Three-dimensional reconstruction of binocular stereo vision based on improved SURF algorithm and KD-Tree matching algorithm

Jie Cheng
Informatization Center
Tianjin Polytechnic University
Tianjin, China
cj0363@163.com

Journal of Digital Information Management

ABSTRACT: *With the constant development of computer science and technology, binocular stereo vision as a special form of computer vision plays a vital role in implementing computer change detection, image correction, three-dimensional reconstruction, and is widely applied in computer vision fields such as aerial mapping, visual navigation, motion analysis and industrial inspection. However, as it is difficult to perform binocular stereo vision, and precise parallax principle and mathematical method are required. Thus this study rapidly and effectively realizes three-dimensional reconstruction of binocular stereo vision by feature description vector generated by improved surface (SURF) algorithm combined with K-dimension tree (KD-Tree) searching. Effective combination of improved SURF algorithm and KD-Tree significantly enhances sense of reality of three-dimensional scene.*

Subject Categories and Descriptors
I.3.7 [Three-Dimensional Graphics and Realism]: Three-dimensional Reconstruction; **I.6 [Simulation and Modeling]**: Vision Model

General Terms: Gamma correction, Three-dimensional reconstruction

Keywords: SURF algorithm; KD-Tree matching; binocular stereo vision; three-dimensional reconstruction

Received: 27 August 2015, Revised 5 September 2015, Accepted 12 September 2015

1. Introduction

Binocular stereo vision as an approach to recognize three-dimensional world has become an important way to acquire scene information of objective world with the boom ing computer software. At the current stage, a wealth of scholars have warmly discussed and studied acquiring scene three-dimensional information and reconstructing vivid three-dimensional model by computer stereo vision. For instance, Yoon and Kweon proposed a self-adaptive weight matching algorithm based on Euclidean distance between pixel and color information deciding the weight between pixel [1], and the algorithm remarkably improves matching precision of local matching algorithm. Macro Vanettl, Gallo et al. studied stereo matching using relevant theories of neural network, reducing the excessive dependence on gray level of pixel and obtaining good robustness [2]. A matching algorithm based on minimum spanning tree put forward by Q. Yang from City University of Hong Kong achieves sound effect in calculating both precision and time [3]. This study researches an image matching algorithm based on improved K-dimension tree (KD-Tree) searching and SURF feature. The algorithm firstly extracts SURF features of image acquired and generates feature description vector, and then establish KD-Tree searching for these feature description vector. Improved surface (SURF) algorithm works faster in detecting features, and approximate nearest neighbor searching (ANNS) of KD-Tree searching greatly reduces computation burden and speeds up matching of SURF algorithm. Thus the algorithm is proved to have high practical and reference value, which lies solid basis for studying three-dimensional reconstruction of binocular stereo vision.

2. Mathematical Model of Binocular Stereo Vision

2.1 Parallel Bbinocular Stereo Vision Model

Stereo vision system formed when optical axises of two cameras are parallel is called parallel binocular stereo

vision. Model is shown in figure 1. Two cameras used in parallel binocular stereo vision system usually have same parameters.

In the figure, optical axises of two cameras are parallel, imaging planes are on the same focal plane, O1 and O2 are optic centers; line between O1 and O2 is as base line L, f stands for focal distance of two cameras. Suppose that projection point of a point p (X, Y, Z) is P1 (x1, y1) and P2 (x2, y2) on two cameras, and imaging process is a process of a object point converting from three-dimensional space to two-dimensional image.

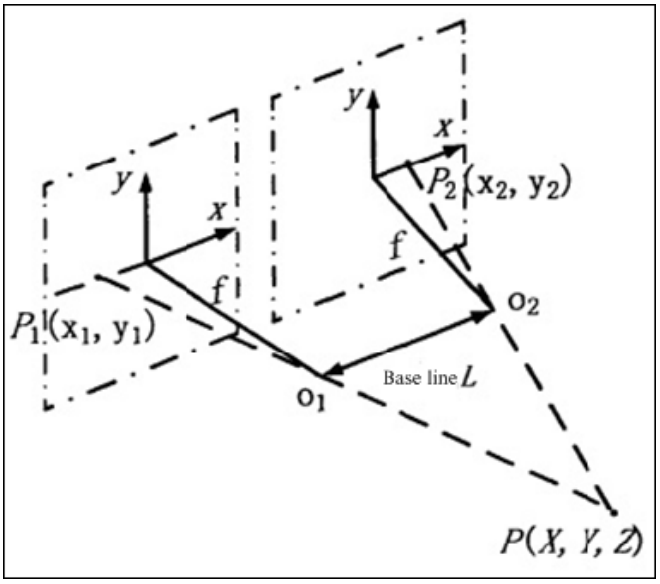

Figure 1. Parallel binocular stereo vision model

2.2 General model of binocular stereo vision

Parallel binocular stereo model is a special vision model. In common system models, two cameras cannot be parallel as expected, but binocular distance measurement can be realized by mathematical method, i.e., acquisition of depth information. General model of binocular stereo vision is shown in figure 2.

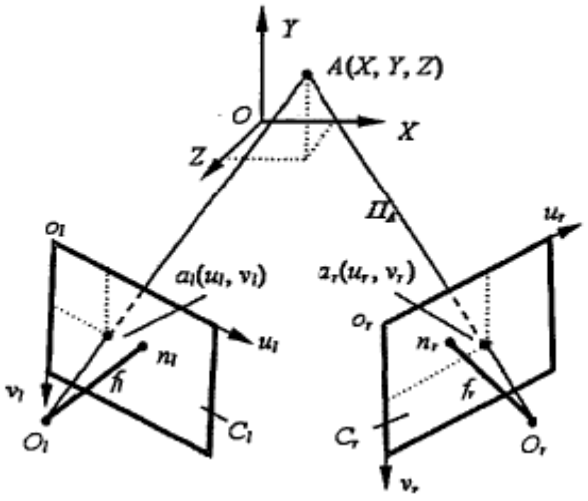

Figure 2. General model of binocular stereo vision

3. Stereo Matching Algorithm Based On Improved Surf

3.1 Stereo matching standard

(1)Unique constraint: one feature point of a matched image is that there is only one matching feature point on the other matching image.

(2) Epipolar [4] constraint: any matching point on a matched image may lies in the epipolar in the other image. Epipolar constraint urges the searching scope of feature point shrink from two-dimensional image to a straight line, thus to improve matching efficiency.

(3) Similarity constraint: if spatial point is on the outline object, then the projection point on the image is also on the outline and meanwhile, the projection points and fields are similar in gray and gradient change.

(4) Sequential constraint: order of projection of objects on two images remains unchanged.

(5) Continuity constraint: projection point of some point on the surface of an object is continuous and parallax along the boundary is also continuous.

(6) Parallax scope constraint: human vision restrains parallax in stereo vision, thus limit searching scope.

(7) Left and right consistency: some pixel point P in left image has a corresponding pixel point P" in right image; in turn, the corresponding point of point P" in right image is P in left image. Usually, this constraint condition is used when the area is covered.

3.2 Feature Description Vector of Improved SURF

To realize the invariance of local feature in rotation, a directional reference is required for feature points detected. First, we design a 4s*4s Haar wavelet [5] template, perform Haar wavelet computation in direction X and Y within a circular neighborhood of some feature point with a

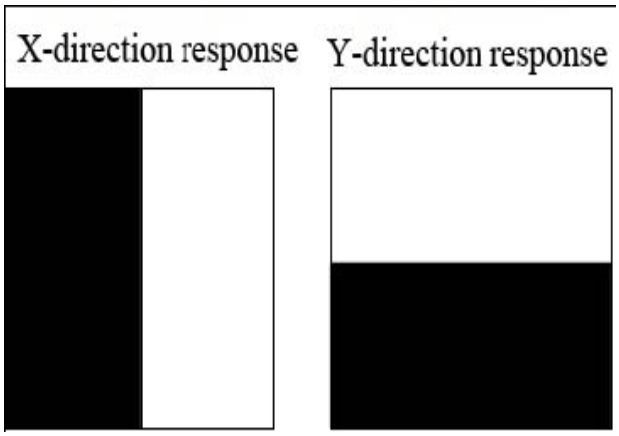

Figure 3. Response template of Haar wavelet in X and Y direction

radius of 6s (s is the scale of a feature point) (figure 3), and make Gaussian weighting ($\sigma = 2s$) on response value. Next, a fan-shaped template with 60-degree central angle is designed to rotate the circular neighborhood of the feature point taking the center of the circular as the axis.

Different Haar response can be obtained from sliding windows during rotation and the sum forms a vector. Rotating angle corresponding to the maximum of the sum of Haar response is taken as the reference direction of the feature point.

After the reference direction of the feature point is conformed, a rectangular region (20s*20s) is selected along the reference direction taking feature point as the center, as shown in figure 4. Then the rectangular region is divided into 16 small regions (4*4) and the side length of every small region is 5s. A Haar template (2s*2s) is designed to calculate response value of every small regions in X and Y reference direction (dx, dy). Gaussian weighting (σ=3.3s) is calculated for response values as well. Finally, four gradient values Σd_x, Σd_y, $\Sigma|d_x|$ and $\Sigma|d_y|$ of every small region are calculated. Four-dimensional vector of every small region is as follows:

$$V_{subkblock} = |\Sigma d_x, \Sigma d_y, \Sigma|d_x|, \Sigma|d_x|| \quad (1)$$

Every feature points has 16 small regions (4*4) whose dimension is 4; therefore, SURF feature descriptor formed by feature vector in 4*(4*4)=64 dimensions is developed.

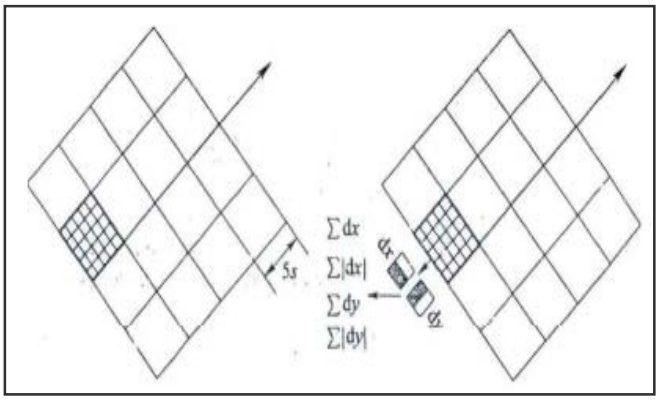

Figure 4. Generation of SURF feature descriptor

Besides scale and rotation invariance, SURF can keep good invariance under the condition of changed light. Generally, longer feature vector carries larger information amount and has SURF descriptor with better peculiarity; therefore, dimension of SURF descriptor is extended.

4. Overview of KD-Tree based Matching Algorithm

4.1 KD-Tree based Matching Algorithm

KD-Tree is one of balanced binary search tree [6]. Matched data structure is established by searching the most similar pair with descriptors. Time complexity of KD-Tree is 0. Every node stands for one point in K-dimensional space, nodes of left subtree are all less than or equal to the values they represent, while that of right subtree are more than the values. If farther node standards for segmentation of ith dimension, then child node represents segmentation of i+1th dimension. The segmentation is stopped, when points of some node are less than the largest point number given.

A premise of constructing KD-Tree is to select a sample $ex\ (d, r) \in \text{exset}\ (d_{split}(split := 0)$ meet the condition that divide the sample into two parts with same number). Make $\text{exset1} = \text{exset} - \{e|e = ex\}$. The remaining part is divided into two parts according to the following formula, and then tree construction is carried out. At this moment, split: = 1, i.e., selecting the second dimension as dividing criteria. It continues until exset = null or solit: = K-1.

$$exsetleft := \{(d', r') \in exset' \mid d'_{split} \le d_{split}\} \quad (1)$$

$$exsetright := \{(d', r') \in exset' \mid d'_{split} \rangle d_{split}\} \quad (2)$$

Where $\text{ex}\ (d, r)$ is one point in d dimensional space, r is its value in the space, exset is the whole sample set, $\text{dsplit}(split: = 0)$ stands for the value of sample in split dimension (currently, it is 1st dimension); one dimension is one attribute of multi-dimensional space; exsetleft is left subtree and exsetright is right subtree.

Specific construction procedures are as follows:

(1)Sort n K-dimension vector (X11,Xl2,....,X1k),(X21,X22,...,X2k),…,(Xnl,Xn2,...,Xnk) in sample space as per some dimension (i=1,2…k);

(2)Then divide all points into two parts based on median, to make the number of one part more than median and the other part less than median;

(3)Establish KD-Tree of left and right point set for next dimension according to the procedures above.

4.2 Matching Algorithm of Improved KD-Tree

KD-tree matching algorithm is still defective though nearest neighbor can be found rapidly through it, which greatly speeds up searching. But usually near point can be matched instead of the nearest point when the sample set is of large amount. As KD-Tree focuses on node checking and only a small part of nodes meet the requirements, which greatly lowers efficiency of the algorithm.

In view of matching accuracy and algorithm efficiency improvement, KD-Tree can be improved as follows: after some dimension is obtained, if the sample size is not large, then we can directly traverse the remaining sample points to prevent error dividing; next priority ranking is performed on dimension taking correlation between dimensions as selection criteria (points having little influence on other dimensions are ranked in latter part while points with large influence rank in front); unnecessary searching works are reduced by restraining the number of leaf nodes in KD-Tree.

Nearest neighbor searching by improved KD-Tree is as follows:

(1) Obtain K-dimensional data point set and establish KD-

Tree;

(2) Next is the priority ranking of dimension. The first step is to scan root node. If a node is searched in some direction, then it is put in priority queue and position of the node and its distance with detection node are recorded as node information; then delete points with the smallest distance in current dimension and other branch child nodes of the nearest neighbor node are searched until the scan times reach upper limit or there is no data in queue;

(3) The final step is determination. Compared the distance of nearest neighbor point and next nearest neighbor with the target point, if it is less than the given threshold, then the nearest neighbor node is taken as the result of searching; otherwise, return null.

4.3 Experimental Contrast Analysis

Searching method	Time (second/s)	Accuracy (%)
KD-Tree searching method	1.748	93.11
Improved KD-Tree searching method	1.607	93.68

Table 3. Comparison of searching performance

It can be seen from table 3 that, when influence on matching accuracy is not large, improved KD-Tree searching method achieves the best effect, with time less than the other method. It proves that, improved KD-Tree searching method shortens time by reducing computation burden, which improves matching speed on the premise of ensuring precision.

5. Three-dimensional Reconstruction of Binocular Stereo Vision

5.1 Basic Principle of Three-dimensional Reconstruction

Three-dimensional reconstruction of binocular stereo vision is an inverse process mapping spatial point in three-dimensional scene to two-dimension image. Coordinate of mapping point in three-dimensional space is obtained through reciprocal transformation of parallax principle and definition of world coordinate system.

(1) Reconstruction of Spatial Point

Spatial point is the basic unit in three-dimensional space. In three-dimensional stereo space, points can be connected to be a line, and then lines can form a plane. If coordinates of all points on the surface of space object can be obtained, surface shape of three-dimensional object and position of object in the space can be reconstructed through connecting points together.

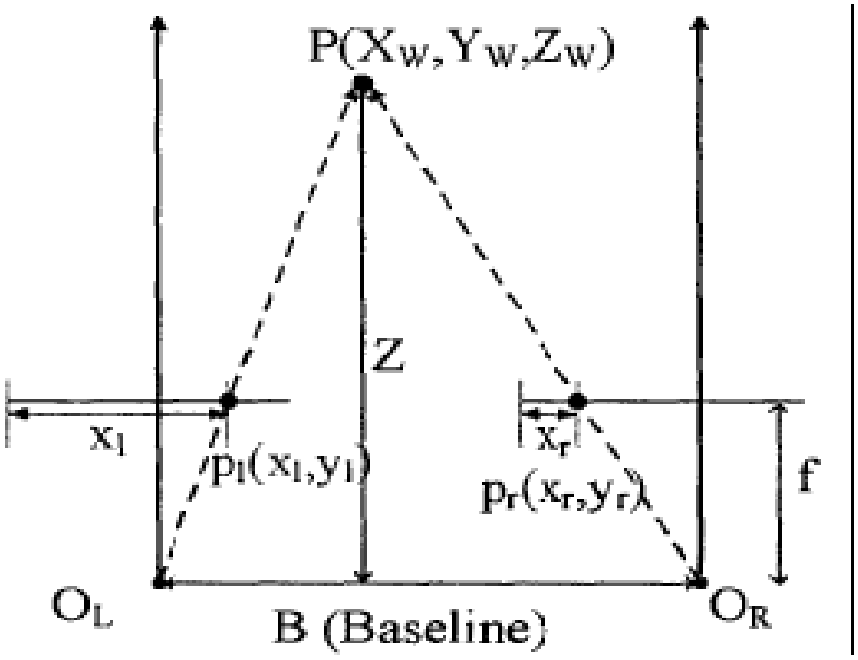

Figure 4. Parallax principle

Internal parameters of camera and external parameters of binocular camera can be obtained by camera calibration, as shown in figure 4. Due to epipolar constraint of stereo matching, parallax only exists in horizontal direction when stereo matching is performed on images processed by binocular correction.

(2) Texture Mapping

Three-dimensional model restored by disparity map is composed of single hue without any texture color. Such three-dimensional model lacking of reality cannot meet the vision demand of people on three-dimensional model. Thus it is necessary to make texture mapping on the reconstructed three-dimensional model [8].

In three-dimensional reconstruction model of binocular stereo matching system, standard color two-dimension image matched is the texture picture of the reconstructed three-dimensional mode. Every pixel point in two-dimension image is a texture primitive and every texture primitive preserves color information of every corresponding point in three-dimensional model. In the process of three-dimensional reconstruction of binocular stereo matching system, color texture of the image is directly assigned to three-dimensional spatial point and meanwhile the plane of model is reconstructed by interpolation smoothing processing. By doing that, color texture mapping of three-dimensional model is completed.

5.1 Three-Dimensional Reconstruction Effect of Standard Image

In reconstruction effect picture of OpenGL three-dimensional point cloud, every three-dimensional coordinate is the confirmed coordinate of a point in world coordinate system, and three-dimensional reconstruction effect of scenes can be observed in different observation location and angle. In this study, three-dimensional reconstruction model is obtained by performing interpolation smoothing processing on neighboring spatial points. Simulation experiment demonstrates that, three-dimensional structure of scene can be clearly observed from simulation result picture, but relative stereo position between spatial points is directly confirmed by parallax values of the image and real position relationship is lack of between spatial points. However, as the viewpoint changes, three-dimensional reconstruction vision picture has the same effect with which is observed by eyes of human, with high reality and strong stereoscopic sensation, which scientifically verifies the effectiveness of three-dimensional reconstruction.

6. Conclusion

Computer vision is an emerging subject developed from various subjects such as image processing, computer graphics and pattern identification. Binocular stereo vision as an important branch of computer vision directly simulates scenery processing method of human vision. It is of wide application prospect in robot vision navigation, aerospace, geological testing, machine manufacturing and virtual reality. Thus three-dimensional reconstruction tech nology research using theory and method of binocular stereo vision is found with important theoretical and application value. Based on deep understanding of SURF algorithm, this study improves it, then uses the improved one to complete feature point detection and expands KD-Tree to high-dimensional data level by BBF (Best Bin First) searching algorithm based on KD-Tree. Matching algorithm combining SURF and KD-Tree is realized in this study. Experiments prove that, the improved algorithm has high precision and sound real-time performance, which provides a new solution for three-dimensional reconstruction research.

References

[1] Yoon, K J., Kweon, I S (2006). Adaptive support-weight approach for correspondence search, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28 (4) 650-656.

[2]Vanetti, M., Gallo, I, Binaghi, E (2009). Dense two-frame stereo correspondence by self-organizing neural network. Image Analysis and Processing–ICIAP 2009. Springer Berlin Heidelberg, 1035-1042.

[3] Yang, Q (2012). A non-local cost aggregation method for stereo matching. C. Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 1402-1409.

[4] Ma, Songde., Zhang, Zhengyou (1998). Computer vision – Computation theory and algorithm basis. Science Press.

[5] Cui, Zhenxing., Zeng, Wei., Yang, Mingqiang (2014). A improved SURF fast matching algorithm. *Journal of Jiangsu Normal University* (Natural Science Edition), 32 (3) 41-46.

[6] Du, Zhenpeng., Li, Dehua (2012). Image matching algorithm research based on KD-tree search and SURF features. *Computer & Digital Engineering*, 40 (2) 96-98.

[7] Xiong, Yunyan., Mao, Yijun., Huaqing Min, Application of the ordered KD-tree on the image features matching. Control and Instruments In Chemical Industry, 37 (10) 84-87, 2011.

[8] Chen, Qiang (2015). 3D reconstruction based on binocular stereovision. *Modern Computer,* (1).