www.s2is.org
International Journal on Smart Sensing and Intelligent Systems
ISSN 1178-5608

MOVING TARGET DETECTION BASED ON GLOBAL MOTION ESTIMATION IN DYNAMIC ENVIRONMENT

GAO Jun-chai, LIU Ming-yong, XU Fei
School of Marine Science and Technology
Northwestern Polytechnical University,
XI'an 710072, China
Emails: gaomuyou@126.com

Submitted: Oct. 30, 2013 ***Accepted: Feb. 6, 2014*** ***Published: Mar. 10, 2014***

Abstract- AUV localization is not accurate based on sequence images if moving target is as landmark, so the moving target detection algorithm is studied based on global motion estimation, which detects and eliminates moving target according to the motion inconsistency of the moving target. Generally grid block matching is used in the global motion estimation, it can't effectively dispose the dynamic background, and the gradient direction invariant moments descriptors method of free circular neighborhood based on feature points is proposed, which is effective for the background rotating and light changing in two adjacent images. For the matching points, the parameters of global motion are estimated robustly combined with normalized linear estimation method and least median squares. Experiments show that the designed algorithm can effectively estimate parameters of global motion, and eliminate the motion target as mismatch.

Index terms*:* Moving target, global motion parameters, gradient direction, SURF operator, robust estimation

I. INTRODUCTION

While AUV navigating in an unknown environment, environment modeling and AUV (Autonomous Underwater Vehicle) localization are the basic problem to realize autonomous. For which, most researches are based on the assumption of static environment, and the real environment is changing dynamically, such as the waterweeds float and biological swimming， that will bring noise for the environment modeling and AUV localization, so the moving target need to be detected and eliminated.

For the moving target detection in dynamic background, researchers have carried out a lot of research works on the video monitoring system and processing technology, and made a lot of achievements [1-3]. Such as the Computer Department of Cornell University in United States designs a set of video detecting and tracking system of aerial photography[4], the system tracks feature point based on Kanade-Lucas-Tomasi algorithm and estimates robust affine parameters based on M-estimation, and then detects moving objects with the three frames differential. In which there are two main problems, the global motion estimation algorithm has large amount of calculation and parameters estimation deviation is caused by the moving objects in the video. In order to solve these, many scholars have done more researches. For the first problem, Konrad [5] proposed global motion estimation algorithm based on the Levenberg-Marquadet (LM) method, the noise removed with histogram of residuals method, the computational load is also relatively large, it is still difficult to achieve real-time, and the calculation results are sensitive to noise. Amir Averbuch and Yosi Keller [6] improved motion estimation method based on gradient, reduced the amount of calculation about two dozen times. And the G Sorwar [7] estimated global motion parameters based on the hypothesis that moving foreground are usually distributed in the center of the image. Finally those methods improve the calculation speed and exclude the interference of moving objects in the foreground in some degree. For a large number of scholars researches, moving target detection methods have got good performance under the specific conditions.

In recent years, a large number of researchers devoting to the research of motion object detection and tracking technology under complex background have emerged in China, many research institutions and universities also invested considerable research effort. Such as national key laboratory of pattern recognition of the institute of automation in the Chinese Academy of sciences, the graphic image Institute of Tsinghua University and the image processing & pattern recognition research Institute of Jiaotong University, they research on the contour detection algorithm to realize motion object tracking.

In this paper, AUV is located with monocular vision. During the AUV moving, static background changes largely in image, and meanwhile, the moving target is changes with the AUV too. If it locates with moving target, its position is not accuracy, so how to detect and ignore the moving target in the global motion estimation has been difficulty for locating AUV in dynamic background. The aim of moving target detection is not to track, but is to identify and no longer participates in AUV positioning, then a method based on global motion estimation is proposed, in which moving object matching is as false matches, and no longer need to be distinguished from the false matches.

The organization of the paper is as follows. After a general introduction of the moving target detection, the moving target detection principle based on global motion estimation has been discussed in section II. The gradient direction invariant moment extraction of feature point circular neighborhood and feature point matching strategy have been discussed in section III. In section IV, motion model, and robust model parameters estimate methods have been presented. The analysis and simulation results of moving target detection have been discussed in section V. The paper has been concluded in section VI.

II. MOVING TARGET DETECTION PRINCIPLE IN DYNAMIC BACKGROUND

During the AUV motion, the camera is moving, there is about half view field motion between adjacent frame’s static backgrounds, optical flow of static background is generated by camera motion, i.e. global motion. Optical flow of moving target is generated by the relative motion of the camera and moving target, i.e. global and local motion [8-10]. The optical flow of static background and moving target is different, so optical flow estimation is effective to detect moving target.

Optical flow estimation methods based on matching include grid block matching and the feature matching. Grid block matching method firstly segments the region regularly, and then calculates the optical flow with corresponding regional displacement [11-14]. This method has been widely used in video coding, which is valid only for a little static background translation. While the feature matching method is insensitive for the static background translation, In order to enhance the rotate invariant of feature matching algorithm, the circular neighborhood invariant moments descriptors of feature point is used to match. With the matched feature points, the affine matrix of two images is estimated, which represents the global motion, moving target is wiped off as mismatch, and then is detected. The algorithm of moving target detection is indicated in figure 1.

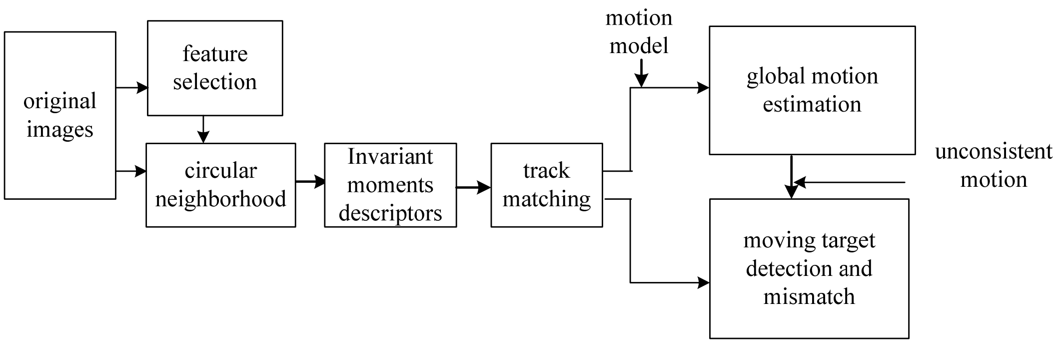


Figure 1. The moving object detection based on global motion estimation

III. INVARIANT MOMENTS MATCHING BASED ON FEATURE POINTS CIRCULAR NEIGHBORHOOD

For region matching, the smooth region is easy mismatch [15-16], so the feature point is detected firstly. For the light changing in adjacent frames, the neighborhood gradient direction is insensitive, it is computed as current pixel attribute. For the rotation in adjacent frames, the circular neighborhood invariant moments' descriptor of feature point is used to match.

a. Feature points extraction based on SURF

In order to reliably match region, the smooth region which is easy causing mismatch is not participate in the match [17-20], and the feature point is detected out. 2006 Bay proposes SURF feature, based on the Harr feature and integral image which is stable to match and greatly speeds up the running time of process. In order to detect the stable feature points of different scale, the pyramid image of different scale need be constructed, in SURF it is the approximation image of Hessian matrix. Because the feature points is scale independence, before construction of a Hessian matrix, the original image $I(x)$ is filtered with the Gauss function $g(\sigma)$, after that it is calculated with Hessian matrix $H(X,\sigma)$, the process is as following:

$$G_{xy}(\sigma)=\frac{\partial^2 g(\sigma)}{\partial x\partial y} \tag{1}$$

$$L_{xy}(X,\sigma)=G_{xy}(\sigma)\bullet I(X) \tag{2}$$

$$H(X,\sigma)=\begin{bmatrix} L_{xx}(X,\sigma) & L_{xy}(X,\sigma) \\ L_{xy}(X,\sigma) & L_{yy}(X,\sigma) \end{bmatrix} \tag{3}$$

$L(X,\sigma)$ is an expression of an image in different resolution, σ is Gauss variance. With this method every pixel can be discriminated, it can be expressed as:

$$\det(H)=L_{xx}L_{yy}-(0.9L_{xy})^2 \tag{4}$$

With non-maxima suppression, each pixel that is treated with Hessian matrix compare with the 26 points of its 3 dimensional neighborhood, which is retained if it is the maximum value or the minimum value of these 26 points, otherwise it is removed.

b. Invariant moments descriptors of feature point circular neighborhood

For the light changing slightly in adjacent frames, the gray and the gradient modulus of pixels is changed too, while the gradient direction keeps invariance in some degree. Because the transformation from the gradient direction ratio of horizontal direction and vertical direction to angle is nonlinear, to keep the property of gradient direction ratio, the ratio don't transform to angle, pixels are represented by the gradient direction ratio.

For the rotation in adjacent frames, in order to have the same neighborhood, the circular neighborhood of the feature point is used of 25*25 pixels size. the feature points is as the center point, square neighbor region of 25*25 pixels is as the template, and then the position weight of the pixels is 1 for whose distance to the center is less than or equal to 25 pixel, otherwise weight is 0. To have the same descriptors, here invariant moments are described in a circular neighborhood; it is introduced based on algebraic invariants. Among them, the Hu invariant moment based shape is the most widely used. Image moment function has been widely used in pattern recognition, target classification. Through the nonlinear combination of geometric moments, a group for the image translation, scale, and rotation invariant moments are derived. Gradient direction ratio image of the original image is $Ratio(x, y)$, for the feature neighborhood in the $Ratio(x, y)$, its geometric moments of $(p+q)$ order is defined as

$$m_{pq} = \int\int x^p y^q Ratio(x, y)dxdy \ (p, q = 0,1,\cdots\infty) \tag{5}$$

Zero moment m_{00} of objects expresses the "quality" of the image, the zero order moment with one order moments m_{10}、m_{01} are used to determine the image centric, if the origin of coordinate shifts to the centric, the center moment of image displacement invariant is got:

$$M_{pq} = \int\int (x-\bar{x})^p (y-\bar{y})^q Ratio(x, y)dxdy \tag{6}$$

The normalized center moment:

$$N_{pq} = \frac{M_{pq}}{(M_{00} + M_{00})^{\frac{p+q}{2}+1}} \ (p+q = 2,3,\cdots\infty) \tag{7}$$

From the two orders and three orders normalized central moments, the central moments of 7 area normalization with translation, rotation and scale invariance for the image are calculated. The 7 invariant moments for the region is expressed with T_1、$T_2\cdots$, and so on as following:

$$T_1 = N_{20} + N_{02} \qquad T_2 = (N_{20} - N_{02})^2 + 4N_{11}^2$$

$$T_3 = (N_{30} - 3N_{12})^2 + (3N_{21} - N_{03})^2 \quad T_4 = (N_{30} + N_{12})^2 + (N_{21} + N_{03})^2$$

$$T_5 = (N_{30} - 3N_{12})(N_{30} + N_{12})[(N_{30} + N_{12})^2 - 3(N_{21} + N_{03})^2] +$$

$$(3N_{21} - N_{03})(N_{21} + N_{03})[3(N_{30} + N_{12})^2 - (N_{21} + N_{03})^2]$$

$$T_6 = (N_{20} - N_{02})[(N_{30} + N_{12})^2 - (N_{21} + N_{03})^2] + 4N_{11}(N_{30} + N_{12})(N_{21} + N_{03})$$

$$T_7 = (3N_{21} - N_{03})(N_{30} + N_{12})[(N_{30} + N_{12})^2 - 3(N_{21} + N_{03})^2] +$$

$$(3N_{12} - N_{30})(N_{21} + N_{03})[3(N_{30} + N_{12})^2 - (N_{21} + N_{03})^2] \tag{8}$$

c. The search strategy of the invariant moments descriptor matching

The invariant moments matching is to find the nearest neighborhood in two adjacent frames, the most simple method is linear search, that is the full search algorithm (FS),which is also known as the exhaustive search method. It search all possible candidate block in whole region defined, finally found the corresponding candidate block with minimum distances of invariant moment, the block is the best matching block. The algorithm is simple, reliable, the matching block is global optimal. The only drawback of the FS is the highly time complexity, to speed up the searching time, the faster search algorithm is proposed, such as two-dimensional logarithmic search method, three step search method, diamond search method, etc.

Because the actual data generally distribute with cluster shape, the fastest matching algorithm is based on data index. The structure of index tree is in a tree shape, which divides the search space with levels, K-d tree is a typical method of the index tree. To speed up the searching of the matching region, nearest and nearer neighbors search with K-D tree is employed; its purpose is to search two data points whose distance is the shortest in the k-d tree. When the Euclidean distance ratios of the nearest and the nearer distance less than the threshold, the matching region is found.

Nearest neighbor search is a special case of k nearest neighbors, which is also the one nearest neighbor. The one neighbor expanding to two nearest neighbors is very easy. Here is the most simple k-d tree nearest neighbor search algorithm.

(1) Search binary tree, and follow the "path" to find approximate nearest neighbor very quickly, which is the leaf node in the same sub-space of the query point;

(2) Backtracking search path, and judge the other sub-nodes of searching path nodes might have more recent distance to the query point;

(3) If possible, need to jump to other nodes in the sub-space to search;

(4) Repeats this process until the search path is empty.

If matching block descriptor of reference frame is: $R_i = (r_{i1}, r_{i2}, \cdots r_{i7})$, matching block descriptor of current frame is: $s_i = (s_{i1}, s_{i2}, \cdots s_{i7})$, The Euclidean distance of them is:

$$d(R_i, S_i) = \sqrt{\sum_{j=1}^{7} (r_{ij} - s_{ij})^2} \tag{9}$$

When the Euclidean distance satisfied $d(R_i, S_j) < threshold1$ and $\dfrac{d(R_i, S_j)}{d(R_i, S_{j-1})} < threshold2$, namely the matching block of the current frame is found in the reference frame.

IV. GLOBAL MOTION ESTIMATION BASED ON MODEL

A difficulty of global motion estimation is that a pixel may not experience the global motion [21-23], which is exist in pure rotate. In the AUV navigation, all the pixels of the same image must have global motion, as the global motion is dominant compared with other local motion, theoryly we can employ the same motion model for the whole image, which can be solved with a robust estimation method.

a. The search strategy of the invariant moments descriptor matching

(1) The motion parameter model of translation

Translation model is the simplest motion parameters model in the image plane [24]. The model assumes that each image block move translation. Type (10) describes the transformation from the $k-1$ frame to the k frame. A target block B is selected from the $k-1$ frame image which is located in (x_{k-1}, y_{k-1}), in the k frame, whose center of the block B moves to position (x_k, y_k), and all the pixels follow the same transform relations as:

$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} = \begin{bmatrix} x_{k-1} \\ y_{k-1} \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \tag{10}$$

The two parameters Δx and Δy, respectively expresses the translational motion components along the x axis and y axis direction of the image.

(2) The affine motion parameter model

The two-dimensional spatial translational motion model can be extended to the affine motion model. Affine motion model has six parameters, which can not only describe the translation, rotation, and can process the deformed motion of block.

$$\begin{bmatrix} x_k \\ y_k \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ 1 \end{bmatrix} = A \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ 1 \end{bmatrix} \tag{11}$$

(3) Projecting motion parameter model

Projection motion parameter model has eight parameters, which can handle the deformation from rectangle to arbitrary quadrilateral, and can also handle rotation.

$$\begin{bmatrix} x_k \\ y_k \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & 1 \end{bmatrix} \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ 1 \end{bmatrix} \tag{12}$$

If only the translational motion is between background frames, then the translation model is accurate enough for motion estimation. When the static background moves with rotation changes between frames, there will produce large estimation error with the translation model. If the AUV is sailed in fixed depth level without tilt and roll, there exists translation and rotation transforms relationship between adjacent frames, it is affine transformation between two planes, so the model of affine motion parameters is used.

b. Robust estimation of motion parameter

Parameters solving for the motion model is to search for several groups of matching region which is satisfied the motion model in the two frame images, and then the parameters is solved with center matching points of the corresponding region. For the affine motion model has six parameters, at least three pairs of matching correctly region, and any three center points non-collinear.

When the motion parameter of static target is estimated, the correct match of moving target feature point will as mismatch. How to eliminate the correct feature point matching of moving target and the error matching? The robust parameters estimation method is studied how to obtain the optimal parameter estimation in the condition of observation data errors, M estimation, and least square median and random sample consensus (RANSAC) are common robust parameters estimation method.

In order to effectively handle the correct matching of moving target and the error matching, the robust estimation algorithm of linear estimation method and least median squares is put forward.

The basic idea of normalized linear method is to transform measurement data appropriately, so that the transformed measurement matrix has good condition number, and improve the stability of numerical calculation. Data normalized is mainly the displacement transformation and scaling transformation, the new origin of coordinates is $(\bar{x}, \bar{y})$, data scale ratio is d_x and d_y.

$$\bar{x} = \frac{1}{N}\sum_{j=1}^{N} x_j \quad , \quad \bar{y} = \frac{1}{N}\sum_{j=1}^{N} y_j \ , \quad d_x = \frac{1}{\max\{|x_j - \bar{x}|\}} \ , \quad d_y = \frac{1}{\max\{|y_j - \bar{y}|\}} \tag{13}$$

Let r_j is the residual error of the j data, the parameters estimate with least squares median method is to solve the nonlinear minimization as following:

$$\min Median\{r_j^2(x_j, p)\} \tag{14}$$

As r_j^2 is the regression variance, $Median(\cdot)$ is the median operation.

This method is very robust for data of non Gauss noise and error data, But this method don’t consider the "sick" of the measurement matrix and the nonlinear of measurement matrix elements and measurement data during the linear estimation with the sample, so normalization factor linear estimation would be applied to the sample, global motion estimation performance is improved with the samples processing.

Specific steps of the fusion robust estimation are as following:

(1) With the Monte Carlo method, m pairs of samples are randomly selected from the matching; each sample is composed of three pairs matching regions, because the affine matrix is determined with three pairs of non-collinear regional center;

(2) To avoid too close for random sampling, for matching data, it is blocked and random selected with the rules, for any sample, if three points are collinear, it is reselected until not collinear;

(3) For any 3 pairs’ non-collinear samples, the transformation matrix A_j is estimated with normalizing linear method;

(4) For the transformation matrix A_j, the regression variance is calculated for the matching point set, and the median is found $M_j = \underset{j=1,\cdots,n}{med}\, r_j^2$

(5) Repeat steps (3) m times, the minimum value $M_{\min}$ and the corresponding A_j is found in the m piece of median value;

(6) Then the transformation matrix A_j is the movement parameter matrix with the robust linear estimation.

For the random sample number m, assuming that the percentage of error matching is ε in the data set, then the probability of at least one completely correct matching sampling in m times sample is:

$$P = 1 - \left(1 - (1-\varepsilon)^3\right)^m \tag{15}$$

While $P = 0.99, \varepsilon = 0.2$ Then:

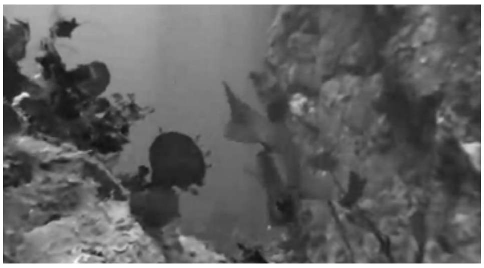
$$m = \frac{\ln(1-P)}{\ln\left(1-(1-\varepsilon)^3\right)} \tag{16}$$

We can guarantee a sample is completely correct when $m \approx 7$. But when the data noise distributes as Gauss, the efficiency of the minimum median method is very poor, in order to make up this shortcoming, we first calculate the robust standard variance with the type of $\hat{\sigma} = 1.4826(1 + 5/(n-3))\sqrt{Median}$, in which, n is all of the matching point pairs, the constant 1.4826 is the correction coefficient for the least square method reaching equally effective with the Gauss noise. For matching pairs, if $r_j^2 \le (2.5\hat{\sigma})^2$, it is the correct matching, the rest is of wrong matching, so most of the false matches can be successfully removed. The correct matching can estimate the global motion and exclude moving target, false matching.

V. EXPERIMENT RESULTS AND ANALYSIS

For the two adjacent frames, according to the algorithm proposed in this paper, the experiments of global motion estimation and the moving target detection is as following, the original image is shown in figure 2, f_{k-1} and f_k are respectively the $k-1$ and k frames in two video image sequences, the f_{k-1} is as the reference frame, the f_k is as the current frame.Figure1 (a) and 1(b) are for the first video image sequence, in which the fish is the moving target, the others are the static background. Figure 1 (c) and 1 (d) are for the second video image sequence.

(a) The f_{k-1} frame image



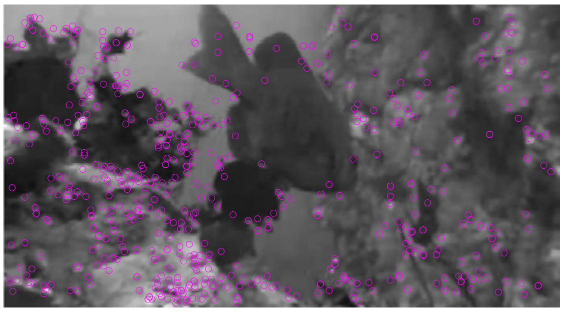
(b) The f_k frame image

(c) The f_{k-1} frame image

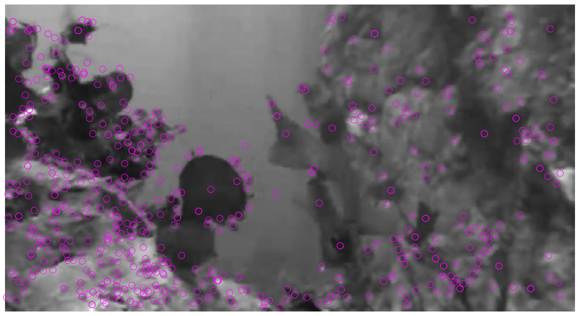
(d) The f_k frame image

Figure 2 Two pairs adjacent frame images

Feature point extraction of the above two pairs adjacent frame images is shown in figure 3. The details of first pair image are more abundant, more feature points are detected, which are shown in figure 3(a) and 3(b), while the second pair image has more smoothing region, less feature points are detected, which are shown in figure 3(c) and 3(d).



(a) The feature extraction of f_{k-1} frame image



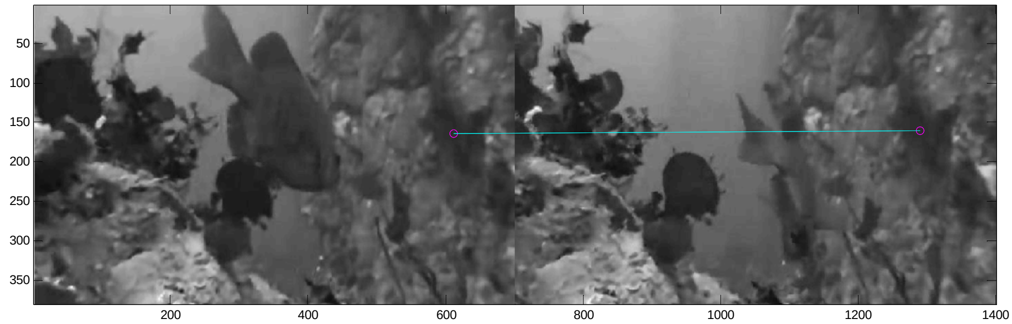
(b) The feature extraction of f_k frame image

(c) The feature extraction of f_{k-1} frame image

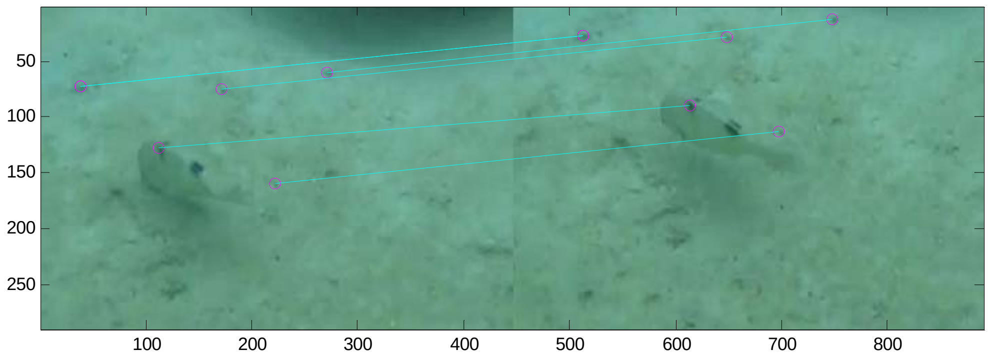
(d) The feature extraction of f_k frame image

Figure 3. The feature points of two consecutive images

As the lighting source mobility and the moving target motion, light irradiation angle and object reflection brightness would change; there would be some differences between corresponding gray regions of adjacent frames. Matching feature point pairs are shown in figure 4, one pair matching feature point of the first pair image is shown in figure 4(a), five pairs matching feature points of the second pair image are shown in figure 4(b).



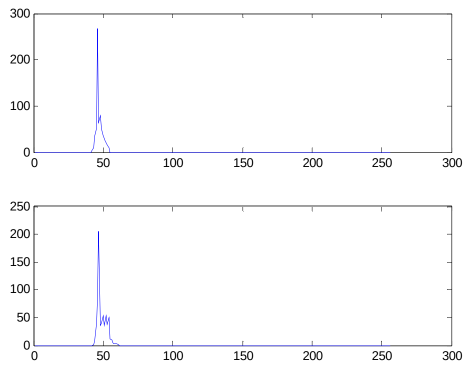
(a) One pair matching feature point



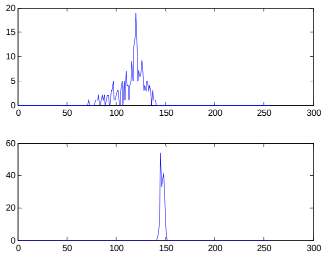
(b) Five pairs matching feature points

Figure 4. Pairs matching feature points of two adjacent frame images

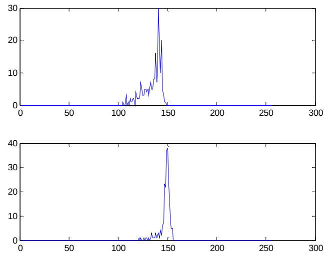
The histograms in the circular neighborhood of pairs matching feature points are shown in figure 5, the histograms of one pair matching feature point neighborhood are shown in figure 5(a), the histograms of five pairs matching feature point neighborhood are shown in figure 5(b), there is clearly difference not only in the range of gray, but also in the gray levels distribution shape.



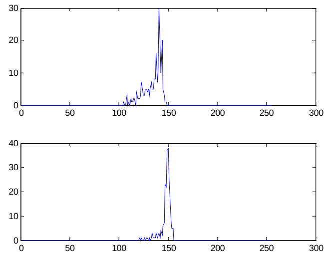
(a) The gray histograms of one pair matching feature point neighborhood



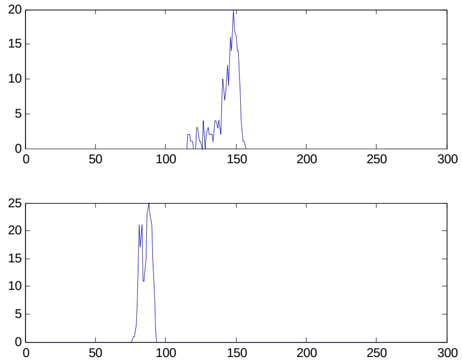
(b1)



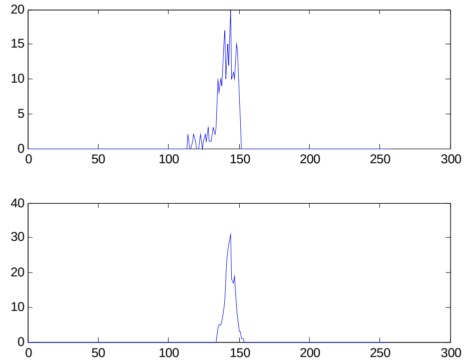
(b2)



(b3)



(b4)



(b5)

(b1) The first pair matching feature point neighborhood histograms. (b2) The second pair matching feature point neighborhood histograms. (b3) The third pair matching feature point neighborhood histograms. (b4) The fourth pair matching feature point neighborhood histograms. (b5) The fifth pair matching feature point neighborhood histograms

(b) Gray histograms of five pairs matching feature point neighborhood

Figure 5. Gray histograms of matching feature point circular neighborhood

For the change of gray level, it produced by the same light in the different position, whose differential would weaken the change to a certain degree. The matching feature point circular neighborhood is transformed with Haar wavelet, the gradient latitude histogram is shown in figure 6. The gradient latitude histograms of one pair matching neighborhood are shown in figure 6(a), there is small difference. The gradient latitude histograms of five pairs matching neighborhood are shown in figure 6(b), the difference of gradient latitude is smaller than gray levels, but there are still obvious differences, and the distribution shapes also have obvious difference.

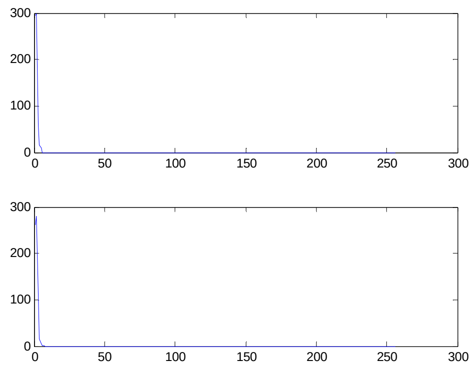
When the gray levels change, the gradient directions keep stable to a certain degree. The distribution histogram of Haar wavelet direction are respectively shown in figure 7, The gradient direction histograms of one pair matching neighborhood are shown in figure 7(a), the gradient direction histograms of five pairs matching neighborhood are shown in figure 7(b). The range and distribution shapes are of the similar in essential, except the latitude of maximum direction value is different, so maximum direction value is instable and eliminated. Then it is conducive to region matching as the illumination changing.

For the feature points neighborhood extracted, the invariant moment descriptor is calculated, and matched based on the K-D tree nearest search strategy, the global motion parameter matrix A_1 and A_2 are estimated according to the center feature point matching with fusion robust estimation method.

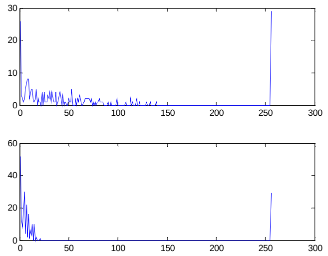
$$A_1 = \begin{bmatrix} 0.9106 & -0.0938 & 57.1758 \\ -0.0102 & 1.0013 & 7.3992 \\ 0.0000 & 0.0000 & 1 \end{bmatrix} \tag{17}$$

$$A_2 = \begin{bmatrix} 1.2941 & -0.0959 & -40.1622 \\ 0.0698 & 1.6005 & 37.2467 \\ 0.0000 & 0.0000 & 1 \end{bmatrix} \tag{18}$$

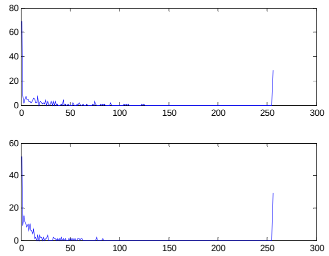
The matching points of unconformity with the A_1 and A_2 matrix include the error matching points and matching points of moving target, they are not easy to distinguish in a complex environment, and the life cycle of the moving target observed is only a few frames, it is of little significance if segmented to keep track for the AUV location of fast motion, which would be eliminated as positioning of the AUV, so as to ensure the accuracy of AUV positioning.



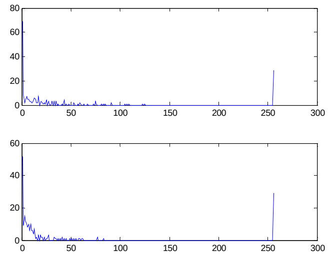
(a) The gradient latitude histograms of one pair matching neighborhood



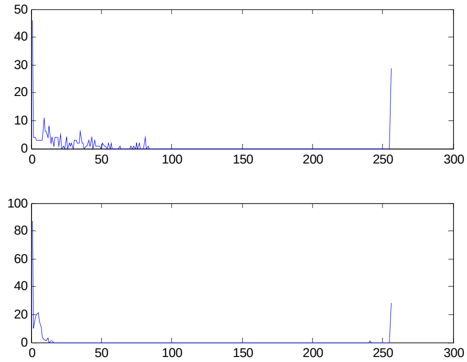
(b1)



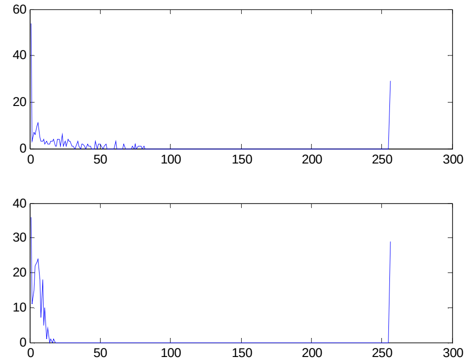
(b2)



(b3)



(b4)

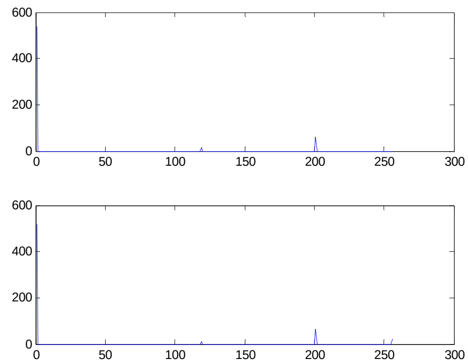


(b5)

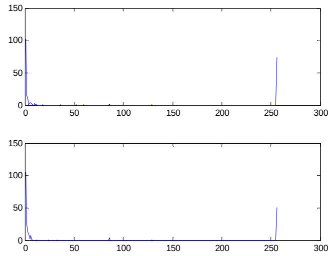
(b1) The gradient latitude histograms of first pair matching neighborhood; (b2) The gradient latitude histograms of second pair matching neighborhood; (b3) The gradient latitude histograms of third pair matching neighborhood; (b4) The gradient latitude histograms of fourth pair matching neighborhood; (b5) The gradient latitude histograms of fifth pair matching neighborhood

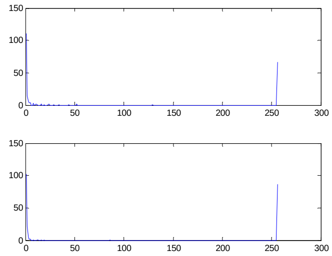
(b) The gradient latitude histograms of five pairs matching neighborhood

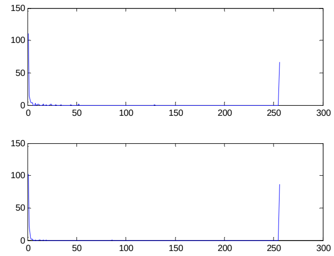
Figure 6. The gradient latitude histograms of circular neighborhood



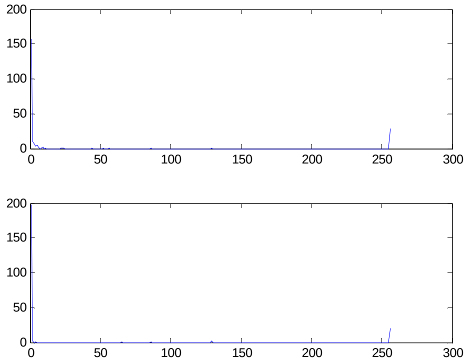
(a) The gradient direction histograms of one pair matching neighborhood

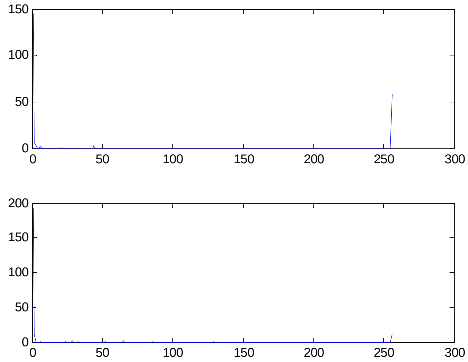






(b1) (b2) (b3)





(b4) (b5)

(b1) The gradient direction histograms of first pair matching neighborhood; (b2) The gradient direction histograms of second pair matching neighborhood; (b3) The gradient direction histograms of third pair matching neighborhood; (b4) The gradient direction histograms of fourth pair matching neighborhood; (b5) The gradient direction histograms of fifth pair matching neighborhood

(b) The gradient direction histograms of five pairs matching neighborhood

Figure 7. The gradient direction histogram of circular neighborhood

VI. CONCLUSIONS

According to the motion characteristics of mobile robot in unknown dynamic environment, the moving target detection is based on the global motion model estimation, and the global motion model estimation is based on gradient direction invariant moments. For the unknown target without prior knowledge, and the changes of imaging angle, the feature neighborhood extracted and the region descriptor employed not only can effectively express the environment with the affine invariance, but also can be reliably matching. For the moving object, its matching is as false matching, and the global motion parameters are estimated with the fuse robust parameter estimation method. Experiments show that this method is stable for motion target detection and AUV localization in dynamic environment.

REFERENCES

[1] M. Betke, E. Haritaoglu and L. S. Davis, "Real-time multiple vehicle detection and tracking from a moving vehicle", Machine Vision and Applications, December 2000, pp. 69-83.

[2] H. Tao, H. S. Sawhney and R. Kumar, "Object tracking with Bayesian estimation of dynamic layer representation", IEEE Trans. PAMI,Vol. 24,No. 1,2002, pp. 75-89.

[3] R. Pless, T. Brodsky and Y. Aloimonos, "Detecting independent motion: The statistics of temporal continuity", IEEE Trans. PAMI, Vol. 22, No. 8, 2000, pp. 768-773.

[4] W. Bell, P. Felzenszwalb and D. Huttenlocher, "Detection and Long Term Tracking of Moving Object in Aerial Video", http://www.cs.cornell.edu/vision/wbell/ identtracker, March 1999.

[5]J. Konrad, "Motion detection and estimation", Handbook of Image and Video Processing (A. Bovik, ed.), ch. 3.10, Academic Press, 2000, pp. 207-225.

[6] Yosi Keller and Amir Averbuch, "Fast motion estimation using bidirectional gradient methods", IEEE Transactions on Image Processing, Vol. 13, No. 8, August, 2004,pp.1042-1054.

[7] G. Sorwar, M. Murshed and L. Dooley, "A Fully Adaptive Distance-dependent Thresholding Search (FADTS) Algorithm for Performance-management Motion Estimation", IEEE Transactions on Circuits and Systems for Video Technology, Vol. 17, No. 4,2007, pp.429-440.

[8]R. T. Collins, A. J. Lipton and T. Kanada, et.al, "A System for Video Surveillance and Monitoring",VSAM Final Report, Technical report, CMU-RI-TR-00-12, Robotics Institute， Carnegie Mellon University, 2000.

[9]Ali Saad and Shah Mubarak,"COCOA-Tracking in Aerial Imagery", SPIE, Airborne Intelligence, Surveillance, Reconnaissance (ISR) Systems and Applications, 2006, pp.110-118, Orlando.

[10] K. Haris, S. N. Efstratiadis, N. Maglaveras and C. Pappas, “Image noise reduction based on local classification and iterated conditional modes”， Proc. IWISP, Manchester, U.K, November 1996.

[11] R. Guo and S. M. Pandit, “Automatic Threshold Selection based on Histogram Modes and a Discriminant Criterion”, Machine Vision and Application, No. 10, 1998, pp. 331-338.

[12] N. Otsu, “A Threshold Selection Method from Gray-level Histograms”, IEEE Transactions on Systems, Man and Cybernetics, Vol. 9,No. 1,1979, pp. 62-66.

[13] E. R. Dougherty and C. R. Giardina, “Mathematical Methods for Artificial Intelligence and Autonomous Systems”, Prentice-Hall, 1988, pp.319.

[14] C. Anderson, P. Burt and W. G. van der, “Change detection and tracking using Pyramid transformation techniques”, Proceedings of SPIE-Intelligent Robots and Computer Vision, Vol. 579,1985, pp.72-78.

[15] A. J. Lipton, H. Fujiyoshi and R. S. Patil, “Moving Target Classification and Tracking from Real-Time Video”, Proc. Fourth IEEE Workshop on Application and Computer Vision, 1998, pp.8-14.

[16] Y. H. Yang and M. D. Levine, “The background Primal sketch: An approach for tracking moving objects”, Machine Vision and Application, No. 5, 1992, pp.17-34.

[17] J. B. Kim and H. J. Kim, “Efficient region-based motion segmentation for video monitoring system”, Pattern Recognition Letters, Vol. 24, No.1, 2003, pp.113-128.

[18]J. Barron, D. Fleet and S. Beauchemin, “Performance of optical flow techniques”, International Journal of Computer Vision, Vol. 12, No.1, 1994, pp. 42-77.

[19]S. Brandt, “Maximum likelihood robust regression with known and unknown residual models”, Proc. of the ECCV2002, pp.97-102.

[20] A. Smolic and I. R.Ohm, “Robust global motion estimation using a simplified M-estimator approach”, Proceedings of the IEEE International Conference on Image Processing, 2000, pp.868-871.

[21]P. Torr and D. Murray, “The development and comparison of robust methods for estimating the fundamental matrix”, Int. Journal of Computer Vision, Vol. 24, No.3, 1997, pp. 271-300.

[22] Z. Y. Zhang, “Determining the epipolar geometry and its uncertainty: A review”, Int. Journal of Computer Vision, Vol. 27, No.2, 1998, pp. 161-195.

[23] Kim Yeon-Ho and A. C. Kak, “Error Analysis of Robust Optical Flow Estimation by Least Median of Squares Methods for the Varing Illumination Model”, IEEE Trans. PAMI, Vol. 28, No.9, 2006, pp. 1408-1435.

[24] R. Subbarao and P. Mee, “Beyond RANSAC: User Independent Robust Regression”, Computer Vision and Pattern Recognition Workshop, 2006, pp.101-105.