TELKOMNIKA, Vol.11, No.3, September 2013, pp. 597~602
ISSN: 1693-6930, accredited A by DIKTI, Decree No: 58/DIKTI/Kep/2013
DOI: 10.12928/TELKOMNIKA.v11i3.1042

Wide Baseline Matching Using Support Vector Regression

Haifeng Xi, Weihua Cui*
Chongqing University of Posts and Telecommunications
No.2 ChongWen Road,NanAn District, Chongqing 400065, China
*Corresponding author, e-mail: elenacui123@163.com

Abstrak

Pada tulisan ini kami membahas metode baru untuk menyelesaikan pencocokan dasar luas dengan menggunakan regresi dukungan vektor (SVR). Tingginya rasio pertandingan awal yang benar digunakan untuk melatih hubungan SVR, yang diperoleh dengan cara mencocokkan fitur SIFT skala besar dan membuang beberapa ketidaksesuaian dengan perbaikan skema penyaringan topologi yang kami tawarkan, dan pertandingan baru akan dicarikan yang paling dekat nilai prediksi yang diberikan oleh hubungan SVR yang terbaik. Dalam kasus ini, baik indoor maupun outdoor pada lingkungan yang dipasang gambar di bawah kondisi baseline pengujian, hasil eksperimen menunjukkan bahwa algoritma kami secara otomatis mendapatkan sejumlah besar titik korespondensi yang lebih akurat.

Kata kunci*: dasar lebar pencocokan, penyaringan topologi, dukungan regresi vektor.*

Abstract

In this paper, we newly solve wide baseline matching using support vector regression (SVR). High correct ratio initial matches are used to train SVR relationships, obtained by matching large-scale SIFT features and discarding some mismatches by our improved topological filtering scheme; and new matches are searched near the prediction given by trained SVR relationships. Both indoor and outdoor environments image pairs under wide baseline condition are tested, experiment results show that our algorithm automatically gain large numbers of accurate point correspondences.

Key words*: wide baseline matching; topological filtering; support vector regression*

1. Introduction

Finding correspondences in a set of images of a scene is a fundamental part of most computer vision tasks, such as camera self-calibration, registration, structure and motion recovery, object recognition and so on[1].

Most of current wide baseline matching approaches obey this general scheme: 1. detect interest points or distinguish regions and extract their local features; 2.initial features matching; 3.discarding mismatches and expanding correct matches iteratively or not;4.getting final matches and geometry relations.The third step:Discarding mismatches and extending correct matches is the most challenging work in the current study about wide baseline matching approaches.RANdom SAmple Consensus(RANSAC) was applied to estimate global geometry relationship--the fundamental matrix or homography between image pairs[2], outliers are eliminated, and the inliers are used to refine estimation, the refine estimation derived from inliers can guide matching so as to extend matches. Some works developing new robust estimators were reported[7]. However, the initial matches are possibly sparse and with high ratio false matches, too early to get the global geometry relationship estimation or even bring on a bad result. Ferrari et al.[3]refined each initial match by looking for the local affine transformation that optimized the similarity, and the matches with refined similarity above a similarity threshold were kept. New matches were searched guided by local affine transformations, and a sidedness topological constraint for triple matches was used to remove bad matches.Tuytelaars and Gool gave two local constraints[5], and the matches with sufficient other matches obeying the same local constraints were reserved. Steele and Egbert [9] expanded the matches guided by local similarity transformation estimated previously, and constrained the matches by epipolar geometry with RANSAC estimation using a high inlier error tolerance.

Received March 11, 2013; Revised July 10, 2013; Accepted July 25, 2013

Using global geometric relationships-the basic matrix [2] or local geometric relations-local affine transformation to guide the match[2,9]. However,the fundamental matrix just provides a bilinear constraint, if there is no adequate match, then it can not be a better estimate.Local geometric relations can provide prediction image, but it is very sensitive. As support vector regression (SVR) can construct good classification rules in high dimensional space, this paper proposes a Wide baseline matching algorithm based on SVR , and using an improved topology filter to weed out false match,by the experimental simulation, trying to solve a wide baseline image pairs matching problem.

We present a novel wide baseline matching algorithm using support vector regression (SVR) in Section 2 and Section 3. As a part of this algorithm, an improved topological filter has been invented for discarding mismatches, called Random Planes Topological Filter (RPTF) in Section 4. Experiment results of indoor and outdoor image pairs are shown in Section 5, and the paper is concluded in Section 6 finally.

2. Support Vector Regression (SVR)

The support vector machine (SVM) invented by Vapnik [12]and his research groups is a popular classification approach.It minimizes the upper bound of the generalization error. Furthermore, with the introduction of Vapnik's $\varepsilon-$insensitive loss function, the SVM is extended to solve nonlinear regression estimation problem, called support vector regression (SVR). Here we describe the standard SVR algorithm:

Let $(\vec{x}_1, y_1)\ldots(\vec{x}_i, y_i)\ldots(\vec{x}_l, y_l)$ be the training data points, where $\vec{x}_i \in \Re^n$ and $y_i \in \Re$. The SVR maps the input vector $\vec{x}$ into a high-dimensional feature space F by $\vec{z} = \Phi(\vec{x})$. In this new space, we consider a linear function, which corresponds to a nonlinear one in the original space:

$$f(\vec{x}, \vec{\beta}) = \langle \vec{z}, \vec{\omega} \rangle + b = \sum_{i=1}^{l} \beta_i \langle \vec{z}, z_i \rangle + b = \sum_{i=1}^{l} \beta_i \langle \Phi(\vec{x}), \Phi(\vec{x}_i) \rangle + b = \sum_{i=1}^{l} \beta_i K(\vec{x}, \vec{x}_i) + b \quad (1)$$

where $\langle \cdot, \cdot \rangle$ is the inner product, and $K(\cdot, \cdot)$ is the kernel function. The goal is to find the linear function $\langle \vec{z}, \vec{\omega} \rangle + b$ that minimizes the empirical risk function (2) in feature space

$$R_{emp} = \frac{1}{l}\sum_{i=1}^{l} L_\varepsilon(y_i - \langle \vec{z}_i, \vec{\omega} \rangle - b) = \frac{1}{l}\sum_{i=1}^{l} L_\varepsilon(y_i - f(\vec{x}_i, \vec{\beta})) \quad (2)$$

constrained to

$$\langle \vec{\omega}, \vec{\omega} \rangle \le \gamma \quad (3)$$

where γ denotes the regularization constant and $L_\varepsilon(\cdot)$ represents the $\varepsilon-$insensitiveloss function, and is defined as

$$L_\varepsilon(e) = \begin{cases} 0, & for|e| \le \varepsilon \\ |e| - \varepsilon, & otherwise. \end{cases} \quad (4)$$

By using the Lagrange multiplier method minimizing(2)yields the following dual-optimization problem; minimize

$$\psi(\alpha, \alpha^*) = \varepsilon \sum_{i=1}^{l} (\alpha_i + \alpha_i^*) - \sum_{j=1}^{l} y_j (\alpha_j^* - \alpha_j) + \frac{1}{2} \sum_{i,j=1}^{l} (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(\vec{x}_i, \vec{x}_j) \quad (5)$$

subjected to the constraint

$$\sum_{i=1}^{l} \alpha_i = \sum_{i=1}^{l} \alpha_i^*, 0 \le \alpha_i, \alpha_i^* \le \gamma, i = 1,\ldots, l \quad (6)$$

The final solution is:

$$f(\vec{x}) = \sum_{i=1}^{l} \beta_i K(\vec{x}, \vec{x}_i) + b, \quad (7)$$

$$\beta_i = \alpha_i^* - \alpha_i,\ i = 1, \ldots, l, \quad (8)$$

$$b = \frac{1}{l}\sum_{k=1}^{l}\left(\varepsilon.sign(\alpha_k^* - \alpha_k) + y_k - \sum_{i=1}^{l} (\alpha_i^* - \alpha_i) K(\vec{x}_k, \vec{x}_i)\right). \quad (9)$$

Only some of $\left(\alpha_k^* - \alpha_k\right)$ are not equal to 0, the corresponding vectors $\vec{x}_k$ are called Support Vectors. Usually, the kernel function is polynomials $K(\vec{x},\vec{y}) = \left(\langle \vec{x},\vec{y}\rangle + 1\right)^d$ or radial basis functions (RBF) $K(\vec{x},\vec{y}) = \exp\left\{-\frac{\|\vec{x}-\vec{y}\|^2}{2\sigma^2}\right\}$.

3. Wide baseline matching using SVR

We change the problem of expanding accurate matches to a prediction problem: based on the available matches with noise, for unmatched points in one image, we want to get correspondence pixel coordinate predictions in the other image[10], and then search matches near the prediction. This prediction can be derived by SVR as following description:

Suppose $\begin{pmatrix} u \\ v \end{pmatrix}$, $\begin{pmatrix} u' \\ v' \end{pmatrix}$ are pixel coordinates of the image pairs, and their correspondence relationship can be represented as:

$$\begin{cases} u' = u'(u,v) \\ v' = v'(u,v) \end{cases}. \tag{10}$$

Suppose initial matches are: $\left(u_i, v_i, u_i', v_i'\right), i = 1,\ldots,n$, using the training data $\left(u_i, v_i; u_i'\right), \left(u_i, v_i; v_i'\right),\ i = 1,\ldots,n$ for SVR, we get:

$$\begin{cases} u'(u,v) = \sum_{i=1}^{l} \beta_i^{u'} K\left(u,v;u_i,v_i\right) + b^{u'} \\ v'(u,v) = \sum_{i=1}^{l} \beta_i^{v'} K\left(u,v;u_i,v_i\right) + b^{v'} \end{cases} \tag{11}$$

Because $\varepsilon-$insensitive loss function is used, we should search for the new match in a neighborhood of the prediction in the next image. This algorithm is summarized as follows:

Step 1: Detect interest points or distinguish regions and extract their local features in both image pair. Step 2: Get initial feature matches by nearest neighbor based matching strategy. Step 3: Execute (a)-(c): (a) Discard some mismatches, (b) Use current matches for SVR to get relationship (11), (c) Search new matches near the prediction given by relationship (11). Step 4: Estimate the fundamental matrix or homography between the image pair by the normalized eight-point method .

4. Discarding mismatches

The key to use SVR to expand correct matches is that we should obtain high correct ratio initial matches as training data. Ferrari et al. [3] invented a topological filter for removing mismatches.

Considering three points (X_1, X_2, X_3) in one image and their correspondence points $\left(Y_1, Y_2, Y_3\right)$ in the other image, the function

$$side\left(X_1, X_2, X_3\right) = sign\left(\left(X_2 \times X_3\right)\cdot X_1\right) \tag{12}$$

takes value –1 if X_1 on the right side of the directed line from X_2 to X_3, or value 1 if it's on the left side. The sidedness constraint

$$side\left(X_1, X_2, X_3\right) = side\left(Y_1, Y_2, Y_3\right) \tag{13}$$

holds for all the coplanar points.

A triple match (such as $\{(X_1, Y_1), (X_2, Y_2), (X_3, Y_3)\}$) including one, or more mismatches has higher chances to violate equation (13). Based on this assumption, Ferrari et al.[3] design a voting scheme called topological filter for removing mismatches:

for each match $X_i \leftrightarrow Y_i$, computing

$$h(i) = \frac{1}{(n-1)(n-2)} \sum_{j,k \neq i,\, j>k} \left| side\left(X_i, X_j, X_k\right) - side\left(Y_i, Y_j, Y_k\right) \right| . \qquad (14)$$

Obviously, $0 \leq h(i) \leq 1$. At each iterative step, remove the $\arg\max_i h(i)$ match and update $h(i)$ of the reserved matches, until $h(i)$ of each match is lower than a given threshold t_{topo}.

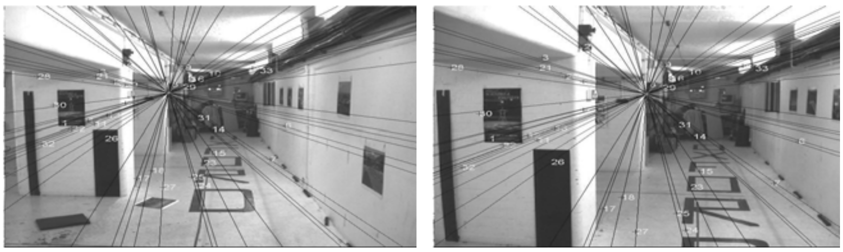
Actually, all of the detected interesting points or distinguishing regions can't be coplanar; otherwise the fundamental matrix estimation is a degenerate problem [12]. The threshold t_{topo} is actually determined by both camera motion and scene structure, so setting an appropriate value for it is a very hard problem. However, there is no need to use this topological filter on the whole match set, if dividing the match set into subsets consisting of coplanar matches at first, we can use this filter simply with $t_{topo} = 0$ on each subset [11]. When the camera has translation between image pair, coplanar points and their correspondence points obey the homography relationship [12]. Considering these, we iteratively estimated some homographies by four randomly sampled matches and take their inliers as coplanar match subset. It's a bit different from the RANSAC method introduced in [12]: at each iterative step, we only searched the homography that has adequate number (more than six for example) of inliers rather than the homography that has largest number of inliers as RANSAC, so the method is efficient. We called this improved topological filter "Random Planes Topological Filter" (RPTF).

5. Experiments

We tested our algorithm on images acquired from a variety of indoor and outdoor environments, such as: Lab-desk (Fig.1), Corridor (Fig.2), Gloriette (Fig.3), House (Fig.4), and Park (Fig.5). The camera underwent significant translation or ration for each image pair, and illumination changed for some pairs. Repeating things like books or leaves made matching more difficult.

(a) (b)

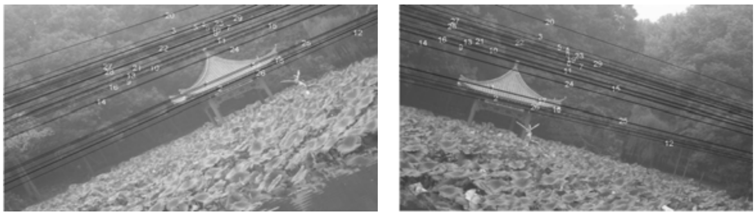
Figure 1. Test Bech Map



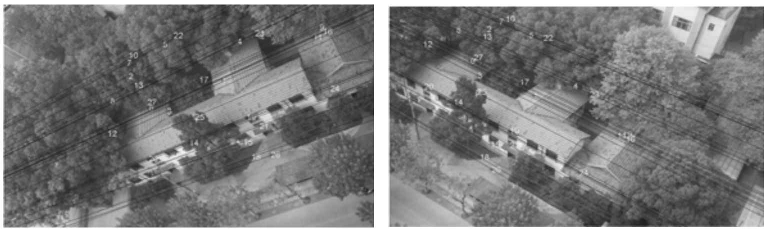
(a) (b)

Figure 2. corridor map

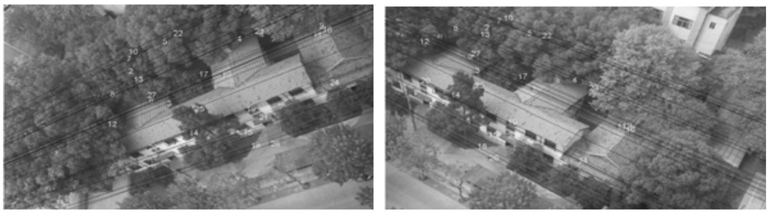
For every image pair, we extracted SIFT features[4],There were about 5000 to 8000 from an image, showed in Table 1. However, we just used large-scale features extracted from the six coarsest layers to get initial matches for the accuracy and the efficiency.



(a) (b)
Figure 3. pergola map



(a) (b)
Figure 4. house map



(a) (b)
Figure 5. House Map

Table 1.Features extracted from image pairs and for initial matching

Image pairs for test	Num of SIFT features extracted	Num of layers	Num of features for initial matching	Num of initial matches
Lab-desk	5474×5813	8	814×800	57
Corridor	4953×4432	9	322×313	79
Gloriette	7944×7743	8	1133×1075	89
House	8050×8103	8	980×1010	107
Park	6168×6950	8	908×1036	105

We set similarity threshold as 0.1 and the ratio threshold 0.7.After discarding some mismatches by RPTF, we trained the SVR to get the relationship (2). Radial basis functions (RBF) was selected as the kernel function for its local property, and parameters were $\sigma = \max_{i=1}^{n} \min_{j=1, j\neq i}^{n} dist\left(\begin{pmatrix}u_i\\ v_i\end{pmatrix},\begin{pmatrix}u_j\\ v_j\end{pmatrix}\right)$, $\varepsilon = 2$, $\gamma = 10^5 \varepsilon$. The obtained SVR relationship (2) was used to guild the searching for new matches,The search widow radius is3ε, and the similarity threshold is 0.8. After the expansion step, the fundament matrix was estimated automatically using RANSAC and inliers were reserved as the final matches.

To measure the accuracy of the matches,we handpicked some matches in each image pair to compute an accurate fundamental matrix F (manual F in Table 2), and computed the average match distances ($d_{\perp}$ in Table 2) from their respective epipolar lines against this manual F for both handpicked match sets and final match sets automatically acquired by our algorithm.We also computed the average match distances from their respective epipolar lines against the automatically estimated F (auto F in Table 2) for final matches. From Table 1 and Table 2, we can see that final matches were expanded about 7 to 10 times of initial matches, while $d_{\perp}$ of final matches against manual F were less than 2 times of handpicked matches. The time performance of the proposed approach was presented in Table 3. The algorithm was tested on the matlab 7.9 platform with Intel CPU 2.4GHz. From Table 3, one can see that, the training time of SVR is a small part of the running time of the whole algorithm.

Table 2. Matching expansion result and evaluation

Image pairs for test	Num of initial matches	Num of final matches	Manual matches $d_{\perp}$ (pixel) against manual F	Final matches $d_{\perp}$ (pixel) against manual F	Final matches $d_{\perp}$ (pixel) against auto F
Lab-desk	57	604	0.6738	1.3369	1.1997
Corridor	79	828	1.0041	1.1900	1.2004
Gloriette	89	673	0.6374	0.7273	0.6572
House	107	999	0.6700	0.5186	0.5458
Park	105	797	0.8723	0.8184	0.5825

Table 3. The time performance of the proposed approach

Image pairs for test	Size of image	Training time of SVR (second)	Running time of the whole algorithm(second)
Lab-desk	430-by-645	2.9	94.2
Corridor	512-by-512	10.5	112.1
Gloriette	430-by-645	1.6	73.3
House	430-by-645	3.0	91.5
Park	430-by-645	1.9	64.4

6. Conclusion

The paper demonstrated how to use SVR to accomplish wide baseline matching. As a part of this novel algorithm, we invented an improved topological filter for discarding mismatches. And Experiment results showed that our algorithm can automatically obtain large numbers of accurate matches for indoor or outdoor image pairs under wide baseline condition.

Essentially, when using SVR, the selection of kernel function is crucial. Finding or designing some kernel functions, which have better learning and generalization capability for the relationship, is a valuable work as further research.

Acknowledgments

The National Natural Science Foundation of China (No.61102131)

References

[1] Yue Zhao,Jianchong Lei.A Camera Self-Calibration Method Based on Plane Lattice and Orthogonality. *TELKOMNIKA,*2013,11(4).
[2] Hartley, A.Zisserman. Multiple View Geometry in Computer Vision.Cambridge University Press.2000.
[3] Ferrari.V,Tuytelaars.T,Van Gool.L.*Wide-baseline Multiple-view Corre-spondences.*IEEE Comp.Vis.and Patt.Rec.2003:718-725.
[4] Bing Chen,Yigong Zhao,Xin Li.A New Method for Wide Basline Matching.*Journal of Xidian University*.2011,38(2):116-123.
[5] Tuytelaars.T,Van Gool.L.Matching Widely Separated Views Based on Affine Invariant Regions. *Computer Vision.*2004,1(59):61-85.
[6] Baumberg.A. Reliable feature matching across widely separated views,*Computer Vision and Pattern Recognition, Hilton Head, SC, USA.*2000:774–781.
[7] Myatt.D,Bishop.J.M,Craddock.R.NAPSAC:High Noise, High Dimensional Robust Estimation It's in the Bag.*British Machine Vision, Cardiff University, British,*2002:458-467.
[8] Tuytelaars.T,Van.Gool.L.Matching Widely Separated Views Based on Affine Invariant Regions. *Computer Vision,*2004,1(59):61-85.
[9] Steele.K.L, Egbert.P.K.Correspondence Expansion for Wide Baseline Stereo.*Computer Vision and Pattern Recognition, San Diego, CA, USA.*2005:1055 - 1062.
[10] Xiuping Ping.The Reserch of Granular Computing Applied in Image Mosaic .*TELKOMNIKA,* 2013,11(3).
[11] Vapnik, V. N.*The Nature of Statistical Learning Theory,*Springer-Verlag, New York, 1995.
[12] R.Hartley and A.Zisserman,Multiple View Geometry in Computer Vision,Bratain:Cambridge University Press,2004.