visually pleasing composites. Thus we add an optional orientation energy  $E_o$  to maintain image orientations. This term requires the orientation changes of



mesh edges are small. A mesh edge  $\mathbf{e} = [e_x, e_y]^T \in \mathbf{E}$  is defined as the vector formed by two neighboring vertices. Let  $\tilde{\mathbf{e}}$  be the deformed edge.  $E_o$  is then defined as

$$\sum_{\mathbf{e} \in \mathbf{E}_h} \tilde{e}_y^2 + \sum_{\mathbf{e} \in \mathbf{E}_v} \tilde{e}_x^2, \quad (10)$$

where  $\mathbf{E}_h$  and  $\mathbf{E}_v$  are the horizontal and vertical edge set respectively. Figure 6 shows an example for constraining image orientations.

The third option is to use a homography energy  $E_h$ , which constrains the deformation-induced transformation of a selected region to be consistent with a single homography. The region is specified by users, and may cover multiple quads. The single-homography constraint has the benefits that there is only a single perspective change for the region and lines in the region remain straight. This term is especially useful for refining scenes which have planar regions with strong line structures. We define  $E_h$  as

$$\sum_{k \in \mathbf{K}} \|\tilde{\mathbf{v}}_k - H(\mathbf{v}_k)\|^2, \quad (11)$$

where  $\mathbf{K}$  is a set of vertex indices that lie within the region.  $H$  is the homography that will be solved by our method. Figure 7 gives an example for this option.

The total energy can be written as follows

$$E = \lambda_c E_c + \underbrace{\lambda_d E_d + \lambda_s E_s}_{\text{shape}} + \underbrace{\lambda_p E_p + \lambda_o E_o + \lambda_h E_h}_{\text{user refinement}}. \quad (12)$$

The last three terms are for optional user refinement. Without  $E_h$ , the energy function is still a linear least squares problem and can be solved by a sparse solver. When including  $E_h$ , Equation 12 becomes nonlinear and we have to resort to iterative optimization. The process alternates between two steps: (1) fix  $H$  and optimize for the deformed vertices; (2) fix deformed vertices and optimize for  $H$ . The first step is linear. The second step is reduced to a homography estimation problem. In our experiments, we set  $\lambda_c = 5000$ ,  $\lambda_d = 1$ ,  $\lambda_s = 100$ ,  $\lambda_p = 500$ ,  $\lambda_o = 100$  and  $\lambda_h = 1$ .

### III. RESULTS

The proposed method involves heavy linear system solving and we implemented it using Matlab. Depending on the number of images to stitch, it took two to ten minutes for the results shown in this section. We tested the proposed method mainly on the Scene Collage dataset [6]. Figure 3 shows some collage results by stitching 5 to 35 images with the proposed method. The first one was taken on our own and the other two were taken from the Scene Collage dataset. We want to emphasize again that, for all results, we did not use any blending to better demonstrate the effectiveness of alignment.

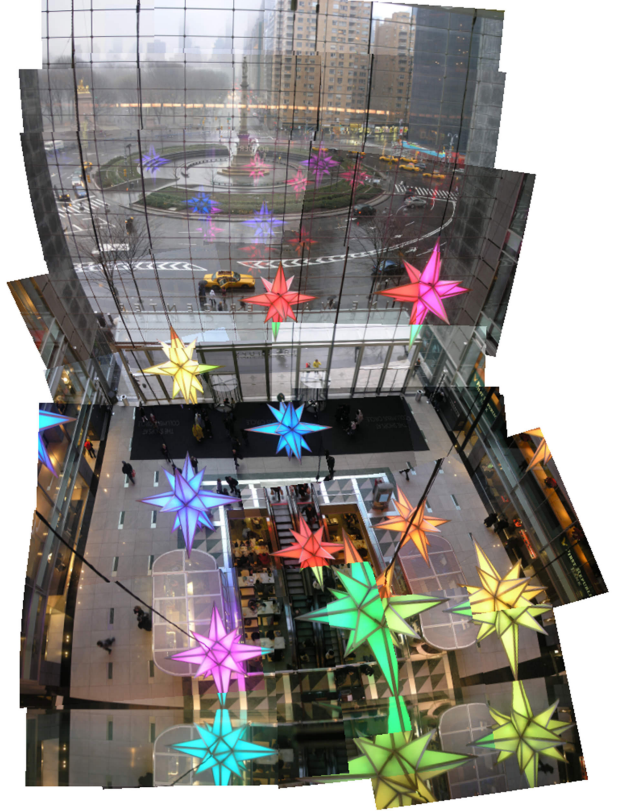
Figure 4 shows results for three image sets with comparisons to the Scene Collage method [6]. The first two image sets were taken from the Scene Collage dataset. The highlighted areas show that our method yields better alignment than Scene Collage, with less discontinuity along seams. The last example was taken on our own. We captured the input image by panning the camera and shooting a long paint pinned on a wall. The scene collage method generates a curved result while our method aligns more accurately. Our method allows



(a) stitching of 5 images with 4 matched image pairs.



(b) stitching of 15 images with 22 matched image pairs.



(c) stitching of 35 images with 88 matched image pairs.

Fig. 3. Results on image stitching. The first one (a) was taken on our own and the other two (b, c) were taken from the Scene Collage dataset.

spatially varying warps and thus offers more flexibility for more accurate alignment than global transforms.

Figure 5-7 show how to incorporate various types of user hints for creating more visually pleasing results. Figure 5 controls the overall shape of the final composite by placing control points. The composite better matches reality by having

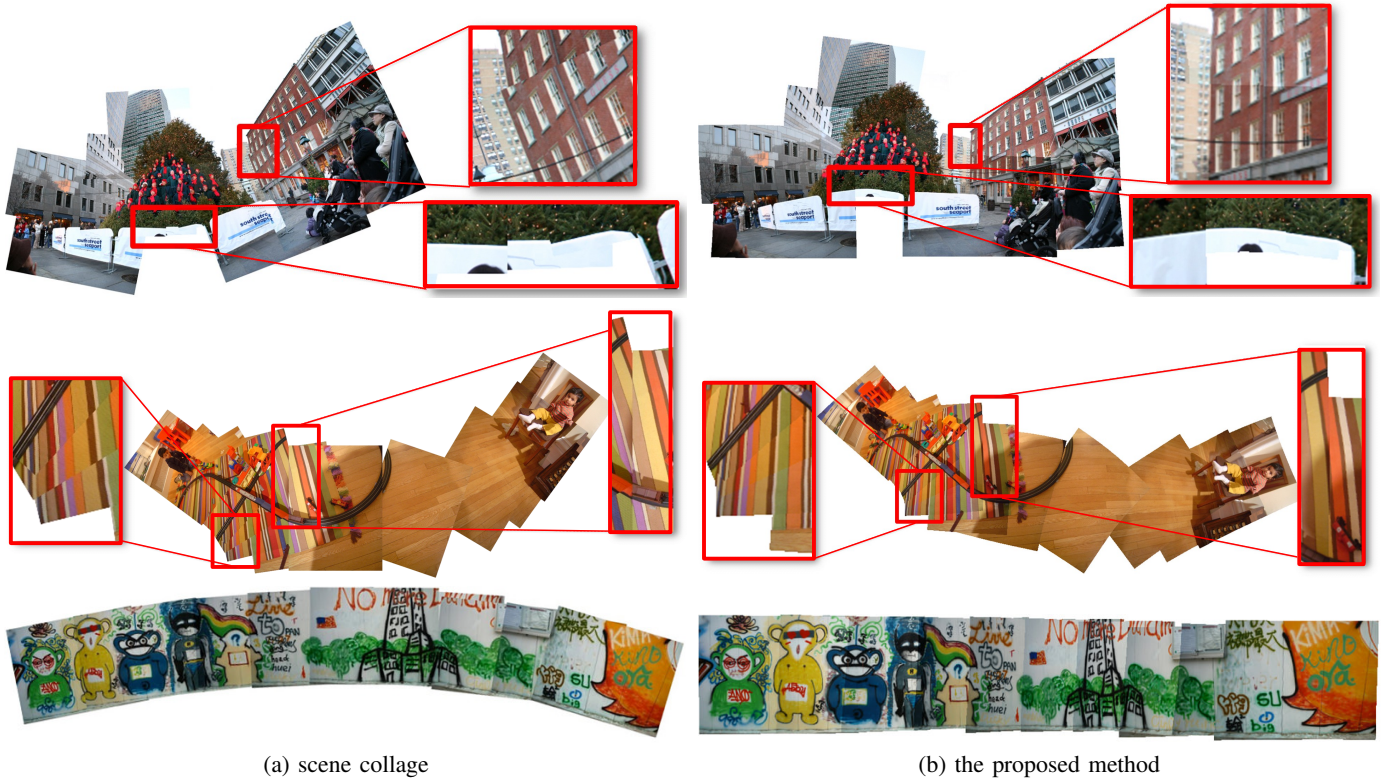


Fig. 4. Comparisons with the Scene Collage method [6] on three image sets. The first two image sets were taken from the Scene Collage dataset. The proposed method provides better alignment accuracy than the scene collage method. The last example was taken on our own. We captured the input image by panning the camera and shooting a long paint pinned on a wall. The scene collage method generates a curved results while our method aligns more accurately.



Fig. 5. By moving control points to the desired positions as shown on the right, users can adjust the overall shape of the composite so that it looks more realistic. The composite (b) looks better because the kids sit upright and the floor is level.

kids sit upright and keeping the floor level. Figure 6 again shows how to create a more pleasing result, this time, by constraining the orientations of two images on the bottom. Finally, Figure 7 shows that, by requiring the carpet to be deformed by a single homography, distortion of the carpet is greatly reduced, and the composite looks more natural.

#### IV. CONCLUSIONS

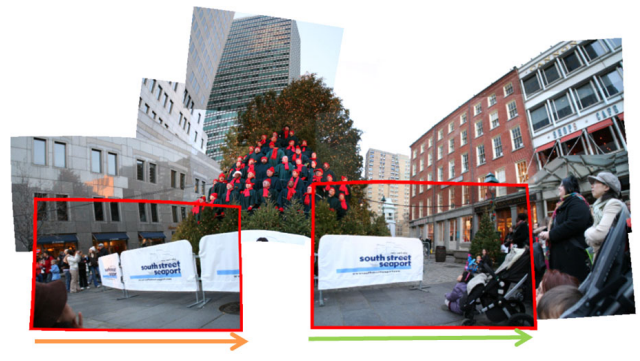
We present an scene alignment method for aligning a set of images from multiple view points. Instead of warping images with global transforms, our method uses mesh deformation to

induce spatially-varying warps. It provides great flexibility for aligning images with different optical centers and shows improved results against a global-warp-based method. In addition, we provides refinement tools for users to obtain more visually pleasing results. With the proposed tool, users can combine images from multiple viewpoints to produce a compelling and informative representation of scene with more freedom and ease. Although the aggregation stage compensates errors by shape preservation, our results rely on the performance of pairwise alignment. In addition, we would like to offer more editing options such as line structure preservation.



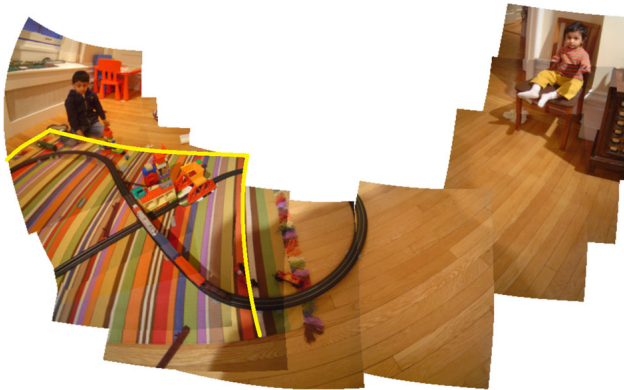


(a) before editing

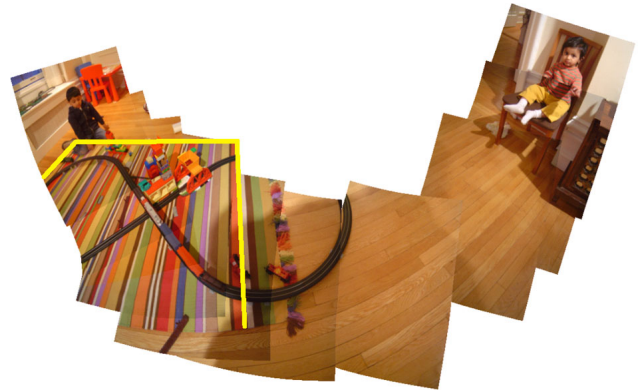


(b) after editing

Fig. 6. By constraining orientations of the highlighted images on the bottom, the composite (b) is visually more pleasing.



(a) before editing



(b) after editing

Fig. 7. By requiring that deformations of the carpet fit well with a single homography, distortion of the carpet can be greatly reduced in (b).

#### ACKNOWLEDGMENT

The authors would like to thank the reviewers for their helpful comments. This work was partly supported by grants NSC101-2628-E-002-031-MY3 and NSC102-2622-E-002-013-CC2 from The Ministry of Science and Technology of Taiwan, and grants 103R7761 and 103R7609-5 from National Taiwan University.

#### REFERENCES

- [1] L. Zelnik-Manor and P. Perona, "Automating joiners," in *Proceedings of Non-Photorealistic Animation and Rendering (NPAR)*, 2007.
- [2] W.-Y. Lin, S. Liu, Y. Matsushita, T.-T. Ng, and L.-F. Cheong, "Smoothly varying affine stitching," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 345–352.
- [3] J. Gao, S. J. Kim, and M. S. Brown, "Constructing image panoramas using dual-homography warping," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 49–56.
- [4] J. Zaragoza, T.-J. Chin, M. S. Brown, and D. Suter, "As-projective-as-possible image stitching with moving DLT," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013, pp. 2339–2346.
- [5] R. Szeliski, "Image alignment and stitching: a tutorial," *Foundations and Trends in Computer Graphics and Vision*, vol. 2, no. 1, pp. 1–104, 2006.
- [6] Y. Nomura, L. Zhang, and S. Nayar, "Scene collages and flexible camera arrays," in *Proceedings of Eurographics Symposium on Rendering (EGSR)*, 2007.
- [7] M. Brown and D. Lowe, "Recognising panoramas," in *Proceedings of International Conference on Computer Vision (ICCV)*, 2003, pp. 1218–1225.
- [8] A. Agarwala, M. Agrawala, M. Cohen, D. Salesin, and R. Szeliski, "Photographing long scenes with multi-viewpoint panoramas," *ACM Transactions on Graphics*, vol. 25, no. 3, pp. 853–861, July 2006.
- [9] A. Rav-Acha, G. Engel, and S. Peleg, "Minimal aspect distortion (MAD) mosaicing of long scenes," *International Journal of Computer Vision*, vol. 78, no. 2-3, pp. 187–206, July 2008.
- [10] E. Zheng, R. Raguram, P. Fite-Georgel, and J.-M. Frahm, "Efficient generation of multi-perspective panoramas," in *Proceedings of 2011 International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT)*, May 2011, pp. 86–92.
- [11] C.-H. Chang, Y. Sato, and Y.-Y. Chuang, "Shape-preserving half-projective warps for image stitching," in *Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, 2014.
- [12] G.-X. Zhang, M.-M. Cheng, S.-M. Hu, and R. R. Martin, "A shape-preserving approach to image resizing," *Computer Graphics Forum*, vol. 28, no. 7, pp. 1897–1906, 2009.
- [13] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *Proceedings of the 7th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 674–679, 1981.
- [14] S. Baker and I. Matthews, "Lucas-Kanade 20 years on: A unifying framework," *International Journal of Computer Vision*, vol. 56, no. 3, pp. 221 – 255, March 2004.
- [15] R. Carroll, A. Agarwala, and M. Agrawala, "Image warps for artistic perspective manipulation," *ACM Transactions on Graphics*, vol. 29, no. 4, 2010.