Learning Correspondence View with Support Vector Machine

Xiangru Li
Shandong University of Science and Technology, Southern Medical University
Xiaoming Li
Shanxi University
Huarong Xu
Xiamen University

Abstract

Correspondence view (CV) is recently introduced for rejecting outliers in computer vision. The fundamental idea of CV is that, for given two images of a scene, the corresponding points constitute a manifold in joint-image space R^4, and outliers can be detected by checking whether they are consistent with the upward views of the manifold. This work studies CV learning and outliers rejecting by Support Vector Machine. Experiments on real image pairs demonstrate the excellent performance of our proposed SVM+CV learning method and its superiority over the available robust methods in literature, especially the widely used RANSAC.

1. Introduction

Finding reliable correspondence points is a fundamental problem in computer vision [1][2]. For given images, correspondence points are the projection of the same point in the scene. The putative correspondences are usually established by matching the interest points with local information, for example, the intensity in a small region around the interest points, or some kind of local descriptor [3][4]. However, usually a large proportion of the putative correspondences are outliers due to viewpoint change, occlusion, local ambiguousness, etc. And the outliers are usually enough to ruin the traditional estimation methods. Therefore, much of the endeavor in computer vision community is to overcome or alleviate this problem by rejecting the outliers, and many methods are introduced, for example, RANSAC (RANdom SAmple Consensus) [5][6], LMedS (Least Median of Squares) [7][8], MLESAC [9][10].

Our interest here is in rejecting outliers by learning Correspondence view (CV)[11]. For given two images I and I' of a scene, fundamental idea of CV is that corresponding points between them constitute a manifold $M_c : F(p, p') = 0$ in joint-image space R^4, where $p \in I$ and $p' \in \mathrm{I}'$. CVs are two upward views f and f' of M_c, which are the scenerio when we stand on I and I' respectively to observe M_c. The important is that correct corresponding points should be consistent with at least one of the two CVs and outliers can be detected by checking whether they are not consistent with both CVs. This work studies CV learning and outliers rejecting by support vector machine. Experiments on real image pairs demonstrate the excellent performance of our proposed SVM-CV learning method and its superiority over the available robust methods in literature, especially the widely used RANSAC.

The remainder of this paper is organized as follows: in the next section, the related works on correspondence view (CV) will be briefly reviewed. In section 3, we give our method to learn CV with support vector machine. In section 4, we study the performance of our proposed method experimentally. Conclusion is made in section 5.

2. Correspondence view

Given two images $\mathrm{I} : U \times V$ and $\mathrm{I}' : U' \times V'$. Correspondence points are the point pair in two images that are the projection of same scene point. Based on the theory of correspondence manifold, correspondence points between two images lie on a manifold M_c (correspondence manifold [11]) in joint image space R^4 [12]. Correspondence view

$$f(p) = (g_1(p), g_2(p)) = p' \quad (1)$$

and

$$f'(p') = (g_1'(p'), g_2'(p')) = p \quad (2)$$

are the upward views of correspondence manifold M_c, where $p \in \mathrm{I}$ and $p' \in \mathrm{I}'$. Correct

978-0-7695-3571-5/09 $25.00 © 2009 IEEE
DOI 10.1109/GCIS.2009.402

corresponding points should satisfy at least one of the two views f and f'. Mismatches can be rejected by checking whether they are consistent with the correspondence view f or f', and the consistency, for example a putative correspondence $(p, p') \in \mathrm{I} \times \mathrm{I}'$ with correspondence view f, can be defined as: if $p \in \mathrm{I}$ and $p' \in \mathrm{I}'$ satisfy g_1 and g_2, then we say (p, p') is consistent with f, and the consistency of $(p, p') \in \mathrm{I} \times \mathrm{I}'$ with correspondence view f' can be defined similarly. However, we do not know real $\{g_i, g_i', i = 1,2\}$ in practice. Therefore, the key is how to estimate them in applications.

3. Learning correspondence view by support vector machine

Support Vector Machine (SVM) is a typical method for estimation. The fundamentals of this method are Structural Risk Minimization (SRM) and kernel technique [13]. By the SRM principle, SVM balances generalization performance and empirical error. By the kernel technique, SVM implicitly project observations from given input space X into a high dimensional space F, and solve the original nonlinear regression problem linearly in F. SVM has been studied widely and successfully in theory and applications [13][14].

Given a set of putative point correspondences

$$S = \{(p_i, p_i') = (u_i, v_i, u_i', v_i'), i = 1, \cdots, n\} \subset \mathrm{I} \times \mathrm{I}' : U \times V \times U' \times V'. \quad (3)$$

If a given correspondence point pair

$$(p, p') = (\mathrm{u}, \mathrm{v}, u', v') \in \mathrm{I} \times \mathrm{I}'$$

satisfies $f = (g_1, g_2)$, then its projection $(\mathrm{u}, \mathrm{v}, u')$ on space $U \times V \times U'$ is consistent with g_1. Therefore, the projection

$$S_{U \times V \times U'} = \{(u, v, u') | (u, v, u', v') \in S\} \quad (4)$$

of S can be regarded as a sample set from $u' = g_1(u, v)$. In this work, we will estimate $u' = g_1(u, v)$ from $S_{U \times V \times U'}$ with SVM. Similarly, we can define

$$S_{U \times V \times V'} = \{(u, v, v') | (u, v, u', v') \in S\}, \quad (5)$$

$$S_{U \times U' \times V'} = \{(u, u', v') | (u, v, u', v') \in S\}, \quad (6)$$

$$S_{V \times U' \times V'} = \{(\mathrm{v}, u', v') | (u, v, u', v') \in S\}, \quad (7)$$

and estimate $v' = g_2(u, v)$, $u = g_1(u', v')$ and $v = g_2(u', v')$ from $S_{U \times V \times V'}$, $S_{U \times U' \times V'}$ and $S_{V \times U' \times V'}$ respectively with SVM.

4. Experiments

In this section, we study the performance of our proposed method by real image pairs. The SVM is implemented by LIBSVM [16].

There are two images of a relievo in Figure 1 a), and we want to establish point correspondences between them. The putative correspondences S in Figure 1 b) are computed from the SIFT keypoints by Nearest Neighbor method [4]. Due to local ambiguousness, there are many mismatches in S, approximately 79.61%. Therefore, the corresponding projection $S_{U \times V \times U'}$, $S_{U \times V \times V'}$, $S_{U \times U' \times V'}$ and $S_{V \times U' \times V'}$ are all contaminated with a larger percentage of outliers. We estimated $u' = g_1(u, v)$, $v' = g_2(u, v)$, $u = g_1'(u', v')$ and $v = g_2'(u', v')$ from them with SVM and reject the mismatches from S by checking consistency of the putative correspondences with the correspondence views, the consistent putative correspondences and the inconsistent ones are presented in Figure 1 c) and d) respectively. The experimental results show that SVM can be successfully used in correspondence view learning for mismatch rejecting.

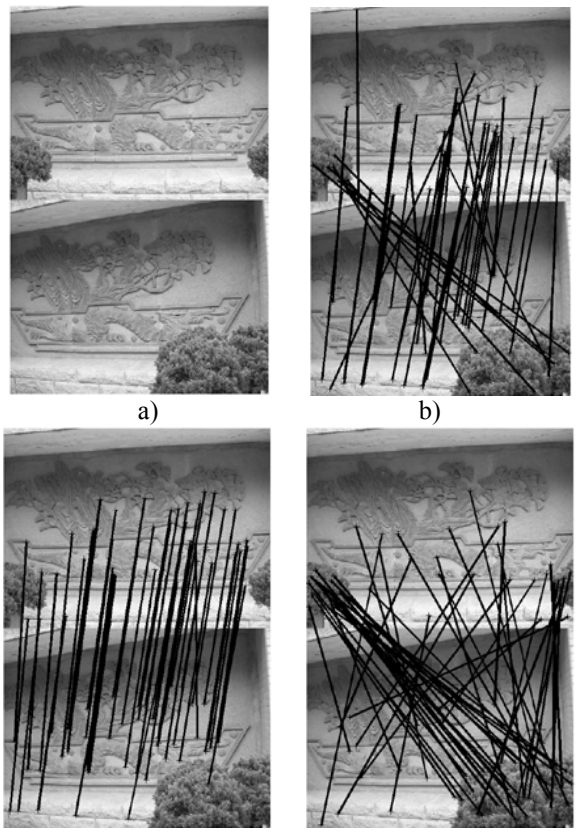
In this work, we compared our proposed SVM-CV learning method with RANSAC on rejecting outliers. We choose RANSAC as the benchmark in this comparison, because it is the most popular method in computer vision community till now [1][15]. Since the coordinates of putative corresponding points are usually corrupted by noise, the observed corresponding points usually do not strictly satisfy the estimated correspondence view in SVM-CV, and fundamental matrix in RANSAC. Therefore, a tolerance parameter is needed in both methods, and for convenience we label them α_{CV} and α_{RA} respectively. To be comparable, the tolerance parameters are chosen such that same number of point pairs can be preserved as suspect correct matches in two experiments, and the results are presented in Figure 1, Figure 2 and Table 1.

Table 1. **Comparing the proposed SVM+CV with RANSAC on efficiency. Putative correspondence sets are those in the**

experiments of Figure 1. Second is the used time unit in this table.

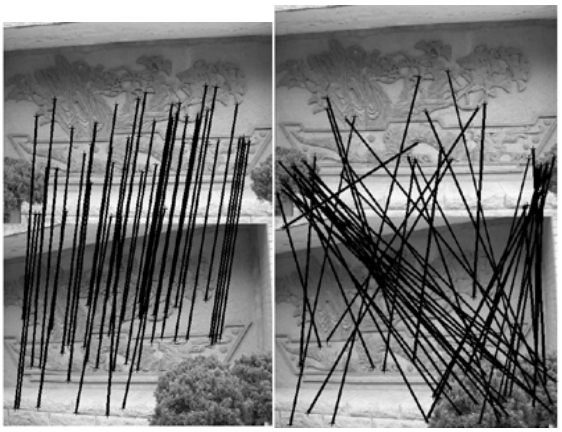
Methods	Time
SVM+CV	0.375
RANSAC	2.593

The experimental results in Figure 1 and Figure 2 show that, although RANSAC is slightly more accurate than the proposed SVM-CV learning method, the performance of SVM+CV is still very good and the two methods are comparable on accuracy. The most important is that it only takes SVM-CV 0.375 seconds to reject the outliers, and RANSAC 2.593 seconds. Therefore, our proposed SVM-CV learning method is much more efficient than RANSAC.



a) b) c) d)

Figure 1. **An image pair of a relievo: refining putative correspondences by learning correspondence view with SVM. a) original image pair; b) 399 putative correspondences with 174 outliers(43.61%); c) 210 identified matches (outliers 3.33%); e) the identified mismatches. For visibility concern, only 50 randomly selected point pairs are presented in b), c), d).**



a) b)

Figure 2. **Refining putative correspondences by RANSAC. a) the identified matches (mismatches is reduced to 0.95%); b) 210 identified mismatches. For visibility concern, only 50 randomly selected point pairs are presented in a), b).**

5. Conclusion

Outlier rejecting is an important problem in computer vision. And correspondence view (CV) is a recently introduced concept for rejecting outliers. We studies CV learning and outliers rejecting by support vector machine. Experiments on real image pairs demonstrate the excellent performance of our proposed SVM-CV learning method and its superiority over the 'state-of-the-art' RANSAC.

Acknowledgments

This work is supported by the Natural Science Foundation of Shandong Province under grant No. Q2007G02, National Natural Science Foundation of China under the grant No. 60773132, and Opening Task-fund for National Laboratory of Pattern Recognition.

References

[1] R. Hartley, A. Zisserman. Multiple View Geometry in Computer Vision, Second edition. Cambridge University Press, 2003.

[2] M.Z. Brown, D. Burschka, G.D. Hager. Advances in Computational Stereo. IEEE Trans on Pattern Analysis and Machine Intelligence, 25(8): 993-1008,2003.

[3] M. Tico, C. Rusu, P. Kuosmanen. A geometric invariant representation for the identification of corresponding points. In Proceedings of International Conference on Image Processing, 2:462-466,1999.

[4] D.G. Lowe. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 60(2): 91-110, 2004.

[5] M.A. Fischler, R.C. Bolles. Random sample consensus: a paradigm for model fitting with application to image analysis and automated cartography. Communication Association and Computing Machine, 24(6): 381-395, 1981.

[6] R. Subbarao, P. Meer. Beyond RANSAC: user independent robust regression. In: Workshop on 25 Years of RANSAC (in conjunction with CVPR'06), New York, NY, pp. 101-108, 2006.

[7] P J. Rousseeuw. Least median of squares regression. Journal of the American Statistical Association, 79:8 71–880, 1984.

[8] Z. Zhang, R. Deriche, O. Faugeras, and Q. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar Geometry. Artificial Intelligent, 78(1-2): 87-119,1995.10.

[9]P.H.S. Torr and A. Zisserman. MLESAC: A new robust estimator with application to estimating image geometry. Computer Vision and Image Understanding, 78:138–156, 2000.

[10]B. Tordoff, D.W. Murray. Guided-MLESAC: Faster Image Transform Estimation by Using Matching Priors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(10): 1523-1535,2005.

[11]X.R. Li, X.M. Li, H.L. Li, M.Y. Cao. Rejecting Outliers based on Correspondence Manifold. Acta Automatica Sinica, In Press.

[12]W. Triggs. The geometry of projective reconstruction I: Matching constraints and the joint image. In Proceeding of international conference on computer vision, pp.338-343, 1995.

[13]V.N. Vapnik. The Nature of Statistical Learning Theory, 2nd ed. Springer-Verlag, New York, 2000.

[14]A.J. Smola, B. Schölkopf. A tutorial on support vector regression. Statistics and computing, 14(3): 199-222, 2004.8.

[15]P.H.S. Torr and D.W. Murray. Outlier detection and motion segmentation. In Sensor Fusion VI, SPIE Vol. 2059, P. Schenker (Ed.), Boston, pp. 432–443,1993.

[16]C.C. Chang and C.J. Lin. LIBSVM: a library for support vector machines, 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm.