Farhan, Erez

Highly accurate matching of weakly localized features. (English) Zbl 1434.68586

SIAM J. Imaging Sci. 12, No. 4, 1833-1863 (2019).

Summary: Matching corresponding local patches between images is a fundamental building block in many computer vision algorithms. It is extensively used in applications like three-dimensional structure estimation, model fitting, superresolution, image retrieval, and image recognition. As opposed to global image registration approaches, local matching reduces the high-dimensional challenge of recovering geometric relations between images to a series of relatively simple and independent tasks, defined in a limited local context. This approach is geometrically very flexible, simple to model, and has clear computational advantages. But it also has two fundamental practical shortcomings: (1) sparsity: the need to rely on repeatable features for matching limits the potential coverage of the scene to highly textured locations; (2) reliability: the limited spatial context in which those methods work often does not contain enough information for achieving reliable matches. These shortcomings also tend to trade off. While highly textured features, such as corners and blobs, often produce reliable matches, they are also relatively rare and thus very sparse in the image. And conversely, more common textures, such as edges or ridges, are largely discarded due to their low reliability for matching. We observe that while classic methods avoided using poorly localized features (e.g., edges) as matching candidates, these features contain highly valuable information for matching. We show how, given the appropriate geometric context, reliable matches can be produced from these features, contributing to a better coverage of the scene, while also producing highly accurate geometric transformation estimation. We present a statistically attractive framework for encoding the uncertainty that stems from using these features into a coupled geometric estimation and match extraction process. We examine the practical application of the proposed framework to the problems of guided matching and affine region expansion and show significant improvement over preceding methods. We compare our approach to state-of-the-art point matching methods and show its attractiveness in a variety of geometrically challenging scenarios.

MSC:

68T45 Machine vision and scene understanding
68U05 Computer graphics; computational geometry (digital and algorithmic aspects)
68U10 Computing methodologies for image processing

Keywords:

image registration; geometric context; match expansion; low-texture matching; transformation estimation; covariance propagation; image matching; local matching; affine transformation; perspective transformation; guided matching; geometric feedback; match verification

Software:

TILDE; SURF; EpicFlow; SIFT; ORB-SLAM; Vlfeat; LF-Net; HPatches; ORB-SLAM2

Full Text: DOI

References:

[1] V. Balntas, K. Lenc, A. Vedaldi, and K. Mikolajczyk, Hpatches: A benchmark and evaluation of handcrafted and learned local descriptors, in CVPR, IEEE, Piscataway, NJ, 2017, pp. 3852-3861.

[2] C. Barnes, E. Shechtman, D. B. Goldman, and A. Finkelstein, The generalized patchmatch correspondence algorithm, in European Conference on Computer Vision, Springer, Berlin, 2010, pp. 29-43.

[3] H. Bay, T. Tuytelaars, and L. Van Gool, Surf: Speeded up robust features, in Computer Vision-ECCV 2006, Springer, Berlin, 2006, pp. 404-417.

[4] J. Bentolila and J. M. Francos, Homography and fundamental matrix estimation from region matches using an affine error metric, J. Math. Imaging Vision, 49 (2014), pp. 481-491. · Zbl 1361.68293

[5] A. Bódis-Szomorú, H. Riemenschneider, and L. Van Gool, Fast, approximate piecewise-planar modeling based on sparse structure-from-motion and superpixels, in Proceedings CVPR 2014, 2014, pp. 469-476.

[6] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, Brief: Binary robust independent elementary features, in Computer Vision-ECCV 2010, Springer, Berlin, 2010, pp. 778-792.

[7] E. Farhan and R. Hagege, Point matching via affine region expansion, in Image Processing, ICIP, IEEE, Piscataway, NJ, 2014, pp. 383-387.

[8] E. Farhan and R. Hagege, Geometric expansion for local feature analysis and matching, SIAM J. Imaging Sci., 8 (2015), pp. 2771-2813. · Zbl 1352.68261

[9] E. Farhan, E. Meir, and R. Hagege, Local region expansion: A method for analyzing and refining image matches, IPOL J. Image Process. Online, 7 (2017), pp. 386-398.

[10] M. A. Fischler and R. C. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Comm. ACM, 24 (1981), pp. 381-395.

[11] W. Förstner, Uncertainty and projective geometry, in Handbook of Geometric Computing, Springer, 2005, pp. 493-534.

[12] W. Förstner, Minimal representations for uncertainty and estimation in projective spaces, in Asian Conference on Computer Vision, Springer, 2010, pp. 619-632.

[13] D. Fortun, P. Bouthemy, and C. Kervrann, Optical flow modeling and computation: a survey, Computer Vision and Image Understanding, 134 (2015), pp. 1-21. · Zbl 1410.94003

[14] D. Fortun, P. Bouthemy, and C. Kervrann, Aggregation of local parametric candidates with exemplar-based occlusion handling for optical flow, Computer Vision and Image Understanding, 145 (2016), pp. 81-94. · Zbl 1410.94003

[15] Y. HaCohen, E. Shechtman, D. B. Goldman, and D. Lischinski, Non-rigid dense correspondence with applications for image enhancement, ACM Trans. Graphics (TOG), 30 (2011), 70.

[16] C. Harris and M. Stephens, A combined corner and edge detector, in Alvey Vision Conference, Vol. 15, Manchester, UK, Butterfield, London, 1988, p. 50.

[17] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, Cambridge, 2003. · Zbl 0956.68149

[18] K. Lenc and A. Vedaldi, Learning covariant feature detectors, in European Conference on Computer Vision, Springer, Cham, Switzerland, 2016, pp. 100-117.

[19] D. G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis., 60 (2004), pp. 91-110.

[20] J. Matas, O. Chum, M. Urban, and T. Pajdla, Robust wide-baseline stereo from maximally stable extremal regions, Image Vis. Comput., 22 (2004), pp. 761-767.

[21] N. Mayer, E. Ilg, P. Hausser, P. Fischer, D. Cremers, A. Dosovitskiy, and T. Brox, A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, IEEE, Piscataway, NJ, pp. 4040-4048.

[22] J. Meidow, C. Beder, and W. Förstner, Reasoning with uncertain points, straight lines, and straight line segments in 2d, ISPRS J. Photogram. Remote Sensing, 64 (2009), pp. 125-139.

[23] K. Mikolajczyk and C. Schmid, Scale and affine invariant interest point detectors, Int. J. Comput. Vis., 60 (2004), pp. 63-86.

[24] D. Mishkin, F. Radenovic, and J. Matas, Repeatability is not enough: Learning affine regions via discriminability, in Proceedings of the European Conference on Computer Vision (ECCV), Cham, Switzerland, 2018, pp. 284-300.

[25] R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, Orb-slam: A versatile and accurate monocular slam system, IEEE Trans. Robot., 31 (2015), pp. 1147-1163.

[26] B. Ochoa and S. Belongie, Covariance propagation for guided matching, in Proceedings of the Workshop on Statistical Methods in Multi-Image and Video Processing (SMVP), Vol. 83, 2006.

[27] Y. Ono, E. Trulls, P. Fua, and K. M. Yi, Lf-net: learning local features from images, in Advances in Neural Information Processing Systems, Curran Associates, Red Hook, NY, 2018, pp. 6234-6244.

[28] J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid, Epicflow: Edge-preserving interpolation of correspondences for optical flow, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ, 2015, pp. 1164-1172.

[29] J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid, Deepmatching: Hierarchical deformable dense matching, Int. J. Comput. Vis., 120 (2016), pp. 300-323.

[30] C. Tomasi and J. Shi, Good features to track, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1994, IEEE, Piscataway, NJ, pp. 593-600.

[31] B. Triggs, Joint feature distributions for image correspondence, in Proceedings of the Eighth IEEE International Conference on Computer Vision, vol. 2, IEEE, Computer Society, Los Alamitos, CA, 2001, pp. 201-208.

[32] A. Vedaldi and B. Fulkerson, Vlfeat: An open and portable library of computer vision algorithms, in Proceedings of the 18th ACM international Conference on Multimedia, MM '10, New York, ACM, New York, 2010, pp. 1469-1472, https://doi.org/10.1145/1873951.1874249

[33] Y. Verdie, K. Yi, P. Fua, and V. Lepetit, Tilde: a temporally invariant learned detector, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ, 2015, pp. 5279-5288.

[34] T. Whelan, S. Leutenegger, R. Salas-Moreno, B. Glocker, and A. Davison, Elasticfusion: Dense slam without a pose graph, Robotics: Science and Systems, 2015.

[35] C. Wu, Towards linear-time incremental structure from motion, in 3D Vision-3DV 2013, IEEE, Piscataway, NJ, 2013, pp. 127-134.

[36] J. Yang and H. Li, Dense, accurate optical flow estimation with piecewise parametric model, in Proceedings of the IEEE

Conference on Computer Vision and Pattern Recognition, IEEE, Piscataway, NJ, 2015, pp. 1019-1027.

[37] J. Yang, J. Wright, T. S. Huang, and Y. Ma, Image super-resolution via sparse representation, IEEE Trans. Image Process., 19 (2010), pp. 2861-2873. · Zbl 1371.94429

[38] K. M. Yi, E. Trulls, V. Lepetit, and P. Fua, Lift: Learned invariant feature transform, in European Conference on Computer Vision, Springer, Cham, Switzerland, 2016, pp. 467-483.