

# A novel and fast blurred image matching method

Zhi Huang\*, Yaran Yang

*The Department of Image Processing, Suzhou North America High School, SuZhou, 215101, China*

**Abstract:** *In this paper, a novel image matching method is proposed in order to improve the performance of image registration, especially for blur images. Firstly, A set of Scale Invariant Feature Transform (SIFT) points are extracted. Secondly, in order to further improve the distinctiveness of the SIFT descriptors, three scale invariant concentric circular regions are applied to produce descriptors. Thirdly, for the purpose of decreasing the high dimensional and complexity of SIFT descriptors, The Local Preserving Projection (LPP) technic is applied to reduce the dimensions of the descriptors. Lastly, the Euclidean distance similarity measurements are used to obtain the results of matching feature points. The experimental results show that the novel image matching method can not only reduce the data amounts, but also improve the matching speed and the matching precision.*

**Keywords:** *Image Match, SIFT descriptors, LPP, Blur images*

## 1. Introduction

Image registration is an extremely important and fundamental step in the computer vision field, such as panoramic stitching Szeliski [1]; Brown and Lowe [2], object recognition Ferrari et al.[3]; Lowe [4], wide baseline matching Schaffalitzky and Zisserman [5]; Tuytelaars and Van Gool[6], 3D recognition Kratochvil et al. [7], recognition of object categories Dorko and Schmid[8]; Fergus et al.[9]; Leibe and Schiele[10]; Opelt et al.[11], and so on. One of the most popular methods to solve the image registration problem is the local correspondence approaches, which includes two essential steps: local feature detection Lowe[12]; Morel and Yu[13]; Mikolajczyk et al.[14] and local feature description Mikolajczyk and Schmid[15]. In an comparative study on local feature descriptors Mikolajczyk and Schmid[15] including SIFT Lowe[4], steerable filters Freeman and Adelson (1991)[16], differential invariants Koenderink and van Doorn[17], moment invariants Van Gool et al.[18], complex filters Schaffalitzky and Zisserman[5], and cross-correlation of different types of interest points Harris and Stephens[19]; Mikolajczyk and Schmid[20], the authors found that the SIFT descriptors performs best.

Although their best matching results were obtained using SIFT descriptor, the high dimensionality of this descriptor (128 dimensions) causes computational inefficiency when there are a large number of points to be processed. Then dimensionality reduction techniques can find a place here, and can also be applied to design features as well. As the first attempt, PCA-SIFT Ke and Sukthankar[21], which used the principal components of gradient patches to form local descriptors, was proposed. Despite the widespread use in various fields, the validity of PCA is limited by its priori assumption that the relationship among data is linear. However, in real-world applications, it is common where the relation among variables is nonlinear. In this case, nonlinear techniques, such as Isomap Tenenbaum et al.[22], Locally Linear Embedding (LLE) Roweis and Saul[23]; Saul and Roweis[24], and Laplacian Eigenmap (LE) Belkin and Niyogi[25] might be more appropriate, which are proposed to discover the submanifold structure hidden in high dimensional ambient space. Though these methods have been successfully applied to some benchmark artificial datasets, the yielded mappings are only defined on the training data points and it is unclear how to extend the mapping for new test data points. Therefore, the nonlinear manifold learning techniques Tenenbaum et al.[22]; Roweis and Saul[23]; Saul and Roweis (2003)[24]; Belkin and Niyogi[25]; Brand[26]; Zha and Zhang[27] are limited in applicability for information comparison tasks. In contrast, the Locality Preserving Projections(LPPs) Niyogi[28] which definitely considers the structure of the manifold may be expediently and reliably applied to any new data point to locate it in the intrinsic low dimensional submanifold. LPP Niyogi [28] is a local structure preserving method, which can preserve the intrinsic geometric relationships of the data and share many important properties with nonlinear techniques such as LLE Roweis and Saul[23] or LE Belkin and Niyogi[25]. LPP builds a graph maintaining neighborhood relationship of the given data set, and then uses the notion of the Laplacian of the graph to compute a projection matrix. This projection matrix can map the high dimensional data points to a subspace, and has the property that local

neighborhood information is well preserved. This property makes the algorithm insensitive to outliers and noises to some extent. Since it is likely that a nearest neighbor seek in the locality preserving low dimensional submanifold will yield corresponding results to that in the high dimensional ambient space, the locality preserving quality of LPP is to be of effective and credible use in the information comparison applications. On the other hand, though the SIFT descriptors can accurately describe invariant image characteristic around a keypoint, the description region need be improved to generate more accurate description, so it can avoid mismatches between two keypoints, which are geometrically far away but with similar local image information. Thus we consider three scale invariant concentric circular regions, which are applied to produce more discriminative descriptors.

This paper attempt at developing more compact descriptors, which are suitable for faster matching while still retaining their outstanding performance. To do this, we propose a novel matching method based on SIFT and LPP. The experiment results indicate that the proposed method is robust against all image transformation, especially for the image blur.

## 2. A novel matching method based on SIFT and LPP

### 2.1 Scale invariant feature points detection and localizations

As noted, the first step in point correspondence is feature point detection. the scale invariant feature points are detected and localized as introduced in Lowe[4].

As noted, the first step in point correspondence is feature point detection. In this paper, the scale invariant feature points are detected and localized as the way introduced in Lowe[4]. To obtain true scale invariance Lindeberg (1998), the normalized Laplacian is approximated by the Difference of - Gaussian (DOG) scale space. The DoG scale space is sampled by blurring an image with successively larger Gaussian filters and subtracting each blurred image from the adjacent (more blurred) image. Three levels of scale are created for each octave by blurring the image with incrementally larger Gaussian filters with scale steps of  $\sigma = 2^{1/3}$ . After completing one octave, the image with twice the initial scale is resampled by taking every other row and column and the process is repeated for the next octave, and thus reducing computation. That is, stable keypoint locations in scale space are detected by using scale - space extrema in the DOG function convolved with the image,  $D(x,y,\sigma)$ , which can be computed from the difference of two nearby scales separated by a constant multiplicative factor  $k$ :

$$D(x,y,\sigma) = (G(x,y,k\sigma) - G(x,y,\sigma)) \otimes I(x,y) = L(x,y,k\sigma) - L(x,y,\sigma) \quad (1)$$

where  $G(x,y,\sigma)$  denotes a two-dimensional Gaussian kernel with standard deviation.  $L(x,y,\sigma)$  denotes the scale-space representation, and is produced from the convolution of  $G(x,y,\sigma)$  with an input image  $I(x,y)$ . Interest points are characterized as the extrema (maxima or minima) in the 3D  $D(x,y,\sigma)$  space. As such, each pixel is compared with its 26 neighbors in scale space and a pixel is selected as a feature point if its value is larger or smaller than all of its neighbors. Subsample accurate position and scale is computed for each extrema point by fitting a quadratic polynomial to the scale space function  $D(x,y,\sigma)$  and finding the extremum, which is given by

$$\hat{x} = \frac{\partial^2 D^{-1}}{\partial x^2} \frac{\partial D}{\partial x} \quad (2)$$

The feature point candidates obtained above need to be refined by eliminating some unstable ones, which are low in contrast and sensitive to changes in illumination. In addition, some features points that have a strong response along edges are also eliminated for higher stability. The accurate point localization can be obtained by a quadratic interpolation. After the above operations, the 3D coordinate of a feature point can be obtained, i.e.,  $(x,y,\sigma)$ , where  $(x,y)$  is the spatial coordinate, and denotes the scale.

**2.2 Triple regions based SIFT description**

In this stage, three scale invariant concentric circular regions is adopted to assign orientations as well as produce descriptors (see Fig. 2). Orientation is determined by building a histogram of gradient orientations from the key points neighborhood, weighed by a Gaussian and the gradient magnitude. Every peak in the histogram with a height of 80the maximum produces a key point with the corresponding orientation. For each pixel positioned in (x,y) , the gradient magnitude m(x,y) and gradient orientation (x,y) are pre - computed using pixel differences. Note that the assigned orientation, together with the scale above, provides a scale and rotation invariant coordinate system for the descriptor. The SIFT descriptor computes the gradient vector for each pixel in the feature points neighborhood and builds a normalized histogram of gradient directions. The SIFT descriptor creates a 1616 neighborhood that is partitioned into 16 subregions of 44 pixels each. For each pixel within a subregion, SIFT adds the pixels gradient vector to a histogram of gradient directions by quantizing each orientation to one of 8 directions and weighting the contribution of each vector by its magnitude. Each gradient direction is further weighted by a Gaussian of scale  $\sigma=n/2$  where n is the neighborhood size and the values are distributed to neighboring bins using trilinear interpolation to reduce boundary effects as samples move between positions and orientations. That is, the proposed triple regions based SIFT descriptor can be given by

$$PD = \alpha_1 L_1 + \alpha_2 L_2 + \alpha_3 L_3 \tag{3}$$

where  $L_i$  ( $i=1,2$  and  $3$ ) is the 128 - dimensional local SIFT descriptor, and thus PD is also 128 dimensions.  $\alpha_1, \alpha_2$  and  $\alpha_3$  are three weight coefficients, and satisfy  $0 \leq \alpha_1 \leq 1, 0 \leq \alpha_2 \leq 1, 0 \leq \alpha_3 \leq 1$ , and  $\alpha_1 + \alpha_2 + \alpha_3 = 1$ .

**2.3 LPP -based dimensionality reduction**

Suppose the data set of SIFT descriptions is  $(x_1, x_2, \dots, x_m)$  and let  $N(x_i)$  ( $N(x_j)$ ) denote the k nearest neighbors of  $x_i$  ( $x_j$ ),  $N(x_j)$  denote the k nearest neighbors of  $x_i$ . Then use  $y_i = w^T x_i$  to denote the one - dimensional representation of  $x_i$  with the transformation vector w, and define the similarity matrix  $S(s_{ij} = s_{ji})$  as follows:

$$s_{ij} = \begin{cases} e^{-\|x_i - x_j\|^2 / t} & \text{if } x_i \in N_{x_j} \text{ or } x_j \in N_{x_i}; \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

The criterion for choosing a reasonable projection is to minimize the objective function as follows:

$$f = \frac{1}{2} \sum_{ij} (y_i - y_j)^2 s_{ij} \tag{5}$$

This objective function undergoes a severe penalty if the neighboring points  $x_i$  and  $x_j$  are mapped far apart (i.e.,  $(y_i - y_j)^2$  is large). Therefore, minimizing f can ensure that if  $x_i$  and  $x_j$  are adjacent then  $y_i$  and  $y_j$  are close as well. Exercising some simple algebraic deduction, f can be rewritten to

$$\begin{aligned} f &= \frac{1}{2} \sum_{ij} (y_i - y_j)^2 s_{ij} \\ &= w^T X L X^T w \end{aligned} \tag{6}$$

where  $D$  is a diagonal matrix  $D_{ii} = \sum_j S_{ij}$  and  $L = D - S$  is the Laplacian matrix. As the bigger value of the  $d_{ii}$  corresponds to the more important  $y_{ii}$ , there is a natural constraint:

$$Y^T D Y = w^T X L X^T w = 1 \quad (7)$$

This minimization problem can be predigested to finding:

$$\begin{aligned} \arg \min w^T X L X^T w \\ w^T X L X^T w = 1 \end{aligned} \quad (8)$$

which can be translated as the generalized eigenvalue problem:

$$X L X^T w = \lambda X D X^T w \quad (9)$$

It is easy to show that the matrices  $X L X^T$  and  $X D X^T$  are symmetric and positive semidefinite. Let the column vectors  $w_i$  ( $i=0,1,\dots,l-1$ ) be the solution of the above generalized eigenvalue problem, ordered according to their eigenvalues,  $\lambda_0 < \dots < \lambda_{l-1}$ . The final  $n \times l$  projection matrix  $W_{LPP}$ , which projects the  $n$ -dimensional descriptive vector to the lower  $l$ -dimensional representation, is constructed as  $W_{LPP} = (w_0, w_1, \dots, w_{l-1})$ . Therefore the 128-dimensional local descriptor is transformed into

$$TPD = PD \cdot W_{LPP} \quad (10)$$

where  $TPD$  is a 1-dimensional local descriptor. In this paper, we simply set  $l=60$ .

#### 2.4 Matching strategy for feature points

Given two or more images, a set of feature points that can be reliably detected in each image, and robust descriptors for those features, we next match feature points between images. Given the definition of our feature descriptor in Eq. (10), and two descriptors,  $TPD_i$  and  $TPD_j$ , the distance metric is a simple Euclidean distance metric

$$D = |TPD_i - TPD_j| = \sqrt{\sum_l (TPD_{i,l} - TPD_{j,l})^2} \quad (11)$$

Consequently, we compare the ratio of the nearest neighbor distance  $D_{nn}$  to the second nearest neighbor distance  $D_{2nn}$  with a threshold  $T$  on the match, and discard matches with a ratio above the threshold. We simply set  $T=0.8$  in this paper.

### 3. Experimental results and analysis

In this section, several sets of experiments are presented to demonstrate the validity of the proposed method. Figure 1 shows six sets of test images with five geometric and photometric transformations for different scene types, i.e., viewpoint changes [(a) and (b)], image blur [(c) and (d)], lighting change (e), zoom-rotation (f). Each set of test images contains six images, in which the first image, i.e., the correspondingly leftmost one in Fig. 1, is the reference image and the others are the training ones with various degrees of geometric or photometric transformations. Besides, 100 additional real-world images, which are randomly sampled from a public data set, is used to train the projection matrix using LPP. In this process, keypoints in different images are first extracted and described with the obtained projection matrix. Then their matches are identified to see whether the descriptor is robust enough to find correspondences in various conditions. Besides the proposed method, another matching algorithms,

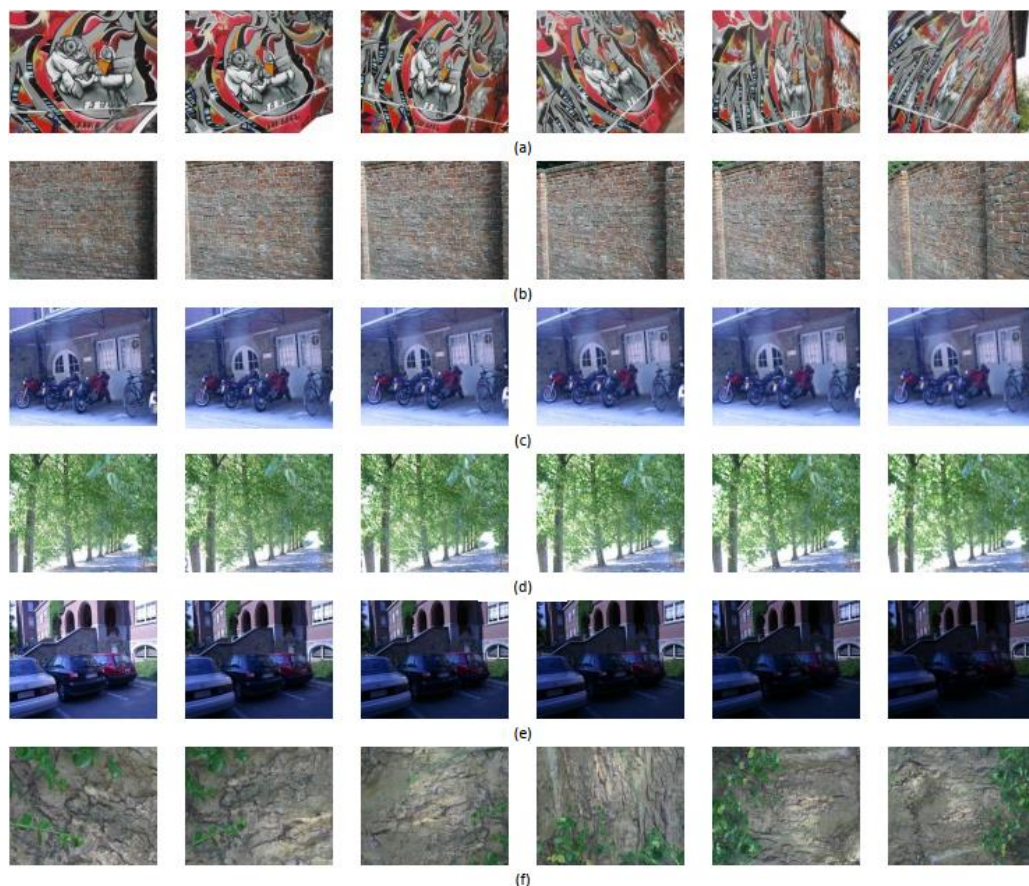
including the SIFT, are performed for comparisons. In addition, two metrics, i.e., the number of correct matches and correct matching rate, are employed to evaluate different matching algorithms. A match is deemed as the correct one if the distance between the predicted location is less than 4 pixels by the provided homography for each pair of relative images. The correct matching rate is defined as the ratio between the number of correct matches and the overall number of matches between the pair of images.

The performance is measured for images with a significant amount of blur by using bikes and trees datasets. The bikes dataset comes from structured scenes, and the trees one textured ones. Blur was introduced by changing the camera focus. Fig. 2 shows the results for the structured scene and Fig. 3 for the textured scene. The images are displayed in Figs. 1(c) and 1(d), respectively. The results show that the proposed method gains much higher scores than SIFT for most of parameter combinations of 1, 2 and 3 (except  $1 = 1, 2 = 0$  and  $3 = 0$ )

The performance is measured for the remaining five sets of images, the results show that the proposed method gains higher scores than SIFT for most of parameter combinations of 1, 2 and 3 (except  $1=1, 2=0$  and  $3=0$ ) for both types of images.

#### 4. Conclusion

This paper proposes a blur image processing method combining SIFT and LPP, which makes full use of the advantages of SIFT and LPP. It improve the differentiability of SIFT descriptor with the aid of three scale invariant concentric circular regions, but also improves the matching efficiency with the help of LPP. Experimental results show that this method can not only get higher matching point pairs, but also has higher correct matching rate for images with blur changes.



*Fig. 1 Data set.*

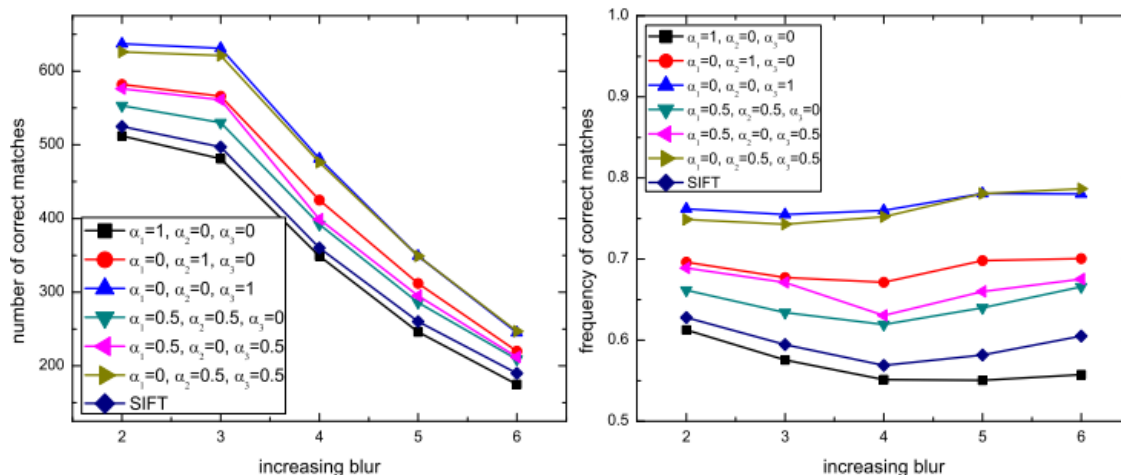


Fig. 2 (bike sequence). (a) number of correct matches. (b) correct matching rate.

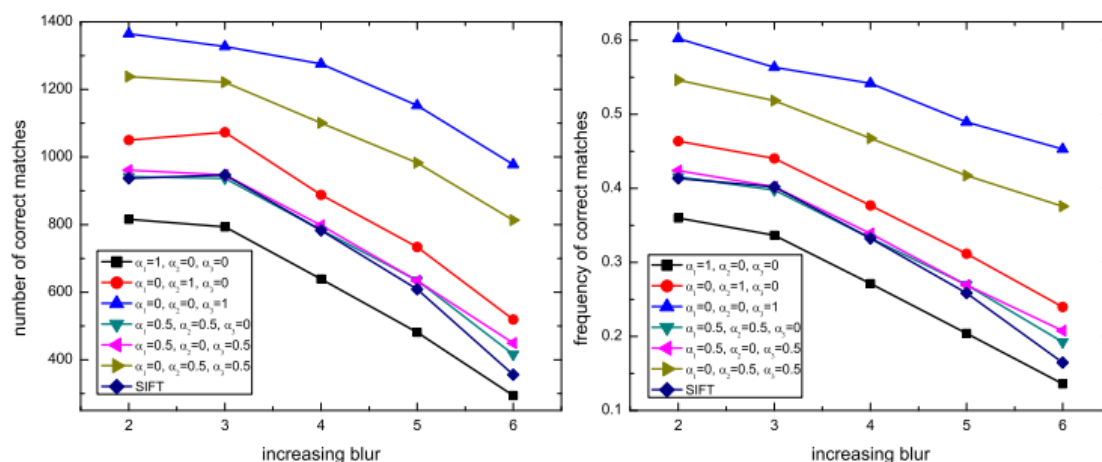


Fig.3(tree sequence). (a) number of correct matches. (b) correct matching rate.

## References

- [1] Szeliski, R., *Image alignment and stitching: A tutorial. Foundations and Trends in Computer Graphics and Vision* 2 (1), 1–104(2006).
- [2] Brown, M., Lowe, D. G., *Automatic panoramic image stitching using invariant features. International Journal of Computer Vision* 74 (1), 59–73(2007).
- [3] Ferrari, V., Tuytelaars, T., Van Gool, L., *Simultaneous object recognition and segmentation by image exploration. In: Computer Vision-ECCV 2004. Springer, pp. 40–54(2004).*
- [4] Lowe, D. G., *Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision* 60 (2), 91–110(2004).
- [5] Schaffalitzky, F., Zisserman, A., *Multi-view matching for unordered image sets, . In: Computer VisionECCV 2002. Springer, pp. 414–431(2002).*
- [6] Tuytelaars, T., Van Gool, L., *Content-based image retrieval based on local affinity invariant regions. In: Visual Information and Information Systems. Springer, pp. 656–656(1999).*
- [7] Kratochvil, B., Dong, L., Zhang, L., Nelson, B., *Image-based 3d reconstruction using helical nanobelts for localized rotations. Journal of Microscopy* 237 (2), 122–135(2010).
- [8] Dorko, G., Schmid, C., *Selection of scale-invariant parts for object class recognition. In: Computer Vision, Proceedings. Ninth IEEE International Conference on. IEEE, pp. 634–639(2003).*
- [9] Fergus, R., Perona, P., Zisserman, A., *Object class recognition by unsupervised scale-invariant learning. In: Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on. Vol. 2. IEEE, pp. II–264(2003).*
- [10] Leibe, B., Schiele, B., *Interleaving object categorization and segmentation. Springer( 2006).*
- [11] Opelt, A., Fussenegger, M., Pinz, A., Auer, P., *Weak hypotheses and boosting for generic object detection and recognition. In: Computer Vision-ECCV 2004. Springer, pp. 71–84(2004)*
- [12] Lowe, D. G., *Object recognition from local scale-invariant features. In: Computer Vision, 1999.*

- The Proceedings of the Seventh IEEE International Conference on. Vol. 2. Ieee, pp. 1150–1157(1999).*
- [13] Morel, J.-M., Yu, G., *Asift: A new framework for fully affine invariant image comparison. SIAM Journal on Imaging Sciences* 2 (2), 438–469(2009).
- [14] Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T., Gool, L. V., *A comparison of affine region detectors. International Journal of Computer Vision* 65 (1), 43–72(2005).
- [15] Mikolajczyk, K., Schmid, C., *A performance evaluation of local descriptors. Pattern Analysis and Machine Intelligence, IEEE Transactions on* 27 (10), 1615–1630(2005).
- [16] Freeman, W. T., Adelson, E. H., *The design and use of steerable filters. IEEE Transactions on Pattern analysis and machine intelligence* 13 (9), 891–906(1991).
- [17] Koenderink, J. J., van Doorn, A. J., *Representation of local geometry in the visual system. Biological cybernetics* 55 (6), 367–375(1987).
- [18] Van Gool, L., Moons, T., Ungureanu, D., *Affine/photometric invariants for planar intensity patterns. In: Computer VisionECCV'96. Springer, pp.642–651(1996).*
- [19] Harris, C., Stephens, M., *A combined corner and edge detector. In: Alvey vision conference. Vol. 15. Manchester, UK, pp. 50–80(1988).*
- [20] Mikolajczyk, K., Schmid, C., *Indexing based on scale invariant interest points. In: Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. Vol. 1. IEEE, pp. 525–531(2001).*
- [21] Ke, Y., Sukthankar, R., *Pca-sift: A more distinctive representation for local image descriptors. In: Computer Vision and Pattern Recognition, on. Vol. 2. IEEE, pp. II–506*
- [22] Tenenbaum, J. B., De Silva, V., Langford, J. C., *A global geometric framework for nonlinear dimensionality reduction. (2000), 2319–2323.*
- [23] Roweis, S. T., Saul, L. K., *Nonlinear dimensionality reduction by locally linear embedding. Science* 290 (5500), 2323–2326(2000).
- [24] Saul, L. K., Roweis, S. T., *Think globally, fit locally: unsupervised learning of low dimensional manifolds. The Journal of Machine Learning Research* 4, 119–155(2003).
- [25] Belkin, M., Niyogi, P., *Laplacian eigenmaps for dimensionality reduction and data representation. Neural computation* 15 (6), 1373–1396(2003).
- [26] Brand, M., *Charting a manifold. In: Advances in neural information processing systems. pp. 961–968(2002).*
- [27] Zha, H., Zhang, Z., *Isometric embedding and continuum isomap. In:ICML. pp. 864–871(2003).*
- [28] Niyogi, X., *Locality preserving projections. In: Neural information processing systems. Vol. 16. pp. 153–(2004).*
- [29] Lindeberg, T., *Feature detection with automatic scale selection. International Journal of Computer Vision* 30 (2), 79–116(1998).
- [30] Cao P , Rui T , Zhang J L , et al. *An improved SIFT matching algorithm based on locality preserving projection LPP[C]// Proceedings of the 4th International Conference on Internet Multimedia Computing and Service. ACM, (2012).*