MOTION—STEREO DOUBLE MATCHING RESTRICTION IN 3D MOVEMENT ANALYSIS

ZHANG Chun-sen

Dept of Survey, Xi'an University of Science and Technology, No.58 Yantazhonglu, Xi'an 710054,China -zhchunsen@yahoo.com.cn

Commission III

KEY WORDS: Image matching, Stereo-motion, Sequence image, 3D reconstruction, 3D motion analysis

ABSTRACT:

In order to realize the correspondence of arbitrariness object-side feature points of movement object in the different time in 3D motion analysis based on binocular sequence images, Aiming at the fact that motion and stereo matching exists simultaneously, motion and stereo matching of images in the process is studied. The algorithm of double restriction matching combining motion and stereo image matching is presented. The basic image matching based on point feature is completed by correlation coefficients and relaxation algorithm, and the feature point's correspondence of movement object is performed by motion—stereo double matching restriction of binocular sequence image. Combining the results of the camera calibration, using the triangulation process for reconstruction feature points of moving object 3D coordinate from binocular sequence images, the method guarantees the correspondence of arbitrary object-side feature point of the move object at different time. And using these object-side sequence "image" (coordinate) accomplishes 3-D object tracking location. A set of experimental results of real data are presented, it shows that the accuracy of the final correspondence is about 76.5%,which can meet the requirements of the 3D motion object tracking location based on point feature.

1. INTRODUCTION

The problems of 3D (three-dimension) motion estimation from visual data are among the most challenging problems in computer vision. The way of 3D motion estimation includes the method of monocular sequence image and stereo (or binocular) sequence image, but there are distinctions in the complexity of computing and accuracy of computing result. When monocular sequence image estimation is being used, only relative moving information can be gotten, and it's exist a scale factor related to the structure information. In order to get the structure information, the traditional approach is having the aid of the equipment of range finder. But it is difficulty to integrate the different equipment or eliminate the system error. This paper presents an approach of integrating the two cues of motion and stereo when two cameras that takes pictures of moving object repeatedly. Multi-ocular cues or binocular sequence images approach, compared with single sequence images, not only the computation is simple, but also the absolute translation in space (structure information) can be acquired. But it requires solutions of these sub-problems: the *images matching problem*, referred to stereo matching in different sequence which at the same time and motion (sequence) matching in same sequence which at the different time, the *reconstruction problem*, in which 3D information is to be reconstructed from the correspondences, and the *features correspondence problem*, referred to different time correspondence of feature points in object side sequence "image". All of these are the most important and difficult things in computer vision.

Aiming at the fact that motion and stereo matching exists simultaneously in 3D movement estimation based on binocular sequence images, the author studies the motion and stereo matching of images in the process. The algorithm of double matching restriction combining motion and stereo image matching is presented. Matching of images based on point feature is completed by correlation coefficients and relaxation algorithm. Feature point's correspondence of movement object is performed by motion-stereo double matching restriction of binocular sequence image. Combining the results of the camera calibration, using the triangulation process for reconstruction feature points of moving object 3D coordinate from binocular sequence images, the method guarantees the correspondence of arbitrary object-side feature point of the move object at different time. And using these object-side sequence "image" (coordinate) accomplishes 3-D object tracking. An experimental result of real data by means of this algorithm is presented in the article. The result indicates that the accuracy of the final correspondence is about 76.5%. It can satisfy the 3D motion object tracking location based on point feature.

2. DOUBLE MATCHING RESTRICTION

2.1 Object feature points extracting and initial matching

Extracting feature points from image is the first step of feature image matching. This paper uses Harris operator to extract arbitrariness feature points on sequence image. The experiment shows that this operator is simple, stable, and insensitive to noise, illumination and so on. It can also extract as ration. And the distribution of feature points extracted is rational. The aim of initial matching is gotten a matching candidate set T. The correlation coefficient method is used in this paper [2]. Namely, for each feature point which $m_1 \in$ image *a*, $m_2 \in$ image *b*. Their image coordinates are supposed to be (u_1,v_1), (u_2,v_2).If the difference between the coordinate of m_1 and the coordinate of m_2 dose not exceed a certain threshold, calculates correlation coefficient of $(2n+1)\times(2n+1)$ window which is m_1, m_2 as centre. A pair of points is presented, if they are considered as matching points candidate, the correlation coefficient must be greater than a certain threshold. The matching candidate relation between a certain feature point in image and some feature points in image b is established. And the pair of points is added to matching candidate set T.

2.2 Relaxation algorithm based on matching support measure

The law of relaxation is to allow the candidate match pair in T to dismiss oneself and to automatically match each other through iterative so as to make the "continuity" and "uniqueness" to obtain biggest satisfaction. The continuity refers to the massive other correct match pair usually existing in the neighborhood of correct match pair; Uniqueness refers to the identical feature point existing in only one matched pair. Or it can be expressed as the phenomenon that if candidate matching is right, there must be many candidate matching around it, while if candidate matching is wrong, there are less candidate matching around it. Matching support is defined as the degree that the neighbour candidate supports the candidate matching. It means that the strongest the matching support is, the more possible that the candidate matching is true.The detailed calculation is as below[2], [3]:

Suppose there are two feature points sets: $P = \{P_1, P_2, \cdots P_m\}$ and $Q = \{Q_1, Q_2, \cdots Q_n\}$, Define relative excursion between the two feature points sets for each paired points (P_i , Q_j). $\delta_{ij}(h,k)$ is relative distance between P_i, P_h and Q_j, Q_k when P_i and Q_j partner (only shift).

$$\delta_{ij}(h,k) = \frac{|\, d(P_i, P_h) - d(Q_j, Q_{kl})\,|}{dist(P_i, P_h; Q_j, Q_{kl})} (l = 1,2...) \qquad (1)$$

Here:

$d(P_i, P_h) = \| P_i - P_h \|$ is the Euclid distance between P_i and P_h.

$d(Q_j, Q_{kl}) = \| Q_j - Q_{kl} \|$ is the Euclid distance between Q_j and Q_{kl}.

$dist(P_i, P_h; Q_j, Q_{kl})$ = $[d(P_i, P_h) + d(Q_j, Q_{kl})]/2$ is the average distance of the two pairing. Suppose $|\, \delta_{ij}(h,k) = 0 \,|$ that means Q_k relative to Q_j and P_h relative to P_i have the same meaning. So points (P_h , Q_k) should sustain (P_i , Q_j) as it the maximum. Along with $|\, \delta_{ij}(h,k) \,|$ increase its support measure reduces.

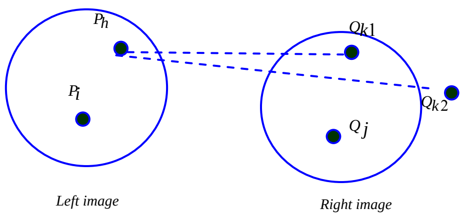

Figure1.Illustration of matching support measurein relaxation matching

From the figure 1, Q_j is the matching candidate feature point of P_i. It is reasonable require when (P_i , Q_j) is a good match. If P_i and Q_j are the correct matching, the feature points distribution in their adjacent field should be similar. If there is a point P_h in the adjacent field of P_i, in the matching candidate point Q_{k1} , Q_{k2} of P_h there will be some point or all points in the adjacent field of Q_j. Defining support measure of (P_h , Q_k) to (P_i , Q_j) is

$$\phi(|\, \delta_{ij}(h,k) \,|) = \frac{1}{1 + |\, \delta_{ij}(h,k) \,|^2} \qquad (2)$$

When P_i partner Q_j , P_h partner only with Q_k .that is relative with P_h and the maximize support measure to (P_i , Q_j) is Q_k alone. The support measure from formula:

$$\max_{k \neq j} \phi(|\, \delta_{ij}(h,k) \,|) \qquad (3)$$

$$\phi(|\, \delta_{ij}(h,k) \,|) = \begin{cases} e^{-\delta_{ij}(h,k)/\varepsilon_r} \\ 0 \end{cases}$$

Here: Q_j is one of P_i matching candidate points, and $\left|\delta_{ij}(h,k)\right| < \varepsilon_r$, $\phi(|\, \delta_{ij}(h,k) \,|) = e^{-\delta_{ij}(h,k)/\varepsilon_r}$ in other case $\phi(|\, \delta_{ij}(h,k) \,|) = 0$. ε_r is the threshold of relative distance change. When accounting use the experiential value.

Because P_h does not have only one matching candidate point Q_{kl} , there will be more value of $\phi(|\, \delta_{ij}(h,k) \,|)$. $\max_{k \neq j} \phi(|\, \delta_{ij}(h,k) \,|)$ as the support measure of point P_h and its matching points (P_i, Q_j). In the actual account, there is not only one point in adjacent field of P_i. If $N(P_i)$ expresses the points set in adjacent field of P_i (without P_i), calculates the support measure that points of $N(P_i)$ to points (P_i, Q_j) one by one. Finally the average value after accumulative is the total initial support measure:

$$S^0(P_i, Q_j) = \frac{1}{m} \sum_{h \neq i} \max_{k \neq j} \phi(|\, \delta_{ij}(h,k) \,|) \qquad (4)$$

Where, m is the number of points in $N(P_i)$.

When calculating $S^0(P_i, Q_j)$, every points (P_h, Q_k) should be treated equally at first. Because there is no priori knowledge at beginning. After iteration for r times (r>0), the support measure that (P_h, Q_k) to (P_i, Q_j) does not only relies on difference of position between P_h and Q_k , but also on their value of $S^{r-1}(P_i, Q_j)$ which is the feedback of permission local support measure. The two factors can be combined together in different way. The least minimum is taken.

$$S^r(P_i, Q_j) = \frac{1}{m} \times \sum_{h \neq i} \max_{k \neq j} \min[S^{r-1}(P_i, Q_j), \phi(|\, \delta_{ij}(h,k) \,|)] \qquad (5)$$

This iteration continue until except the most possible point the support measure of rest points less than threshold which has already given to every P_i.

2.3 Double matching restriction

Stereo and sequence match simultaneously exist in the 3D feature correspondence movement analysis. The method and goal of these two matches are the same from the aspects of image processing. During the process of object-side 3D feature point's correspondence, both stereo-sequence match and sequence-stereo match can be applied. However, the different matching order will has different matching effect to the final moving object-side 3D feature point correspondence. In actual operation, the adjacent images took by one camera are similar, because the time interval between two adjacent images is very short. Thus, the sequence image match from same camera is easy, but because its base line between viewpoints is relatively short, 3D reconstruction is difficult. Therefore, the estimated depth is not precise in the situation when noise exists (it is even impossible when baseline is fairly short). With this correspondence, there is usually a certain baseline between different cameras, the object-side 3D reconstruction precision is fairly high among stereovision because the distance between viewpoints is large, but stereo matching is difficult, especially when huge disparity and image distortion exist. The double match restraint namely first extracts random feature points on moving objects from different-time but same-sequence images, determining the corresponding relationship of feature points among the same sequence images, establish the sequence match of binocular image sequences. If the image sequence sample density is appropriate, the reliability of the feature match in sequence image can be guaranteed. Then according to match corresponding point coordinates which obtained from the sequence match result matching with corresponding images of same-time different sequence (left and right images stereo match). Therefore, the difficulty of random stereo match can be decreased to a great level through this double match restraint. As a result, the whole correspondence of moving object random three-dimensional feature points can be obtained preferably.

The process of double matching restriction as figure 2 shows. The left and right images at time t_i are denoted by I_i and I'_i. A corresponds point in images plane I_i and I'_i are denoted by m_i and m'_i. The point Mi in 3D space expressed in the coordinate system attached to m_i and m'_i. $R_{i,i+1}$,$t_{i,i+1}$ is the rotation matrix and the translation vector describing movement object from at time t_i to t_{i+1}.R_{LR},,t_{LR} is the rotation matrix and the translation vector between the left and the right camera. The images matching between points m_i and m'_i is stereo matching, while the images matching between points m_i and m_{i+1}(or between points m'_i and m'_{i+1}) is motion matching. The correspondences between points Mi and M_{i+1} are features correspondences. Here i=3.

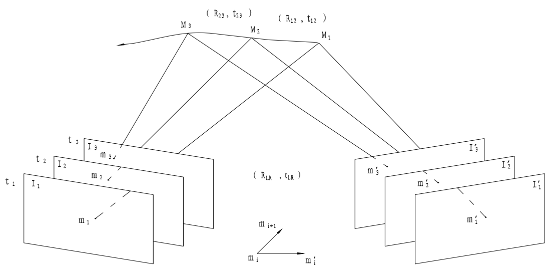

Figure2 motion-stereo double matching restriction

In order to finish double matching restriction, firstly the feature points are extracted from moving images at different the time (front and rear) simultaneously, and original matching table is created between the features of two images, the possible candidate match points of a feature are found in the other image. Based on the above matching methods and some hypothesizes, considering compatibility of feature matching in a certain range, finding out the best feature as the final matching result, confirming corresponding connection of feature points in sequence image and finishing move matching of double sequence image. As long as sampling consistency of sequence image is suitable, we can obtain the credible feature matching of moving sequence image; Secondly, carry on stereo matching between different sequence images and the corresponding feature point is searched on the other image for matching according to the preceding match result (coordinates). In order to advance calculation speed and matching precision, the strategy which examine the dynamic moving object, and limit the matched object on moving object is used before moving matching (screen the static background of moving object).

3. EXPERIMENTAL RESULTS OF REAL DATA

Using the established binocular stereo vision system in doors for this research, obtains binocular sequence motion images of model car in a certain way to move (as figure3). According the request of 3D movement object's location and tacking, the feature point's correspondence of movement object is performed by motion-stereo double matching restriction of binocular sequence image. Combining the results of the camera calibration, using the triangulation process for reconstruction feature points of moving object 3D coordinate from binocular sequence images, the method guarantees the correspondence of arbitrary object-side feature point of the move object at different time. The concrete steps as follows:

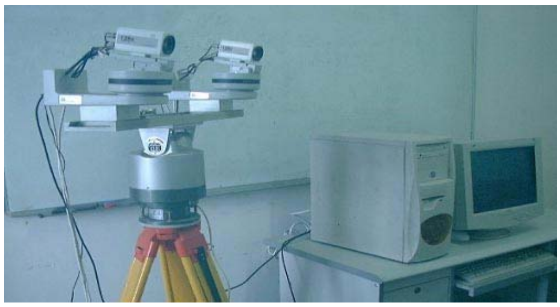

Figure3. Binocular vision system

①Doing motion matching for movement object binocular sequence images can obtain motion matching result which in the same sequence at time t_1 and t_2 (as figure4).The coordinate matching file as table 1. Where X, Y and X′, Y′ is the matching coordinate of the same feature point at time t_1 and t_2 respectively.

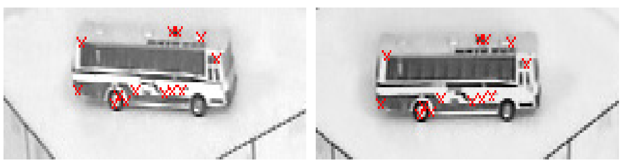

Figure4 Same sequence motion matching at time t_1 and t_2

Pixel	X	Y	X′	Y′
1	129.0	150.0	139.0	151.0
2	133.0	152.0	143.0	153.0
3	127.0	153.0	137.0	154.0
4	160.0	126.0	170.0	127.0
5	157.0	126.0	167.0	127.0
6	180.0	136.0	190.0	137.0
7	153.0	149.0	163.0	150.0

8	110.0	130.0	121.0	131.0
9	138.0	148.0	148.0	149.0
10	162.0	148.0	171.0	149.0
11	157.0	148.0	167.0	149.0
12	183.5	135.5	194.5	136.5
13	108.0	148.0	118.0	149.0
14	187.5	131.5	197.5	132.5
15	172.0	128.0	182.0	129.0

Table1 Image matching coordinate at time t_1 and t_2

② Using the matching result of step① to do stereo matching with corresponding images respectively. according to the matching result (coordinate) in the former moment to search corresponding feature points of the other sequence images at same time with the purposefully. Double matching restriction will be come true. Figure 5 and6 are the stereo matching result of corresponding images in different sequence at time t_1 and t_2 apart. Table 2 and 3 are the corresponding matching coordinate files apart.

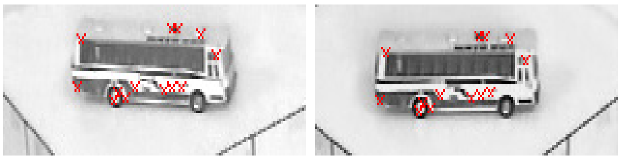

Figure5 Different sequence stereo matching at time t_1

Pixel	X	Y	X′	Y′
1	129.0	150.0	158.0	148.0
2	133.0	152.0	162.0	150.0
3	127.0	153.0	156.0	151.0
4	160.0	126.0	190.0	123.0
5	157.0	126.0	187.0	123.0
6	180.0	136.0	211.0	132.0
7	153.0	149.0	183.0	146.0
8	110.0	130.0	139.0	129.0
9	138.0	148.0	168.0	146.0
10	162.0	148.0	193.0	144.0
11	157.0	148.0	187.0	145.0
12	108.0	148.0	136.0	147.0
13	172.0	128.0	203.0	124.0

Table2 Stereo image matching coordinate at time t_1

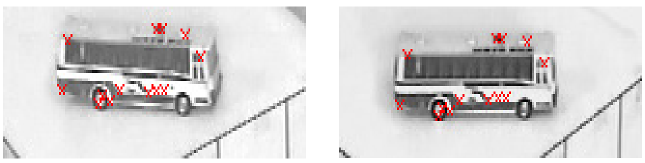

Figure6 Different sequence stereo matching at time t_2

Pixel	X	Y	X′	Y′
1	139.0	151.0	169.0	149.0
2	143.0	153.0	173.0	150.0
3	137.0	154.0	167.0	152.0
4	170.0	127.0	200.0	123.0
5	167.0	127.0	197.0	123.0
6	190.0	137.0	222.0	132.0
7	163.0	150.0	193.0	146.0
8	121.0	131.0	151.0	129.0
9	148.0	149.0	178.0	146.0
10	171.0	149.0	202.0	145.0
11	167.0	149.0	198.0	145.0
12	118.0	149.0	147.0	148.0
13	182.0	129.0	213.0	125.0

Table3 Stereo image matching coordinate at time t_2

③Based on the camera calibration parameters to calculate object-side spatial coordinates at time t_1 and t_2 which correspond with table 2 and 3. The result is shown in table 4 and 5.

Number	X (dm)	Y (dm)	Z (dm)
1	1.499	1.007	-0.128
2	1.435	1.074	-0.078
3	1.516	0.957	-0.065
4	1.041	1.626	-0.681
5	1.089	1.573	-0.682
6	0.613	1.896	-0.514
7	1.075	1.418	-0.178
8	1.809	0.713	-0.597
9	1.291	1.139	-0.209
10	0.867	1.546	-0.249
11	1.016	1.493	-0.198
12	1.860	0.642	-0.159
13	0.770	1.788	-0.688

Table4 Spatial coordinates of
moving object feature points at time t_1

Number	X (dm)	Y (dm)	Z (dm)
1	1.266	1.146	-0.143
2	1.213	1.222	-0.097
3	1.284	1.097	-0.079
4	0.886	1.803	-0.664
5	0.935	1.752	-0.664
6	0.374	2.022	-0.548
7	0.926	1.601	-0.156
8	1.587	0.887	-0.605
9	1.152	1.328	-0.181
10	0.719	1.698	-0.226
11	0.784	1.629	-0.226
12	1.639	0.791	-0.166
13	0.602	1.953	-0.669

Table5 Spatial coordinates of
moving object feature points at time t_2

Feature points' matching is get from correlation method, relaxation method in the process of finishing the move object feature points. Those methods use the elicitation knowledge such as the similar gray. At the same time because there are certain rotate angle and camera measure errors between left and right images, that the same point position in left image is different from the right image, and it can cause perspective projection errors. Moreover there are little difference between parameter in left camera and right camera, so the related points' gray are not completely the same, it cause no-adjust errors etc.

Because of the above factor it inescapable to cause the error matching. Although it takes lot of matching method to ensure the correspondence of moving object feature points such as taking the image correlation based feature, relaxation matching method based matching sustain, image relation based gray, and matching in local area etc, but there are still some error matching points. For example, in the above experiment, from the step ① to step③, the correct feature points' amount which obtained is decrease as the process. For example the feature points in ① feature extraction from same sequence image, there are17 feature points, the right matching points are 15, that is the number, in the process②, of different sequence stereo matching feature points are15, but after finishing the 3D object-side feature point correspondence, the correct number of feature points only are 13, the final correspondence is about 76.5%. So it add feature matching robust arithmetic in the program, such as using the mature bundle adjustment theory in photogrammetry field, to auto discover the error in feature matching or correspondence, and form a feed back control process, it worth to further study.

4. CONCLUSIONS

In order to realize the location and tacking of the three-dimensional object based on points feature, this paper discusses the most challengeable problem in this process that is the features correspondence problem. The algorithm of double restriction matching combining motion and stereo image matching is presented after analysis the character of sequence (motion) matching and stereo matching in the process. Take the advantages of high overlap of motion video frequency images, The final correspondence is about 75.6%, which is validated by real data experiment. It can meet the requirements of the 3D motion object tracking location. It will be the future study to give the further optimizing and robust algorithm of feature point correspondence.

REFERENCES

[1]R.Hartley.A.Zisserman Multiple View Geometry in Computer Vision [M] Cambridge University Press 2000.238-241

[2]ZHANG Chun-sen. Research of 3-D Motion Object Tracking Location Based on Stereo Sequence Images [D]. PhD Dissertation of Wuhan University, 2004(in Chinese)

[3]Z.Zhang, R. Derche, O. Faugeras and Q. T. Luong, A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry[J].Artificial Intelligence 1995,78(1-2): 87-119.

[4]MA Song-de,ZHANG Zheng-you. Computer Vision: Fundamentals of Computational Theory and Algorithms [M]. Beijing: Sciences Press, 1998:151-173(in Chinese)

[5]Pui-Kuen Ho and R. Chung. Stereo-Motion with Stereo and Motion in Complement [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000,2(2):215-220

[6]F. Dornaika and R. Chung. Cooperative Stereo-Motion: Matching and Reconstruction [J]. Computer Vision and Image Understanding, 2000, 79(3): 408-427

[7]Zhengyou Zhang. Motion and Structure of Four Points from One Motion of a Stereo Rig with Unknown Extrinsic Parameters [J].IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(12):1222-1227

Project supported by the National Science Foundation of China (No.40171080)