Spatial and Temporal Interpolation of Multi-View Image Sequences

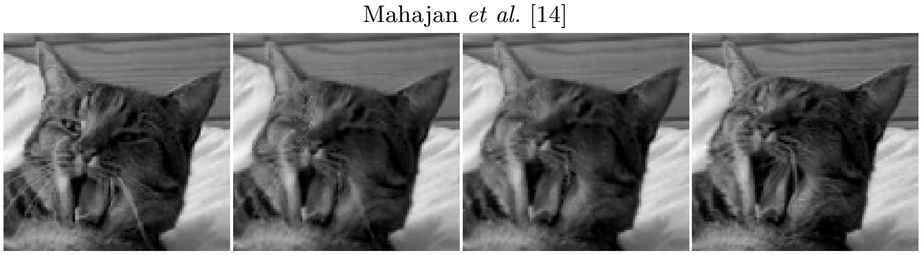
Tobias Gurdan[1,2], Martin R. Oswald[1], Daniel Gurdan[2], Daniel Cremers[1]

[1]Department of Computer Science, Technische Universität München*
[2]Ascending Technologies GmbH, Krailing, Germany

1 Further Results

In this supplementary material we present further experimental results of our proposed framework. Fig. 3 shows a qualitative comparison of the methods by

Mahajan *et al.* [14]



Werlberger *et al.* [25]

Proposed sparse approach

Fig. 1: Comparison of our sparse approach to two state-of-the-art temporal interpolation methods. *From left to right*: Input image, interpolation at $t = 1/3$, interpolation at $t = 2/3$, target image. Our approach yields comparable results.

Mahajan *et al.* [14], Werlberger *et al.* [25] and our proposed sparse approach.

* This work was supported by the ERC Starting Grant “Convex Vision”.

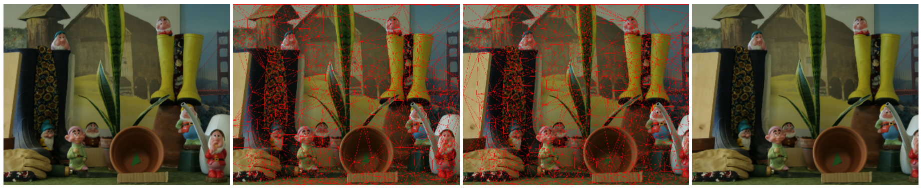
Note that these two state-of-the-art methods have mostly been tailored for and evaluated on small-baseline camera motion sequences, exhibiting small image deformations as typically found in temporal image interpolation problems of nearly static cameras.

In Fig. 2 we illustrate the performance of the proposed sparse approach on a sequence with very large camera rotations. The view interpolations are of high quality and the small distortions in the area of the lamp are barely noticeable. Please see the supplementary video for smooth interpolations and further results on a variety of sequences.

Fig. 2: Uniform transition on the Room dataset between two frames. *Highlighted*: Details of the two input images. Our framework can be applied to any sequence of temporally or spatially related images. The image quality is well preserved despite the large camera rotation except for slight distortions of the lamp.

In Fig. 3 we show an extended version of the 'Dwarves' sequence [18] of Fig. 7 in the paper. Here we further show the computed triangulation of features and the ground truth image along with our results for both dense and sparse approaches.

Fig. 4 illustrates the application of virtual avatars generated from a few images. Especially distortions on human faces are easily detected by human observers. However, our proposed approach yields convincing view interpolations also on this sequence. Please see supplementary video for an interactive illustration.



Source (left), target (right) images along with overlayed triangulated features (middle).

Fig. 3: Extended version of the 'Dwarves' sequence [18] evaluation in the paper (Fig. 7). *Top row:* input images with and without overlayed triangulation. *below, from left to right:* ground truth image, method output, difference to ground truth. *Middle row:* Dense, TV-L1 flow, $e_{RMSE} = 11.19$, *Bottom row:* Sparse, guided matching, $e_{RMSE} = 6.08$

References

1. Ballan, L., Brostow, G.J., Puwein, J., Pollefeys, M.: Unstructured video-based rendering: interactive exploration of casually captured videos. ACM Trans. Graph. 29(4) (2010), `http://dblp.uni-trier.de/db/journals/tog/tog29.html#BallanBPP10`
2. Chen, K., Lorenz, D.A.: Image sequence interpolation based on optical flow, segmentation, and optimal control. IEEE Transactions on Image Processing 21(3), 1020–1030 (2012), `http://dblp.uni-trier.de/db/journals/tip/tip21.html#ChenL12`
3. Debevec, P.: The Campanile Movie. In: SIGGRAPH 97 Electronic Theater (1997), `http://www.debevec.org/Campanile/`, (visited: May 2014)
4. Fehn, C.: Depth-Image-Based Rendering (DIBR), Compression and Transmission for a New Approach on 3D-TV. In: Proceedings of SPIE Stereoscopic Displays and Virtual Reality Systems XI. pp. 93–104 (2004)
5. Fischler, M., Bolles, R.: Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM 24(6), 381–395 (1981)

Fig. 4: Avatar scene. The graphic shows interpolations at $t = 0.5$. *From left to right*: Source image, simple cross-dissolve, Dual TV-L1 based interpolation, target image. Without warping, obvious ghosting artifacts are introduced. Our framework on the other hand produces smooth and plausible transitions.

6. Fragneto, P., Fusiello, A., Rossi, B., Magri, L., Ruffini, M.: Uncalibrated view synthesis with homography interpolation. In: 3DIMPVT. pp. 270–277. IEEE (2012), http://dblp.uni-trier.de/db/conf/3dim/3dimpvt2012.html#FragnetoFRMR12
7. Germann, M., Hornung, A., Keiser, R., Ziegler, R., Würmlin, S., Gross, M.: Articulated billboards for video-based rendering. Comput. Graphics Forum (Proc. Eurographics) 29(2), 585–594 (2010)
8. Goesele, M., Ackermann, J., Fuhrmann, S., Haubold, C., Klowsky, R., Steedly, D., Szeliski, R.: Ambient point clouds for view interpolation. ACM Trans. Graph. 29(4), 95:1–95:6 (Jul 2010), http://doi.acm.org/10.1145/1778765.1778832
9. Hasler, N., Rosenhahn, B., Thormählen, T., Wand, M., Gall, J., Seidel, H.P.: Markerless motion capture with unsynchronized moving cameras. In: CVPR. pp. 224–231 (2009)
10. Inamoto, N., Saito, H.: Free viewpoint video synthesis and presentation from multiple sporting videos. In: ICME. pp. 322–325. IEEE (2005), http://dblp.uni-trier.de/db/conf/icmcs/icme2005.html#InamotoS05
11. Lipski, C.: Virtual Video Camera: a System for Free Viewpoint Video of Arbitrary Dynamic Scenes. Ph.D. thesis, TU Braunschweig (Jun 2013)
12. Lipski, C., Linz, C., Berger, K., Magnor, M.A.: Virtual video camera: image-based viewpoint navigation through space and time. In: SIGGRAPH Posters. ACM (2009), http://dblp.uni-trier.de/db/conf/siggraph/siggraph2009posters.html#LipskiLBM09
13. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. In: International Journal of Computer Vision. vol. 20 (2003)
14. Mahajan, D., Huang, F.C., Matusik, W., Ramamoorthi, R., Belhumeur, P.N.: Moving gradients: a path-based method for plausible image interpolation. ACM Trans. Graph. 28(3) (2009)
15. Morel, J.M., Yu, G.: Asift: A new framework for fully affine invariant image comparison. SIAM J. Imaging Sciences 2(2), 438–469 (2009), http://dblp.uni-trier.de/db/journals/siamis/siamis2.html#MorelY09
16. Muja, M., Lowe, D.G.: Fast approximate nearest neighbors with automatic algorithm configuration. In: Ranchordas, A., Arajo, H. (eds.) VISAPP (1). pp.

331–340. INSTICC Press (2009), http://dblp.uni-trier.de/db/conf/visapp/visapp2009-1.html#MujaL09
17. Replay Technologies Inc.: freeD^{TM} technology (2013), http://replay-technologies.com/, (visited: May 2014)
18. Scharstein, D., Pal, C.: Learning conditional random fields for stereo. In: CVPR (2007)
19. Seitz, S.M., Dyer, C.R.: Physically-valid view synthesis by image interpolation. In: In Proc. IEEE Workshop on Representations of Visual Scenes. pp. 18–25 (1995)
20. Seitz, S.M., Dyer, C.R.: View morphing. In: SIGGRAPH. pp. 21–30 (1996), http://dblp.uni-trier.de/db/conf/siggraph/siggraph1996.html#SeitzD96
21. Snavely, N., Garg, R., Seitz, S.M., Szeliski, R.: Finding paths through the world's photos. ACM Transactions on Graphics (Proceedings of SIGGRAPH 2008) 27(3), 11–21 (2008)
22. Strecha, C., Tuytelaars, T., Gool, L.J.V.: Dense matching of multiple wide-baseline views. In: ICCV. pp. 1194–1201 (2003)
23. Vedula, S., Baker, S., Kanade, T.: Image-based spatio-temporal modeling and view interpolation of dynamic events. ACM Trans. Graph. 24(2), 240–261 (2005)
24. Vlad, A.: Image morphing techniques. JIDEG 5(1) (JUNE 2010), http://www.sorging.ro/ro/member/serveFile/format/pdf/slug/image-morphing-techniques
25. Werlberger, M., Pock, T., Unger, M., Bischof, H.: Optical flow guided tv-l1 video interpolation and restoration. In: Energy Minimization Methods in Computer Vision and Pattern Recognition (2011)
26. Wolberg, G.: Image morphing : A survey. Visual Computer 14, 360–372 (1998), http://ci.nii.ac.jp/naid/80010827845/en/
27. Zhang, Z., Deriche, R., Faugeras, O.D., Luong, Q.T.: A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artif. Intell. 78(1-2), 87–119 (1995), http://dblp.uni-trier.de/db/journals/ai/ai78.html#ZhangDFL95
28. Zitnick, C.L., Kang, S.B., Uyttendaele, M., Winder, S.A.J., Szeliski, R.: High-quality video view interpolation using a layered representation. ACM Trans. Graph. 23(3), 600–608 (2004), http://dblp.uni-trier.de/db/journals/tog/tog23.html#ZitnickKUWS04