Proceedings of 2016 IEEE
International Conference on Mechatronics and Automation
August 7 - 10, Harbin, China

The Application of Image Mosaic in Information Collecting for an Amphibious Spherical Robot System

Shuxiang Guo[1,2], Shan Sun[1]

[1]*Tianjin Key Laboratory for Control Theory & Applications in Complicated Systems and Biomedical Robot Laboratory*
Tianjin University of Technology
Binshui Xidao Extension 391, Tianjin,300384,China
guoshuxiang@hotmail.com; suushan@163.com

Jian Guo[1*]

[2]*Intelligent Mechanical Systems Engineering Department*
Faculty of Engineering
Kagawa University
Takamatsu, Kagawa, Japan
*Corresponding Author: jianguo@tjut.edu.cn

***Abstract* - Amphibious spherical robots that achieve the correct navigation need to collect surround information and clear perception. In addition to equip various sensors, the robot collecting information can also rely on the video images, which include rich texture and information of color. In this paper, we present a video image mosaic method to collect surrounding information of the amphibious spherical robot. We extract key frame from video image based on K-means clustering, and control the number of key frame by changing threshold. Meanwhile, pre-processing key frame and extracting the image key points based on SIFT (Scale-invariant Feature Transform). Then, use RANSAC (Random Sample Consensus) to eliminate the mismatch, and choose suitable value of scaling parameter of SIFT to improve match accuracy. Finally, stitching these key frames up, we got a complete image. Experimental results indicated that image mosaic not only conclude key points of video but also reduce fuzzy phenomenon, enhance the perception of direction. Compared with the direct splicing of video frames, the computation is greatly reduced and still retains important information.**

Index Terms - Amphibious Spherical Robot; K-means clustering; SIFT; RANSAC; Video image mosaic

I. INTRODUCTION

As a tool for detecting ocean resources, amphibious robots, become more and more important. These robots could be used for various tasks, such as pipeline cleaning, data collection, inspection, construction and maintenance of underwater equipments [1]-[4].

No matter under the water or on the land, amphibious spherical robots could complete these tasks very well. In order to control accurately, amphibious robots collect surrounding information and clear perception that make them become extremely effective. Therefore, the data processing such as stitch image is necessary to observe the object over the wide rang [5]-[7]. We focus on the image mosaic, which broadens our view greatly. It has been widely used in various fields, such as remote sensing image processing, virtual reality technology, image reconstruction and other areas [8]-[12].

In the field of medicine, the field of view (FOV) would be restricted when it steers the gastro scope to screen. The narrow FOV increases the operation inefficiency and misdiagnosis rate. In this way, a broader FOV surveillance method which is based on global and local panoramic views is desired to improve the efficiency of gastro copy and guide endoscopes as they manipulate surgical instruments around anatomical structures [13].

UAV (Unmanned Aerial Vehicles) low-altitude remote sensing become improves popular. In order to expand the field of vision, better uniform processing, interpretation, analysis and study of UAV image information, we often need to stitch adjacent image to a panoramic image [14]. Such as in the field of agriculture, obtaining frequent aerial images of their farms by UAV allow the farmers to make informed decisions related to fertilizer application, irrigation, pesticide, application and other farm practices. The images are Geo-tagged and a mosaic is built to inform farm decisions [15].

In the file of virtual environment, 360-degree panoramas could provide an immersive experience for observers through simultaneous camera shots and reliable automated stitching [16]. Such as Google Street View let users explore places around the world through panoramic bubbles or strips [17]. And Microsoft Street Slide employs multi-perspective strip panoramas to generate a visual summary of continuous street sides [16].

In the file of military, such as air combat environment navigation, image mosaic is very reliable, safe, economical, and non-destructive, can be repeated many times, etc. Splicing can be just done between two images, which save time, thereby enhances the quality of real-time virtual navigation [18].

In order to improve the effect of image monitoring, image mosaic technology is introduced into remote video monitoring system. After the completion of mosaic, the whole status of the transformer and the switch can be observed. If they are abnormal, operation and maintenance personnel can easily determine their abnormal position, judge more accurately, and give a timely and effective treat on the defects. So the security and reliability of unattended substation's operation can be improved [19].

From the domestic and foreign research present situation, we can find that the research in the field of image mosaic has made significant breakthrough. In this section, we propose apply image mosaic to amphibious spherical robot information collected.

This paper consists of four parts. Section II introduces the video module of amphibious spherical robot. Then section III illustrates the method of video image processing. In section IV, land movement experiences are carried out; analysis of experimental results is given. Finally, we come to conclusions and bring forward future work in section V.

II. The Video Module of an Amphibious Spherical Robot

In order to adapt to the complex ocean environment, our amphibious spherical robot is designed not only to walk on land but also to move under the water [20]-[23]. It should also realize more degrees of freedom movement. The prototype of the robot is shown in Figure 1.

The video module of spherical amphibious robot consists of two parts. One part of it is video acquisition module. The Charge-coupled Device (CCD) HD camera we used with high definition, based on NTSC system. Its rated voltage is 12V, with the characteristics of waterproof, shockproof, high temperature resistant.

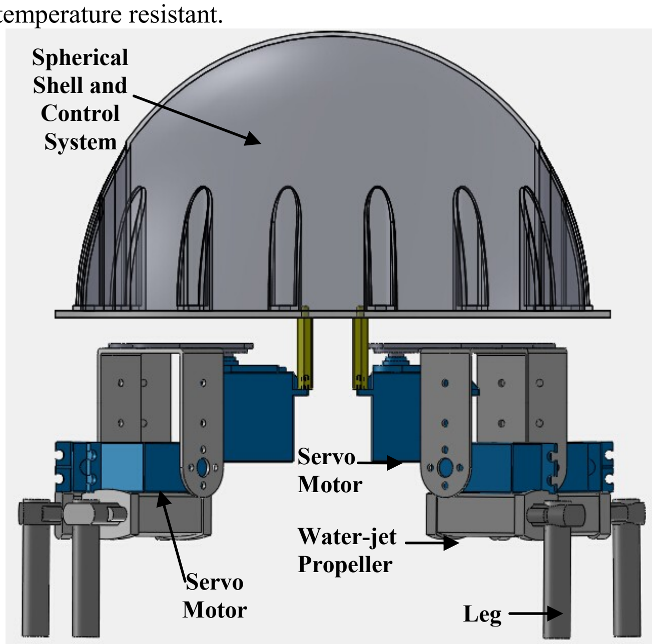


Fig.1 The 3-D model of the robot in Solidworks

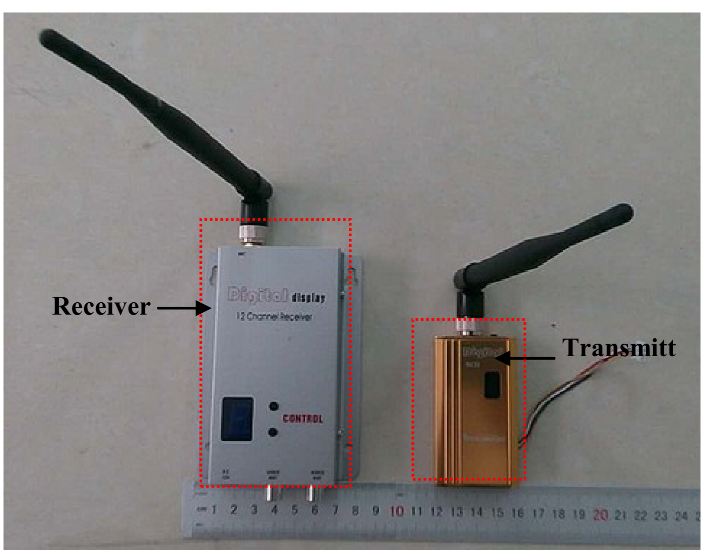


Fig.2 The wireless transmitter and the wireless receiver of video module

The other part of the video module contains a wireless transmitter and a wireless receiver, as shown in Figure 2. The modulation type of wireless transmitter is FM, its operating frequency is 1.1/1.2/1.3GHz and rated voltage is 12V. The wireless receiver's working frequency is 1.2/1.3GHz and working on 12V. The information transferred between the wireless transmitter and the wireless receiver through microwave transmission. The maximum transmission distance is 1000m in the open environment. When the spherical amphibious robot movement underwater the signal transfer would be attenuation, it is also proportional to the distance.

III. Video Images Processing

Figure 3 shows the main process of the method we proposed, which consists of three steps: video key frame extract, key frame pre-processing and image mosaic. (1) Video key frame extract. To reduce the computationally and adapt to the visual content, we extracted video image key frame based on K-means clustering. It needs to choose a suitable threshold value to control the number of key frames, avoid key frame excessive or insufficient. (2) Key frame pre-processing. In this paper, key frame processing based on image enhancement. It could highlight the desired features of the image, in order to reduce the noise and recognize the image characteristics. (3) Image mosaic. Extract feature points of key frame using the SIFT algorithm and get initial matching points from these key points by the method is that two key points distance is less than scaling parameter times the distance to the second closest match. Then use RANSAC to eliminating mismatch. Finally, realize image mosaic.

A. Video Key Frames Extract

The theory of video is according to the characteristic of "short stopping" of human eyes to show some pictures, make people product a feeling of movement. To remove redundant information of video, we present an approach to extract key frame based on K-means clustering.

Comparing with conventional general methods, such as based on shot boundary, based on motion analysis, and based on image information, etc., this shot clustering method could reflect the significant information of the video easily.

Extract a video all frames $f_1, f_2, \cdots$ and select each frame colour histogram in YUV colour space as visual features to gathered S frames divided into M clusters: $C_1, C_2, \cdots$.

First of all, set a threshold as TH . Make the first frame as the first cluster. According to the colour histogram difference between the first frame and one other frame calculate their similarity $X\left(C_i, C_j\right)$, which marks the clustering centre of mass of the distance. Make this similarity $X\left(C_i, C_j\right)$ compare with the threshold value TH , if X is bigger than TH , then fall it into the first clustering also adjust the centre of mass. On the contrary, become another new clustering $M = M + 1$. In the end, S frames in the video can be divided into M clusters.

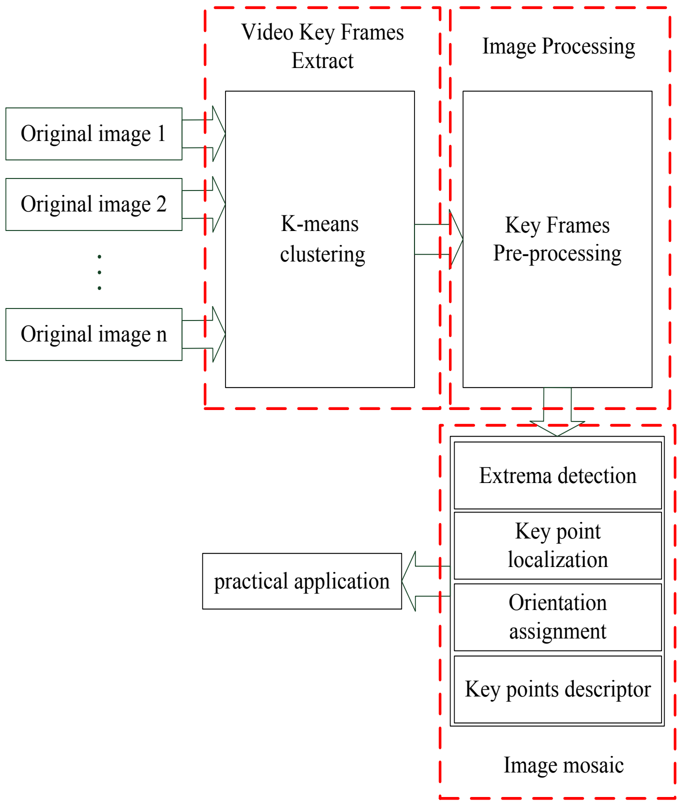


Fig.3 The main process of image musical

In the next step, we will choose the key frame from M clusters. Select the centre of the cluster frame as a key frame. The selection of threshold value influences the number of cluster. The smaller the threshold value, the less key frames. Setting appropriate threshold can be ensuring the speed of stitching under the premise of without losing important information.

B. Key Frames Pre-processing

In this paper, key frames processing based on image enhancement. This method could highlight the desired features of image, also could reduce the noise and recognize the image characteristics. Under the water, amphibious robot uses light cause the phenomenon of uneven illumination.

A physical model for underwater-imaging has been built, and it could be simplified as [24]:

$$Z(x,y) = I_\infty \rho(x,y) e^{-\beta d} + I_\infty \left(1 - e^{-\beta d}\right) \qquad (1)$$

Where I_∞ is light irradiation intensity, $\rho(x,y)$ is object radiance which be observed, β is attenuation coefficient, d is the distance between CCD and viewpoint, $Z(x,y)$ is the brightness of location (x,y).

Underwater image restoration actually to according (1) estimate $I_\infty \rho(x,y)$ which representative recovery image:

$$\lambda = I_\infty \left(1 - e^{-\beta d}\right) \qquad (2)$$

$$I_\infty \left(1 - e^{-\beta d}\right) = \left[Z(x,y) - \lambda\right] e^{\beta d} \qquad (3)$$

The clearness processing of underwater degraded image as follows:

(1) Transform original RGB image to YUV colour image.
(2) Recovery the signal Y through the wavelet transforms.
(3) Adjust the brightness of signal Y on the low frequency sub-band of wavelet.
(4) Combine with signal Y and chrominance signals U, V to form RGB colour image.

C. Image mosaic

1) The algorithm of SIFT

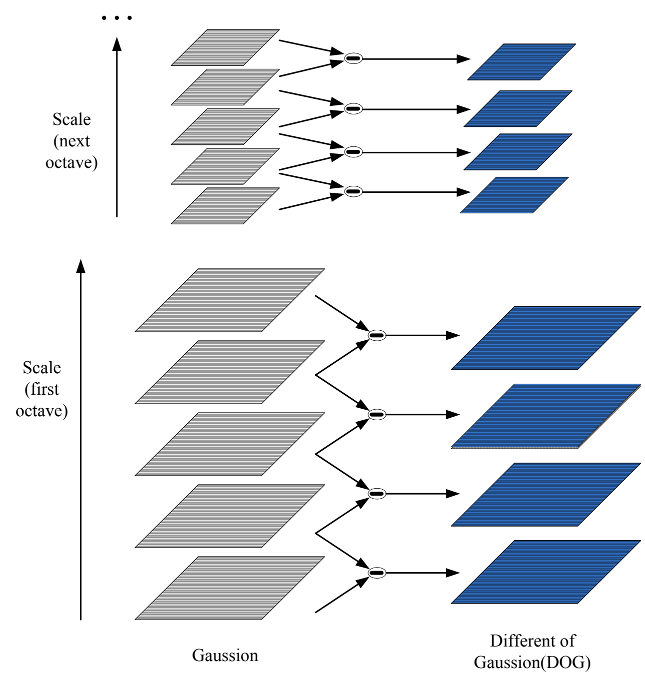


Fig.4 The product of difference-of-Gaussian

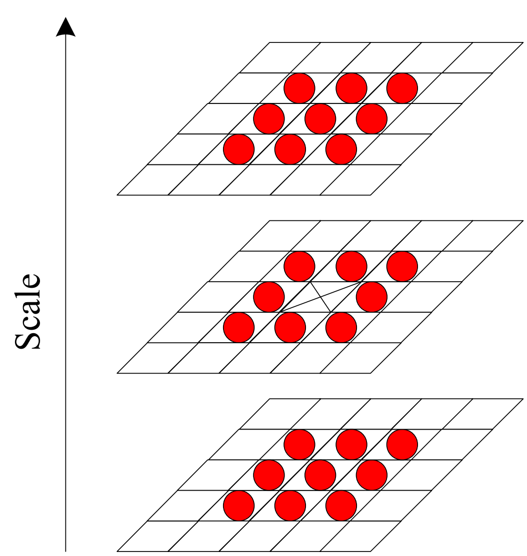


Fig.5 Detected the Maxi-ma and mini-ma by pixel

Fig.7 Extracted key frames

Image mosaic could reduce fuzzy phenomenon of video recording and enhance orientation perception. SIFT has advantages of scale scaling, rotating, radiation transform and unchanged brightness changes on image [25].

The algorithm of SIFT contain following four steps:

(i) Scale-space extrema detection

The first stage of computation searches over all scales and image locations [26]. It is implemented efficiently by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation. The process of product difference-of-Gaussian was shown in Figure 4. For each octave of scale space, the initial image is repeatedly convolved with Gaussians to produce the set of scale space images shown on the left. Adjacent Gaussian images are subtracted to produce the difference-of-Gaussian images on the right. After each octave, the Gaussian image is down-sampled by a factor of 2, and the process repeated.

Once the interest points are defined, the pixel of each sample point is compared to its eight neighbours in the current image and nine neighbours in the scale above and below, as shown in Figure 5. It is selected only if it is larger than all of these neighbours or smaller than all of them.

(ii) Key point localization

At each candidate location, a detailed model is fit to determine location and scale. Key points are selected based on measures of their stability.

(iii) Orientation assignment

One or more orientations are assigned to each key point location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations.

(iv) Key points descriptor

The local image gradients are measured at the selected scale in the region around each key point. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

2）The region based image stitching algorithm

The algorithm of the region based image stitching, a conventional general image mosaic method, according the pixel to select a template in an image. And finding the closest point or region in the other image, rely on the evaluation function.

The algorithm of region based image stitching contain following four steps:

(i) Delimit the image of template,

(ii) Set the search scope in the matching graph, and find the position with the maximum similarity with the template graph,

(iii) Seamless splicing of the images to be stitched according to the maximum similarity.

In the first step of the region-based image stitching algorithm, the select of the template have a significant impact on. The maximum time consumption is in the second step.

3）Comparison of two algorithms

The principle of the region based image stitching is extracting feature points from two adjacent images. If two images have amount of similar regions and the overlap is small, the quality of the mosaic image will be affected by the randomness of the template selection.

The algorithm of SIFT is one method based on feature matching. We firstly extracted the special feature of images and then marching them, so the stitching images can retain the feature of displacement, rotation, scale. In the process of mosaic, the data quantity of images could greatly compress. Registration image with the algorithm of SIFT, the quantity of calculation is become small and the speed of it is rising. Also have very good robustness. As long as there are three or more than three points to be matched, we can get the registration.

4）Feature match and eliminated mismatch

After extract the key points, the next step is matching them. A more effective measure is to obtained by comparing the distance of the closest neighbour to that of the second-closest neighbour. Then we often still need to identify correct subsets of matches, in this paper we used RANSAC to reduce mismatching. In the end, through the perspective to stitch image.

IV. EXPERIMENTS AND RESULTS

A. Extract Key Frames of Amphibious Robot on Land Video

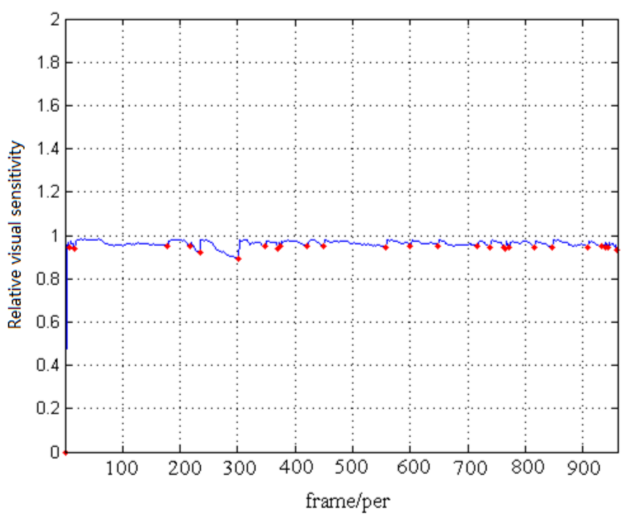


Fig.6 The location of key frame in whole video

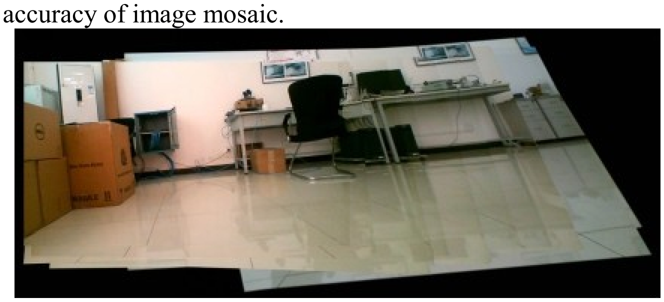
Under accessibility condition, the video module of amphibious spherical robot maximum trans-missive distance is 1000 m on the land. The link between working speed and frequency of amphibious spherical robot is different. When working frequency changes from 0 HZ to 1.25HZ, working speed increase with it. However, when working frequency exceed 1.25Hz, working speed decrease with it. In this case, we perform experiment under the 1.25HZ of working frequency. Different rode surface the amphibious spherical robot has different maximum speed. In this experiment, we drive robot on the tile floor which maximum speed is 7.6 cm/s.

Through the wireless transmitter on the amphibious spherical robot, we received the surrounding information of robot from wireless receiver. The length of video we collected is 39 s. Processed it by the method we refer to in section III.

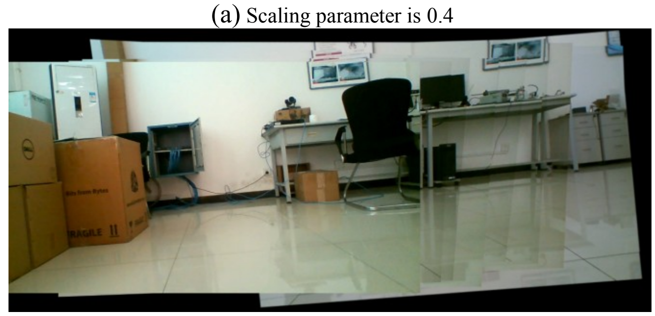
There are totally 983 frames in the video we collected. We extracted the number of key frame is 28, it cost 0.7 s. As shown in Figure 6, we can see the location of key frame in whole video. The number of key frame increase when threshold change bigger, it influence image mosaic quality, information integrity and extraction time. The key frames we extracted as shown in Figure 7.

B. Result Analysis of Image Mosaic

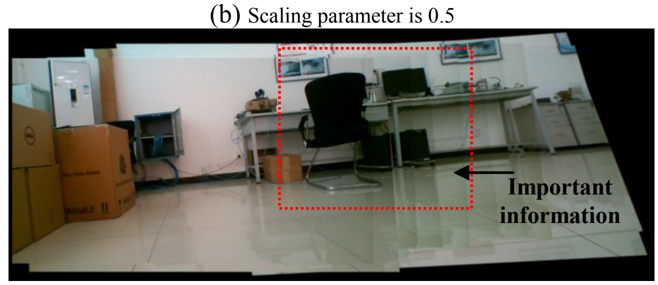
We stitched these key frames through the method we proposed in section Ⅲ, the result as shown in Figure 8. Because of the spherical robot's center of gravity are changing constantly with its four legs put up or down. In this case, the image we stitched would be existed the defect phenomenon of uneven on image edge. Figure 8 shows some deferent results, when we select deferent scaling parameter. The phenomenon of uneven edge could be relieved more when the scaling parameter is grown from 0.4 to 0.6, and the situation of image distortion becomes better. When scaling parameter change form 0.6 to 0.7, the quality of images are changing fewer. Under the situation of 0.7, the stitch time becomes longer. In this case, we set the scaling parameter as 0.6 to insure accuracy of image mosaic.



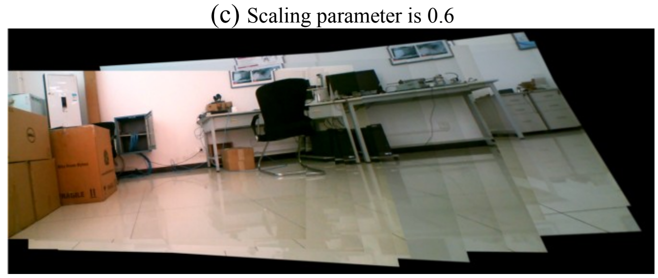
(a) Scaling parameter is 0.4



(b) Scaling parameter is 0.5



(c) Scaling parameter is 0.6



(d) Scaling parameter is 0.7

Fig.8 Results of the image mosaic

Experiments results show that the complete process of extracting key points from key frame, matching key points, eliminating mismatch and stitching key frames cost 17.73 s. Compared with the direct splicing of video frames, the computation is greatly reduced and still retain important information [27]. In the Figure 8, we can easily see that some important information which around the robot also retain. Image transmission is much faster than the transmission of video and easy to save.

V CONCLUSION

In this paper, we proposed a method to collect surrounding information with video module for amphibious spherical robot. This robot could collect surrounding information when it works in the narrow environment.

Image mosaic is used as a method of collect surrounding information of the amphibious spherical robot. Firstly, we use the method of K-Means clustering to extract the key frames from video image which we collected before. And then we processed these frames use the method of image enhancement. Based on the key frames extracted from the video, the features of key points are calculated. Furthermore, the methods of SIFT and RANSAC are used and then the image matching and stitching are completed. The simulation results show that amphibious spherical robot used image mosaic could get better quality image, more interactivity and also reduce the redundant information. Compared with the direct splicing of video frames, the computation is greatly reduced and still retains important information. Image mosaic improve the performance when amphibious spherical robot working in limitations environment.

ACKNOWLEDGMENTS

This research is partly supported by National Natural Science Foundation of China (61375094), Key Research Program of the Natural Science Foundation of Tianjin (13JCZDJC26200), and National High Tech. Research and Development Program of China (No.2015AA043202).

REFERENCES

[1] Chunfeng Yue, Shuxiang Guo, Maoxun Li, Yaxin Li, Hideyuki Hirata, Hidenori Ishihara, "Mechantronic System and Experiments of a Spherical Underwater Robot: SUR-II," Journal of Intelligent and Robotic Systems, vol.99, no.2, pp.325-340, 2015.

[2] Yaxin Li, Shuxiang Guo, Chunfeng Yue, "Preliminary Concept of a Novel Spherical Underwater Robot," International Journal of Mechatronics and Automation, vol.5, no.1, pp.11-21, 2015.

[3] Shaowu Pan, Liwei Shi, Shuxiang Guo, "A Kinect-based Real-time Compressive Tracking System for Amphibious Spherical Robots," Sensors, vol.15, no.4, pp.8232-3252, 2015.

[4] Lin Bi, Jian Guo, Shuxiang Guo，Zhendong Zhong, "Kinematic Analysis on Land of an Amphibious Spherical Robot System," Proceedings of International Conference on Mechatronics and Automation, pp.2082-2087, 2015.

[5] Zhendong Zhong, Jian Guo, Shuxiang Guo, Lin Bi, "Characteristic Analysis in Water for an Amphibious Spherical Robot," Proceedings of International Conference on Mechatronics and Automation, pp.2088-2093, 2015.

[6] Fu Wan, Shuxiang Guo, Xu Ma, Yuehui Ji and Yunliang Wang, "Characteristic Analysis on Land for an Amphibious Spherical Robot," Proceedings of International Conference on Mechatronics and Automation, pp.1945-1950, 2014.

[7] Lin Bi, Jian Guo, Shuxiang Guo, "Virtual Prototyping Technology-based Dynamics Analysis for an Amphibious Spherical Robot," Proceedings of the 2015 IEEE International Conference on Information and Automation, pp.2563-2568, 2015.

[8] Jian Guo, Guoqiang Wu, Shuxiang Guo, "Fuzzy PID Algorithm-based Motion Control for the Spherical Amphibious Robot," Proceedings of 2015 IEEE International Conference on Mechatronics and Automation, pp.1583-1588, 2015.

[9] Chunfeng Yue, Shuxiang Guo, Liwei Shi, "Design and Performance Evaluation of a Biomimetic Microrobot for the Father-son Underwater Intervention Robotic System," Microsystem Technologies, Vol 22, No. 4, 2015.

[10] A Ran, Zhang Xiulin, Liu Yu, "A fast cylindrical panoramic image mosaic algorithm," Computer development and application, vol.568, no.8 pp.4-7, 2013.

[11] Zhang Baolong, Li Hongrui, Li Dan, *et al.*, "Simulation of the image mosaic algorithm for vehicle mounted panoramic system," Journal of electronic and information, vol. 5, no.5, pp.1149-1153, 2015.

[12] Li Jia, Sheng Yehua, Zheng Liping, et al., "Panorama target detection technique based on image mosaic," Ordnance industry automation, vol.33, no.2, pp.7-10, 2014.

[13] R. Johnson *et al.*, "Hierarchical structure from motion optical flow algorithms to harvest three-dimensional features from two-dimensional neuro-endoscopic images," J. Clin. Nwurosci., vol.22, no.2, pp.378-382, 2015.

[14] Gu Qiang, "A Study on the Digital Mosaic Technique of UAV Aerial Photographs," Image Technology, no.1, 2009.

[15] Zhengqi Li, Volkan Isler, "Large Scale Image Mosaic Construction for Agricultural Applications," Proceedings of Robotics and Automation Letters, pp.2377-3766, 2016.

[16] J. Kopf, B. Chen, R. Szeliski, *et al*., "Street slide: browsing street level imagery," ACM Transactions on Graphics, vol.29, no.4, pp.96, 2010.

[17] L. Vincent, "Taking online maps down to street level," Computer, vol.40, no.12, pp.118-120, 2007.

[18] Jun Chen, Xiong Shi, Xiaofeng Feng, "The Research on Air Combat Environment Navigation Based on Spherical Panoramic Images," Proceedings of Information Engineering and Computer Science, pp.2156-7379, 2010.

[19] Zhao Zhenbing, Wang Rui, Zhang Tiefeng, "An Effective Method of Automatic Image Mosaic for Remote Video Monitoring System," Proceedings of Optoelectronics and Image Processing, pp.196-199, 2010.

[20] Yuan Han, Tang Jin, Wang Wenzhong, "A method for automatic measurement of large vehicle body length based on image," Computer and modernization, vol.5, no.5, pp.108-112, 2014.

[21] Xin Li, Jian Guo, Shuxiang Guo, "OFDM-based Micro-Signal Communication Method for the Spherical Amphibious Underwater Vehicle," Proceedings of International Conference on Mechatronics and Automation, pp.2094-2099, 2015.

[22] Chunfeng Yue, Shuxiang Guo and Liwei Shi, "Hydrodynamic Analysis of the Spherical Underwater Robot SUR-II," International Journal of Advanced Robotic Systems, vol.80, no.2, pp.1-12, 2013.

[23] Shaowu Pan, Liwei Shi, Shuxiang Guo, "A Kinect-based Real-time Compressive Tracking System for Amphibious Spherical Robots," Sensors, vol.15, no.4, pp.8232-3252, 2015.

[24] Erickson Nascimento, Mario Campos, Wagner Barros, "Stereo based structure recovery of underwater scenes from automatically restored images," Proceedings of SIB GRAPI 2009 Brazilian Symposium on Computer Graphics and Image Processing, pp.330-337, 2009.

[25] David G Lowe, "Distinctive Image Features from Scale-Invariant Key points," International Journal of Computer Vision, vol.60, no.2, pp.91-110, 2004.

[26] Lindeberg T, "Scale-space theory: A basic tool for analyzing structures at different scales," Journal of Applied Statistics, vol.21, no.1-2, pp.225-270, 1994.

[27] Zhigang Liu, Guiming Liu, Shilin Zhou, "Video Image Mosaic based on Coarse Matching," Proceedings of International Conference on Multimedia Technology, pp.3785, 2011.