Feature Points Matching Algorithm based on Homography Constraint and Gray Scale Truncation Number

Ruo Wu¹, a, Kun Wang¹, b and Jiquan Ma¹,2, c

¹College of Computer Science and Technology, Heilongjiang University
²Bioinformatics Lab, Heilongjiang University, China

E-mail: a2161389@s.hlju.edu.cn; b2171450@s.hlju.edu.cn; cmajiquan@hlju.edu.cn

Abstract: A measure function is used to judge whether match in boost matching. In this paper, we offer a kind of new measure function called GST. We simply use the relationship between gray difference s and threshold of the points of domain tests the matching degree. We use homography and GST to improve the number and precision of correspondences of SIFT matching. Firstly, homography is estimated using the result of SIFT matching by VRANSAC. Then, wrong correspondences are eliminated by homography. Finally, we use GST and homography to boost matching the wrong correspondences. Experiments with the Oxford datasets show that our method can get better performance than the existed methods.

1. Instruction

Feature points matching establishes the relationship between two images of one scene, which is a very important issue in many fields of computer vision, such as image mosaic, distance measurement, pose estimation and image classification[1-3].

In the year 1988, C.Harris[4] used the eigenvalues of the autocorrelation matrix of a point in an image to find feature points. Use Euclidean distance of feature vector to match correspondences. In the year 1999, Lowe extracted feature points in scale space, building Scale Invariant Feature Transform[5, 6] (SIFT), which can achieve the purpose of anti-scaling. Fixed scale[7] in feature point matching can reduce certain amount of computation effectively. And then in the year 2006, Herbert Bay came up with an algorithm for local feature extraction, called Speeded Up Robust Features[8] (SURE), which is not only have scale invariance, but also fast. Yuanhui Yan[9] compared SIFT and SURE in complex background images.

In the year of 2015, S.M.Simth and J.M.Brady[10] came up with a feature points matching algorithm based on geometric criteria, which can extract edge information effectively. It is simple and effectivity. Wavelet transform[11, 12] are often used in feature points matching. Amfred[13] proposed two kind of no geometric constraints feature point matching algorithms.

To increase the number and precision of correspondences, many studies concern on boot matching. Measure functions[14, 15] measure the similarity of correspondences in boosting matching. As we know, many feature points matching algorithms deal with gray scale image, for the material for measure functions is gray scale image.

Some of the previous measure functions used difference of single pixel value, such as Squared Difference[16] (SD), Absolute Intensity Difference[17] (AD) and Truncated Absolute Differences[18] (TAD). Among them, TAD brought the thought of threshold of the gray value differences.

Difference of single pixel value is not enough robust. Many algorithms pay attention to the points
of domain. To improve AD, Sum of Absolute Differences[19] (SAD) calculates the gray differences between points around. Zero mean SAD[20] cuts gray value with average gray value before calculating. Sum of Truncated Absolute Difference [21](STAD) develops TAD in calculating window. Some measure functions concern on the probability models of correspondences, such as Normalized Cross-Correlation[22] (NCC) and zero mean NCC[23] (ZNCC).

There are also some non parametric transform measure functions. Rank[24] uses the difference of the number of points whose gray values are less than the center point. Census[24] uses whether a point has a bigger gray value than the center point coding points of domain, and calculates the Hamming distance of them. Our method belongs to non parametric transform matching functions.

In this paper, we propose a kind of measure function which directly uses the relationship of the points of domain between gray difference and threshold to judge whether the correspondence matches. We call it Gray Scale Truncation Number (GST). We use homography and GST improve the number and precision of SIFT matching. Firstly, use Voting RANSAC (VRANSAC) to estimate homography. Secondly, we use homography to eliminate outliers. Finally, we use homography and GST to boost matching to get final correspondences.

The structure of this article is that some basics of homography, homography estimation and how to culling error matching with homography show in section 2. Section 3 show our measure function GST. Boot matching with homography and GST shows in section 4. Experiments show in section 5. Discussion and conclusion show in section 6.

2. Culling Error Matching with Homography

Homography[25] is a 3*3 homogeneous matrix defined as a projection mapping from one plane to another. If both of them are image planes, we call it plane induced homography. Let m, m’ are a pair of correspondence from two images, Homography H expresses as:

\[
m’ = Hm
\]  

Here, we use VRANSAC to estimate homography. Firstly, initialize the votes of correspondences. Here 1 is assigned to all the elements. Secondly, choose candidate correspondences and use them to estimate homography according to the voting situation and use the inliers to vote the correspondences. Then, repeat above steps until enough number of inliers. Finally, complete the estimation of the final homography using the inliers of the homography with biggest size.

We use final homography \( H_f \) to evaluate correspondences \( P \). If a correspondence \( p \) has no more than 2.0 Sampson error[26] under \( H_f \), we call it inlier and keep it as a right correspondence. On the contrary, we call it outline and keep it as a wrong correspondence. At least, we get \( N_r \) right correspondences and \( N-N_r \) wrong correspondences. We use the wrong correspondences to boot matching with homography and GST in section 4. We also get the biggest symmetric transfer error[26] \( E \) of inliers.

3. Gray Scale Truncation Number (GST)

In theory, a pair of right correspondence is in common of gray value. For the points around, their gray values can be similar. However, in the real environment, due to noise and the change of view, the gray values of correspondences will not be exactly the same. The gray value differences of points of domain will be even larger. We define threshold \( t \), the differences of two points below which we regard as same gray value level in and regard as matching. According to experiments, 11 is best.

For a point to match, we build a search window, which we search right correspondences in. For each correspondence, we build a calculate window around. We calculate measure function value in it. For each point in search window of search image, define a counter \( c \) whose initial value is defined as 0. For each point, if gray different between two points in calculate windows are below the threshold \( t \), \( c \) plus one. At least, we choose the correspondence having biggest \( c \) as the right one. The process of calculating \( c \) shows in Figure 1.
4. **Boot Matching with Homography and GST**

As we know, homography builds relationship between two images. However, it limits the points in space from the same plane. In the real environment, it is almost impossible. Homography can still describe relationship between two images in a certain extent, so we use the points around projected by homography building a search window, and use GST search the best matchings in it.

For $N-N_1$ wrong correspondences, we deal with every individual point. For a point $p$ of the $2^{*}(N-N_1)$ unmatched points, we use homography project to another image, getting points $p'$. We use GST to find the right correspondence. Around $p'$, we build a search window who is a square with a length of $4\times\sqrt{E^2+1}$. And we build a calculate window of a square with a length of $32\times\sqrt{E^2+1}$. We get $N+N_1$ correspondences at least.

5. **Experiment**

We use images of the Oxford[27] datasets. It has 8 groups of images with 5 different changes in imaging conditions that are viewpoint changes, scale changes, image blur changes, JPEG compression changes and illumination changes.

| Table 1. Same Experimental Parameter |
|--------------------------------------|
| The length of search window          |
| The length of calculate window       |
| $4\times\sqrt{E^2+1}$                |
| $32\times\sqrt{E^2+1}$               |

We validate GSD with NCC[22], SAD[20], ZNCC[23], rank[24], census[24] and only SIFT[5]. To make the comparison fair, all methods apply same size of search window and calculate window, which show in Table 1. The process of SIFT, homography estimation and culling error matching with homography of every algorithms except SIFT are same, so they have the same number of correspondences.
Table 2. Right Correspondences of SIFT Matching Covers over 40%

| Correspondences of SIFT | Correspondences of others | SIFT | NCC  | SAD  | ZNC  | rank | censuses | GST |
|------------------------|---------------------------|------|------|------|------|------|----------|-----|
| H1to2b                 | 558                       | 607  | 0.894| 0.829| 0.866| 0.880| 0.798    | 0.794| 0.953|
| ikes                   |                           |      | 265  | 889  | 455  | 763  | 092      | 913 | 871  |
| H1to3b                 | 397                       | 443  | 0.841| 0.711| 0.758| 0.773| 0.675    | 0.677| 0.939|
| ikes                   |                           |      | 31   | 019  | 836  | 389  | 676      | 755 | 052  |
| H1to2u                 | 2327                      | 2403 | 0.965| 0.958| 0.975| 0.973| 0.939    | 0.942| 0.989|
| bc                     |                           |      | 191  | 575  | 145  | 488  | 105      | 833 | 596  |
| H1to3u                 | 1528                      | 1638 | 0.922| 0.903| 0.928| 0.926| 0.851    | 0.865| 0.974|
| bc                     |                           |      | 775  | 477  | 657  | 859  | 319      | 108 | 969  |
| H1to4u                 | 677                       | 745  | 0.881| 0.806| 0.839| 0.830| 0.768    | 0.768| 0.944|
| bc                     |                           |      | 832  | 658  | 949  | 986  | 246      | 246 | 966  |

Table 3. Right Correspondences of SIFT Matching Covers less than 40%

| Correspondences of SIFT | Correspondence of others | SIFT | NCC  | SAD  | ZNC  | rank | censuses | GST |
|------------------------|--------------------------|------|------|------|------|------|----------|-----|
| H1to4t                 | 89                       | 123  | 0.247| 0.167| 0.181| 0.120| 0.147    | 0.140| 0.276|
| trees                  |                           |      | 191  | 785  | 208  | 805  | 651      | 94  | 423  |
| H1to5t                 | 62                       | 108  | 0.145| 0.184| 0.157| 0.175| 0.087    | 0.096| 0.296|
| trees                  |                           |      | 161  | 211  | 895  | 439  | 719      | 491 | 296  |

From Table 2 and Table 3, we can see that boost matching can significantly increase the number of correspondences. GST can get a better boost matching result than others, especially in the changes of image blur and JPEG compression. Also, our method only compares the gray differences and threshold, which has less computational complexity than others.

When the result of SIFT matching is quite bad, our method can also get a better result. For the limit of homography result, the promotion of our measure function to the precision is limited.

6. Discussion & Conclusion

In this paper, we propose a measure function called GST which directly uses the relationship between gray differences and threshold of points of domain to judge whether correspondences match. We use homography to boost match testing our method and others. Our method can get a better precision, especially when the result of SIFT matching is quite bad.

In theory, the larger calculate window, the more precision of our result. In fact, if the calculate window is oversize, the computational complexity is too large and the precision reduces. \(32 \times \frac{E}{\sqrt{2}} + 1\) is best.

We use settled threshold of gray difference. It can not fit for all situations. If the gray values vary little in image, our method can not get a good result. But adaptive threshold can hardly get. Also, if appears specular highlights, the original gray distribution can be destroy and our method can hardly effective.

Reference

[1] Q. Chen 2016 Target pose estimation based on contour feature points matching Zhejiang University
[2] A. Faktor and M. Irani 2014 “Clustering by Composition”—Unsupervised Discovery of Image
Categories IEEE Transactions on Pattern Analysis & Machine Intelligence, vol36, no6 p1092-1106

[3] Y. Yan, H. Xia, S. Huang, and W. Xiao 2014 An improved matching algorithm for feature points matching IEEE International Conference on Signal Processing, Communications and Computing, p292-296

[4] C. J. Harris 1988 A combined corner and edge detector Proc Alvey Vision Conf, vol1988, no3 p147-151

[5] D. G. Lowe and D. G. Lowe 2004 Distinctive Image Features from Scale-Invariant Keypoints International Journal of Computer Vision, vol60, no2, pp. 91-110

[6] D. G. Lowe 1999 Object recognition from scale-invariant keypoints Iccv

[7] J. Li 2013 An image feature point matching algorithm based on fixed scale feature transformation Optik - International Journal for Light and Electron Optics, vol124, no13, p1620-1623

[8] A. Leonardsis, H. Bischof, A. Pinz, H. Bay, T. Tuytelaars, and L. Van Gool 2006 SURF: Speeded Up Robust Features Computer Vision – ECCV

[9] X. Wang, D. Liu, and L. Wang 2015 A Feature Point Matching Algorithm for Complex Background Image IEEE Fifth International Conference on Big Data and Cloud Computing, p243-247

[10] S. M. Smith and J. M. Brady 2015 SUSAN - A New Approach to Low Level Image Processing in International Journal of Computer Vision p45--78

[11] S. Mohammadzadeh and H. Farsi 2014 Image retrieval using color-texture features extracted from gabor-walsh wavelet pyramid Sid Ir

[12] S. Lahmiri and M. Boukadoum 2013 Hybrid Discrete Wavelet Transform and Gabor Filter Banks Processing for Features Extraction from Biomedical Images Journal of Medical Engineering, 2013, (2013-4-15), vol2013, no7, p53 - 56

[13] A. J. T. W. S, and S. S 2013 Mirror Match: Reliable Feature Point Matching without Geometric Constraints

[14] X. Yang, D. Miao, F. Cao, and Y. Ma 2005 Study on the Matching Similarity Measure Method for Image Target Recognition in International Conference on Fuzzy Systems and Knowledge Discovery, p289-292.

[15] D. N. Bhat and S. K. Nayar 1998 Ordinal Measures for Image Correspondence IEEE Transactions on Pattern Analysis and Machine Intelligence, vol20, no4, p415-423

[16] A. Hosni, M. Bleyer, and M. Gelautz 2013 Secrets of adaptive support weight techniques for local stereo matching Elsevier Science Inc

[17] T. Kanade, H. Kano, S. Kimura, A. Yoshida, and K. Oda 2002 Development of a video-rate stereo machine Ieee/rsj International Conference on Intelligent Robots and Systems 95. 'human Robot Interaction and Cooperative Robots', Proceedings, p95-100 vol3

[18] M. J. Black and P. Anandan 1993 A framework for the robust estimation of optical flow Brown University Press, p154-155

[19] F. Yu 2013 Research on stereo vision technology under variant illumination Ocean University of China

[20] G. Luo 2012 Research on depth perception and 3D reconstruction of binocular stereo vision Central South University

[21] H. Meng 2012 Stereo matching algorithms in 3D reconstruction Harbin Institute of Technology

[22] D. I. Barnea and H. F. Silverman 1972 A Class of Algorithms for Fast Digital Image Registration IEEE Computer Society , p179-186

[23] J. S. Ortega 2007 Towards Visual Localization, Mapping and Moving Objects Tracking by a Mobile Robot: a Geometric and Probabilistic Approach Bibliogr

[24] R. Zabih and J. Woodfill 1994 Non-parametric local transforms for computing visual correspondence, p151--158

[25] A. Agarwal, C. V. Jawahar, and P. J. Narayanan 2005 A Survey of Planar Homography Estimation Techniques in Technical Reports, International Institute of Information Technology, p1-5

[26] R. Hartley and A. Zisserman 2000 Multiple view geometry in computer vision p1865 - 1872

[27] http://www.robots.ox.ac.uk/~vgg/research/affine/