Supplementary document for

PANORAMA FROM REPRESENTATIVE FRAMES OF UNCONSTRAINED VIDEOS USING DIFFEOMESHES

PAPER ID: 722

A. RESULTS ON PANORAMA CREATION FROM UNCONSTRAINED VIDEOS

To compare the performance of our proposed method based on DiffeoMeshes with a few recent methods, we formed a dataset consisting of 20 videos, 10 of which are snippets from publicly available Youtube videos and 10 of them taken from the stabilization dataset [S1]. The videos are named as Vid1-Vid10, Simple1-3, Parallax1-5, Run1-2. Thumbnails of all the unconstrained videos collected in the dataset are shown in fig. S1 below:

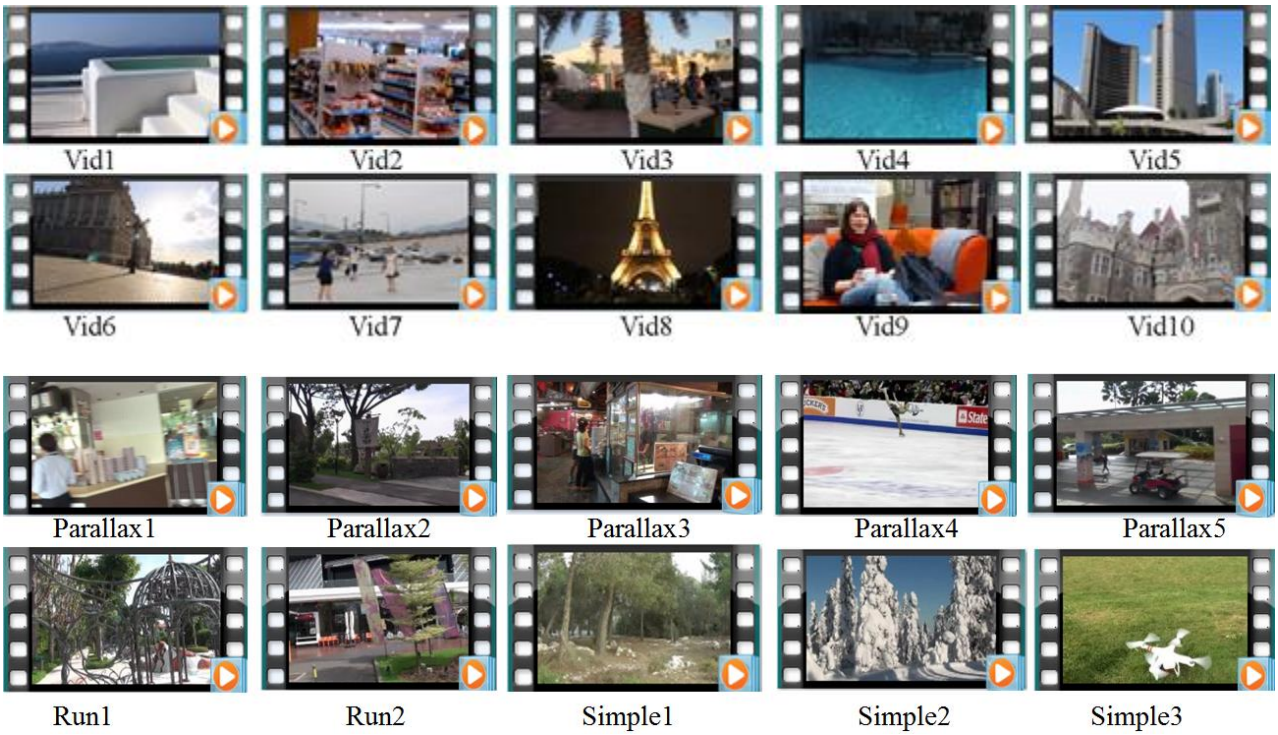


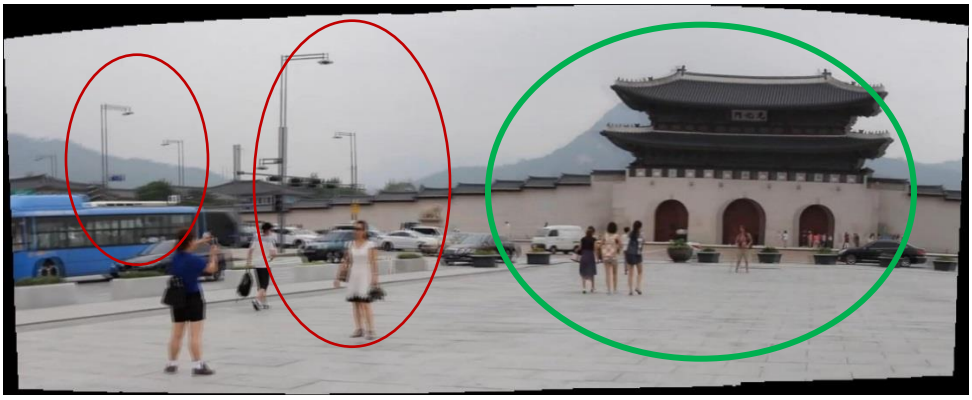
Fig. S1: Thumbnails of videos in dataset for comparing with the performance of different methods for panorama creation from unconstrained videos. The samples ***Vid1-Vid10*** are taken from Youtube videos (first two rows) and the samples Parallax1-5, Vid9, Run1-2, Simple1-3 are taken from the stabilization dataset in [S1].

A1. ABLATION STUDIES

To study the influence of each module of the proposed approach, we have performed some experiments as follows:

Effect of balancing factor in the dissimilarity score (Eqn 1): The balancing factor α decides how much importance should be given to the blur score. Higher values of α ensures that less blurry frames are used for stitching. Among the videos in the dataset, *Vid7* has many blurry frames. The panoramas generated with different values of α are shown:

$\boldsymbol{\alpha = 0.3}$



$\boldsymbol{\alpha = 0.5}$

$\boldsymbol{\alpha = 0.7}$

Fig. S2: Effect of balancing factor in dissimilarity score. As the balancing factor increases, the blurriness of the panorama reduces due to the selection of less blurry frames.

Statistics of representative frame selection in the videos of our dataset: Table S1 shows number of frames after selecting **every third frame** of the video, number of frames selected after the sparse frame selection approach using the proposed dissimilarity score, minimum and maximum amount of overlap and the minimum IFTS score among the adjacent frames of the representative frames.

Video Sample	Total no of frames (3rd)	No of selected frames	Min. Overlap	Max. Overlap	Min. IFTS score
Vid 1	18	08	83%	95%	0.9938
Vid 2	26	13	76%	92%	0.8975
Vid 3	24	12	70%	97%	0.8011
Vid 4	37	08	75%	98%	0.9450
Vid 5	63	31	82%	97%	0.9265
Vid 6	23	18	83%	97%	0.7101
Vid 7	15	6	85%	96%	0.9639
Vid 8	35	5	79%	97%	0.9954
Vid 9	162	42	78%	93%	0.8710
Vid 10	78	40	82%	97%	0.9285
Parallax1	186	55	74%	99%	0.7891
Parallax2	60	26	76%	95%	0.8124
Parallax3	46	23	67%	96%	0.9102
Parallax4	56	24	61%	90%	0.9819
Parallax5	155	35	42%	95%	0.8561
Run1	47	14	75%	99%	0.9977
Run2	34	21	87%	99%	0.4784
Simple1	100	31	74%	98%	0.9671
Simple2	22	10	85%	97%	0.9510
Simple3	20	12	31%	96%	0.9678

Table S1: Statistics of the representative frames selected from our method.

As seen in the last column of the table, the frames selected have good aligning capability because of the high IFTS score for every video, except for the video Run1. This video has large parallax error and hence the number of frames selected is also high. Also, the minimum overlap of the frames can vary based on the IFTS score. For example, for the video *Simple3*, the minimum overlap is 31% whereas for the video *Simple2*, the minimum overlap is 85%. For both the cases, the minimum IFTS score is almost the same. Hence, our approach of representative selection can select frames that are less overlapped and better aligned.

Effect of using Closeness Centrality score for selecting the reference frame: The aim of MST chaining is to obtain an optimal chain of ordering. In some unconstrained videos, the motion of the camera may be arbitrary. For example, the example *Vid9* has to and fro camera motion. Following figure shows the reference frame selected by the naïve method of selecting the middle frame of the video and the central frame generated by our proposed method. The figure also shows the panorama formed. It is seen that the middle frame of the video lies towards the end of the panorama, whereas the reference frame selected lies central to the panorama.

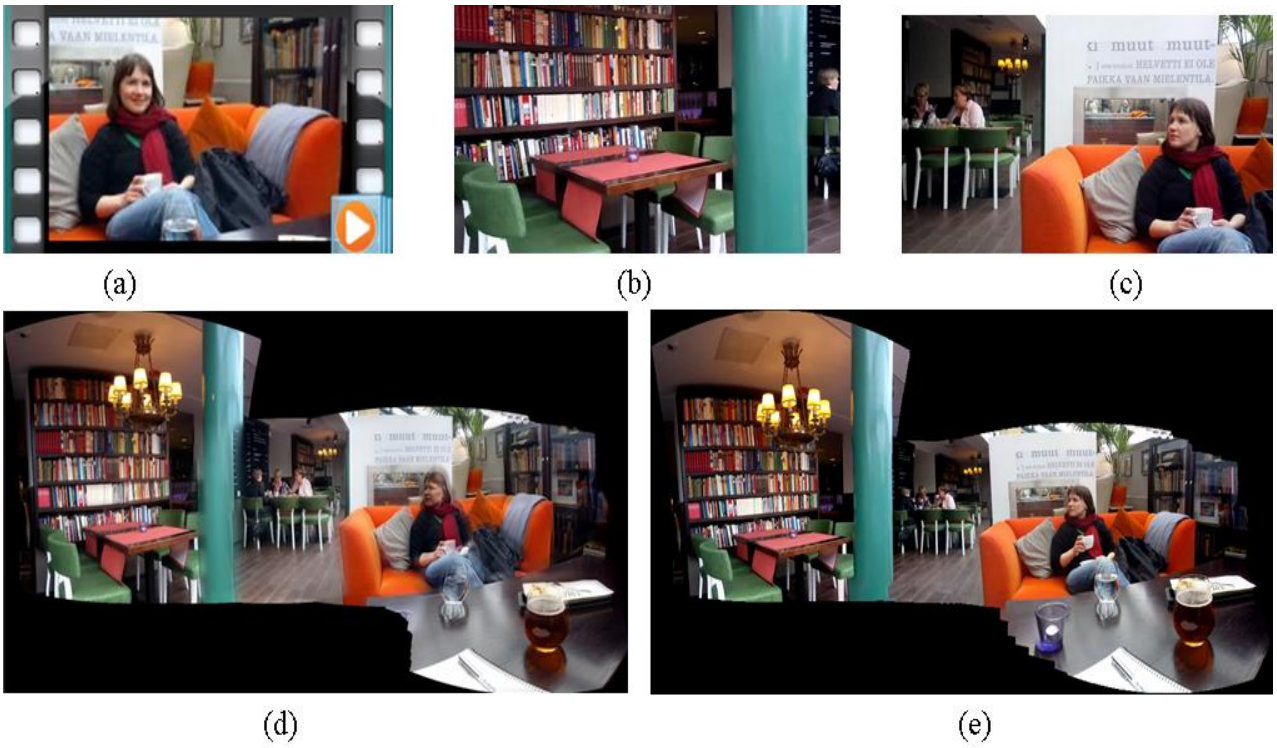


Fig S3: (a) Thumbnail of *Vid9*, (b) middle frame of the video (naïve method of selection of frame), (c) Reference frame selected by our method, (d) panorama with (b) as reference frame and sequential order of stitching vs (b) panorama with (c) as the reference frame and MST chain of ordering. Severe ghosting and misalignment artifacts can be seen in (d).

A2. Qualitative comparison of the performance with AutoStitch [S2] (also reference #1 in the main manuscript)

AutoStitch [S2] generated ghosting artifacts since blending is used to hide the alignment errors and moving objects. In the figs. S2-S14, the regions marked in red shows the ghosting and alignment artifacts caused by AutoStitch. In some cases as in fig. S3 and S6, AutoStitch fails to generate the full panorama.

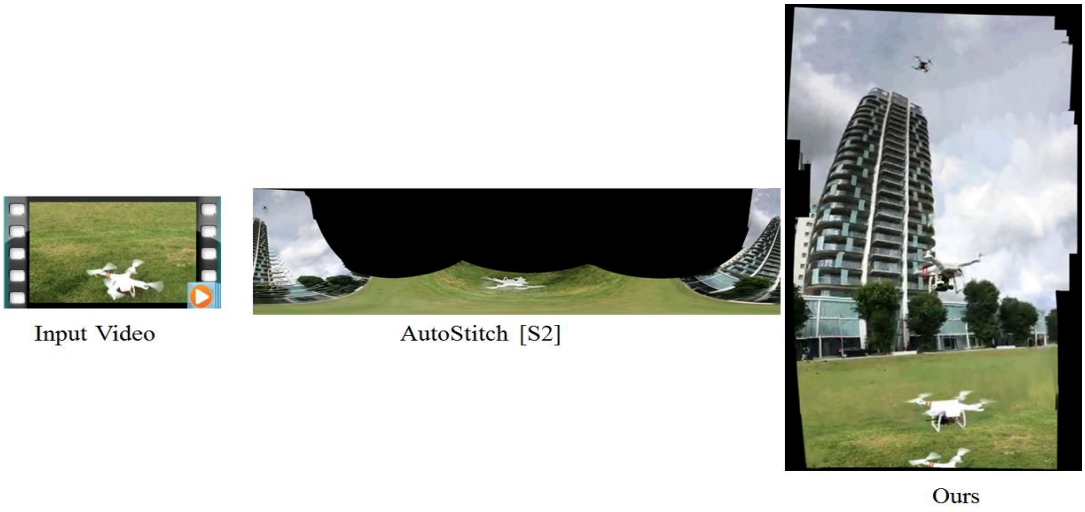


Fig. S4: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Simple3***. The regions marked in red shows the ghosting artifacts caused by AutoStitch. The results for our method is clearer and sharper.

Fig. S5: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Vid2***. AutoStitch fails to generate the whole panorama from the input video, while ours generates the full panorama.

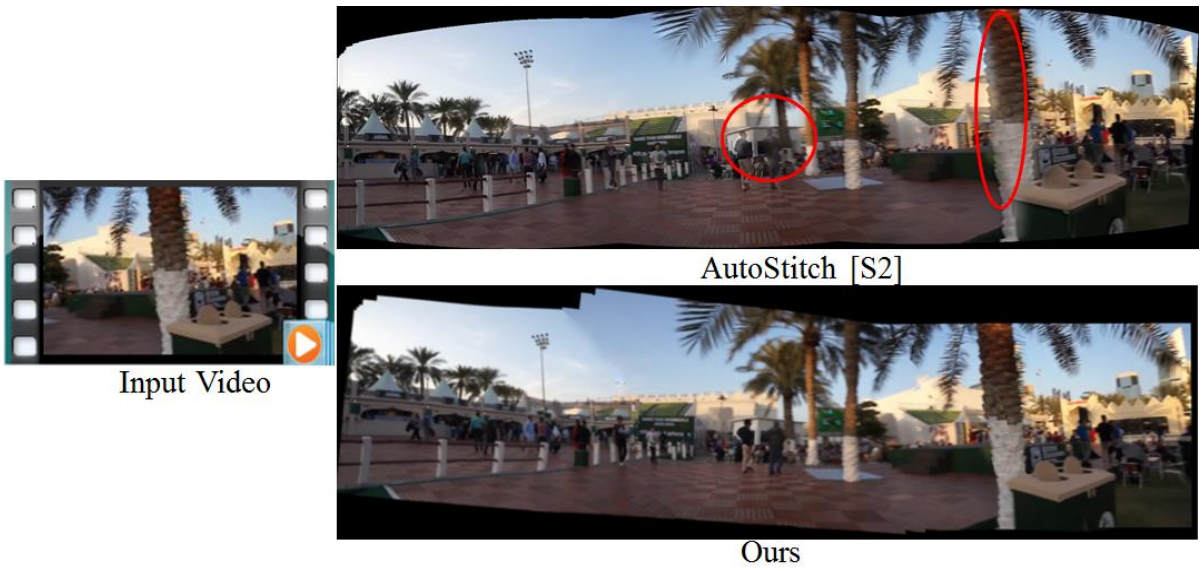


Fig. S6: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Vid3***. AutoStitch causes ghosting artifacts as shown in regions marked in red oval shapes.

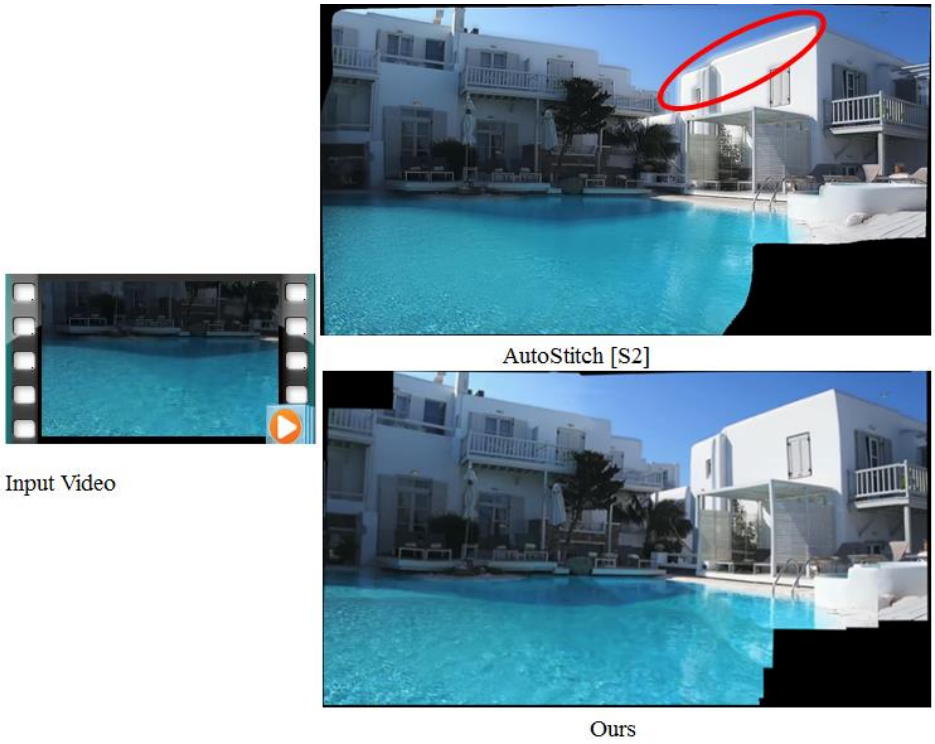


Fig. S7: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Vid4***. AutoStitch causes ghosting artifacts as shown in regions marked in red oval shapes.

Fig. S8: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Vid5***. AutoStitch fails to generate the whole panorama from the input video, while ours generates the full panorama.

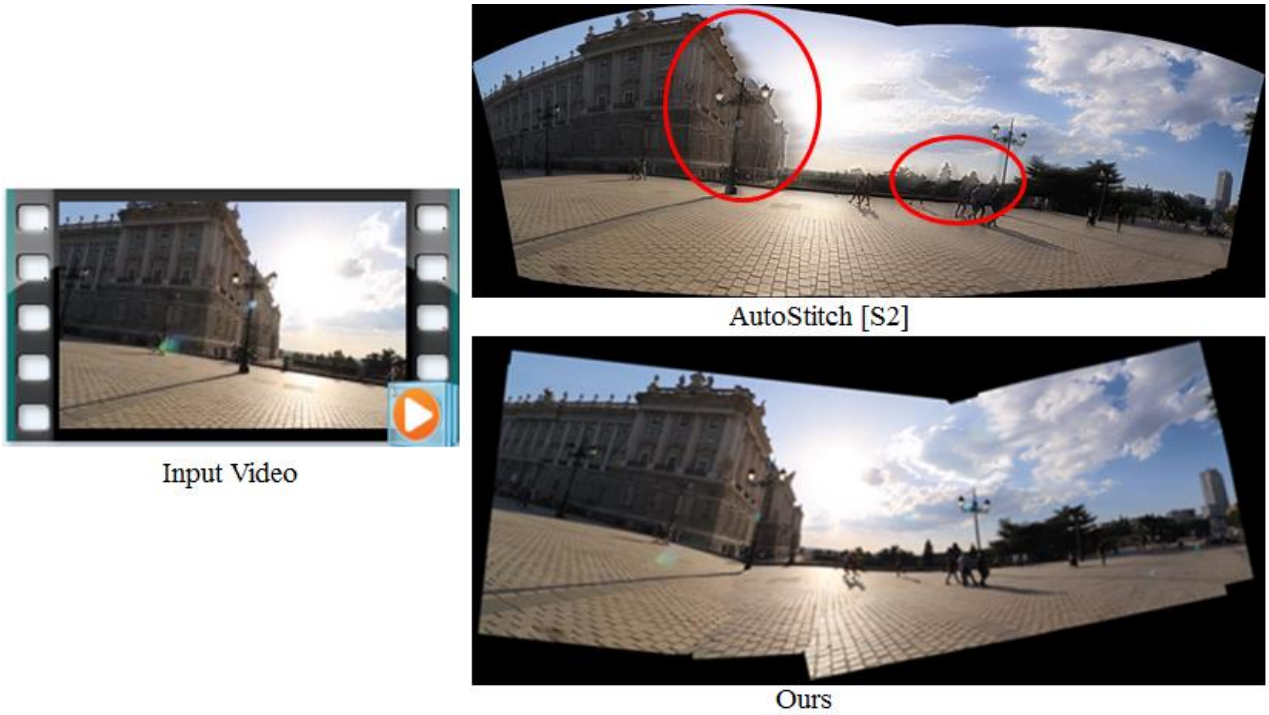


Fig. S9: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Vid6***. AutoStitch causes ghosting artifacts as shown in regions marked in red oval shapes.

Fig. S10: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Vid8***. In AutoStitch output, the region marked in red shows the ghosting effects produced in the panorama. Also the lights in the image appears smudged in the case of the output of AutoStitch, which is not the case with the output of our proposed method. The outputs have been artificially enhanced (brightness, contrast and saturation) equally to clearly show the dark foreground of the stitched images, which is otherwise not clearly comparable.

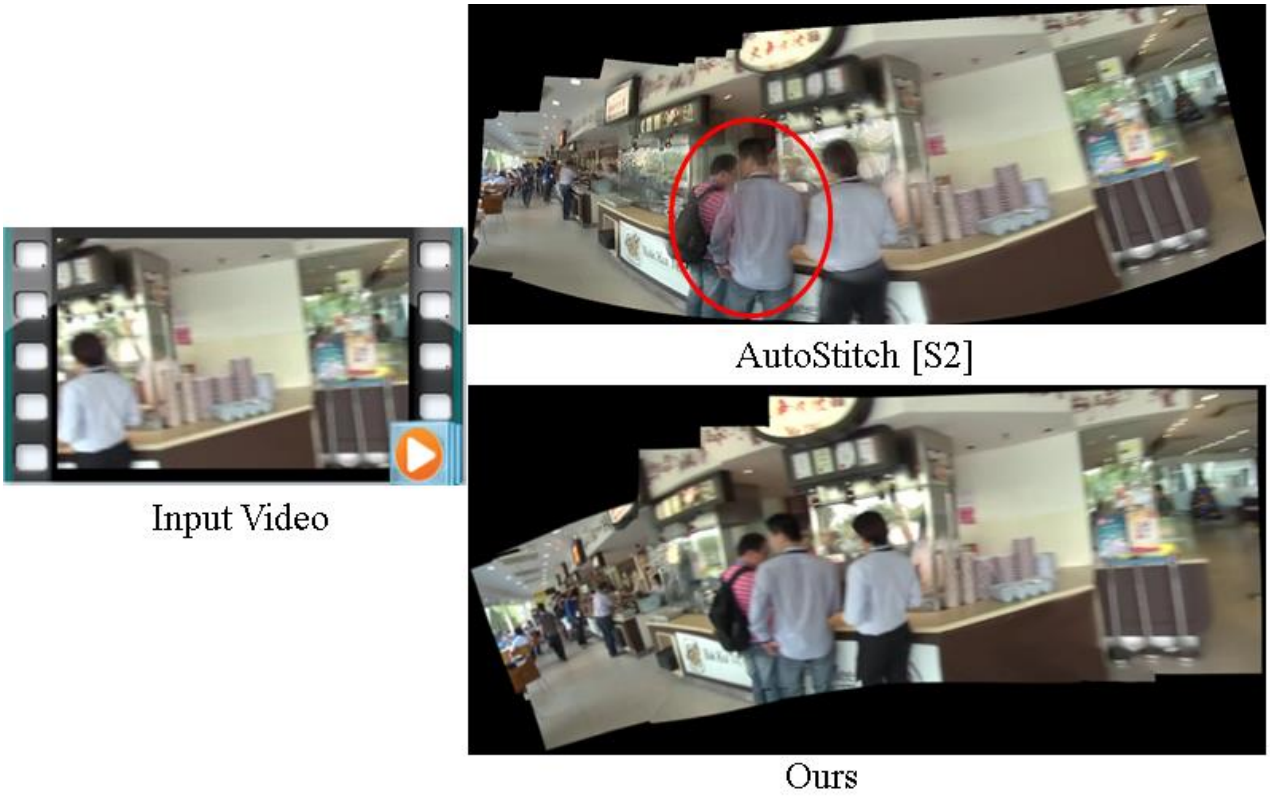


Fig. S11: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Parallax1***. AutoStitch causes ghosting artifacts in regions marked in red oval shapes.

Fig. S12: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Parallax2***. AutoStitch causes ghosting artifacts throughout the stitched image. The red oval boundary highlight the ghosting effects.

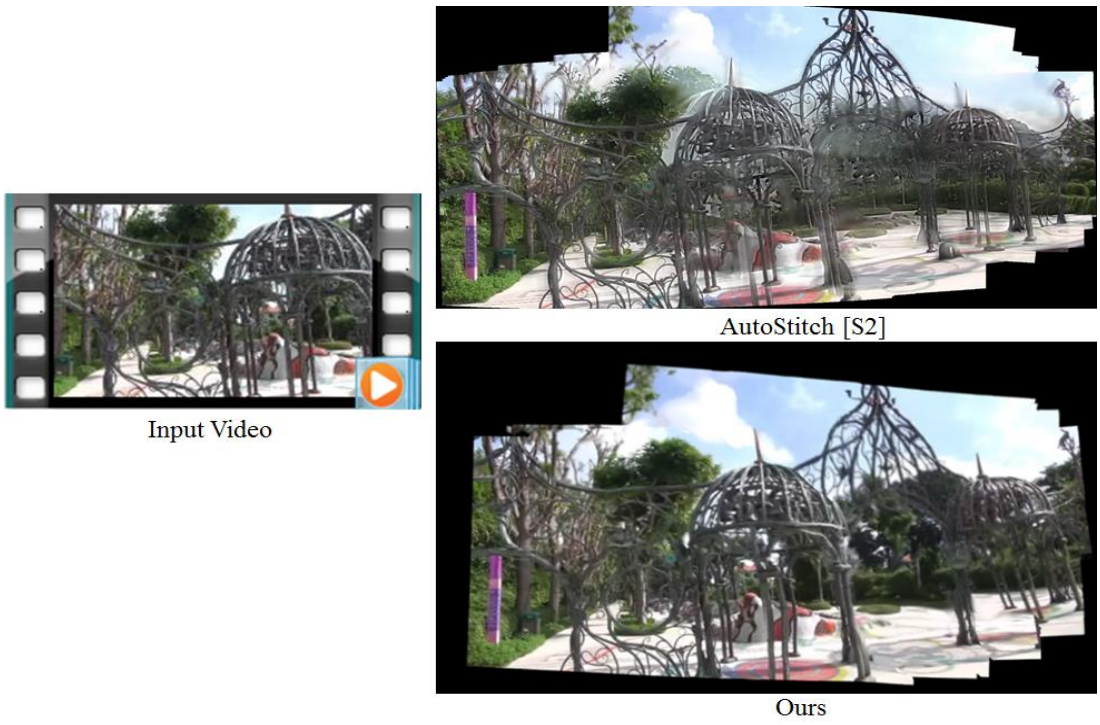


Fig. S13: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Run1***. AutoStitch causes ghosting artifacts throughout the stitched image.

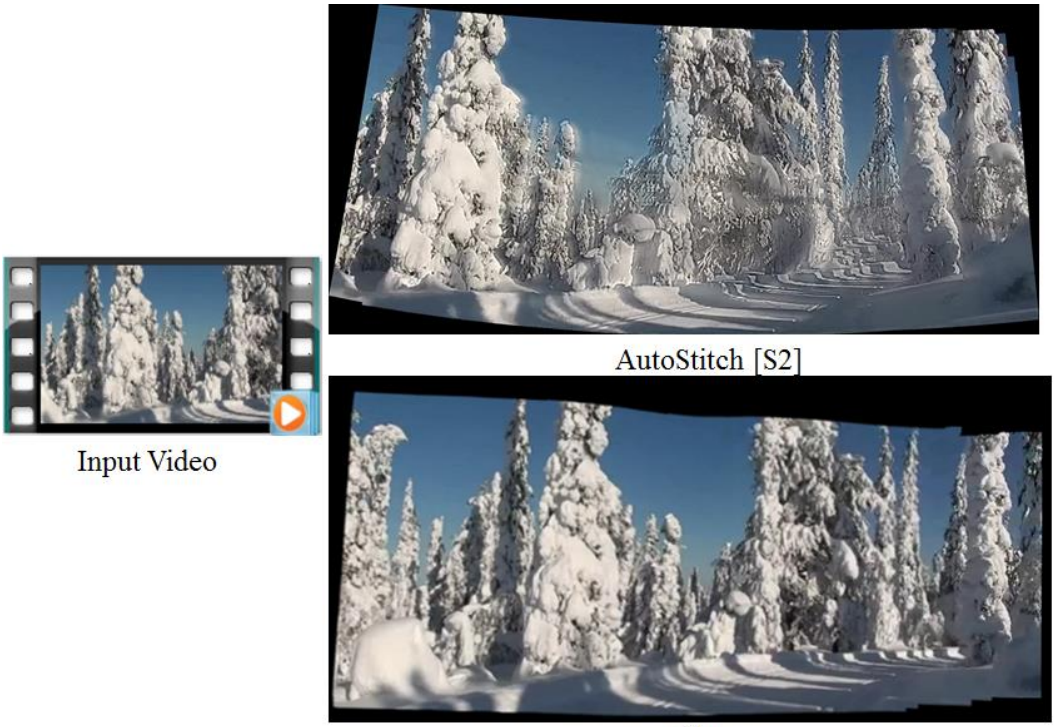


Fig. S14: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Simple2***. AutoStitch causes ghosting artifacts throughout the stitched image.

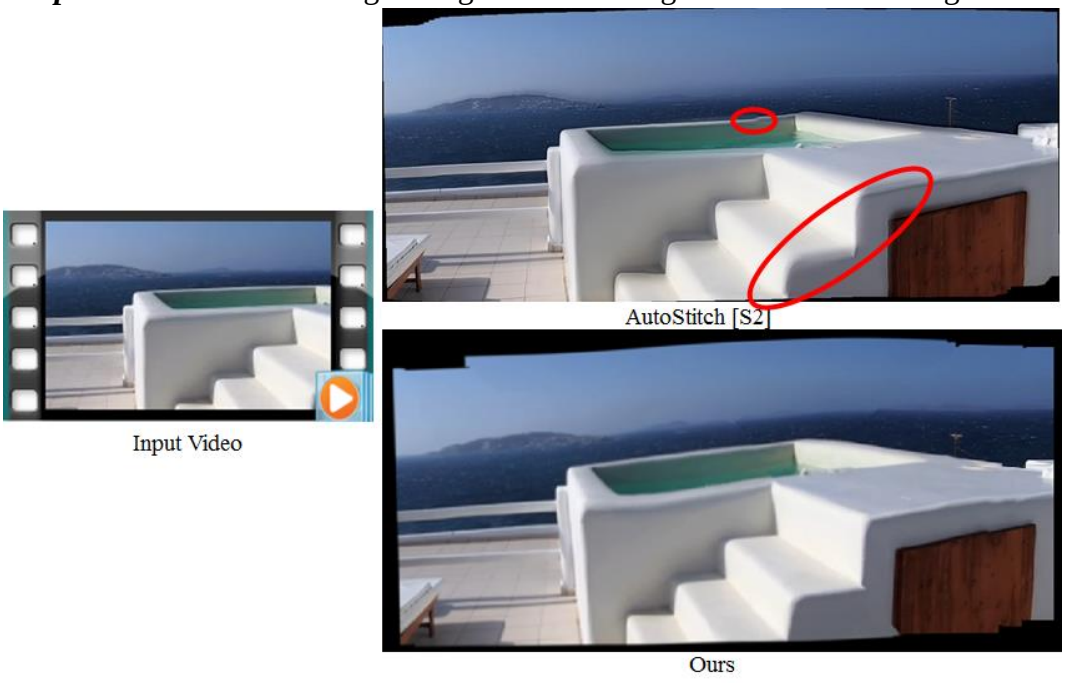


Fig. S15: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Vid1***. AutoStitch fails to generate the panorama, whereas our method produces the desired output.

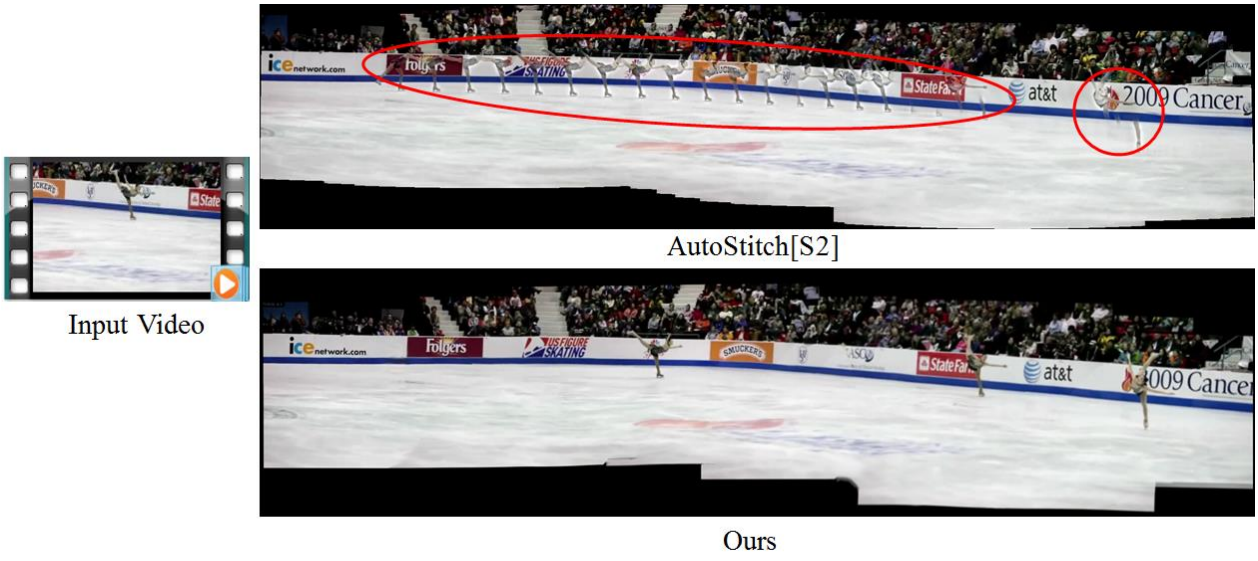


Fig. S16: Qualitative Comparison for panorama generation of our method with AutoStitch [S2] for ***Simple3***. AutoStitch causes ghosting effects in the stitched image due to the moving object (shown in red oval boundaries).

A3. Qualitative comparison of the performance with Adobe Photoshop [S3]

The experiments were performed in Adobe Photoshop CS6 [S3]. The panoramas generated by Photoshop are better than that of AutoStitch and have minimal ghosting artifacts. However, for frames of unconstrained videos, the output is distorted towards the ends (left and/or right) of the panorama as seen in figs. S15-S22.

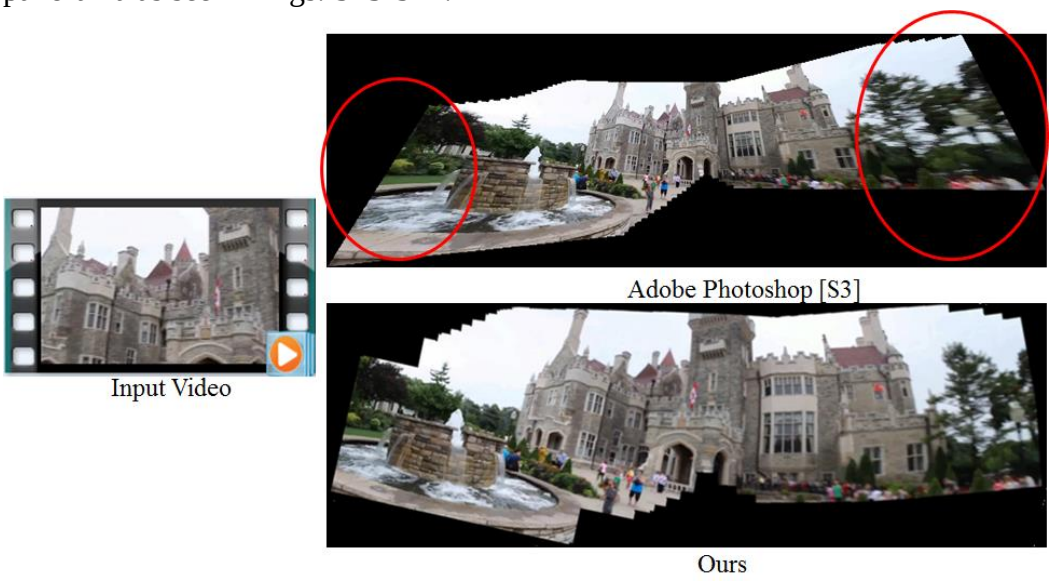


Fig. S17: Qualitative Comparison for panorama generation of our method with Adobe Photoshop [S3] for ***Vid10***. The elongated regions in the output of Photoshop are marked in red oval shape. Our method provides the desired output.

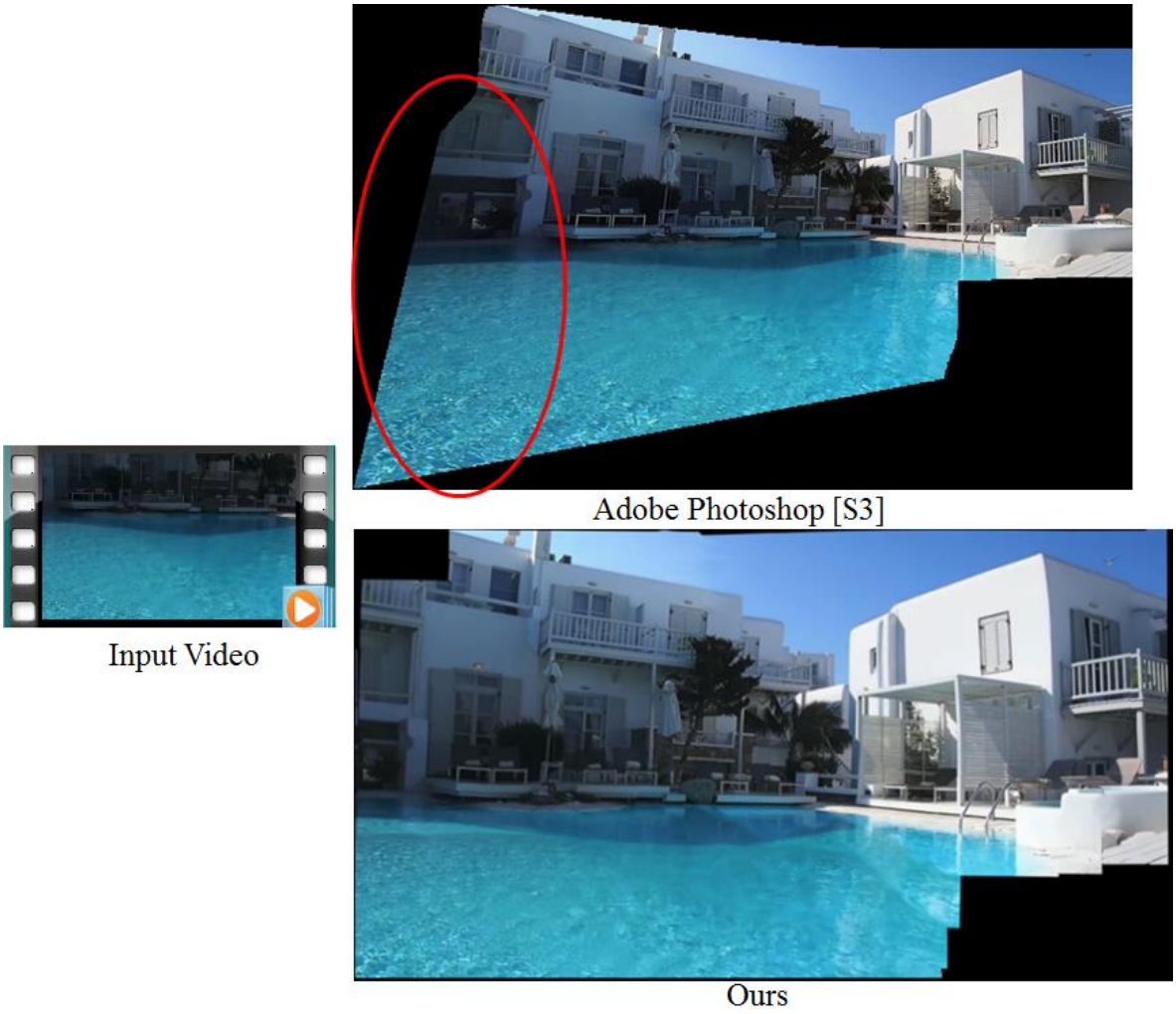


Fig. S18: Qualitative Comparison for panorama generation of our method with Adobe Photoshop [S3] for ***Vid4***. The elongated regions in the output of Photoshop are marked in red oval shape. Our method provides the desired output.

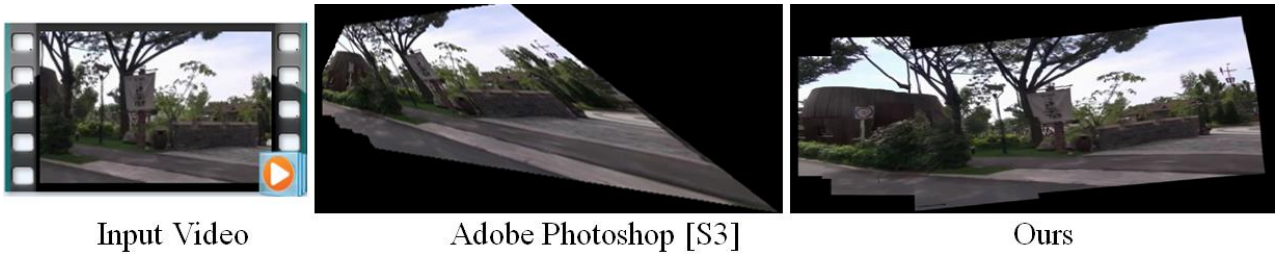


Fig. S19: Qualitative Comparison for panorama generation of our method with Adobe Photoshop [S3] for ***Parallax2***. The panorama generated by Photoshop is distorted at the right end.

.

Fig. S20: Qualitative Comparison for panorama generation of our method with Adobe Photoshop [S3] for ***Parallax1***. The elongated regions in the output of Photoshop are marked in red oval shape. Our method provides the desired output.

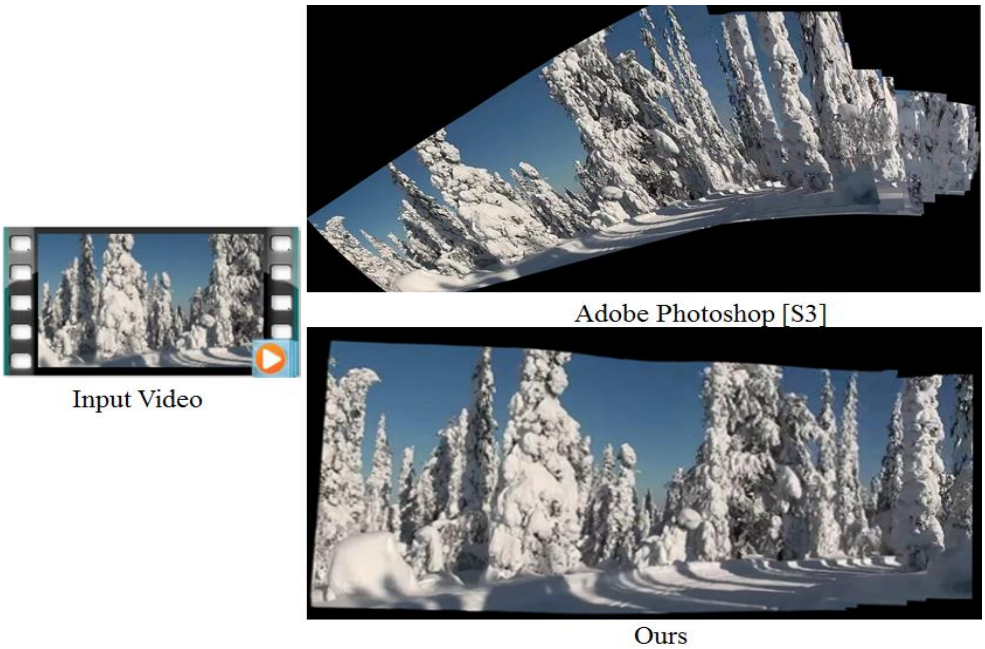


Fig. S21: Qualitative Comparison for panorama generation of our method with Adobe Photoshop [S3] for ***Simple2***. The panorama generated by Photoshop seems distorted, whereas our method provides the desired output.

Input Video

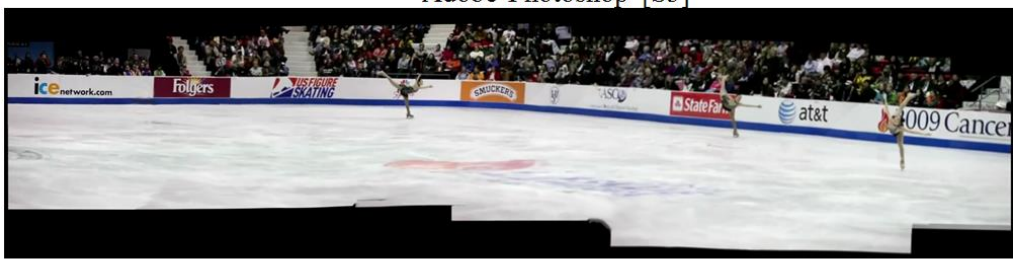
AutoStitch [S2]

Ours

Fig. S22: Qualitative Comparison for panorama generation of our method with Adobe Photoshop [S3] for ***Run1***. The artifacts in the output of Photoshop are marked in red oval shape. Our method provides the desired output

Input Video

Adobe Photoshop [S3]



Ours

Fig. S23: Qualitative Comparison for panorama generation of our method with Adobe Photoshop [S3] for ***Parallax4***. The distortion is Adobe Photoshop is seen in the left region of the panorama.

B. ALIGNMENT RESULTS

B1. Qualitative comparison with SEAGULL [S4] (also reference # 20 in main manuscript)

We have also compared the performance of our method with the method in [S4]. We have taken the results published in the author's webpage [S8]. As seen in most of the examples below in Figs. S24-S30, SEAGULL elongates the stitched image towards the ends while warping, whereas our method produces the desired output without elongation, and all key artifacts (edges, contrast, etc.) of the scene are preserved. All the examples in figs. S23-S29 are taken from SEAGULL dataset [S4].

Fig. S24: Qualitative comparison of performance with SEAGULL [S4], for an example taken from SEAGULL dataset [S4].

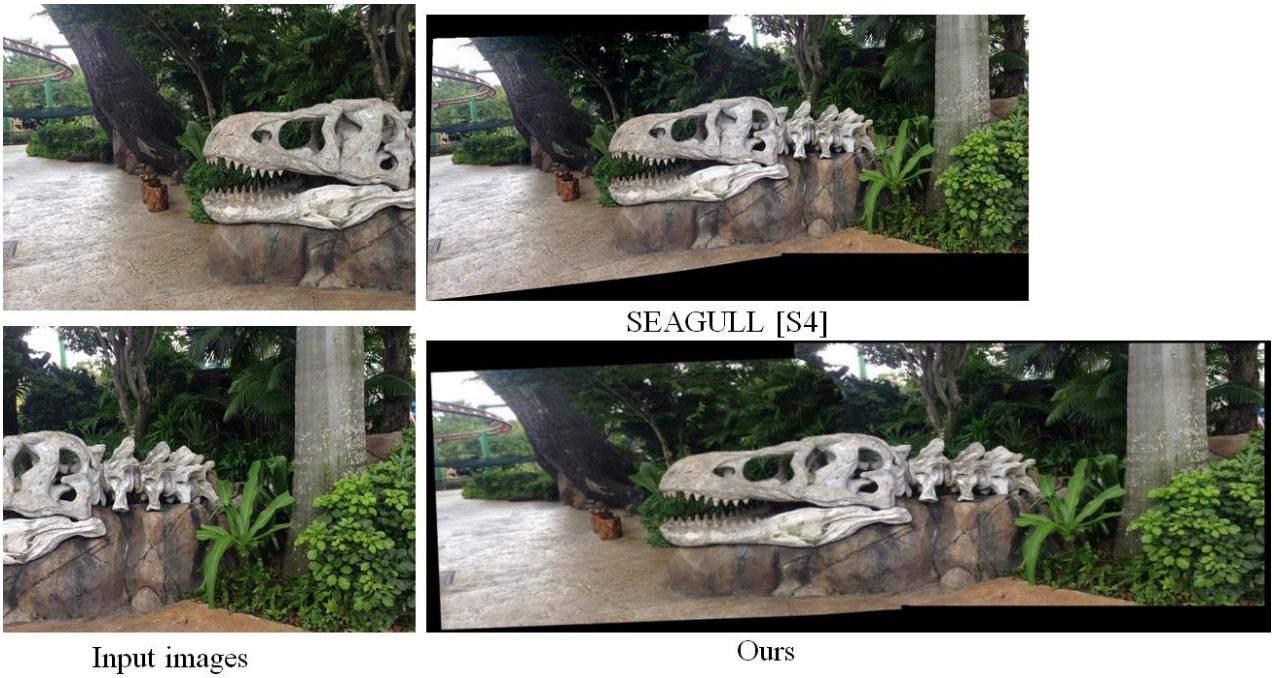


Fig. S25: Qualitative comparison of performance with SEAGULL [S4], for an example taken from SEAGULL dataset [S4].

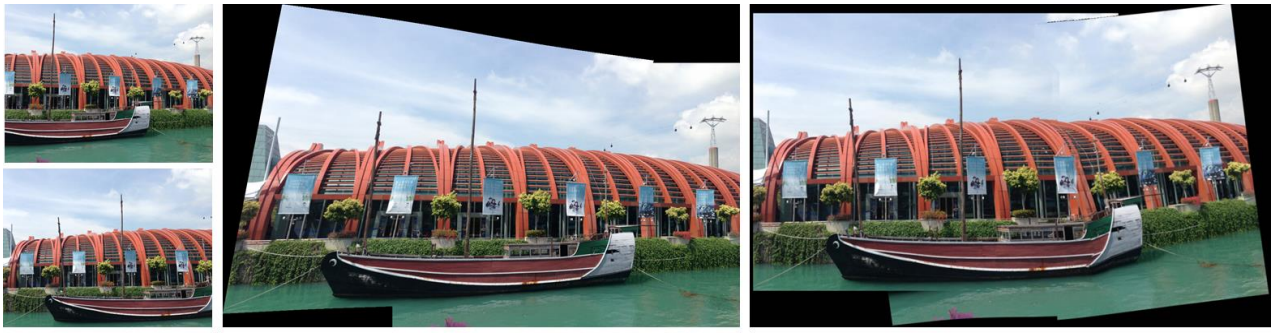


Fig. S26: Qualitative comparison of performance with SEAGULL [S4], for an example taken from SEAGULL dataset [S4].

Fig. S27: Qualitative comparison of performance with SEAGULL [S4], for an example taken from SEAGULL dataset [S4].

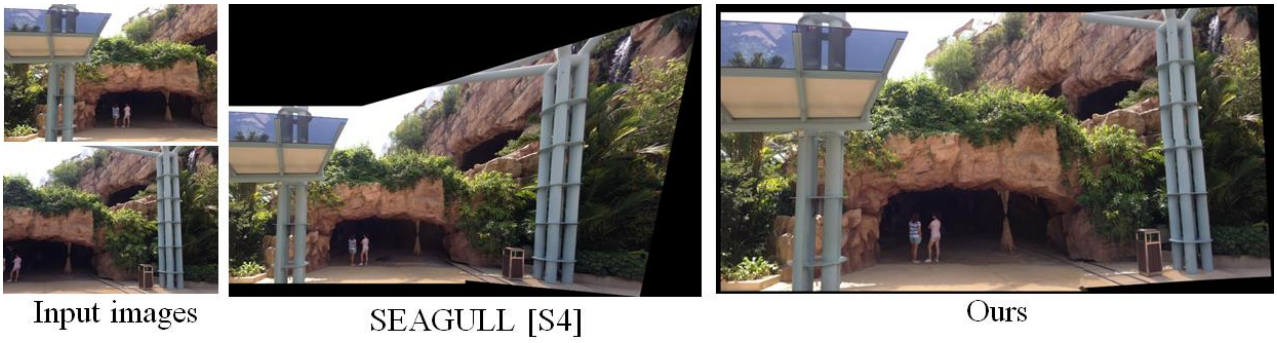


Fig. S28: Qualitative comparison of performance with SEAGULL [S4], for an example taken from SEAGULL dataset [S4].

Fig. S29: Qualitative comparison of performance with SEAGULL [S4], for an example taken from SEAGULL dataset [S4].

Fig. S30: Qualitative comparison of performance with SEAGULL [S1], for an example taken from SEAGULL dataset [S4].

B2. Qualitative comparison with SPHP [S7] (also reference # 15 in main manuscript)

We have compared the performance of our method with that published in [S7]. The code for this method is obtained from author's website [S10]. The example in fig. S30 shows a failure case of SPHP in SEAGULL dataset, whereas the input images in the examples of figs. S31-32 are frames of videos of our dataset (Section A of the supplementary document).

Fig. S31: Qualitative comparison of performance with the method SPHP [S7], for an example taken from SEAGULL dataset [S4]. The alignment fails when SPHP is used, whereas our method produces the desired result.

Fig. S32: Qualitative comparison of performance with the method SPHP [S7], for a pair of frames taken from an unconstrained video, ***Parallax2***, of our dataset (Section A). SPHP produces distortions towards the right top end of the stitched image as seen in the region marked in red oval boundary, whereas our method produces the desired output.

Fig. S33: Qualitative comparison of performance with the method SPHP [S7], for a pair of frames taken from an unconstrained video, ***Run2***, of our dataset (Section A). SPHP produces distortions towards the right top end of the stitched image as seen in the region marked in red oval shape, whereas our method produces the desired output.

B3. Qualitative comparison with APAP [S6] (also reference # 14 in main manuscript)

We have compared the performance of our method with the outputs of the method in [S6]. The code for this method is obtained from author's website [S9]. The examples in figs. S33-S34 are taken from SEAGULL dataset [S4]. APAP causes distortions towards the left or right end in both the examples. The examples in figs. S35-S36 are taken from the dataset in [S11], where our method performs better than APAP.

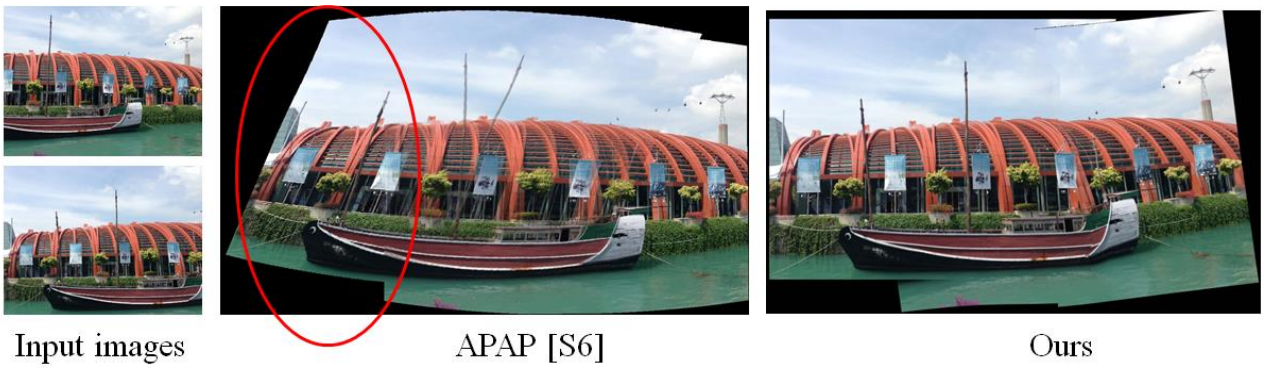


Fig. S34: Qualitative comparison of performance with APAP [S6], for an example of the SEAGULL dataset [S4]. Unwanted distortion in the stitched image is produced by APAP (shown in red oval boundary), whereas our method produces the desired output.

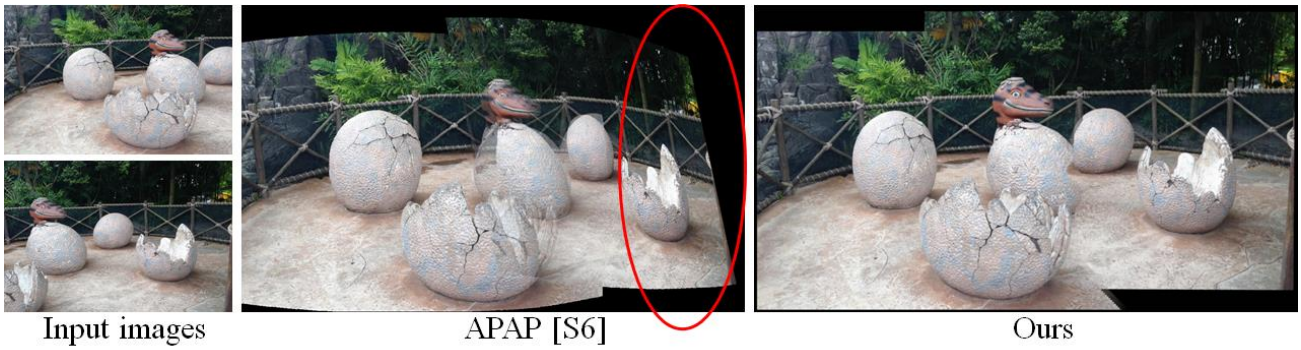


Fig. S35: Qualitative comparison of performance with APAP [S6], for an example of the SEAGULL dataset [S4]. Unwanted distortion in the stitched image is produced by APAP (shown in red oval boundary), whereas our method produces the desired output.

Fig. S36: Qualitative comparison of performance with APAP [S6], for an example of the dataset of the method in [S11]. Unwanted distortion in the stitched image is produced by APAP (shown in red oval boundary), whereas our method produces the desired output.

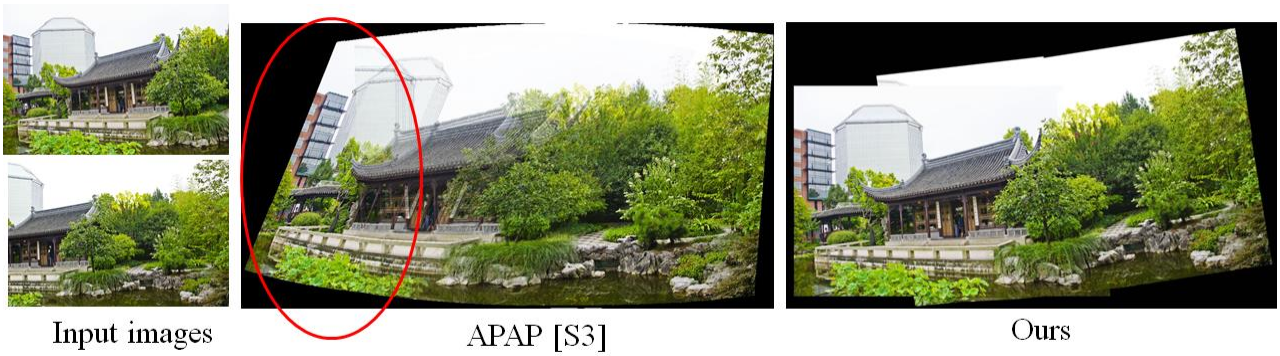


Fig. S37: Qualitative comparison of performance with APAP [S6], for an example of the dataset of the method in [S11]. Unwanted distortion in the stitched image is produced by APAP (shown in red oval boundary), whereas our method produces the desired output.

REFERENCES

[S1] http://liushuaicheng.org/SIGGRAPH2013/database.html/

[S2] Brown, M and Lowe, D.G, "Recognising panoramas", in ICCV, 2003

[S3] https://www.adobe.com/in/products/photoshop.html

[S4] Kaimo Lin, Nianjuan Jiang, Loong-Fah Cheong, Minh Do, and Jiangbo Lu, "Seagull: Seam-guided local alignment for parallax-tolerant image stitching," in ECCV, 2016.

[S5] Feng Liu, Michael Gleicher, Hailin Jin, and Aseem Agarwala, "Content-preserving warps for 3d video stabilization", ACM Transactions on Graphics (TOG), vol.28, no. 3, pp. 44, 2009.

[S6] Julio Zaragoza, Tat-Jun Chin, Michael S Brown, and David Suter, "As-projective-as-possible image stitching with moving DLT," in CVPR, 2013.

[S7] Che-Han Chang, Yoichi Sato, and Yung-Yu Chuang, "Shape-preserving half-projective warps for image stitching," in CVPR, 2014.

[S8] http://www.linkaimo.com/publications/ImageStitching/ImageStitching.html

[S9] https://cs.adelaide.edu.au/~tjchin/apap/

[S10] http://www.cmlab.csie.ntu.edu.tw/~frank/SPH/

[S11] Fan Zhang and Feng Liu, "Parallax-tolerant image stitching," in CVPR, 2014.