Application of Feature Point Detection and Matching in 3D Objects Reconstruction

Jie Yuan, Ting Feng, Yu Zhou, Yao Yu
School of Electronic Science and Engineering, Nanjing University, Nanjing 210093, China
Email: yuanjie@nju.edu.cn, fengnju1988@gmail.com, nackzhou@nju.edu.cn, allanyu@nju.edu.cn

Abstract— **In this paper, we present a novel wide baseline based approach to detect and match feature points in image series. We found the wide baseline correspondence problem with large scale, rotation, illumination and affine transformations is still not tackled very well. We proposed a new matching method which based on multi-scale Harris algorithm and two-way guided matching to achieve large number of accurate point correspondences between un-calibrated image sequences of the same scene for wide baseline. We apply our method in the experiments of 3D object reconstruction with satisfied results. It shows that the guided matching method can be used for severe scene variations and provide evidence of improved performance with respect to the SIFT distance and Harris matchers. It is also useful to the matching in short baseline, and the results of this method are better than that of the traditional method.**

Keywords-feature points; image matching; 3D object reconstruction

I. INTRODUCTION

Wide Baseline Matching (WBM) is one of the most important issues that have been extensively studied in the field of computer vision, as well as the foundation of many computer vision theory and applications [1][2][3] such as object identification, camera calibration, 3D reconstruction, and motion analysis. Meanwhile, WBM is a bottleneck in the field of computer vision research. Therefore, research on WBM is of significant importance. WBM primarily divides into two parts: feature point detection and matching.

The primary methods of feature point detection are Harris feature point extraction algorithm and Scale Invariant Feature Transform (SIFT [14]) feature point extraction algorithm. Both of them have their own advantages and disadvantages. To get better corner detection results, Keju [6] combined the two algorithms during 3D reconstruction on demand. But this method has limited the range of application, which means that it is not applicable if we merely to get more and more accurate feature points. Schmid [5] reported that corner extraction algorithm, being invulnerable to camera pose and sunlight, performs the best currently. However, as for vision systems with large scale changing, this method can hardly guarantee invariability of the feature points. In this case, the paper provides a novel multi-scale corner detection approach which combines scale space theory and Harris feature point detection algorithm.

To match feature points, generally, the relative methods are used to achieve the correspondence of two images' point sets. Considering noise interference, light conditions, and other factors which may result in a great number of mismatches, however, removing mismatches is essential. One of the direct ways to remove mismatches is to find an affine transformation which is applicable to all the feature points in the image, and then use it to pre-estimate the position of these feature points located in the other image [4]. Nevertheless, it is not applicable to complicated scenes. To solve this problem, Ferrari et al. [3] proposed to estimate local affine transformation matrix for every pair of corresponded feature points, using least mean square method. Later it was suggested in another approach that this affine transformation matrix should be compared to the predetermined threshold to gather the most similar points to the affine transformation matrix. Although this method has been proven to be effective in confirming mismatches, it pays a high cost of computational complexity. Currently, a comparatively better method in the field of removing mismatches is to use epipolar geometry restriction proposed by Zhang [9]. This method can produce excellent results on the condition that matching points are less in quantity and parallax is small. However, there are two issues remained to be solved. One is that the quantity of matching points is relatively small. The other is the restriction of disparity. Increasing the disparity means enlarging the match searching window, while enlarged match searching window will probably introduce mismatches.

In response to the above problems, we propose a novel approach combining epipolar geometry, homograph constraint, mismatch detection and guided matching which, to some extent, greatly make up the deficiency in these two areas mentioned above. At first, we use relative method to conduct initial match of the image feature points set. Secondly, we use RANSAC (Random Sample Consensus) method to estimate fundamental matrix and homography matrix and remove mismatches in correspondence. Then, we remove mismatches again according to euclidean distance. Finally, we use optimized fundamental and homography matrix guiding the matching to get more and more accurate matching points.

In 3D objects reconstruction, the quality of WBM will affect the result of reconstruction directly. The purpose of this paper is to get a better algorithm in WBM, and apply it to 3D reconstruction. Then the accuracy of 3D objects reconstruction will be enhanced, and fewer cameras or video cameras will be used in experiment.

 ISBN: 978-1-61208-158-8

II. FEATURE POINT DETECTION

An effective feature point detection algorithm is introduced in this section. We introduce the scale space theory at first.

A. Scale Space Theory

Scale space theory is carried out by scaling the original image to obtain multi-scale sequence of scale space, and extracting the main contour based on the sequence as a feature vector, to achieve edge detection, corner detection and feature extraction on different resolutions. As an important concept in scale space theory written by Lindbergh [12][13] that scale space is describing original image at different levels, each layer has a scale parameter which may be discrete and also can be continuous. All scales of space should have the following properties:

- All the signals should be defined in the same domain.
- With the growth of scale parameters, the output image is increasingly blurred.
- Details contained at the coarse level of a signal are less than that at the fine level. If the local maximum is a measure of smoothness, as the scale blurred, extreme non-increasing, this property is known as the "scale space causality".
- All that is generated by a convolution operator.

Scale space kernel is defined as:

$$L(x, y, \sigma) = K(x, y, \sigma) \times I(x, y) \quad (1)$$

In (1), $I(x, y)$ is the original image, σ is the scale parameter. For all the images I, if the extremes of the image $L(x, y, \sigma)$ obtained after its convolution with transform kernel K is less than the extremes of original image, then we call K the scale space kernel. Generally we only use the Gaussian kernel as the scale convolution. Because in Gauss scale space, fine-scale information on the parameter value with the increase in scale was inhibited in the scale of the change from coarse to fine process, no new structure. However, since Gauss kernel is linear, translation invariant, rotation invariant, has subset features and many other properties, it can be proved that Gauss kernel is the only transform kernel to achieve scale space transformation [8]. Feature points and edges of the same type at different scales have a causal relationship, which means that when scale changes, new feature points may arise while old ones may be displaced or disappear. The ambiguity brought by the causal relationship is inherent and inevitable which should never be expected to be eliminated but it can be decreased.

B. Multi-Scale Harris Feature Point Detection Algorithm

Multi-scale Harris feature point detection algorithm was introduced and the experimental results of this method was given in this section.

1) Harris Operator of Scale Space

Harris operator R can be represented as:

$$R = det(C) - ktr^2(C) \quad (2)$$

In (2), $C(x) = \begin{bmatrix} I_u^2(x) & I_{uv}(x) \\ I_{uv}(x) & I_v^2(x) \end{bmatrix}$, k is empirical value, which usually between 0.04~0.06. To obtain the presentation of Harris operator, I_u, I_v can respectively be represented as：

$$I_u(x, s\sigma_n) = I_u(x) * G_u(x, s\sigma_n) \quad (3)$$

$$I_v(x, s\sigma_n) = I_v(x) * G_v(x, s\sigma_n) \quad (4).$$

Then the $C(x)$ function of Harris algorithm will become

$$\hat{C}(x, \sigma_I, \sigma_D) = \sigma_D^2 G(\sigma_I) * \begin{bmatrix} I_u^2(x, \sigma_D) & I_{uv}(x, \sigma_D) \\ I_{uv}(x, \sigma_D) & I_v^2(x, \sigma_D) \end{bmatrix} \quad (5)$$

In (5), $\sigma_I = \sigma_n$ is the selected scale parameter to calculate feature points; $\sigma_D = s\sigma_n$ is the differential scale; $G(\sigma_I)$ is Gaussian function. Through judging to detect the feature corner under σ_n scale level

$$R = det(\hat{C}) - ktr^2(\hat{C}) > T \quad (6).$$

2) Multi-Scale Harris Feature Point Detection Algorithm

Arithmetic operator $LOG\nabla^2 g$ is forwarded by [7]. Two-dimensional LOG operator can be represented as:

$$\nabla^2 g = \left(\frac{x^2 - \sigma^2}{\sigma^4}\right) exp\left(-\frac{x^2}{2\sigma^2}\right) \otimes f(x, y) \quad (7)$$

Where $f(x, y)$ is the function to be detected. Using a typical template LOG operator in the text to detect whether the corner point measured under a certain specified scale level is the extreme value, which result in an invariant scale feature corner. The procedure of multi-scale Harris feature point detection algorithm procedure is as follows:

- Primarily select scale variables σ_n and the threshold value T, using formula (6) and (7) to calculate the candidate feature corner of each scale level.
- Use iterative algorithm to detect whether the LOG operator of each scale level candidate corner points to obtain the maximum value, and determine the results in the location and scale of the final feature corner. Consider the entire scale space of the image, assumed to detect the corner set C_{opt} under $\sigma_n = \sigma_0$ large-scale level. Decrease the image scale coordinate to $\sigma_n = \sigma_1$, and detect new corner set C_{new} in the neighborhood of the image. If there is C_{new}, regard C_{new} as the corner feature set of the current image. Repeat the above process until there is no change of C_{new}, or until the scale is small enough.

3) Compare Experimental Results of multi-scale Harris Feature Point Detection Algorithm with That of Ordinary Harris Algorithm

During experiment, taking the standard deviation proportional constant of Gaussian kernel function $S = 0.7$, $k = 0.04$ [7], use a typical 5×5 LOG operator.

$$\begin{bmatrix} 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & -2 & -1 & 0 \\ -1 & -2 & -16 & -2 & -1 \\ 0 & -1 & -2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \end{bmatrix}$$

 ISBN: 978-1-61208-158-8

When σ_n is taken as 2, 1 and multi-scale, the results of feature point extracting of the original image which is shown highlighted are shown as follows:

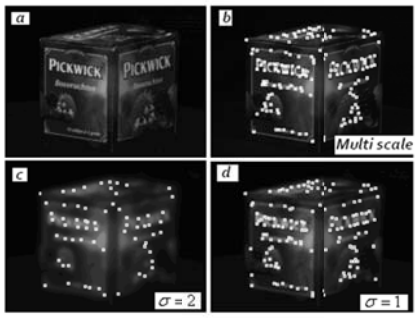

Figure1. Traditional Harris feature point detection of different scales compared with multi-scale Harris feature point detection

Stability criteria, reliability criteria, and anti-noise performance can be applied to evaluate the superior or inferior of the detection algorithm in a certain corner [8][12]. These three criteria are determined by the repetition rate γ of both initial detection of the corner and the corner points that detected after the change of parameters, threshold value, or the increase of the noise. In our experiment, we test the stability criteria of the multi-scale Harris feature point detection by changing the scare.

$$\gamma = \frac{|C_1 \cap C_2|}{min(|C_1|,|C_2|)} \times 100\% \quad (8)$$

In (8), C_i denotes the detected feature points set. $|C_i|$ represents the number of elements in the collection C_i. Formula (8) shows that the bigger the repetition rate γ is, the more stable the algorithm will be. Results of different scales for corner detection to the original image is showed in Figure 1, C_1 is the corner set with fixed scale of $\sigma = 1$, C_2 is the corner points set with the scale of $\sigma_n \in [2,6]$, calculate the repetition rate of C_1 and C_2 respectively. See the experimental results in Figure 2.

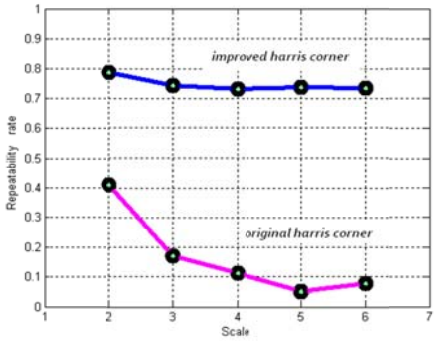

Figure 2. Comparison of 2 kinds Harris feature point repetition rate

It can be seen from the above graph that, when the scale changes, repeatability of the scale space theory which based on Harris feature point detection method for duplicate detection rate is significantly higher than the original Harris method, that is to say, Harris feature point detection method that based on scale space theory is more stable and reliable .

III. FEATURE POINT MATCHING

In this section, initial matching, epipolar geometry constraint and homography constraint are introduced. Two-way guided matching is proposed.

A. Initial Matching

Initial match consists of two steps: general matching and specific matching. General matching is achieved by respectively calculating the correlation coefficient of feature points of the two images. When the correlation coefficient is larger than a given threshold value, both the feature points are considered to be reciprocal, and therefore become the candidate of the corresponding feature points. x Is the feature point of I_1, while x' is the feature point of I_2. First of all, assume (x, x') is a real corresponding feature point, then within its neighborhood there might be more corresponding feature point (y, y'), where y and y' are respectively within the neighborhood of x and x'. The necessary condition that (y, y') is supported by real feature point (x, x') is: the angle betweenx, y and x', y' should be less than the preliminary set value. Search support strength from the neighborhood to accumulate matching strength. Since one feature point can match more than one candidate, each corresponding point has support strength. Therefore, we just consider the maximum support of each neighborhood and the symmetry of such support.

B. Using Epipolar Geometry and Homography Constraint to Eliminate Mismatches

In this section, fundamental matrix F and homography matrix H are calculated, and the mismatches are removed at the same time.

1) Epipolar Geometry Constraint

In two images that view from different view points, epipolar geometry constraint is the certain limit between the corresponding points of the same physical space point when collecting. It can be algebraically described by the fundamental matrix.

$$m'^T F m = 0 \quad (9)$$

In (9), F is fundamental matrix, m and m' is a pair of matching points of two images. Because there are many mismatches among the initial matching points, directly use of usual least squares method to calculate the fundamental matrix does not achieve good results. We choose the RANSAC [10][11] method, which is thought to be more robust. The process of using RANSAC method to get F is also the process of removing the mismatches.

2) Homography Constraint

Homography constraint maps the points of one geometry plane surface to another. It is a reversible mapping and obeys keeping linear. m and m' are corresponding match points of the two images, so that the homography H should obey:

$$m' = Hm \quad (10)$$

During experiment, we apply the RANSAC method again to find the solution for H, at the same time, further eliminate mismatches. The polar line runs through the entire image in epipolar geometry constraint, which still has relatively large

 ISBN: 978-1-61208-158-8

matching space. As for the images with lots of feature points, there are still plenty of mismatches after using epipolar geometry constraint. The mentioned homography constraint method above can further diminish the scope of matching so that matching precision is proved. H is a mapping of points from one geometric plane surface to another, which is inapplicable to the images of significant depth change. Nevertheless, most physical scene images have little depth change, by appropriately broadening the threshold of matching, good results can be achieved in practice.

3) Guided Matching by Using Epipolar Geometry and Homography Constraint

The above matching process has removed a large number of mismatches, but it gets less correct matching points at the same time. In order to solve this problem, we integrated the use of epipolar geometry and homography constraint to guided matching. Redirect and match all feature points of image I_1 and I_2 to get more accurate matching points.

The epipolar geometry constraint show that m_i' which is the matching point of m_i (m_i is the feature points detected in I_2) is necessarily adjacent to the corresponding polar line. According to homography constraint, m_i' is also near to estimation point $\hat{m}_i'$. S3 the intersection of the above two mentioned areas are the approximate area where the matching points located, see Figure 3below.

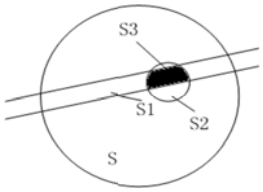

Figure 3. The searching scope of guided matching

Compared with the initial match, selecting the area defined in S3 to guided matching have greatly diminished the searching range of finding matching points and reduced but cannot avoid mismatching. Apply the euclidean distance detection method to calculate the euclidean distance between m_i' and $\hat{m}_i'$, sort the matching points from short to long in accordance with the euclidean distance, and select the top-ranked matching points. Further diminish the threshold for interior points of RANSAC method, and use top-ranked matching points to solve F, H. Then to guided match and select the accurate matching points with the euclidean distance again, see Figure 4, use two cycles guided matching to obtain more accurate F and H and make more accurate matching. The above steps have completed the guided matching of the feature points of I_1. Similarity, apply the same method to guided matching the feature points of I_2, and then take their union set.

Our experiment guides all the feature points for the second round of guided matching rather than keep only top-ranked points, because each round will get even better.

IV. EXPERIMENTS AND RESULT

The procedures and results of our experiment are shown in this section.

A. Experimental Operation

The entire process of experimental test operation is shown in Figure 4:

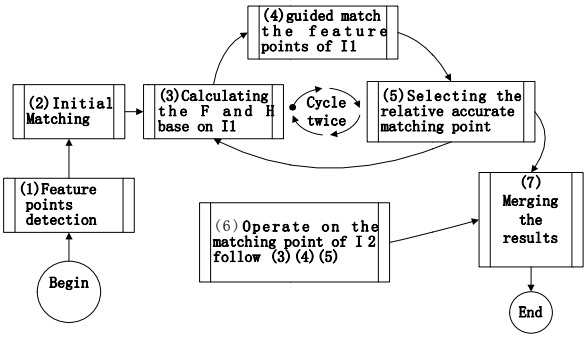

Figure 4. Flow chart of the whole experiment

B. Experimental Results

In order to verify the feasibility of proposed algorithm, this paper conducted experiments with two downloaded images captured by an un-calibrated SLR camera from different view points, and the parallax angle is 85 degrees. The two images we used are as follow:

Figure 5. Original images

The results are showed in Figure 6-Figure 8. Every pair of matching points in I_1 and I_2 is connected by a straight line in the figure. Figure 6 shows the results of initial matching, Figure 7 shows the matching results of epipolar geometry and homography constraint. Figure 8 shows the results of two-way guided matching. The number of accurate matching points before guided matching is 102, and the number of matching points after two-way guided matching is 320 pairs. We can see from the corresponding matching points from Figure 8 that the amount of the feature points is relatively large and match each other accurately, but the matching points which can express the features of the box, such as feature points at the top side of the box are not matched by our method for the epipolar geometry constraint can fit the whole image, but the homography constraint cannot. In our experiment, the method of WBM is even applicable to the image pairs whose parallax angle are 85 degrees, but it is more suitable to that less than 80 degrees. Figure 9-Figure 11 are the results of our approach by using the proposed images whose parallax angle is 65 degrees. We can easily find out that after applying two-way guided matching algorithm, nearly all feature points are matched correctly, and there is no error matching at all.

 ISBN: 978-1-61208-158-8

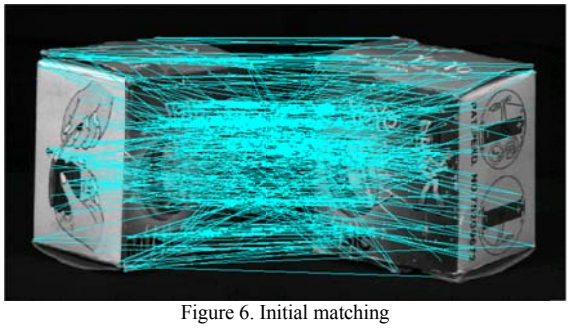

Figure 6. Initial matching

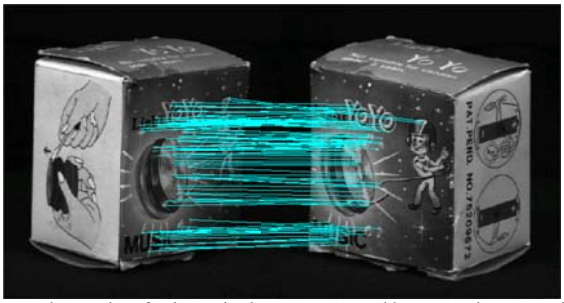

Figure 7. The results of using epipolar geometry and homography constraint to remove mismatches (102 pairs)

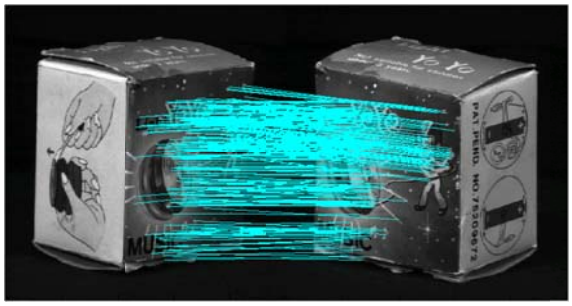

Figure 8. Two-way guided matching (320 pairs)

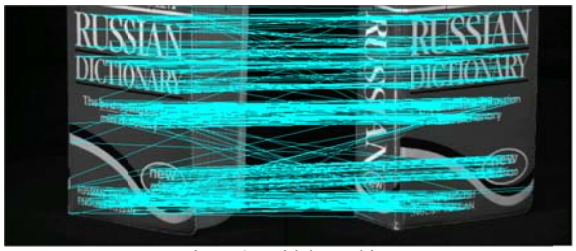

Figure 9. Initial matching

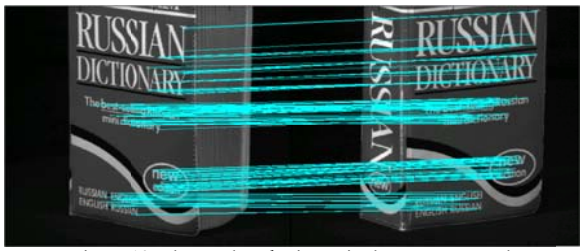

Figure 10. The results of using epipolar geometry and homography constraint to remove mismatches (198 pairs)

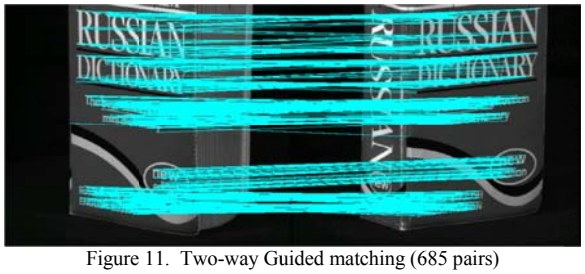

Figure 11. Two-way Guided matching (685 pairs)

To verify the correctness of our algorithm, we calculated the parallax angle of the images and conducted experiments on four groups of error rate testing. The results show that the calculated parallax angle is in accordance with the real parallax angle.

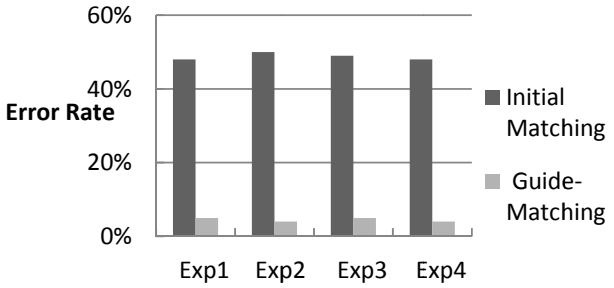

Figure 12. Results Statistics

Figure 12 shows the initial matching and the guided matching error rate. As can be seen from the chart, this method greatly reduces the matching error rate.

V. Conclusion and Future Work

In this paper, we successfully solved the problem of how to obtain sufficient and reliable matching points over two wide baseline image pairs. First, a large number of feature points are detected by using multi-scale Harris corner detection algorithm. Then, we comprehensively apply epipolar geometry and homography constraint to guided matching algorithm, which effectively settle the conflict between wide baseline and sufficient matching points without reducing the accuracy of the matching points. Meanwhile we have applied Euclidean distance to filter the matching points, making the matching more precise. Finally, we have tested a number of wide baseline image pairs under different severe camera motions with illumination changes, self-similarities, and have obtained the excellent results for all of the images. The experiments show that this algorithm can effectively detect and match, and it also has better match property compared with the traditional corner detection and matching algorithm. This method can increase the quantity of common points in 3D and then enhance the accuracy of 3D reconstruction. However, due to the homography constraint of geometric plane surface that used in this algorithm, the pictures with great depth change usually get poor test results, which still needs to be improved in future research.

 ISBN: 978-1-61208-158-8

VI. ACKNOWLEDGEMENT

This paper is supported by the Fundamental Research Funds for the Central Universities, Number: 1107021051, and National Natural Science Fundation of Jiangsu Province of China, Number: BK2010386.

REFERENCES

[1] V. Kolmogorov and R. Zabih, "Computing visual correspondence with occlusions via graph cuts", Proceedings of International Conference on Computer Vision, 2001, pp. 508–515.

[2] O. Chum and J. Matas. "Matching with prosac" Progressive sample consensus". Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2005, pp. 220–226.

[3] P. Pritchett and A. Zisserman., "Matching and reconstruction from widely separated views", Proceedings of the European Workshop on 3D Structure from Multiple Images of Large-Scale Environments, 1998, pp. 78–92.

[4] C. Wang, K.K. Ma, T. Khim, and D. Guo, "Mismatch Removal for Wide baseline Image Matching via Coherent Region-to- Region", Proceedings of Fourth Pacific-Rim Symposium on Image and Video Technology, 2010, pp. 101-106

[5] C. Schimid, R. Mohr, and C. Bauckhage, "Evaluation of interest point detectors", International Journal of Computer Vision, 2000 37(2), pp. 151-172

[6] K. Peng, X. Chen, D. Zhou, and Y.H. Liu, "3D Reconstruction Based on SIFT and Harris Feature Points", Proceedings of International Conference on Robotics and Biomimetics, 2009, pp. 960-964

[7] C. Harris and M. Stephens. "A combined corner and edge detector", Proceedings of 4th Alvery Vision Conference, 1988, pp. 147-151

[8] D. Lowe. "Distinctive Image Features from Scale-Invariant Interest Points". International Journal of Computer Vision 2004, 60(2), pp. 91-110

[9] Z. Zhang, R. Deriche, and O. Faugeras, "A robust technique for matching two uncalibrated images trought the unknown epipolar geometry". Artificial Intelligence, Special volume on computer vision, 1995, 78(1-2), pp. 87-119

[10] L. Tony. "Seale-space A framework for handing image structures at multiple Seales". Proceedings of the conference Eurean organization for Nuclear Research school of computing, Egmond aan zee, The Netherland, 1996, pp. 8-21.

[11] L. Tony. "Scale-space for Discrete signals". IEEE transactions on pattem Analysis and machine intelligence, 1990, pp. 233-253

[12] Y. Yang and T.W. Zhang. "Assessing criterion of corner finders" Journal of Harbin Institute of Technology, 1998, 30(2), pp. 7-10.

[13] R. Hartley and A. Zisserman, "Multiple view geometry in computer vision". Cambridge University Press, 2000, pp. 237-259.

[14] Lowe and G. David G "Object recognition from local scale-invariant features". Proceedings of the International Conference on Computer Vision. 2. pp. 1150–1157

 ISBN: 978-1-61208-158-8